University of Nottingham
UK | CHINA | MALAYSIA

Engineering ‘extracellular matrix factories’ to study how the extracellular microenvironment regulates gene expression.

Hannah Tomlin, MSc

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

List of contents

ABSTRACT

The extracellular matrix (ECM) is a specialised scaffold as well as a biochemical signalling platform for cells. The ECM plays a key role in gene regulation, with components such as tenascin-C (TN-C) being implicated in post-transcriptional regulation of inflammatory genes. However, the lack of physiologically relevant, healthy human ECM models has hindered the establishment of how specific ECM components regulate gene expression. Here, a human BJ skin fibroblast-derived ECM model was therefore developed combining cell-derived matrix (CDM) technologies with CRISPR/Cas9 genome editing. Assessment of fibroblast CDMs at the protein level by immunostaining and western blotting revealed the CDM contained collagen 1, collagen 3, fibronectin and glycosaminoglycans (GAGs). Microscopy techniques demonstrated a three-dimensional, 45 μm thick CDM. In addition, a protocol to effectively remove fibroblasts and DNA was optimised to obtain a non-immunogenic decellularised ECM substrate. FACs analysis confirmed the skin fibroblast CDM supported normal differentiation of primary human monocytes into macrophages and was a suitable culture substrate. Furthermore macrophage expression of LPS-induced miRNAs, including miR-155, was more stringently regulated when cultured on CDMs compared to plastic. Next, *TNC* KO fibroblasts were engineered via CRISPR/Cas9 gene editing to result in a fibroblast-derived ECM model that lacked TN-C. This was confirmed at the DNA, RNA and protein level. This allowed us to establish, for the first time in humans, that ECM-associated TN-C drives macrophage miR-155 expression in response to LPS. Furthermore, we identified that *TNC* expression is linked to normal cell architecture, and that a lack of *TNC* leads to actin cytoskeleton dysregulation, premature cell aging, subsequent cell cycle arrest and senescence in fibroblasts, as confirmed by light microscopy, β-galactosidase staining, F-actin staining and RNA-Seq. Candidate molecules specifically responsible for this effect have been identified but require validation. This model highlights an ECM-dependent regulatory mechanism of inflammation and could highlight novel targets for the treatment of diseases resulting from ECM alteration. Furthermore, this ECM model could be utilised as a biological tool, fully customisable and amenable to genetic modifications of other ECM proteins to allow the elucidation of how ECM components regulate gene expression and cell behaviour in health and disease, whilst also replacing the use of animals in this area of research.

ACKNOWLEDGEMENTS

The past four years have been a truly indescribable experience; challenging, disappointing, and satisfying with peaks and troughs of the three. I think a common misconception is that a PhD is a sole endeavour, however it would not have been possible without the many people along the way who have both contributed to and aided my journey of scientific and personal growth.

Firstly, I would like to express my gratitude to my supervisor Dr Anna M. Piccinini, for her limitless knowledge, guidance, support and fastidious experimental plans. I have developed into a thorough and methodical scientist over the course of my study and fully believe that I can now deal with almost anything, after tackling this PhD.

I would also like to thank Prof David M. Heery, my second supervisor, for his kind words of encouragement over the years and invaluable expertise, particularly with the CRISPR side of the project.

Not a day goes by that I do not miss the utter joy that Nicole brought to my life. Nicole has trained, supported, and entertained me from start to finish. There was no problem too large for us to solve together and Nicole is one of the most cherished friendships I have made during this PhD.

In a similar vein I would like to thank the other fantastic members of the Gene Regulation and RNA biology Lab. Choosing GRRB to work in was one of the easiest decisions I have made, due to the wonderful assortment of people with their individual charms and expertise. Kat, Poppy, Mitchell, Asta, Trudi and the rest of GRRB thankyou!

I would also like to say a special thanks to Dr Hilary Collins, whose reassuring, knowledgeable, approachable, and humorous nature I will not forget.

Infinite thanks goes to my twin sister, Bekii, who has been my partner in crime and a pillar of support from my earliest memories to now; even more so during the PhD. I would not have completed this PhD without her pushing me along. She has celebrated and commiserated alongside me throughout. I also cannot thank my parents enough for supporting and encouraging me in every way possible throughout the entirety of my life so far, who I think have done a wonderful job of nurturing me into the adult I am today.

Lastly, I would like show my gratitude for two people outside of the GRRB lab; Jamie, who has always been there for emotional support and equally importantly, to party. The gin to my tonic. I'd also like to thank Rebecca for her support – particularly in the last few months!

Writing little Haiku's is something I have gradually increased over the years and I'd like to finish with one that I think summarises my experience.

An uphill battle,
The precipice of success,
A rush of relief.

Tables

Figures

Abbreviations

Abbreviation	Full name
ADA	Adenosine deaminase
ADAM	A disintegrin and metalloproteinase
ADAMTS	A disintegrin and metalloproteinase with thrombospondin motifs
ADCY4	Adenylate cyclase 4
ALK-1	Activin-like kinase
AP3D1	Adaptor related protein complex 3
APC	Adenomatous polyposis coli
AZT	Azidothymidine
BAZ1A	Bromodomain adjacent to zinc finger domain 1a
BER	Base excision repair
BM	Basement membrane
BMP	Bone- morphogenic protein
C2ordf62	Ciliogenesis associated ttc17 interacting protein
CAMLG	Calcium modulating ligand
Cas9	CRISPR associated protein
Cas9 RNP	Cas9 ribonuclear protein
Cas9n	CRISPR associated protein nickase
CCLB1	Kynurenine-oxoglutarate transaminase 1
CD11b	Cluster of differentiation 11b
CD14	Cluster of differentiation 14
CD206	Cluster of differentiation 206
CD44	Cluster of differentiation 44
CD47	Cluster of differentiation 47
CD68	Cluster of differentiation 68
CDKi	Cyclin dependent kinase inhibitors
CDM	Cell- derived matrix
COL1	Collagen 1
COL2A1	Collagen type II
COL3	Collagen 3
CRISPR	Clustered regularly interspaced short palindromic repeats
CRISPRa	CRISPR activation
cr-RNA	CRISPR-derived RNA
DAMP	Damage-associated molecular pattern
DDR	Discoidin domain receptor
DG	Dystroglican
DSB	Double strand break
ECM	Extracellular matrix
EDA	Extra domain A of fibronectin
EGFL	Epidermal growth factor-like
EGFR	Epidermal growth factor receptor
EHS	Engelbreth-holm-swarm
ER	Endoplasmic reticulum
EVX1	Even-skipped homeobox1

FACITs	Fibril-associated collagens with interrupted triple helices
FAK	Focal adhesion kinase
FAK	Focal adhesion kinase
FBG	Fibronogen like globular domain
FN	Fibronectin
FNDC3B	Fibronectin type III domain containing 3B
FNDCA	Fibronectin type III domain containing 3A
FN-III	FN-type III repeats
GAG	Glycosaminoglycan
GM-CSF	Granulocyte-macrophage colony-stimulating factor
GMEM	Glioma-mesenchymal extracellular matrix antigen
GMIP	Gem interacting protein
gRNA	Guide RNA
HA	Hyaluronan
HDR	Homology directed repair
HGF	Hepatocyte growth factor
HSPG	Heparan sulfate proteoglycan
ICAM	Intercellular adhesion molecule
IFN	Interferon
IFN- B	Interferon beta
IFN-γ	Interferon gamma
IGF	Insulin-like growth factor
IL-1a	Interleukin 1 alpha
IL-10	Interleukin-10
IL-13	Interleukin-13
Il-1b	Interleukin 1 beta
IL-4	Interleukin-4
Il-6	Interleukin-6
Il-8	Interleukin-8
INDELS	Insertions or deletions
IPF	Idiopathic pulmonary fibrosis
KASH	Klarsicht, ANC-1 and Syne/Nesprin homology
KO	Knockout
LAP	Latency-associated peptide
LLC	Large latent complex
LOX	Lysyl oxidase
LPS	Lipopolysaccharide
LTBP	Latent TGF-β binding protein
M1	Classically activated macrophage
M2	Alternatively activated macrophage
MACITs	Membrane associated collagens with interrupted triple helices
MAGPs	Microfibril-associated glycoproteins
M-CSFR	M-CSF receptor
M-CSF	Macrophage colony-stimulating factor
miRNA	MicroRNA

MLK1	Megakaryoblastic leukemia-1
MMP	Matrix metalloproteinase
MPS	Mononuclear phagocytic system
MSC	Mesenchymal stem cell
MULTIPLEXIN	Multiple triple-helix domains and interruptions
MyD88	Myeloid differentiation primary response 88
NCOA4	Nuclear receptor coactivator 4
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NHEJ	Non-homologous end joining
OPN	Osteopontin
PAI-I	Plasminogen inhibitor-1
PAM	Protospacer adjacent motif
PAMPs	Pathogen-associated molecular patterns
PDCD4	Programmed cell death 4
PDGF	Platelet-derived growth factor
PDL	Population doubling level
PEG	Poly-ethylene-glycol
PGs	Proteoglycans
PHYKPL	5-phosphohydroxy-L-lysine phospho-lyase
PI3K	Phosphatidylinositol-3-kinase
PIGF	Phosphatidylinositol-glycan biosynthesis class F protein
PPP4C	Serine/threonine-protein phosphatase 4 catalytic subunit
pre-miRNA	PrecursormicroRNA
pri-miRNAs	Hairpin-containing primary transcripts
PRR	Pathogen recognition receptor
PRRX1	Paired-related homeobox 1
PTEN	Phosphate and tensin homolog
PTPN14	Tyrosine-protein phosphatase non-receptor type 14
Rb	Retinoblastoma
RISC	RNA-induced silencing complex
ROCK	Rho-associated protein kinase
ROS	Reactive oxygen species
RP5-1014d13	Serine/threonine-protein phosphatase 4 catalytic subunit
RTK	Receptor tyrosine kinase
Sca1/Ly6	Stem cell antigen-1
SHIP1	Src homology 2 (SH2) domain-containing inositol 5’-phosphatase 1
SIRPa	Signal regulatory protein-a
SKOR1	Ski family transcriptional corepressor 1
SLRPs	Small-leucine rich PG
SMC	Smooth muscle cell
SPARC	Secreted protein acidic and rich in cysteine
SUN	Sad1 and UNC-84
TA	Terminal assembly domain

TAZ	WW-domain-containing transcription regulator 1
TEX2	Testis expressed 2
TGF-β	Transforming growth factor beta
TIMPS	Tissue inhibitors of MMPs
TLR2	Toll-like receptor 2
TLR4	Toll-like receptor 4
TMEM204	Transmembrane protein 204
TMEM42	Transmembrane protein 42
TNC	Tenascin-C
TN-C+	Tenascin-C overexpression
TNF-α	Tumour necrosis factor alpha
TNR	Tenascin-R
TN-W	Tenascin-W
tPA	Tissue-type plasminogen activator
TRIF	TIR-domain-containing adapter-inducing interferon-β
TβRII	TGF-β receptor II
uPA	Urokinase type PA
VCAM	Vascular cell adhesion molecule
VEGF	Vascular endothelial growth factor
VEGFR	Vascular endothelial growth factor receptor
VWA1	Von Willebrand factor A domain containing 1
WDR90	WD repeat-containing protein 90
WT	Wildtype
YAP	Yes-associated protein 1

1 INTRODUCTION

1.1 THE EXTRACELLULAR MATRIX (ECM)

The extracellular matrix (ECM) is a highly complex, dynamic network of macromolecules that provides vital cues and support to cellular tenants. Every cell is in direct or indirect contact with the ECM and it is now widely recognized that, as well as providing physical structure and support, the ECM orchestrates most extracellular and intracellular events, modulating cell growth, proliferation, adhesion, migration, and cell fate with paramount roles in tissue morphogenesis and homeostasis [1]. However, whilst the importance of the interactions between ECM macromolecules and cells and their role in modulating gene expression are recognized, these remain largely unknown.

Over 1000 genes encoding ECM and ECM-associated proteins encompass the 'matrisome', the identification of which has been enabled by mass spectrometry-based proteomics [2]. The ECM has unique tissue-specific signatures, resulting from combinations of up to ~ 300 different proteins including glycoproteins, proteoglycans and fiber forming proteins such as collagens I, II, III, V, XI, XXIV and XXVII, elastins, fibronectin (FN) and laminins [3], capable of generating a variety of supramolecular structures. The complexity of these structures and macromolecules is further increased by processes such as alternative splicing, proteolytic cleavage, glycosylation and other posttranslational modifications, thereby increasing and diversifying the structures and functions of these ECM components.

1.2 STRUCTURE OF THE ECM

Whilst the exact structure of the ECM varies based on the function and localisation of the ECM, two major ECM formats exist; the interstitial matrix and the basement membrane (Figure 1). The interstitial matrix refers to the gel-like ECM that fills the intercellular spaces around and between cells whilst the basement membrane is a cell-associated, specialised 2D planar assembly of ECM proteins that functions to separate different tissue types.

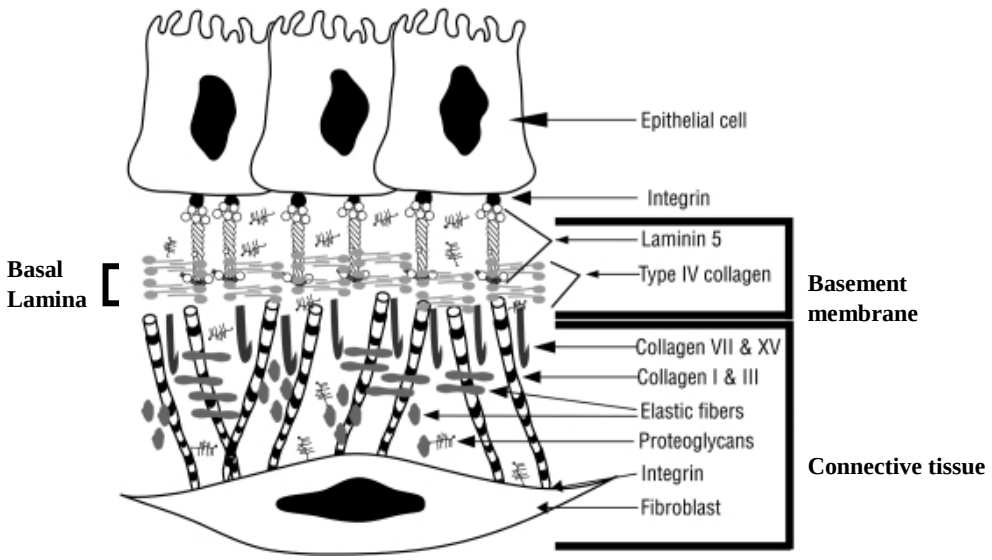


Figure 1. Schematic of structural organisation of the two main ECM structures.
The basement membrane and interstitial matrix (connective tissue) constitute the two main types of ECM, each composed of distinguishing macromolecules. The major components of the basement membrane are collagen type IV, nidogen, laminin and perlecan. The main components of the interstitial matrix include fibronectin and collagen type I. These have vital roles for anchoring and supporting cells. Adapted from Dunsmore *et al*. [4].

1.2.1 The interstitial matrix

The interstitial matrix surrounds cells, in a gel-like format consisting of fibrillar proteins including collagen type I, II, III, as well as non-fibrillar proteins such as fibronectin, proteoglycans and glycoproteins (Figure 2). This is very loosely organised compared to the basement membrane and its composition varies depending on the cellular origin, resulting in highly specialised tissue types whose components are normally excluded from the basement membrane. For example, bone sialoprotein is a phosphorylated glycoprotein that is a major non-collagenous protein in mineralised connective tissues such as bone, dentin (teeth) and calcified connective tissues that binds to hydroxyapatite (calcium and phosphate) [5]. Secreted protein acidic and rich in cysteine (SPARC) also binds hydroxyapatite and collagen type I resulting in the release of calcium ions to enhance collagen mineralisation in bone [6]. Furthermore, interstitial collagen and its fragments resulting from ECM remodelling are required to initiate bone turnover by osteoclasts; approximately 10% of the skeleton is remodelled each year to

replace damaged fractured bone or hyperminerlised bone [7]. Matrillins are disulphide-bonded proteins containing von-Willebrand domains involved in the formation of filamentous networks that are specific to cartilage and connective tissues [8]. Elastin, which together with fibrillin makes up the elastic fiber, is found in a variety of tissues that require flexibility and reversible deformity such as skin, arteries, lungs and the bladder [9]. Fibrillin, the main component of microfibrils, is also found in cardiovascular tissues. These examples highlight the key role of specific proteins in the differing interstitial connective tissue types and how these specific proteins provide the tissues with their mechanical and or biological properties.

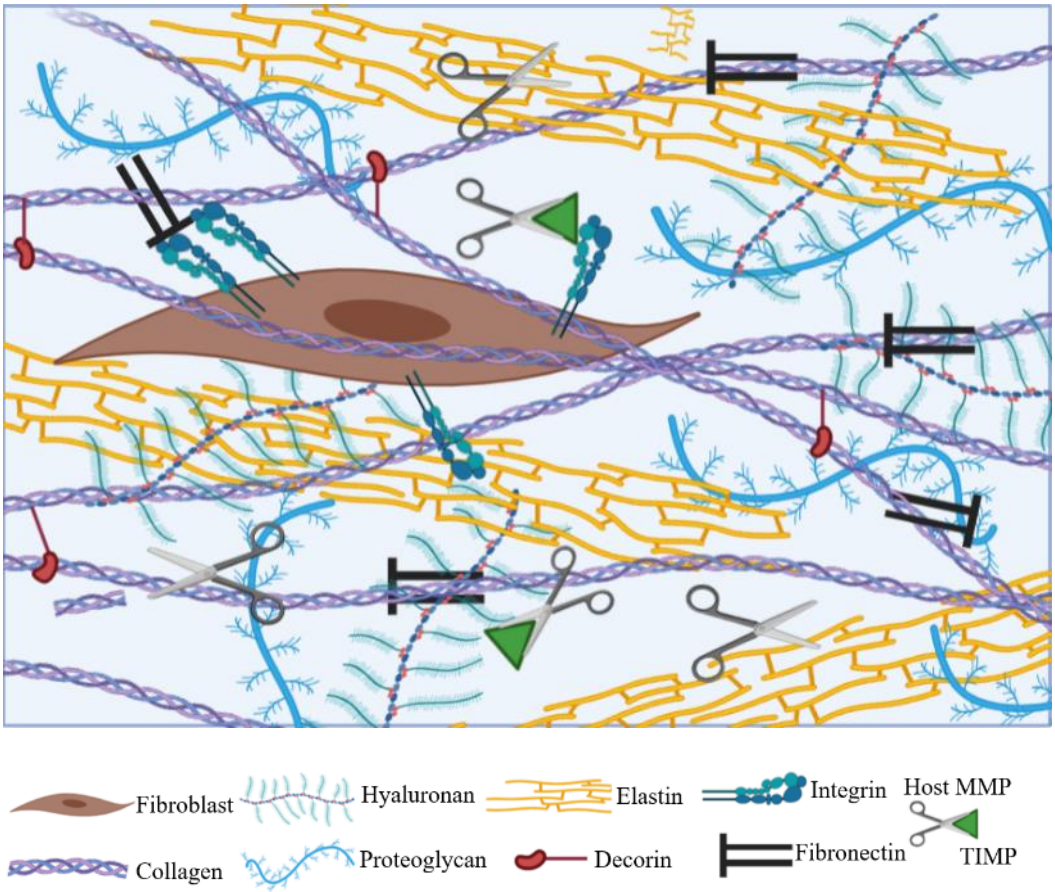


Figure 2. Schematic of the interstitial ECM.
The main components of the interstitial matrix that form a meshwork and gel consisting of collagen type I, elastin, fibronectin, hyaluronan, proteoglycans such as decorin and matricellular proteins such as tenascin-C (discussed later). This, in conjunction with the basement membrane, is vital for cell survival and function.

1.2.2 The basement membrane

The basement membrane (Figure 3) functions akin to flooring and walls to support endothelia, epithelia and fat cells, and as a sheath around muscle and nerve cells outside the central nervous system. The basement membrane separates epithelia and endothelia from the underlying stroma [10], forms the boundaries for organs and mediates specialized interactions in neuromuscular synapses and myotendinous junctions [11, 12]. The term basement membrane is often used interchangeably with 'basal lamina'. However, the basal lamina, together with the reticular lamina, forms the basement membrane (figure 3). Laminin self assembles near cell surfaces forming a provisional matrix [13] of which other molecules attach, such as non-fibrillar type IV collagen polymers attached to the heparan sulphate proteoglycans aggrecan and perlecan, and the sulphated glycoprotein nidogen [14]. In this respect nidogen is a structural adaptor protein linking laminin and collagen polymers to proteoglycans. Partial basement membranes can form in the absence of nidogens, collagen type IV, perlecan and agrin, but not in the absence of laminin [13]. The reticular lamina is formed of fibrillar collagens type VI, VII, which interacts with collagen IV in the basal lamina, XV and XVIII connecting the lamina to the underlying connective tissue to form the basement membrane [12]. The basement membrane arrangement is selectively permeable, but yet allows rapid diffusion of nutrients and metabolites, between the blood, tissue and cells [15]. The basement membrane facilitates selection by both size and charge (the latter property conferred by heparan sulphate side chains) in the kidney glomerular basement membrane [16]. Epithelial cells become anchored to the basement membrane via hemidesmosomes, which are specialised integrin-mediated binding regions, maintaining tissue integrity [17]. The basement membrane is therefore vital for both the structure and function of tissue.

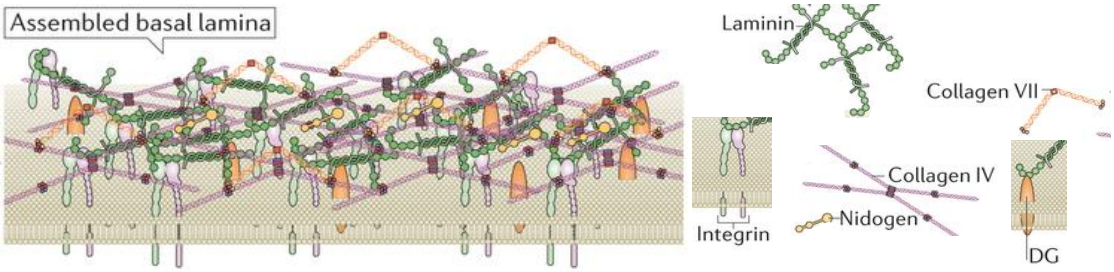


Figure 3. Schematic of the basal lamina composed of non fibrillar ECM components.
The basal lamina forms a thin sheet that is held in place by binding cell surface receptors, dystroglycan (DG) and integrins. Nidogen-dependent collagen IV-laminin binding forms a collagen IV meshwork covering laminin. Collagen VII further strengthens the basal lamina by linking it to the reticular lamina resulting in the formation of the basement membrane. Adapted from Mouw *et al*. [14].

1.3 ECM SYNTHESIS, MAINTENANCE AND REMODELLING

Cells residing in the ECM synthesize ECM proteins in the endoplasmic reticulum (ER) and secrete these components in a temporal and tissue-specific manner. Once properly folded and assembled proteins are transported from the rough ER to the Golgi apparatus where posttranslational modifications take place. Fibroblasts are primarily responsible for secreting ECM components in connective tissue, whilst osteoblasts are responsible for bone formation [18] and chondrocytes secrete cartilage components [19]. The basal lamina ECM is secreted by epithelial cells. The resident cells rely on their interactions with the existing matrix and other cells to direct and regulate their tissue synthesis and maintenance resulting from finely tuned signalling cascades. Cell-cell interactions occur via tight junctions, adherens junctions and desmosomes, which form structural links between cells, as well as gap junctions, which create a pore to facilitate transport between cells [20]. These contacts and interactions are important for synchronising tissue, for example calcium signalling in cardiomyocytes to coordinating beating is dependent on gap junctions. Cell-ECM interactions occur via integrins [21], discoidin domain receptors (DDRs) [22], cell surface proteoglycans (PGs) such as syndecans [23], and the receptor for many ECM proteins, cluster of differentiation 44 (CD44) [24]. For example, both $\alpha_1\beta_1$ and $\alpha_2\beta_1$ integrins bind type IV collagen [25] and laminin [26]. The β1 integrins are particularly important in the assembly of the basal lamina by interacting with the dystroglycan receptor to promote laminin polymerisation [27]. $\alpha_5\beta_1$ also binds fibronectin [28], promoting fibronectin matrix assembly

via stimulation of a Rho GTPase, RhoA [29]. The proper attachments to the substratum are essential for cell survival and the continued synthesis of ECM components to maintain tissue integrity.

Cells generate mechanical forces to organise their secreted ECM products. Both fibroblast and epithelial cells exert tension on the matrix via Rho-associated proteinkinase (ROCK) mediated actin-myosin contractility to result in orientation and alignment of collagen [30-32]. *In vitro*, collagen synthesising cells direct the orientation of collagen deposition [33, 34] and already deposited collagen [35]. Interestingly, fibroblast cell orientation has been demonstrated to be dependent on chemokine gradients towards which fibroblasts migrate, aligning towards growth factors [36, 37]. This only occurred when there was a steep concentration gradient and not a uniform concentration of the growth factor within the fibrin or fibronectin gel. The leading protrusion of the fibroblast itself was in the direction of the growth factor and Phosphoinositide 3-kinases (PI3K) signalling was localised in these protrusions [37]. Linked to this, the ECM itself functions as a growth factor, chemokine and cytokine repository, the effects of which are not fully understood. However, in this way the already deposited ECM must provide cues not only for fibroblast orientation, but subsequent ECM protein deposition. It has been logically speculated that the ECM repository may be to restrict or enhance cryptic ligand binding sites, control the concentrations of certain molecules (i.e. ECM components/fragments) or simply act as storage for appropriate and timely release of growth factors [38]. The best studied example of an ECM-sequestered growth factor is transforming growth factor beta (TGF-β). TGF-β is complexed with its latency-associated peptide (LAP) in the small latency complex (SLC) forming an inactive complex. LAP binds many ECM proteins including fibrillin and fibronectin forming large latent complexes (LLCs). Only upon degradation of the ECM is TGF-β released to activate inflammatory signalling cascades in a regulated manner [39].

Once released, TGF-β also initiates the synthesis of ECM molecules. In the canonical pathway, TGF-β binds to and activates TGF-β receptor II (TβRII), which then recruits and activates TβRI. Both receptors have intrinsic serine/threonine kinase activity. TBRII activates receptor (R) SMAD – 2 and -3 which form a trimer with SMAD-4 (co-SMAD) before translocation to the nucleus where their association with transcription factors forms a regulatory transcription complex [40]. Following target gene

transcription, SMAD complexes are ubiquitinated and targeted for proteasomal degradation [41]. A cDNA microarray approach in human dermal fibroblasts found 90 target genes, including fibrillar and non-fibrillar collagens, laminin, nidogen, fibronectin, decorin, integrins, membrane-bound matrix metalloproteinases (MMPs) and tissue inhibitors of MMPs (TIMPs), amongst others [42]. TGF-β is also able to inhibit the expression of MMP-1 (collagenase) by other inflammatory cytokines such as IL-1β, via SMAD-3 and SMAD-4, thus promoting ECM synthesis [43]. In this way, TGF-β is able to drive fibrosis; ECM accumulation, in particular type I collagen, is a hallmark of fibrosis. In the non-canonical pathway, TGF-β is able to activate the mitogen-activated protein kinase (MAPK) pathway, including ERK, JNK and p38 as well as PI3K/Akt signalling [44, 45]. TGF-β is also able to activate another group of R-SMADs (SMAD-1, -5 and -8) via binding activin-like kinase (ALK-1) [46]. This pathway was previously thought to be mainly activated by bone-morphogenic protein (BMP) – 9 which activates non-canonical SMAD-1/-5/-8, as well as canonical SMAD-2/3 which results in the upregulation of collagen 1 and fibronectin in fibroblasts [47]. Dysregulated TGF-β and BMP signalling results in a variety of bone formation disorders. As well as the formation of bone, TGF-β signalling is crucial for chondrogenesis (the formation of cartilage) by chondrocytes. TGF-β is able to stabilise and phosphorylate Sox9 protein in a SMAD-2/-3 dependent manner. P38 is also able to regulate Sox9 stability and phosphorylation independently of SMAD-2/-3 [48], indicating cross talk between canonical and non-canonical pathways of ECM synthesis.

As the ECM provides structural and biochemical support resulting in tissue homeostasis and normal functioning, fine regulatory systems exist to control the maintenance of the ECM. The ECM is constantly remodelled based on the secretion of matrix from fibroblasts and processing by intra- and extracellular enzymes, also secreted by fibroblasts. Tissue homeostasis is maintained by a variety of molecules; two main families responsible for the degradation and turnover of ECM components, are MMPs and a disintegrin and metalloproteinase with thrombospondin motifs (ADAMTS). This degradation activity is counterbalanced by coordinated expression of tissue inhibitors of matrix metalloproteinases (TIMPs), which can also inhibit a disintegrin and metalloproteinases (ADAMs) [49]. 23 MMPs have been identified in vertebrates that, due to their destructive nature, are synthesised as

inactive zymogens containing a self-inhibitory pro-domain encoding a peptide approximately 80 amino acids in length. This peptide binds to the 170 amino acid long catalytic domain preventing activation [50]. They can be activated by cleavage of this pro-domain by proteases, reactive oxygen species or other MMPs. For example, plasmin can activate the majority of MMPs, and MMP-3 can activate MMP-1, more commonly known as collagenase [51]. Most MMPs are secreted, however some are membrane bound. ADAMs and ADAMTs [52] contain MMP like domains which enable proteolytic activity. Degradation by mechanical stress can also occur, and this has been shown to induce MMP expression in chondrocytes [53] and tenocytes [54]. Whilst the exact molecular mechanism for every remodelling outcome is not completely understood, it is well established that mechanical stimuli are necessary for directing and controlling tissue remodelling and fate, via mechanosensing by cells [55, 56].

Other proteins involved in degradation of extracellular proteins are serine proteases such as the serine proteases, plasmin and cathepsin G, whereas the cysteine, aspartate, and threonine proteases mainly digest intracellular proteins due to pH-dependent activity; neutral pH outside the cells and acidic pH inside the cells [57]. Plasmin is involved in a variety of biological processes and, like MMPs, is secreted as inactive zymogen, known as plasminogen. Its conversion to plasmin is controlled by the activators tissue-type plasminogen activator (tPA) and urokinase type PA (uPA) and their inhibitors (PAI 1-2). tPA-activated plasmin cleaves basement membrane laminin-5 [58] and binding of tPA and plasminogen to laminin-5 results in an enhancement of further plasmin activation [59]. *In vitro*, plasminogen has been shown to be activated by fibroblast-secreted urokinase plasminogen activator and results in fibronectin proteolysis. TGF-β was shown to protect against this proteolysis which resulted from the upregulation of PAI-1 [60]. Plasmin has been shown to activate latent TGF-β, suggesting a negative feedback loop to control its destructive activities [61]. It could be that *in vitro* ECM degradation is limited by the release of stored latent TGF-β.

As well as being subject to cleavage extracellularly, a number of ECM proteins have been shown to be degraded intracellularly in lysosomes, including defective fibril forming collagen I [62], native fibronectin [64] and conformationally-altered vitronectin [63, 64]. These processes are regulated by ECM receptors, including integrins such as alpha v beta 5 regulating endocytosis of vitronectin [63]

and Endo180, an endocytic transmembrane glycoprotein and member of the mannose receptor family regulating native and denatured collagen uptake [62]. In this way accumulation and deposition of native and faulty proteins into the ECM is avoided, emphasising another level of ECM homeostasis regulation. Proteoglycans can also be degraded intracellularly, however this is only facilitated after extracellular cleavage of core proteins by MMPs, ADAMTs and other proteases that allows smaller chains to be endocytosed [65].

As well as degradation and turnover of matrix components, remodelling occurs by modulation of existing ECM proteins. Many posttranslational modifications occur to ECM proteins. Collagen cross-linking is required to stabilise and orient collagen fibrils and provide the elasticity and stiffness to tissues. Lysyl oxidase (LOX) covalently crosslinks collagen fibers. Increased LOX protein expression correlates with increased collagen deposition [66], leading to enhanced cross-linking and the modulation of stiffness that exerts profound changes on cell behaviour [67]. Such stiffness can also lead to fibrosis, thus leading to and propagating pathological changes.

Together, these events create a cell-specific context that differentially modulates the resident cells behaviour resulting in normal tissue homeostasis [68]. The context in which the cell resides is therefore rigidly defined by the structural matrix components. This is illustrated in cartilage, which is enriched in fibril forming, and therefore load bearing, collagens and contrasts blood vessels, which are composed of smooth muscle cells and enriched with elastin, which gives the blood vessels their necessary stretch and flexibility in order to pump blood around the body [10].

1.4 ECM COMPONENTS

1.4.1 Proteoglycans

Proteoglycans are one of the most important ECM components and consist of a core protein covalently linked to one or more glycosaminoglycan (GAG) chains. GAGs are characterised by long, negatively charged repeating disaccharide units composed of an amino sugar (glucosamine that is N-acetylated or N-sulphated or *N*-acetylgalactosamine) and a uronic acid (glucuronic acid or iduronic acid) or galactose

[15] with the exception of hyaluronic acid (HA) which is not noncovalently linked. There are several types of proteoglycans: dermatan sulphate, chondroitin sulfate, keratan sulphate, heparan sulfate, heparin and hyaluronan. With the exception of HA, GAGs are sulphated at various groups; heparin a highly sulphated variant of heparan sulphate [69]. Proteoglycans interact with a variety of ECM proteins (discussed below) through either their core protein or GAG side chains, resulting in roles in the structural organisation of the ECM as well as cell proliferation, adhesion, migration, differentiation and apoptosis [69]. Proteoglycans give the ECM its hydrated, gel-like format due to their carboxyl and sulphate groups which create a net negative charge attracting Na^+, drawing water in osmotically [38]. The specific structure of the GAGs has been shown to confer their biological functions [70]. For example, aggrecan is the major proteoglycan in cartilage and mediates chondrocyte-chondrocyte and chondrocyte-matrix interactions by its ability to aggregate by binding HA via a link protein [71]. Water is drawn into cartilage osmotically, as explained previously, creating a cushioned cartilage environment that allows frictionless motion. The large mass of the proteoglycan aggregates means it's movement is restricted in the aggrecan and collagen cartilage matrix that confers strength and load bearing properties to the tissue [71]. The numerous possible arrangements of core proteins, types and numbers of GAG chains and O- and N-linked oligosaccharides results in a variety of proteoglycans. However, they can be grouped based on their localisation as extracellular, pericellular and cell-surface associated.

1.4.1.1 Extracellular proteoglycans

Secreted, extracellular proteoglycans involve large aggregating proteoglycans and small-leucine rich proteoglycans (SLRPs), including versican, aggrecan, neurocan and brevican, which share common structural features and can all bind to HA via their N-terminal domain. Versican is produced by stromal cells throughout the body and can interact with fibronectin [72] and cell receptors such as CD44 and toll-like receptor 2 (TLR2) on bone marrow derived macrophages resulting in the expression of pro-inflammatory cytokines like tumour necrosis factor alpha (TNF-α) and nuclear factor kappa-light-chain-enhancer of activated β cells (NF-Kβ) [73]. Versican and HA form a scaffold for the recruitment of leukocytes and macrophages in a pro-inflammatory manner [74, 75]. However, whilst stromal cell-produced versican tends to be pro-inflammatory, versican produced by myeloid cells exerts both pro

and anti-inflammatory effects by promoting the production of interferon beta (IFN-β) and interleukin 10 (IL-10) [76].

Aggrecan is another extracellular proteoglycan and the major proteoglycan constituent in cartilage. Aggrecan interacts with tenascin-C (TN-C) [77], tenascin-R (TN-R) [78] and fibulins 1 and 2 [77]. Brevican is located in the central nervous system with neurocan and binds to fragments of fibronectin [79]. SLRPs are ubiquitously expressed and include: decorin, biglycan, osteoadherin, lumican, fibromodulin and mimican (osteoglycin), all characterised by a protein core with leucine-repeats and a minimum of one GAG chain [80]. SLRPs bind to various collagens and regulate collagen fibrillogenesis [81], and are therefore paramount for assembly of the ECM. Decorin is one of the best characterised SLRPs and was so named for its ability to decorate collagen [82] before its role in collagen fiber formation and fibrillogenesis was elucidated [83, 84]. Decorin is also capable of sequestering growth factors including TGF-β [85] and platelet derived growth factor (PDGF) [86]. Decorin directly antagonizes the epidermal growth factor receptor (EGFR) [87], vascular endothelial growth factor receptor 2 (VEGFR2) [88] and induced TIMP-3 [89]. Biglycan is highly expressed by macrophages acting as a pro-inflammatory activator of the innate immune response receptors TLR2/4 [90]. *In vivo,* biglycan is also able to bind and sequester active TGF-β, but with less protective effects than sequestering by Decorin [91]. This may be explained by the different localisations of these proteoglycans; decorin is associated with collagen within the interstitial matrix whilst biglycan is closer to TGFRI/II at the cell surface and therefore may present TGF-β to its receptors via binding to it [92].

1.4.1.2 Pericellular basement membrane proteoglycans.

These proteglycans associate closely with cells and interact with many ECM proteins. These include perlecan, agrin and some non-fibrillar collagen types XV and XVIII that characteristically have heparan sulfate (HS) side chains. Growth factors such a fibroblast growth factor (FGF) and vascular endothelial growth factor (VEGF) become tethered to the basal lamina by areas of negative charge curtesy of the sulphated proteoglycans heparin and heperan sulfate [93, 94], thereby establishing and controlling growth factor gradients and availability. Perlecan links laminin and collagen IV in the basal lamina.

Perlecan and agrin link nidogen and laminin to integrins, dystroglycan and sulphated glycolipids thereby establishing connections between basement membrane proteins and cell surfaces [95, 96].

1.4.1.3 Cell surface associated PGs

Cell surface PGs are divided into syndecans and glypicans. Syndecans (syndecan-1-4) are transmembrane proteins that comprise the main family of heparan sulfate proteoglycans (HSPGs). Through their numerous side chains, syndecans are able to interact with a huge number of soluble and insoluble ECM proteins. As such, syndecans are considered non-integrin receptors for ECM molecules [97]. Many different cell types synthesise syndecans, with different GAG structures which lead to different affinities for collagens, laminins and fibronectin [98]. On fibronectin gels syndecan-4 is able to interact with the a5β1 integrin (integrins are discussed later) to regulate cell adhesion and migration via Rho-GTPase signalling [99]. Linked to this, syndecan-4 co-regulates fibroblast morphology and matrix contraction, in conjunction with tenascin-C during tissue repair [100].

Unlike syndecans, glypicans (of which there are 6 in mammals) are not transmembrane proteins but anchored to the cell surface through a lipid anchor, with numerous HS chains protruding into the extracellular space. As such, these proteoglycans are found in cell membrane regions rich in lipids or cholesterol [101]. Glypicans are differentially expressed during development and i*n vivo* studies indicate that glypicans primary function is to regulate Wnt, Hedgehog, BMP and FGF signalling [102, 103].

1.4.2 Collagens

Collagens are secreted and aligned by fibroblasts and account for 30% of total protein mass in the body. Collagens constitute the main structural element of the ECM, providing mechanical properties, tensile strength, regulating cell polarity and adhesion, supporting chemotaxis and migration, and direct tissue development [38]. The characteristic feature of collagens is a polyproline-II, left-handed triple helical structure composed of three α-chains that can form the majority of their structure to less than 10%, depending on the collagen type [104]. Each α chain has a repeating Gly-X-Y triplet with the X and Y

positions usually occupied by proline and 4-hydroxyproline, respectively, and glycyl residues occupying every third position [105]. This repeating triplet is shared in all collagens except for some disruptions in non-fibrillar collagens such as collagen IV [106]. The integrin binding motif RGD also occur in the triple helical domain of some collagens (i.e. collagen I), allowing binding to cells. The α-chains also contain a trimerization domain that is essential to align the chains together in the triple helical structure [107]. Tight packing of the helices is facilitated by the hydrogen atom on the side chain on each glycine residue within the α chains, with the hydrogen residues in the helix interior and the proline rings on the exterior. Procollagens have non-collagenous, non-triple helical propeptide domains at both the C- and N-terminus that are essential for the formation of fibrils [107].

In vertebrates there are 46 distinct collagen polypeptide α chains which can assemble to form 28 collagen types. The diversity is further increased by different molecular isoforms of the same collagen type (IV/VI) and hybrid collagen types composed of different α-chains [108]. Collagens can be homotrimeric (three identical α-chains) or heterotrimeric (one α-chain that differs) and this has implications in disease and tissue remodelling. For example, in healthy adult tissues type I collagen is a heterotrimer of two α1(I) and one α2(I) chains ($\alpha1_2\alpha2$) and this form is susceptible to MMP cleavage [109]. Type I collagen homotrimers α1(I)3 are found in foetal tissues [110], cancers and cancer cell cultures [109] or result from genetic disorders resulting in the α2 chain deficiency [111, 112] that are completely resistant to collagenase MMP processing [109]. This was shown to be due to increased helical stability, thereby prohibiting MMP active site access [113]. The structure of the collagen therefore has consequences for its function.

This is demonstrated in the many disorders resulting from mutations which affect the formation of the triple helix and collagen network or cleavage of the propeptides that include osteogenesis imperfecta (Collagen I), Ehlers-Danlos syndrome (Collagen I, III, V), Stickler syndrome (Collagen II), Alport syndrome (Collagen IV), Marshall syndrome (collagen XI) and epidermolysis bullosa (collagen VII) [104, 114]. These disorders can also result from an abolishment of collagen interactions with integrins or proteoglycans, as is the case for lethal osteogenesis imperfecta mutations in collagen I [115].

Collagens can be categorised into groups based on their function and supramolecular assemblies: fibrillar forming, network forming, fibril-associated collagens and membrane associated collagens with interrupted triple helices (FACITs and MACITs), anchoring fibrils, beaded collagens, and multiple triple-helix domains and interruptions (MULTIPLEXIN) (Figure 4) [104, 116].

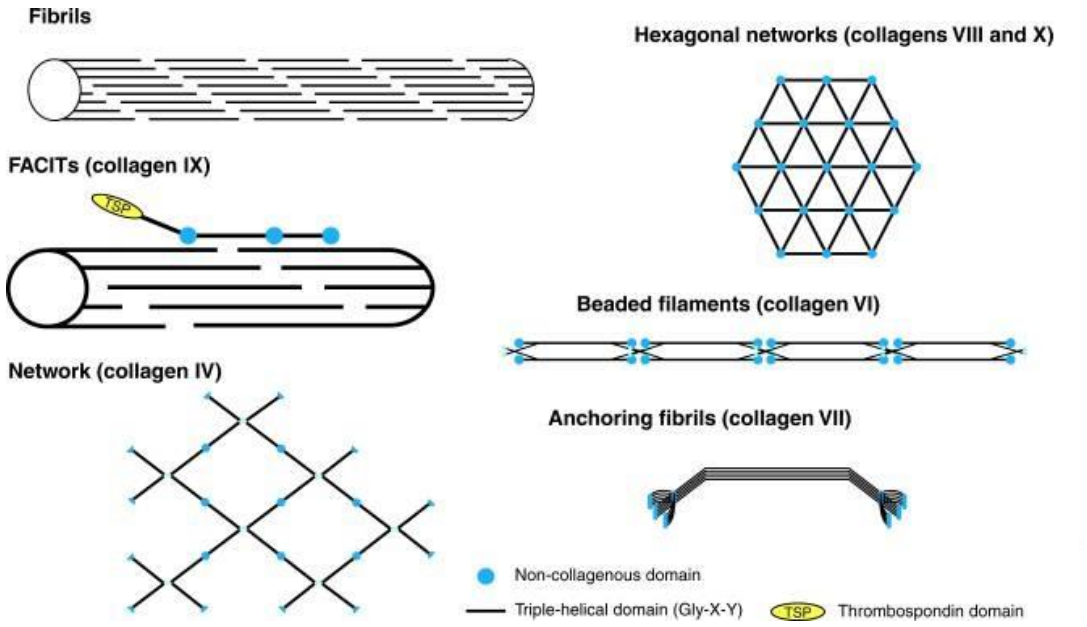


Figure 4. Supramolecular assemblies by collagen.
Collagens are divided into groups based on their assemblies: fibril forming, network forming, fibril or membrane associated collagens with interruptions in their triple helix (FACIT/MACIT) and beaded filaments. Adapted from Blum [104].

1.4.2.1 Fibril forming collagens

Fibril forming collagens encompass types I, II, III, V and XI. Collagen I is the most abundant collagen and fibrils are formed mainly from this type. However, fibers can be heterotypic and thus composed of different collagen types [104]. Collagen I and III fibrils are primarily found in the skin [117] whilst collagen type II and XI are found in cartilage [118]. The characteristic feature of collagen fibrils is that they are *D*-periodic with D = 64 nm in skin and cartilage fibrils, or D = 67 nm in tendons, ligaments and bones, respectively, resulting in an overlapping formation within the collagen fiber [104]. Collagen

I is composed of perfect triple helical repeats which ultimately form super twisted microfibrils of five molecules, which then arrange to form fibers [119].

Collagen fibrillogenesis (Figure 5) occurs in a entropy-driven, self-assembly manner and occurs in two steps: nucleation and fiber growth [105]. Collagen is synthesised as procollagen, a soluble precursor, with N- and C-propeptides. Tropocollagen is generated after cleavage of the N- and C-propeptides to telopeptides by the ADAMTS procollagen N-proteinases and BMP-1/tolloid C-proteinases. The C-terminus is essential for proper triple-helix formation but prevents fibrillogenesis, and cleavage of the C-terminus alone initiates collagen fibrillogenesis *in vitro [120]*. Other proteins participate in fibrillogenesis, for example lysyl oxidase (LOX) oxidizes the lysine side chains in the telopeptides, resulting in the formation of cross-links in the triple helix, stabilising fibrils. It is the cross-linking that gives mature fibers their mechanical properties. L-ascorbic acid (vitamin C) is an essential cofactor for this process [121]. Lysyl-mediated cross linking is tissue-specific, not collagen-specific and can occur between collagens of the same or different types. Cross-linking of collagen type V and XI occurs that collagen type I and II then use as a scaffold to assemble [122]. Cartilage oligomeric matrix protein (COMP), also known as thrombospondin-5, is another protein with a high affinity for collagen I and II that aids in fibrillogenesis by bringing molecules closer together and promoting association of immature collagens [123]. Decorin, which is shaped like an arch, has been shown to bind around the collagen triple helix at the C-terminus of collagen type I near an intramolecular cross-linking site. However, decorin also binds collagen type I, II, III and VI and has been shown to regulate collagen fibril diameter and length [124].

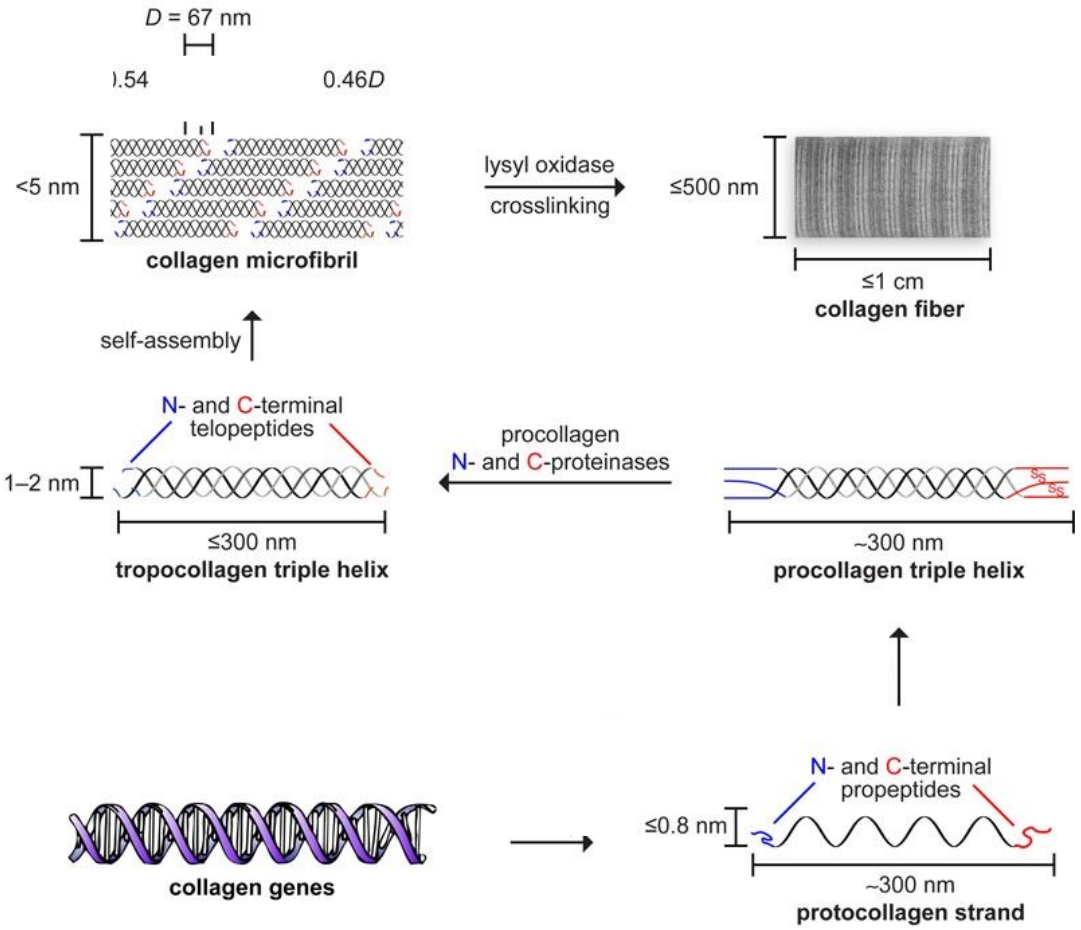


Figure 5. Schematic of the formation of collagen fibers.
Collagen genes are transcribed and translated to protocollagens strands with N- and C-terminal propeptide domains. C-terminal domain-mediated procollagen triple helix assembly occurs, followed by cleavage of the propeptide to leave N- and C-terminal telopeptides. Self-assembly into a collagen micro fibril then occurs with cross-linking by LOX to eventually form the mature collagen fiber. Adapted from Shoulders *et al.* [116].

1.4.2.2 FACITS

As the name suggests these collagens (IX, XII, XIV, XVI, XIX, XX, XXI, XXII, and XXVI) are not fibril forming but associate with fibril forming collagens (figure 5). These molecules have interrupted triple helical domains conferring flexibility and have a thrombospondin domain [125]. Type IX is associated with collagen type II fibrils and is the only FACIT member known to cross-link via LOX [122]. Type XI and XIV are involved in controlling fibril diameter [118, 126] whilst type XII stabilises fibrils during development [127]. A large number of FACITs are also found in the basement membrane. For instance, type VII forms anchoring fibrils between the dermis and epidermis of skin [128] thereby stabilising the basement membrane.

1.4.2.3 Network forming

Network forming collagens consist of type IV, VIII and X. Unlike fibrillar collagens, these types retain their propeptides when incorporated into tissue. Collagen type IV is the predominant network forming collagen and essential for basement membranes formation. This collagen integrates laminin, nidogens and other components into the two dimensional supramolecular basement membrane structure. Knock out of the major type IV collagen isoform a1a1a2 in mice is not embryonic lethal; basement-membrane like structures are assembled, however embryonic lethality occurs at 10.5 days due to a lack of basement membrane structural integrity and therefore function [129]. This type of collagens, like FACITs, also have interruptions in their triple helix structure that confer flexibility. Collagens VIII is found in Descemet's membrane (the membrane for the corneal endothelium) and vascular tissue whilst collagen X is found in cartilage [104]. The hexagonal network of collagen VIII is suggested to resist compression yet retain an open porous design in the former [130], whilst in cartilage the hexagonal collagen X network is thought to facilitate the migration of endothelial cells in angiogenesis and the distribution of proteoglycans within the cartilage growth plate [131]. There does not appear to be any obvious differences between the two apart from their localisation.

1.4.2.4 MACITs

Collagen XIII and XVII comprise this group and are transmembrane proteins. The ectodomain of collagen XIII colocalises with α1β1 integrins and the cytoskeleton in focal contacts [132], whilst collagen XVII interacts with α6β4 and keratin intermediate filaments in hemidesmosomes. Collagen XVII can be cleaved to a soluble basement membrane collagen by ADAM-9 and 10 in a furin-dependent manner [133]. Laminin is a major binding partner of Collagen XVII and this interaction is thought to facilitate linking of cytoplasmic structural elements of hemidesmosomes, which are specialised adhesion junctions between cells and ECM such as the basement membrane [134].

1.4.2.5 Beaded filament forming proteins

Collagen type VI is the archetypal collagen in this group. This has a beaded structure due to a much shorter collagenous domain compared to collagen I and two large N- and C-terminal domains. [135] The exact function of this collagen is unclear, however it is ubiquitously expressed in all tissues and forms networks which are undoubtedly essential for tissue integrity. Collagen VI is found at the basement membrane and interstitial matrix interface [136] and, as such, has many ECM binding partners, including the predominant basement membrane collagen type IV [136] and basement membrane SLRPs, decorin [137], biglycan [138], perlecan [139] and fibronectin [139]. This collagen also interacts widely with cells via its integrin binding motif within the alpha chain [140] and, as it can be synthesised by macrophages, a tissue repair role for collagen VI has been suggested [141].

1.4.3 Elastin and associated proteins

Elastin provides stretch, recoil and flexibility to tissues such as arteries, lungs, tendons, ligaments, skin and cartilage. Elastin forms large networks of two components: elastic fibers and peripheral microfibrils [142]. Elastin is synthesised as a precursor, tropoelastin (60-70 kDa), which associates with multiple tropoelastin molecules (coacervation) during elastogenesis to form elastin fibres with cross linking by LOX. The alignment and cross-linking is directed by alternating hydrophobic, hydrophilic and lysine-containing domains within the tropoelastin monomers [142]. Microfibrils are composed primarily of large (350 kDa) fibrillin glycoproteins (of which there are 5 isoforms) and also microfibril-associated glycoproteins (MAGPs). Fibrillin-1, -2, and -3 are the major structural proteins of the microfibril. The necessity of these proteins for the correct formation of elastin fibres is highlighted in disorders caused by improper elastin formation resulting from mutations in fibrillin-1. Such genetic disorders include Marfan syndrome [143] and arachnodactyly [144]. MAGPs (-1 and -2) are small glycoproteins (20 kDa) which bind elastin and fibulin and have roles in elastogenesis involving the coacervation of tropoelastin [145]. MAGPs also bind collagen VI, decorin and biglycan [146]. Microfibrils are thought to act as a scaffold for tropoelastin to assemble on.

Fibulins (of which there are 7 members) are also involved in the formation of the elastic fiber and associated with the basemement membrane. The short fibulins (fibulin-3, -4 and -5) play an essential role in elastogenesis and contain calcium-binding epidermal growth factor (EGF) like repeats [147]. Fibrillin-3, -4, -5 also interact with heparin; these sites have been defined *in vitro* but *in vivo* these proteins bind to heparan sulfate GAGs (secreted by mast cells into the ECM) found in close proximity to cell surfaces [148]. Fibulins can bind LOX and tropoelastin thereby regulating their crosslinking [145]. Fibulins appear to interact with fibronectin as fibulin-1 deposition is abolished when a fibronectin matrix is not present [149].

1.4.4 Fibronectin

Fibronectin is one of the most ubiquitously expressed fibrous ECM proteins and is necessary for vertebrate development; mice lacking fibronectin die [150]. Fibrous proteins, such as collagens and fibronectin are responsible for the mechanical and structural properties of the ECM [38]. A functional fibronectin molecule consists of a dimer of two subunits (220 kDa to 250 kDa) each composed of a combination of 12 type I repeats, 2 type II repeats, or 15-17 type III repeats, held together by a disulfide bond at their C-termini [151]. Many fibronectin isoforms exist due to alternative splicing of the different repeats [152]. The type III repeats (FNIII) form the largest part of the fibronectin molecule and, although each repeat differs in its 90 amino acids sequence, there is a high level of structural homology with each repeat forming 7 β-strands and two antiparallel sheets [153]. However, these sheets lack disulfide bonding which means this region can undergo conformational changes induced by mechanical forces to expose or hide an integrin binding motif [154] or by the inclusion of the type III alternatively spliced extra domain A or B [155]. This extra domain sequence is specific to cellular insoluble fibronectin incorporated into fibrils, whilst soluble versions (plasma fibronectin), which are secreted by hepatocytes, lack this domain but their biological functions remain similar [151]. However, this circulating plasma fibronectin can be assembled into the ECM and become insoluble [156].

The type I repeats 1-5 and the first type III repeat are necessary for fibrillogenesis as these domains are essential for self-association [157]. This process requires mechanical force to expose homophilic sites

within the type III repeats and trigger homophilic binding events as well as exposing other protein binding sites [158]. Fibronectin forms a provisional matrix for many ECM proteins and is essential for the incorporation of collagen I [105, 157], fibrillin [149], latent TGF-β binding protein (LTBP) [159] and tenascin-C (TN-C) [160] into the ECM. Some proteins associate directly with fibronectin whilst others use fibronectin as a loading scaffold [161]. For example, fibroblast-produced fibronectin positively regulates LOX activity [162] and enhances the proteolytic activity of BMP-1 proteinase [163], thereby regulating collagen cross-linking and maturation. In line with this role as a scaffold, fibronectin expression is upregulated during wound healing and tissue remodelling [164].

Fibronectin has key roles in cell survival via binding to and presenting PDGF growth factors, amongst others, to adherent cells [165]. In this way fibronectin can act as a storage facility, releasing growth factors during ECM remodelling. This further regulates ECM synthesis and deposition via cellular interactions with integrins and the cytoskeleton (mechanotransduction) that are then able to regulate cell behaviour [3, 166].

1.4.5 Laminins

Laminins are large (140 kDa – 400 kDa) glycoproteins and the major constituent of the basement membrane that confer its signalling platform attributes. Laminins are composed of a trimer of three disulphide linked chains; α, β and γ which are distinct at the amino acid level. Laminins vary in size, dependent on isoforms and the degree of post-translational glycosylation [167]. Whilst there are over 50 theoretically possible heterotrimeric combinations between the 3 chains, 16 are known for mouse and humans with 8 distinct chains identified (α1, α2, α3, β1, β2, β3, γ1, γ2) and 7 forms (Figure 6) [167]. The N-terminal moieties of each chain form three short arms which are capped with a globular domain and the three chains are joined in a characteristic triple helical coiled coil [168]. The complexity of the isoforms is further enhanced by regions which can be proteolytically cleaved from the heterotrimers whilst remaining non-covalently bound. The globular domains are responsible for self-assembly [169]. Nidogen-1 and -2 interact with the γ1 chain [170], whilst agrin has high affinity binding with a γ chain sequence only when this is within the coiled coil [171]. Integrins α6β1, α6β4, and α7β1 bind to the C-terminal domain of the heterotrimer and, like agrin, require the coiled-coil to be present

[172]. This integrin binding allows the linkage of the cell to the cytoskeleton and propagation of signals within the cells [13].

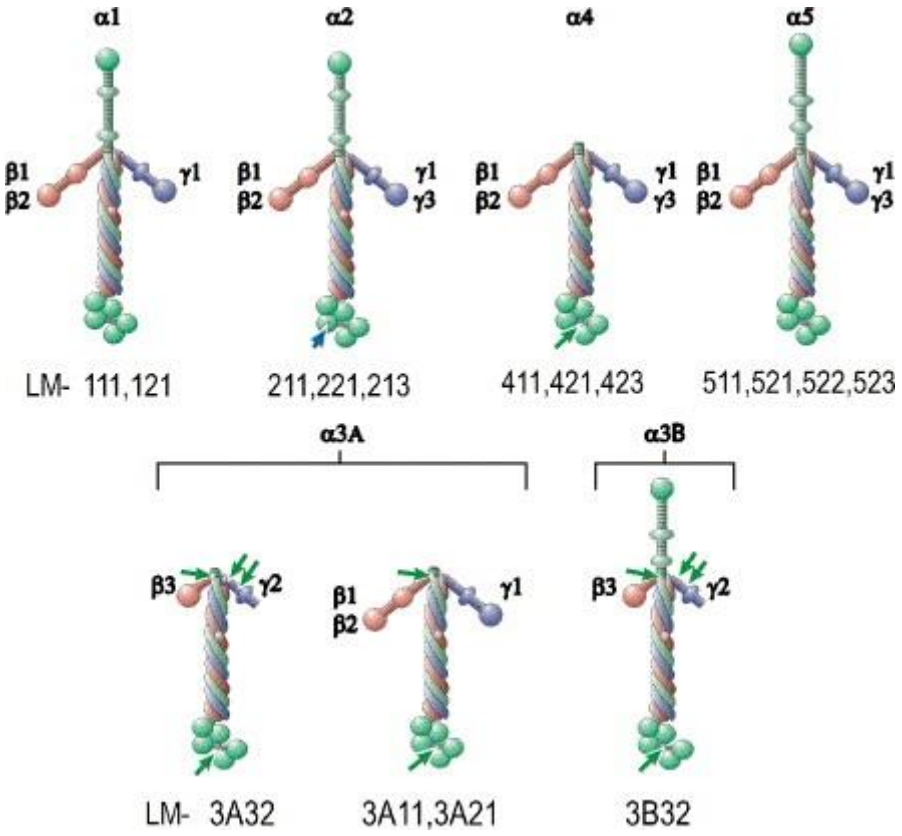


Figure 6. **Laminin trimer configurations.**
Schematic of 16 mouse or human laminin heterotrimers. The trimer composition is indicated below the laminins. Greek letters with numbers indicate chains. Chain types α, β, γ are distinguished by green, red and blue, respectively. Green arrows indicate known cleavage sites which generate free fragments. The smaller blue arrow (at the helical domain of α2) indicates cleavage where the fragment remains non-covalently associated with the trimer. From Aumailley *et al.* [173]

1.4.6 Matricellular proteins

The ECM also contains non classical proteins, termed 'matricellular proteins', which include thrombospondins, tenascins, SPARC (secreted protein, acidic and rich in cysteine), CNNs [174] and osteopontin (OPN) [175]. These are distinguished from classical ECM proteins such as laminins by several key characteristics; they do not confer structural properties to the ECM; their expression pattern is restricted to early development, whilst in adults expression is limited to injury; they are capable of binding to other ECM components, growth factors, proteases and cytokines, and they discourage cell adhesion [175]. Mice lacking matricellular proteins display an apparently normal phenotype, or a

phenotype indistinguishable from that of wild type animals, until injury occurs, due to the role of matricellular proteins in tissue repair (Figure 7) [176]. Most matricellular proteins share structural domains including thrombospondin repeats, EGF-like repeats and Von-Willebrand motifs, thereby allowing a wide range of binding partners.

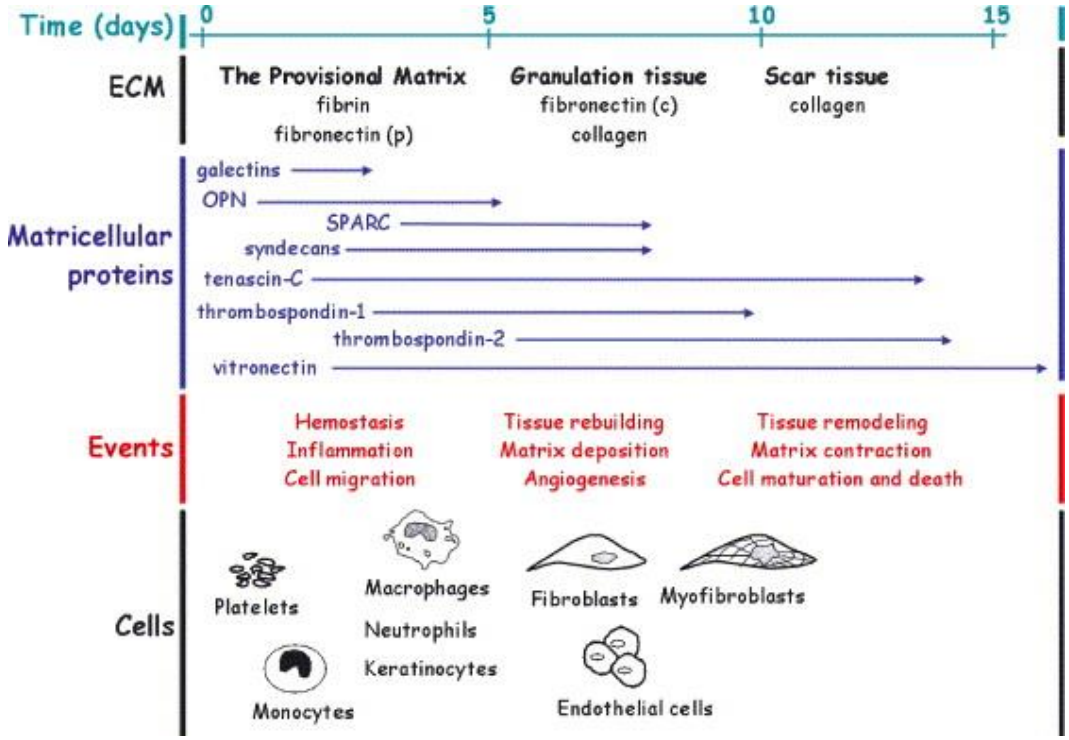


Figure 7. Matricellular proteins and their role in tissue repair over time.
The coordinated timing of major events, matricellular protein expression, ECM protein expression and the cell types involved throughout the tissue repair process. Matricellular proteins are expressed throughout the repair process, starting early on during inflammation and when the provisional fibronectin ECM, that is necessary for the incorporation of other ECMs proteins, is deposited. Plasma (P) fibronectin and fibronectin expressed from cells in the tissue (c) are present. Two types of fibronectin are present at sites of injury, plasma (p) from the blood and cellular (c) expressed by cells in the tissue. Adapted from Midwood *et al.* [177].

1.4.6.1 Tenascin-C

Tenascin–C (TN-C) has previously been known as glial/mesenchymal extracellular matrix protein (GMEM) [178], myotendinous antigen [177], hexabrachion [179], cytotactin [180], neuronectin [181] and finally TN-C [182]. TN-C is the founding member of the highly conserved tenascin family of ECM glycoproteins composed of tenascin-X (-Y in the chicken), -R, –W and –C [183].

TN-C conforms to the defining characteristics of matricellular proteins; being highly expressed during development throughout neural, skeletal, and vascular morphogenesis, and largely expressed in

connective tissue. Expression in adulthood is restricted to the response to tissue injury in wound healing, cancer and inflammation [183-185]. There are some niches where TN-C is present during adulthood, including regions subjected to mechanical/compressional stress, for example in tendons, in which it is theorised to aid in the maintenance of fibrocartilaginous regions by decreasing cell adhesion [186]. TN-C is also present in other niches such as lymph nodes and the spleen [187]. In addition, *TNC* knockout mice also appear normal until stressed or challenged [188]. However, it has been suggested that the similarity of the tenascin family's structure results in functional compensation, although each member exhibits distinct tissue distribution patterns [189]. TN-R is restricted to the central nervous system, TN-X is present in most connective tissues and TN-W is localised in the developing skeleton and adult kidneys [190-192].

1.4.6.2 Tenascin-C gene and protein structure

The *TNC* gene is located at chromosome 9q33 and is composed of 97,680 bp with a total of 29 exons and 27 introns [193] (Figure 8). Exon 1 is not translated but contains the transcription start site and a TATA box promoter region 220 bp upstream. Exon 2 contains the translation initiation codon ATG (Figure 8). Exons 2 and 3 encode the terminal assembly (TA) domain and 14.5 epidermal growth factor-like (EGFL) repeats, each of which are 30-50 amino acids long and contain six cysteine residues important for disulphide bonding between chains [183]. *TNC* also contains 17 fibronectin type III (FNIII) repeats, approximately 90 amino acids long, which are encoded by exons 4-23. However, the repeats encoded by exons 11-17 are subject to alternative splicing, generating a variety of TN-C isoforms [183]. These fibronectin-like repeats form two β sheets. This is a highly elastic structure which may be capable of masking or revealing the integrin binding site embedded within this region. The last structural feature of TN-C is the C-terminal fibrinogen like globular domain (FBG), which is 210 amino acids in length [183] and homologous to the β and γ chain of fibrinogen, encoded by exons 24-28. This region harbours a calcium binding loop.

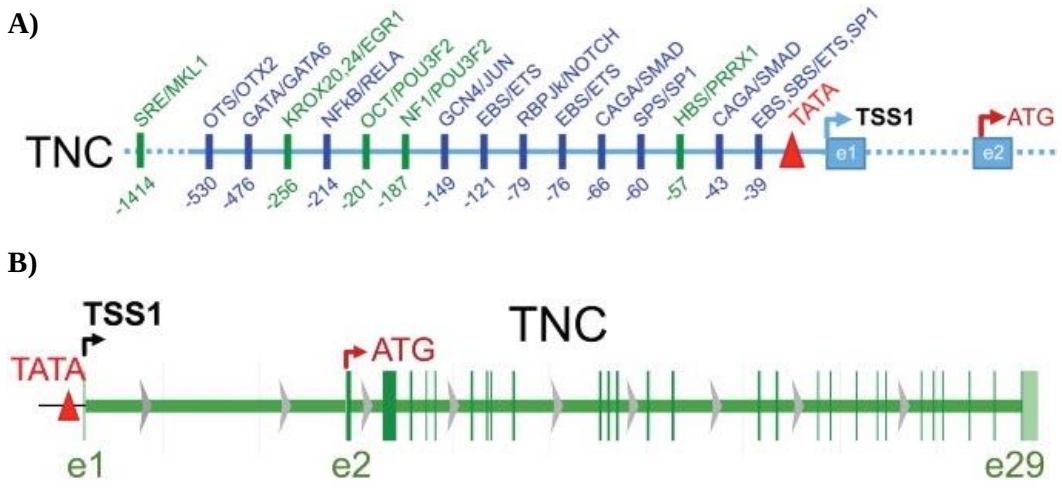


Figure 8. Schematic of the gene promoters of *TNC*.
A) The transcription start site (TSS) is indicated by a blue arrow at the start of the first exon (blue box, e1). The start codon of *TNC* is indicated by a red arrow in the second exon (blue box, red ATG start codon). Upstream promoters are indicated by vertical lines in which experimentally confirmed transcription factor binding sites are marked for humans (dark blue) and mouse (dark green) at the corresponding positions they bind. **B)** Schematic of the *TNC* gene indicating coding exons (green lines starting at exon 2), the start codon (ATG) indicated in the second exon (red arrow), the last coding exon (29) and the TSS and TATA consensus sequence (red text). Adapted from Chiovaro *et al.* [193].

Whilst the mechanism that generates the different splice variants is not understood, it has been suggested that extracellular pH, TGF−β1 and splice factor Sam68 are involved [194]. It has also been demonstrated that shorter and longer isoforms are produced in response to different cytokines. For example, TNF-α induces the expression of short isoforms, containing no alternatively spliced repeats, whilst long isoforms, containing one or more alternatively spliced fibronectin repeats, are induced by interferon- γ (IFN-γ). In contrast, interleukin – 4 (IL-4) results in equal production of long and short isoform transcripts [195]. Whilst many isoforms have been identified at the mRNA level, it still remains uncertain whether all of these transcripts are translated into protein and what role these transcripts have [195].

TN-C is characterised by its unique domain composition of each monomer. TN-C monomers consist of polypeptides ranging from 180 kDa to300 kDa, depending on splice variants of the 17 FNIII repeats [196]. TN-C, -R and -W are able to oligomerise to form trimers, whilst only TN-C and TN-W can subsequently form 6-armed hexabrachion structures (Figure 9). The assembly of trimers and hexamers is facilitated by the N-terminal TA domain, which contains cysteine residues and heptad repeats vital

for monomer interaction. Additional cysteine residues are present in those tenascins able to form hexamers [197]. Hexamer formation occurs in a two-step process. First, a triple stranded coiled coil is formed by three TN-C polypeptides via alpha-heptad repeats, then the TA domain associates with a second trimer at the N-terminus [198]. Such oligomerization occurs very rapidly after translation in the endoplasmic reticulum [199]. The multidomain structure of TN-C allows binding to a vast number of ligands (Figure 9) and accomplishment of diverse, time- and context-specific cellular functions, including cell adhesion, migration, survival, differentiation and polarization [200].

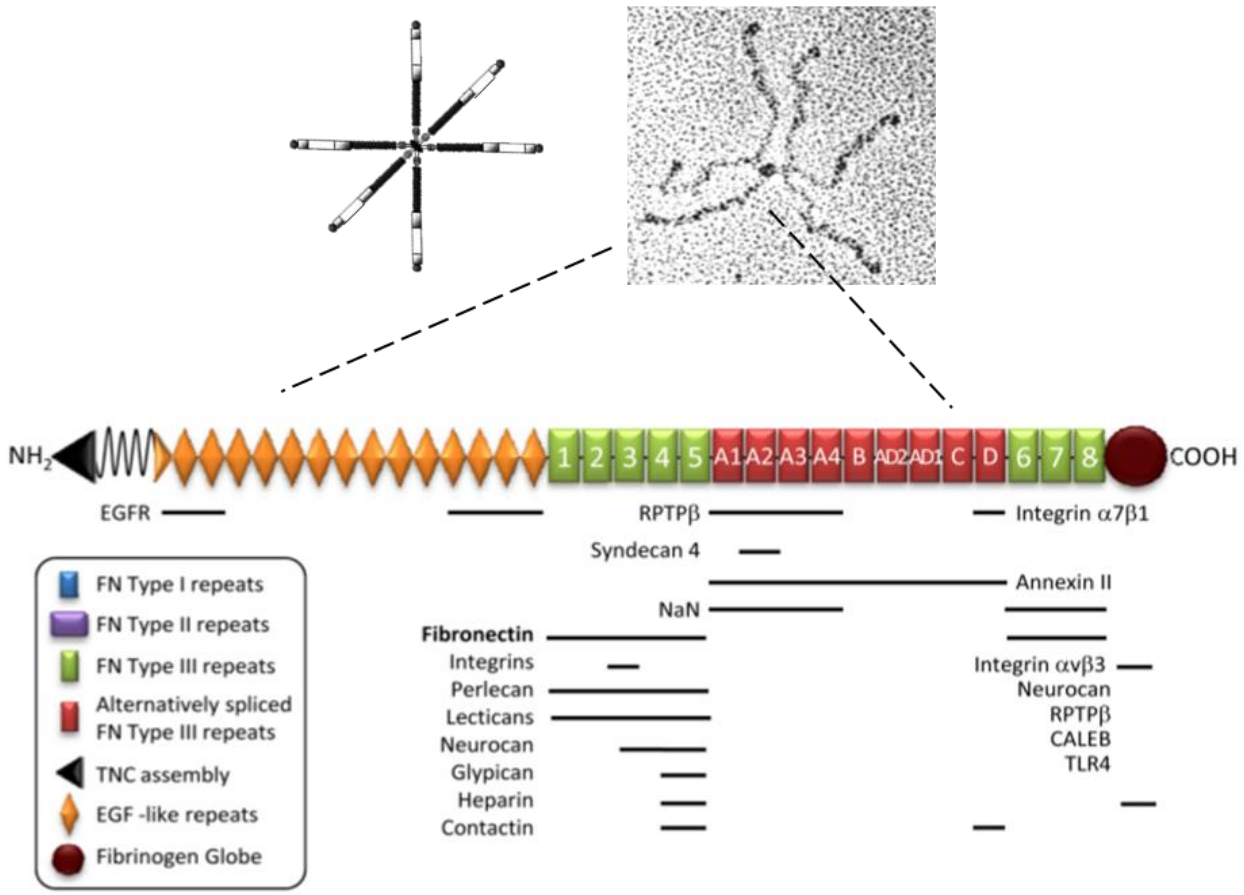


Figure 9. Schematic of TN-C structure.
TN-C forms a 6–armed hexabrachion structure from two trimers, shown is a schematic and electron micrograph of this structure. Conserved and alternatively-spliced fibronectin type-III (FNIII) repeats are represented by green and red boxes, respectively. All species contain the assembly domain (TA) followed by 13 epidermal growth factor (EGF) like repeats (yellow diamonds), whereas mammalian TN-C polypeptides contain 1.5 additional repeats. The C-terminus is made up of a C-terminal fibrinogen like globe domain (FBG) (circle). Each 'arm' is capable of binding a variety of binding partners, including integrins, perlecan, heparin and fibronectin via the multidomain structure. The RGD motif in the FNIII domain facilitates interactions with many integrins. Functionally, the FBG domain is able to activate TLR4 and act as a 'matrikine'. Adapted from Erikson *et al.*, Jones *et al.*, Orend *et al.,* Van obberghen-Schilling *et al.* [185, 201-203].

1.4.6.3 Regulation of Tenascin-C expression

TNC expression is highly regulated at both the transcriptional and post-transcriptional levels to ensure a restricted expression pattern. Among the tenascin family of glycoproteins, the *TNC* promoter is the best studied in several species. It has been established that even-skipped homeobox1 (EVX1) strongly activates transcription and this is specifically due to synergy with JUN/FOS transcription factors that target the AP1 site [204]. The POU-homeodomain transcription factor (POU3F2*)* has also been shown to interact with the mouse *TNC* promoter, which is conserved across the human and chicken genes [205]. Another conserved homeobinding sequence allows the paired-related homeobox 1 (PRRX1) transcription factor to bind and induce *TNC* transcription [206].

Additionally *TNC* expression is present at basal levels in tendon-associated tissues and other regions that are subjected to mechanical loading [207]. Due to this expression it was therefore hypothesised that mechanical load may regulate the induction of *TNC* expression. To test this, Fluck *et al.* loaded the anterior latissimus muscle of chickens by attaching a weight (10% of the total body weight) to the wing [208]. After 4 hours of loading, it was confirmed that *TNC* expression was induced. In support of this, further experiments demonstrated *TNC* expression directly correlated with loading [207]. The molecular mechanism by which mechanical loading induces *TNC* expression has partially been elucidated. Fibroblasts cultured on membranes that are stretched in a repeated pattern (cyclical strain) result in Rho-associated kinase (ROCK) signalling resulting in *TNC* induction [209]. This effect is attenuated by inhibition of ROCK [210]. This induction was later demonstrated to be dependent on β1 integrins [211] and fibronectin [212]. This cyclical strain-induced Rho activation leads to translocation of the transcription factor megakaryoblastic leukemia-1 (*MLK1*) to the nucleus where it can bind to the *TNC* promoter region to induce transcription [213]. Further experiments involving unloading reinforced the theory that mechanical loading results in *TNC* expression. Immobilization of developing tendons in chick embryos by decamethonium bromide resulted in downregulation of *TNC* expression [214]. Similarly, unloading resulted in downregulation of many other ECM components such as collagen that are also directly regulated by mechanical loading, and required for homeostasis. However, the kinetic

of expression of *TNC* differs from the other matrix components indicating different signalling pathways are involved [208] [211].

In further support to the mechanotransduction pathways resulting in the expression of the *TNC*, it has been demonstrated that contracting wound tissue, which is undergoing healing, expresses *TNC*. TN-C stimulates fibroblasts to migrate and results in more efficient re-epithelialization than in wounds with delayed *TNC* expression [177]. Tissue injury and inflammation have been demonstrated to induce *TNC* expression via pro-inflammatory cytokines, including interleukin 1 alpha (IL-1α) [215] and IL-1β [216] as well as anti-inflammatory cytokines such as intelekin-13 (IL-13) [217] and interleukin-4 (IL-4) [218]. Direct functional studies of the effect of cytokines on the *TNC* promoter remain to be carried out. However, there are numerous studies on TN-C during inflammation. Treatment of chick embryo fibroblasts with TGF-β has been shown to stimulate TN-C synthesis [219]. PDGF and TGF-β growth factors, and tensile strain (mechanical loading) were shown to act in an additive fashion promoting *TNC* mRNA expression [209]. The signalling pathways that are involved in the induction of *TNC* are emerging. For instance, PDGF induced *TNC* expression in a dose-dependent manner via PI3K/AKT signalling, which triggers the transcription factors SP1, ETS1 and ETS2 to bind the ETS binding sites (EBS) in the *TNC* promoter [217, 220]. SMAD proteins have two potential binding sites in the *TNC* promoter regions and have been shown to control TGF-β–induced TN-C in rat astrocytes [221]. Glucocorticoids have also been shown to act in an anti-inflammatory fashion and downregulate *TNC* expression in bone marrow stromal cells and fibroblasts. [222]. Furthermore, the zinc finger transcription factor GATA6 has also been shown to repress human *TNC* expression in fibroblasts inhibiting induction by IL-4 and TGF-β [223]. LPS-induced TN-C transcription has also been demonstrated by ChIP of RNA polymerase II. The same study showed that nuclear factor kappa-light-chain-enhancer of activated B cells (NF-kB) and PI3K/Akt signalling pathways mediated this induction [224].

As well as inducing transcription of TN-C, LPS has further reaching influences by determining the isoform which is generated; post-transcriptional regulation occurs mainly by alternative splicing which can generate up to 500 isoforms [200]. Giblin *et al.* [225] demonstrated that in unstimulated dendritic

cells and macrophages, all 9 alternatively spliced FNIII repeats are detectable at low levels comparable with the constitutively expressed FNIII 7-8 domains. However, LPS-activated dendritic cells and macrophages produced a smaller isoform lacking additional domain -1(AD1) and -2 (AD2), but with significant upregulation of all other alternatively spliced FNIII domains and the FNIII 7-8 domain. In contrast, in both unstimulated and IL-1β-activated human dermal fibroblasts, a larger isoform is produced in which FNIII 7-8 was expressed at higher levels than any alternatively spliced domains. Using recombinant proteins and chemotaxis assays, it was determined that the exclusion of AD2 and AD1 is necessary to promote dendritic cell chemotaxis and toll like receptor -4 (TLR-4) dependent cytokine synthesis. Correspondingly, inclusion of these domains supported stromal cell adhesion. The study emphasised the immune-regulatory role that different domains and therefore different isoforms can exhibit [225]. Furthermore, this shorter LPS-induced isoform can contribute to shaping the microenvironment by promoting remodelling via cell migration and destructive pro-inflammatory cytokine synthesis.

However, an additional mechanism by which TN-C is post- transcriptionally regulated is by miR-335, together with SOX4 which significantly abolished metastasis and the invasive ability of LM2 cells, a derivative of MDA-MB-231 breast carcinoma cells [226].

Alternative splicing has further effects on post-translational TN-C regulation given that MMPs target specific sites in the alternatively spliced fibronectin repeats. Therefore, the inclusion or exclusion of certain repeats in ‘small’ or ‘large’ isoforms confers or removes susceptibility to cleavage. A small isoform of TN-C, lacking any alternatively spliced FNIII repeats is resistant to MMP-2, -3, and -9-mediated cleavage. This was in stark contrast to a large isoform, containing alternatively spliced FNIII repeats A1, A2, A3, A4, B, C and D, which was sensitive to these MMPs, in particular within the FNIII A3 repeat (Figure 9). MMP-7 was also shown to be able to degrade both small and large TN-C isoforms by cleaving the C-terminal FBG [227].

Given that smaller isoforms have been indicated in regulating the TLR4 pro-inflammatory activity of activated immune cells [225] and TLR4 activation induces MMP production (discussed later), the

resistance or susceptibility to processing likely provides an additional regulatory step of TN-Cs immunomodulatory activities. Indeed, transient expression of *TNC* is necessary for its biological functions and a role for proteolytic processing of TN-C has been demonstrated in mice. In an asthma mouse model, MMP-19 null mice display elevated TN-C protein compared to wildtype mice, indicating MMP-19 processing of TN-C [228]. This was suggested to obstruct an accumulation of TN-C and appears to be a protective mechanism. In line with this, in another experimental model of asthma, *TNC* deficiency was linked to attenuated inflammation [229]. As with most ECM proteins, posttranslational modifications can diversify TN-C isoforms. TN-C contains two potential glycosylation sites in the TA domain, two in the EFG-like repeats, 18 in the FNIII repeats and one in the FBG domain [183]. Functionally, glycosylation has been shown to protect rat fibronectin from proteolytic cleavage [230]. A similar protective mechanism may occur here to protect TN-C from proteolytic cleavage, particularly in the putative glycosylation sites within the alternative spliced FNIII repeats, that associate with the FNIII domains that are susceptible to degradation by MMPs [227]. Glycosylation may also regulate binding activities of TN-C and regulate hexamer formation [200].

Post-transcriptional and post-translational modulation of TN-C abundance and domain composition is a key regulatory mechanism which allows isoforms with different biological functions and differing susceptibility to tissue degrading enzymes to be regulated in a spatial and temporal manner to result in an appropriate inflammatory response.

1.4.6.4 Tenascin - C function

The expression of TN-C is tightly regulated; restricted to neural, skeletal and vascular morphogenesis during development and, in adults, limited to injury, cancer and inflammatory diseases with the exception of some niches as previously described [186]. Despite the widespread expression of TN-C during development, and theorised importance of this protein due to such expression, knockout mice (*TNC* -/-) display a phenotype apparently indistinguishable to that of wild type mice. This was demonstrated in mice with a disruption to the *TNC* gene by LacZ insertion in exon 2, in which offspring developed normally with no alterations to collagen, fibronectin, laminins or proteoglycans [231]. This

cast doubt if TN-C really was essential for normal development. A second knockout model was generated by the introduction of a neomycin resistance cassette into exon 2 which resulted in a frameshift and premature stop codon [232]. These mice were also fertile and appeared normal, with no alterations in development or life span. The only apparent difference was in wound healing, in which wounds healed slightly slower but normally, with reduced fibronectin. It has been suggested that this null phenotype may be due to compensatory activity of TN-W capability, due to similarities in structure and overlapping expression patterns, but this has yet to be elucidated [233].

Later, it was revealed that *TNC* $^{-/-}$ mice only appear normal until challenged. In *TNC* $^{-/-}$ mouse models of reversible kidney disease, involving three different genetic backgrounds, the disease phenotype in *TNC* $^{-/-}$ mice was more severe compared to wild type animals and varied across the three strains [234]. Cells cultured from these mice were also resistant to stimulation with cytokines such as PDGF until exogenous TN-C was added indicating a role for this protein in the propagation of pro-inflammatory signalling pathways. This aspect will be discussed in detail later on. Supporting this, in models of rheumatoid arthritis, *TNC* $^{-/-}$ mice are protected from joint destruction and prolonged inflammation [235]. However, the role of TN-C is not straightforward. Whilst *TNC* $^{-/-}$ KO can exbibit more severe phenotypes, in mouse models of tumour angiogenesis, the KO appeared to be protective as vascular endothelial growth factor (VEGF) expression and subsequent angiogenesis were reduced [236]. TN-C is heavily linked to angiogenesis in cancer and diabetes [237] [238]. Although this may be linked to TN-C roles in regulating cell adhesion.

Early *in vitro* studies investigating the function of TN-C were primarily focussed on its anti-adhesive function. The culture of mammary tumours on TN-C coated cell culture plates indicated that TN-C had anti-adhesive properties, but pro-proliferative roles [182]. Additionally, TN-C was shown to inhibit the binding of many cell types to fibronectin, including 3T3 mouse fibroblasts, primary human dermal fibroblasts, MRC-5 human lung fibroblasts, immortalised rat cell line-52, L929 mouse fibroblasts, T98G human brain glioblastoma fibroblasts and MDA-MB-435 human mammary gland adenocarcinoma cells [239]. Consistent with this, neurite cells attach poorly to TN-C coated plates, however this is a mechanism by which axonal outgrowth is encouraged [240]. Supporting this,

processing of TN-C by meprin-β (MP), abrogated this anti-adhesive effect [241]. More recent studies have confirmed this role in cell adhesion *in vivo* - monocytes and macrophages have been shown to exhibit increased migration in mammary tumour stroma in *TNC* -/- mice [242].

More recent studies have also demonstrated that the mechanism by which TN-C manipulates cell adhesion behaviour is via integrins. The FBG domain and alternatively spliced domain FNIII D are able to bind αVβ3 and additionally αvβ7 for FNIIID [243], whilst the FNIII3 domain of TN-C harbours binding regions for αVβ3, αVβ1, α9β1, α8β1 and αVβ6 [243, 244]. Studies of lymphocyte migration using *TNC* -/- lymphocytes demonstrated reduced lymphocyte infiltration and migration, and lower expression of TGF-β, IFN-γ, TNF-α, and IL-4 during liver hepatitis compared to wild type mice [245]. Interestingly, this was accompanied by a reduction in collagen I and III expression. Specifically, T-lymphocyte migration was inhibited via TN-C binding to fibronectin and thus blocking of fibronectin binding to integrin α5β1 [244]. Similarly, TN-C can prevent interaction of syndecan-4 and fibronectin, by binding to the syndecan-4 binding site in the 13th FNIII repeat. This causes a downregulation in RhoA and focal adhesion kinase signalling by preventing α5β1 integrin clustering and subsequently abrogates adhesion [246].

Linked to the adhesion and migration properties of cells, TN-C also regulates cell proliferation in the epidermis [247]. From patients with psoriasis and blistering diseases, which are akin to erroneous wound healing and re-epithelialisation, TN-C is upregulated in the epidermis underneath the hyperproliferative region [248, 249]. Furthermore, TN-C controls fibroblast proliferation as the EGF-like domains of TN-C (Ten1-2, Ten 11- 13 and Ten-14) activate EGFR signalling and ERK/MAPK signalling [250]. Differential activation can occur between these repeats - activation of the EGFR by the 14th EGF- like repeat specifically induces the activation of migratory signalling over proliferative signalling pathways [251].

TN-C cleavage products can also become bioactive fragments, known as ‘matrikines’, with an additional range of biological roles to full length TN-C. Fragmented EGF-like TN-C repeats, but not full length TN-C induce apoptosis in smooth muscle cells (SMC) and TN-C itself was found to upregulate MMP production [252]. *In vitro,* FNIII TN-C repeats, but not full length TN-C, was found

to inhibit fibronectin fibrillogenesis by binding to fibrils and blocking intermolecular fibronectin interactions [253]. A further *in* vitro study demonstrated MMP-2 cleavage of the FNIII A3 domain also exposes a cryptic syndecan-4 heparan sulfate side chain binding site in the FNIII A2, which is necessary for stimulating integrin β1-mediated cell adhesion [254]. *In vivo* work in this area is lacking.

The above studies demonstrate the multitude of ways, dependent on TN-C domain composition and size, by which TN-C is able to regulate all aspects of cell behaviour by way of direct interaction with the main ECM receptors, integrins, or indirectly by activating a variety of cell signalling cascades involving growth factor receptor activation.

1.4.7 ECM receptors

ECM receptors are the crucial link for the propagation of signals from ECM components to the nucleus inside the cell, to direct and control cell fate. Indirect regulation occurs by ECM molecules binding several receptors concurrently with growth factors [255]. Direct regulation occurs primarily via integrin binding, and other receptors such as CD44, the HA receptor, syndecans, the fibronectin receptors, and discoidin receptors (DDRs), non-integrin collagen receptors for classical outside-in signalling [256-259]. Cluster of differentiation 47 (CD47) is an integrin-associated protein receptor that protects cells from phagocytosis by binding and activating the signal regulatory protein -α (SIRPα) receptor on macrophages and dendritic cells that blocks integrin activation [260].

DDR 1 and 2 are receptor tyrosine kinases (RTKs). These bind non-structural ECM components such as growth factors and cytokines, including VEGF, FGF, EGF, PDGF, insulin and ephrins to name a few [261-263]. DDR RTKs are central components of most signalling pathways due to their core tyrosine kinase domain with sequences that are subject to autophosphorylation, or phosphorylation by other protein kinases [264].

These communications between cells and ECM form the basis of development, maintenance and repair of organisms [265].

1.4.7.1 **Integrins**

Integrins are a superfamily of cell adhesion proteins principally responsible for cell binding to a variety of ligands, including ECM components as well as other cells via counter receptors such as members of the ADAM family or immunoglobulin-type receptors such as intercellular adhesion molecule (ICAM) and vascular cell adhesion molecule (VCAM). In this way integrins function as regulators of both cell-cell and cell-matrix interactions [266]. By extension, this means that integrins also have fundamental roles in development, and disease pathogenesis as well as all aspects of cell behaviour. Integrins were so named for the integrity they provide the cytoskeleton with. Integrins function as one of 24 transmembrane αβ heterodimers which can be generated from 18 α and 8 β subunits in humans. Alternative splicing of each subunit adds further complexity [267]. Whilst the α and β subunits are distinct, structurally they both contain a cytoplasmic tail, an extracellular domain and transmembrane domain. Each subunit type has ligand specific affinities.

The most extensively studied subset includes integrins αIIb, αv, α5 and α8, which contain the RGD motif (Arg-Gly-Asp). These form heterodimers with β1 and β3 subunits and can promiscuously bind to many ligands that contain an RGD motif including fibronectin, vitronectin, fibrinogen, tenascin–C and some proteolytically processed collagens and TGF-β, amongst others [243, 268, 269]. Another subset of α integrins contains an inserted domain termed 'αI' domain [270], which binds a triple helical sequence in collagens (α1, α2, α10 and α11) [271] and leukocytes (αE, αL, αM, αD and αX) [272]. The β1 integrin contains a sequence homologous to α1 [270]. Integrins that lack the αI domain primarily pair with β1 subunits and bind laminins (α3, α6 and α7), crucially regulating cell contacts with the basement membrane [26]. With these specific ligands in mind it is important to note that a hallmark of integrins is the diverse number of binding ligands from a single heterodimer.

Processing of ECM components can result in the unmasking of cryptic integrin binding sites that function to provide instructions in a temporal and location specific manner. For example, laminin-5 has been shown to play a role in wound healing via alternatively displaying different integrin binding sites. Laminin-5 is able to bind α3β1, α6β1 and α6β4 via a sequence in its globular domain. Binding of α3β1

by the globular domain promotes epithelial cell migration and thus initiation of wound closure. Subsequent processing of laminin reveals an α6β4 binding site, which upon binding to the cell promotes hemidesmosome formation and cell-cell contacts [273].

Mechanical force has also been demonstrated to expose integrin binding sites. This occurs in fibronectin which is highly sensitive to mechanical force, unsurprising given that fibronectin requires exposure of cryptic sites for its polymerisation [158].

The key trademark of integrins is their ability to signal bidirectionally; inside–out and outside–in. Both these signalling processes require the assembly and destruction of multimodular complexes around the cytoplasmic integrin tail which results in conformational changes to their extracellular domains [274]. For both inside-out and outside-in signalling, the binding of talin (and kindlin with less importance) to the β subunit cytoplasmic tail results in a conformational change in the αβ extracellular domain which activates integrin ligand binding abilities [21]. On ligand binding the next crucial step for integrin activation is ligand-bound integrin clustering. Talin and kindlin bind at sites of filamentous actin polymerisation (F-actin) [275] with talin directly binding F-actin, thus providing a mechanism for ECM and cell cytoskeleton linkage [276]. Talin also recruits regulators of F-actin, focal adhesion kinase (FAK) and PIPK1γ90 (a splice variant of phosphatidylinositol (4)-phosphate 5-kinase type) Iγ), along with vinculin, paxillin, a-actinin and Src-kinases [276, 277]. The recruitment of adaptors and cytoskeletal proteins results in the formation of a focal adhesion plaque in which FAK is also able to recruit other talins, thus reinforcing integrin clustering [278]. Generally, integrin activation results in clustering and activation of the MAPK/ERK signalling cascade which is then able to modulate gene expression that in turn regulates all aspects of cell behaviour [279]. Integrins' roles in 'sensing' and propagating signals leading to matrix regulation has led therapeutic strategies targeting these molecules, particularly in autoimmune disorders, and anti-fibrotic strategies [280].

1.5 THE INFLAMMATORY RESPONSE

During the inflammatory response, macrophages (and other innate immune cells such as neutrophils and mast cells) recognise highly conserved structures known as pathogen-associated molecular patterns

(PAMPs), present on pathogens, via their pattern recognition receptors (PRRs) such as Toll-like receptors (TLRs). The most studied PAMP is lipopolysaccharide (LPS), the outer membrane component of gram-negative bacteria [281]. PRRs are also able to recognise endogenous ligands which become available as a result of infection or tissue damage and are termed damage-associated molecular patterns (DAMPs) [282, 283]. Notably, many ECM proteins (matrikines) can act as endogenous ligands and be recognised by specific TLRs (TLR2 and 4) on macrophages, including TN-C [235], citrullinated fibrinogen [284], the extra domain A (EDA) of fibronectin [285, 286], heparan sulfate [287], biglycan [288] and decorin [289]. Citrullinated proteins are of particular interest due to their presence and persistence in a variety of autoimmune disorders [290]. On ligand binding, TLR4 dimerises and recruits a host of adaptor proteins to activate a signalling cascade resulting in the production of pro-inflammatory cytokines, chemokines and tissue degrading enzymes, and inducing the upregulation of MHC II and co-stimulatory molecules on antigen-presenting cells that enhances the antigenic signal to T cells, thus generating an inflammatory reaction and initiating adaptive immune responses [291]. Two distinct signalling pathways can occur, depending on whether the adaptor molecule TIR-domain-containing adapter-inducing interferon-β (TRIF) or the adaptor protein myeloid differentiation primary response 88 (MyD88) is recruited. However, both pathways converge with the production of NF-kB and its translocation to the nucleus to induce the transcription of other inflammatory genes [292-295]. The MyD88-dependent pathway generally results in the production of proinflammatory cytokines, whilst the MyD-88-independent pathway results in proinflammatory cytokine production as well as IFN secretion that is vital for immune cell maturation. In this way, the ECM contributes to gene regulation during the inflammatory response.

Macrophages are specialised cells that are able to detect and phagocytose bacteria, as well as present antigens to T-cells and initiate the inflammatory response by producing cytokines that activate and recruit other immune cells. These cells, along with neutrophils and other dendritic cells can be seen as the first responders of the innate immune response. Macrophages constantly sample the ECM microenvironment in every tissue, patrolling for foreign entities or tissue damage, and, on detection, orchestrate the initiation and resolution of an inflammatory response by secreting pro-inflammatory

cytokines and recruiting other cells [296]. Macrophages are the final lineage in the mononuclear phagocytic system (MPS) differentiating from progenitor monocytes, which originate in the bone marrow and from there circulate in the blood for approximately two days and, if they are not recruited to tissues, apoptosis occurs. Maturation of monocytes into macrophages occurs by exposure to macrophage colony-stimulating factor (M-CSF), which is secreted ubiquitously and constitutively by a variety of cell types, including fibroblasts, stromal cells, endothelial cells, smooth muscle cells and osteoblasts [91]. Circulating levels of M-CSF are regulated by macrophages, which selectively bind M-CSF, via their M-CSF receptor (M-CFS-R), leading to intracellular degradation of M-CSF. This provides a self-regulating mechanism of mature macrophage population [297]. Macrophages are identified by the expression of the markers cluster of differentiation (CD)-11b, -14, -68 and -206 [298]. CD68 is associated with the lysosome compartments of macrophages and is commonly used to identify macrophages [299]. CD11b is an integrin alpha chain member, whilst CD14 is necessary for recognition of LPS [300] and CD206 is a transmembrane protein known as mannose receptor which acts as a pathogen recognition receptor (PRR) able to recognise a broad range of microbial and endogenous ligands [301]. The latter two markers can discriminate between monocytes and macrophages, with CD206 expression associated with 'M2' polarised or alternatively activated macrophages [302].

Macrophages are classified into subpopulations depending on their location and two main groups depending on their activation state: classically activated (M1) pro-inflammatory macrophages resulting from activation by IFN-γ and/or lipopolysaccharide and alternatively activated (M2) anti-inflammatory macrophages stimulated by IL-4 and IL-13. These two classifications term the defence and resolution states of the inflammatory response, respectively. However, this classification is too arbitrary for application *in vivo* due to the complex and varied secretion of cytokines that ultimately decide the plasticity and final macrophage state [298]. Indeed, macrophages have been shown to switch activation states depending on cues from the microenvironment [303, 304]. Whilst M-CSF is a vital signal regulating the populations of monocytes and resident tissue macrophages, granulocyte-macrophage colony-stimulating factor (GM-CSF) alters the activation and polarisation status of macrophages. GM-CSF production requires a stimulus such as infection or LPS in order to be produced during an

inflammatory state, usually by leukocytes [305]. The activation of macrophages by exogenous pathogenic ligands, ECM matrikines and soluble molecules is a tightly controlled process for the generation of an appropriate inflammatory response.

1.5.1 Tenascin- C in the inflammatory response

The role of TN-C in inflammation associated with wound healing and disease is well documented. For instance, TN-C is overexpressed in the stroma of most solid cancers [185, 306], is transiently expressed after human myocardial infarction [307] and in mouse models of myocardial infection [308], and is elevated in autoimmunity and chronic inflammation (e.g. rheumatoid arthritis) [224, 309]. Activation of TLR4 upon infection and sterile inflammation induces rapid TN-C expression through the PI3K/AKT and NF-kB signalling pathways [224]. TN-C is then directly involved in the global inflammatory response acting as a proinflammatory molecule.

1.5.2 Tenascin-C and inflammatory cytokine expression

The inflammatory effects of TN-C are mediated by three main binding partners: TLR4, integrin α9β1 and integrin α5β3. It was demonstrated that TLR4 is activated specifically by three distinct epitopes within the FBG domain of TN-C that resulted in the induction of a host of inflammatory cytokines such as interleukin-6 (IL-6), interleukin-8 (IL-8) and TNF-α in macrophages and IL-6 in human synovial fibroblasts, in a MyD88-dependent manner [235, 310]. This inflammatory cytokine profile induction not only occurs in human macrophages [235, 311], but also in mouse macrophages [312]. Further to this, TN-C-mediated activation of TLR4 has been demonstrated to induce IL-6, IL-8 and TNF-α in many cell types including primary human chondrocytes [313], mouse macrophages [314] and human cardiac myofibroblasts [215] using a variety of methods, including TLR4 knockout mice and antibodies against TLR4 or MyD88.

Further studies using mutant FNIII3 domains and integrin antibodies also revealed inflammatory cytokines were produced in response to the FNIII3 domain of TN-C binding α9β1 integrin in mouse synovial macrophages [315], mouse synovial fibroblasts [315] and mouse dendritic cells [316].

Signalling following binding of α9β1 integrin by TN-C also results in the expression of IL-6, IL-1α, and the chemokines CCL2, CCL4, and CXCL5. It was also was deduced that synthesis of a IL-6, IL-1β, and TNF-α triad occurred in response to TN-C binding αvβ3 integrin on mouse peritoneal macrophages in an NF-kB-dependent manner [317]. However, it is unclear whether either of the known binding regions (FNIII3 [318] of FGB[319]) are responsible for this proinflammatory activity. It is clear that TN-C is able to activate a variety of receptors to generate a relatively homogenous cytokine release profile.

More recently, and further to this specific cytokine profile, TN-C/TLR4 signalling has been linked to IL-1β synthesis via inflammasome priming [320]. Inflammasomes are intracellular, multiprotein complexes that are directly activated following detection of intracellular PAMPs or indirectly following PRR signalling [321]. The inflammasome is synthesised and assembled (including caspase-1) following PRR activation that results in maturation of pro-IL-1β. This proinflammatory cytokine can also induce TN-C expression; stimulation of mouse retinol endothelial cells with IL-1β and TNF-α is associated with an upregulation of TN-C [322]. In human articular cartilage, stimulation with IL-1 also resulted in an increase of TN-C expression in bovine cartilage explants [313]. After myocardial infarction, rat epicardium cells displayed elevated TN-C levels. Stimulation of the epicardium cells with TN-C was found to activate TLR4 and the inflammasome resulting in an upregulation of pro-IL-1β production [320]. Furthermore, in mouse dendritic cells stimulated with TN-C, an upregulation of mature secreted IL-1β was found [323]. This demonstrates the existence of a pro-inflammatory loop in which the inflammasome generates mature IL-1β, which in turn induces TN-C expression, which is then able to promote further IL-1β expression. Whilst no direct studies of the role of the TN-C on the inflammasome in humans have been carried out, these studies in mice likely provide clues as to the role of TN-C in the inflammatory response in humans and indicate a mechanism by which inflammation is propagated.

Crucially, these cytokines result in the recruitment, maturation and activation of innate immune cells. TN-C deficient mice display a diminished ability to steer T-cells towards Th17 polarization due to reduced proinflammatory cytokine synthesis [324]. This effect was resolved by the addition of

exogenous TN-C, and reverted and perpetuated by anti-IL-6 and anti-TLR4 antibodies [323]. The weakened capability to drive T-cells towards the highly proinflammatory Th17 polarisation may be explained by the ability of TN-C to promote the production of a known T-cell polarisation compound and mediator of inflammation - prostaglandin E2.

Other studies have highlighted a different role for TN-C-induced activation of TLR4 in stromal cells, in which TLR4 activation of fibroblasts stimulated synthesis of collagen type 1 and alpha-smooth muscle actin (α-sma) [325]. In human macrophages *TNC*-induced TLR4 activation also stimulated collagen production [311]. Whilst the induction of collagen contributes to fibrosis, this axis may also contribute to wound healing.

The synthesis of ECM components arising from TLR4 activation appears to be balanced with the destructive activities of MMPs, also induced upon receptor activation. TN-C-induced TLR4 activation results in an increase in expression of the collagenase MMP-9 in mouse bone marrow neutrophils [326] and mouse synovial fibroblasts [315], the collagenase MMP-1 and the membrane-type MMP MMP-14, which can activate MMP-2, in human macrophages [311], and MMP3 in human fibroblasts [327].

Just as TN-C can have different functions depending on cell type and location, in addition to its proinflammatory attributes, TN-C is also crucial for the resolution of inflammation. Normally, the induction of TN-C expression is transient, whilst prolonged expression is associated with chronic inflammatory and autoimmune diseases such as rheumatoid arthritis [224, 309] [313] and systemic sclerosis [325]. Correspondingly, mice that lack TN-C are protected from chronic inflammation [235]. However, it is unclear whether an absence of TN-C simply creates a less proinflammatory environment or whether it directly contributes to the resolution of inflammation. It is known that TN-C can directly bind TGF-β [328], possibly sequestering it into the matrix and providing a concentration gradient of this cytokine, localised simultaneously with TN-C in regions of inflammation. Integrins are able to process TGF-β into its active form; it could be that TN-C facilitates this interaction by bringing TGF-β into close proximity to integrins, given both TN-C and TGF-β LAP have the ability to bind αvβ3 [329]. TGF-β has a central yet pleiotropic immunomodulatory role. It can propagate the inflammatory

response depending on the presence or absence of specific cytokines, but is also able to act in an anti-inflammatory fashion (reviewed in [330]). In *in vitro* studies using human peripheral blood monocytes, TN-C was shown to be immunosuppressive and halt T-cell activation [331]. This has been reproduced in mouse models of cancer [332]. It was later established, using recombinant and mutant TN-C, that the domain responsible for this was the FNIII A1A2 domain [333].

1.6 THE ECM AND GENE REGULATION

The microenvironmental context of cells and its impact on gene regulation are becoming increasingly studied. The role of mechanotransduction in TN-C gene regulation was discussed earlier. In the same vein, gene regulation at the transcriptional level by ECM stiffness is a prominent research area. Cells sense the physical properties, with particular regard to force, of the ECM microenvironment and are able to regulate and maintain these properties, else pathologies and disease occur [334]. This involves integrins, which bind both the actin cytoskeleton and ECM proteins, and subsequent activation of downstream signalling pathways [335]. Matrix stiffness has been shown to exert behavioural changes in smooth muscle cells with softer substrates resulting in increased motility, less cell spreading and less actin stress fibre formation, than stiffer substrates. Treatment with a Rho/ROCK inhibitor ablated these effects indicating ECM-regulated Rho-mediated cell contractility [336]. Increased spreading was seen in fibroblasts cultured on stiff substrates with a concurrent increase in DNA synthesis and reduced rate of apoptosis. The authors suggested that the traction induced in the stiffer mechanical environment resulted in increased signalling at focal adhesion sites, increased cell spreading and subsequent regulation of cell shape, which in turn regulates cell cycle progression [337]. However, no specific genes were investigated here. The role of stiffness in controlling apoptosis is likely due to the necessity of cell attachment for survival [335]. Further experiments by Meredith Jr *et al.* treating rat epithelial cells and mouse fibroblasts with tyrosine phosphatase and myosin inhibitors indicated spreading and motility were ECM-regulated in a myosin and phosphotyrosine dependent manner [338]. It was later proved that this is indeed the case, at least in epithelial cells, mouse fibroblasts, mouse osteoblasts and mouse vascular smooth muscle cells, whereby ECM stiffness resulted in activation of the small GTPase Rac1 and induction of cyclin D [339].

Other, more general studies have shown that increased matrix stiffness promoted adenomatous polyposis coli (APC) protein dependent RNA localization in highly contractile protrusions of migrating fibroblasts, by activating Rho-mDia1 signalling pathway that affects microtubule networks [340]. Recently, it has been demonstrated that during inflammation- or injury-induced regeneration of the intestinal crypt, collagen transcriptionally induces expression of stem cell antigen-1 (Sca1/Ly6). These cell surface proteins are usually absent in adult epithelium, thus collagen results in reprogramming adult intestinal epithelial cells to a foetal intestinal phenotype during repair. This was shown to be regulated via the focal adhesion kinase (FAK), the proto-oncogene tyrosine-protein kinase (Src), and the Yes-associated protein 1 (YAP)/ WW-domain-containing transcription regulator 1 (TAZ), indicating these proteins and the HIPPO pathway are involved in the transcriptional effects of matrix stiffness [341].

More specifically, in knockout mouse models, elastin fibers with their associated glycoproteins, promote smooth muscle cells maturation from a migratory and proliferative phenotype associated with development, to a contractile, quiescent phenotype, which is necessary for proper artery/vascular function [342]. Expression of MMPs, TIMPs and collagen type III is also differentially regulated in fibroblasts depending on collagen substrate stiffness [343].

A point to note is that these *in vitro* experiments were carried out using synthetic substrates like hydrogels or polyacrylamide gels coated with collagen or fibronectin with cells seeded at low density to remove any cell-cell interaction effects. These do not typically represent physiological substrates even when experiments are carried out with stiffness in range of specific physiological tissues. Although authors declare that only the property of stiffness was altered, it has been demonstrated that this is not an independent variable as crosslinking density can alter matrix porosity and substrate anchoring [344, 345]. Furthermore, biological tissues stiffen as they are strained, unlike synthetic polymer gels [346].

It is clear that the ECM is able to orchestrate direct changes on cell behaviour and fate by modulation of tissue and cytoskeleton architecture, and the subsequent modulation of signalling cascades. However, the specific mechanism by which this occurs, particularly in a 3D tissue context, is still elusive. More recently, it has been suggested that the ECM–mediated cytoskeleton rearrangements can alter the

arrangement of nuclear components and localisation of transcription factors, which are located in bodies within the nucleus, to promote the transcription of 'ECM-response elements' [347, 348]. Evidence has been presented which links organ and tissue architecture rearrangement with nuclear and chromatin organisation. For example, the cytoskeleton is linked across the nuclear envelope to the nuclear lamina by bridge proteins consisting of highly conserved eukaryotic SUN (Sad1 and UNC-84) and KASH (Klarsicht, ANC-1 and Syne/Nesprin homology) proteins [349]. A number of studies have demonstrated dynamic and reciprocal, physical and biomechanical interaction between the nucleus and ECM. Very early studies utilised reconstituted basement membrane (BM) ECM as a substrate for the culture of human mammary epithelial cells. They induced acinar morphogenesis in culture, in which the cells formed 'berry' structures similar to those found natively *in vivo*. Furthermore, these acini secreted a basement membrane-like ECM in addition to the exogenous BM ECM [350]. The authors used the same model to induce acini morphogenesis and probed nuclear rearrangement by assessing the localisation of the nuclear matrix proteins and the cell cycle regulator retinoblastoma (Rb) protein. They found these proteins were localised in distinct regions dependent on the status of differentiation into acini [351]. Alteration of histone acetylation in the acini altered chromatin structure and induced cell proliferation; acini cells do not proliferate indicating a reversal of differentiated phenotype. Furthermore, blocking structural nuclear protein foci formation with antibodies abrogated basement membrane synthesis from cells, as well as increased ECM destruction via increased MMP secretion. This study demonstrated that chromatin structure upon acini formation was critical for cell differentiation maintenance and tissue architecture maintenance. Taken further, an ECM-response element was identified in the promoter region of the mammary cell specific milk precursor β-casein gene [352], which is associated with differentiated mammary epithelium. Modulation of chromatin structure by drug treatment or with methylation modulators revealed that the expression of β-casein and acinar differentiation were dependent on chromatin structure in the presence of the basement membrane, its component laminin in particular [353, 354]. Together, these studies demonstrate a direct link between the ECM, chromatin structure and gene expression.

Gene regulation by the ECM at the posttranscriptional level is emerging. Primary fibroblasts from idiopathic pulmonary fibrosis (IPF) patients or healthy lung tissue were cultured on decellularised lung ECM from IPF or healthy patients. Notably the IPF ECM origin resulted in the upregulation of IPF associated ECM components, compared to control tissues, indicating a self-sustaining ECM-driven fibrosis. This upregulation was the result of IPF ECM-mediated miR-29 reduction [355].

1.6.1 microRNAs, the ECM and gene regulation

microRNAs (miRNAs) are small 18-25 bp single stranded ribonucleic acid sequences that, once incorporated into the RNA-induced silencing complex (RISC), posttranscriptionally regulate gene expression via binding to complementary 'seed' sequences in the 3' (and less frequently the 5') UTR of mRNAs [356]. miRNAs are first transcribed as longer transcripts (pri-mRNAs) which are then processed by the microprocessor complex, comprised of Drosha, a ribonuclease III, and its cofactor DGCR8 Microprocessor Complex Subunit (DGCR8), into a smaller pre-mRNA ~70 nt sequence [357]. A subsequent cleavage step by the endoribonuclease Dicer processes the pre-mRNA into a smaller duplex, of which one strand is selected for incorporation into the RISC. This is a finely tuned process that is tightly regulated at each step [358]. Whilst the exact mechanism of silencing is still debated, gene silencing by miRNA binding is hypothesised to occur in one of two ways. Perfectly complementary miRNA binding to mRNA is thought to result in mRNA degradation, whilst partial complementarity results in translational repression. In this way, miRNAs play a key role in the regulation of gene expression [359] with up to 1/3 of genes being regulated by miRNAs [360], including many ECM proteins [359]. Recent studies are now focussing on the reciprocal gene regulation of ECM components by miRNAs and miRNAs by ECM components.

By regulating gene expression, miRNAs have a broad range of function in development, homeostasis and disease and are able to regulate virtually every cell process. Recently miRNAs have been shown to regulate organogenesis. miR-17 has been shown to repress the expression of fibronectin and fibronectin type III domain containing 3A (FNDCA) *in vitro* and in transgenic mice, in which miR-17 expression has profound effects on growth and organ size. This miRNA was suggested to have a panel of target

genes [361]. However, as discussed earlier, fibronectin can bind integrins, which confer its abilities in regulating many aspects of cell behaviour. Furthermore, versican expression is restricted specifically to the atrioventricular canal in zebrafish heart by action of miR-138, thereby controlling cardiomyocyte differentiation [362].

The ECM, a contributing factor to the tumour microenvironment, also plays a role in the pathogenesis and metastasis of cancer with its effects on cell proliferation, adhesion and migration. The tumour ECM is usually stiffer due to abnormal matrix deposition of collagen and other ECM proteins. These different compositions augment metastasis. For instance, it was shown that matrix stiffness induced miR-18a expression in breast cancer tissue, subsequently decreasing the levels of a direct target, the tumour suppressor protein phosphate and tensin homolog (PTEN), thus driving tumour progression [363]. Mir-18 was also able to supress PTEN activity indirectly via decreasing Homeobox protein Hox-A9 (HOXA9) levels. This suppression occurred by increased collagen deposition, which in turn resulted in a stiffer ECM in a LOX dependent manner. The increased collagen deposition led to increased integrin activation (the main collagen receptor), clustering and subsequent FAK and PI3K signalling that culminated in the expression of the transcription factor β-catenin which then induced transcription of miR-18. This pathway is able to progress in a self-propagating loop in which the ECM drives miRNA gene expression [363]. In a similar vein, miR-301 also promotes tumour progression in breast cancer by repressing PTEN, and by downregulating collagen type II alpha chain 1 (COL2A1) expression [364]. In a model of lung fibrosis post-transcriptional upregulation of oncogenic miR-21 has been shown to correlate positively with collagen I deposition and reciprocally collagen I is able to upregulate miR-21 expression [365]. Multiple collagens, COL1A1, COL1A2, COL3A1, COL4A1, COL4A2 and COL15A1, and laminin γ1 have been shown to be downregulated by miR-29c. This miRNA is usually downregulated in nasopharyngeal carcinomas and an increase of these proteins is associated with metastatic potential [366]. These findings of miR-29c regulating COL1A1 and COL3A1 were confirmed in bleomycin-induced mouse models of fibrosis, which features an excessive accumulation of collagen and other ECM components, and biopsies and fibroblasts from systemic sclerosis patients [367].

Many miRNAs indirectly regulate the expression of ECM genes by regulating MMPs. mir-29b directly targets MMP-2 which acts to supress tumour angiogenesis, by impairing VEGFR-2 signalling [368]. miR-328-3p inhibited the proliferation and migration and enhanced the apoptosis of osteosarcomas by targeting MMP-16 [369]. Furthermore, miR-616 acts to promote invasiveness of tumour cells by targeting TIMP2, an inhibitor of MMP-2 [370]. During mammary gland development, miR-212/132 is specifically expressed in the mammary gland region to target MMP-9 which would otherwise interfere with collagen deposition [371].

miRNAs have also been shown to regulate integrin expression. The miR-29 family (a, b and c) is able to regulate the expression of integrin β1 in head and neck squamous cell carcinoma. miR-29 therefore functions as a tumour suppressor protein by inhibiting integrin oncogenic signalling [372]. In cancer, integrin α6β4 has been shown to regulate the expression of miR-29a, which is able to regulate a variety of molecules. miR-31 acts to regulate several α subunits of β1 integrins and β3 integrins. Reduction in expression of these proteins, resulting from miR-31 targeting conserved seed regions in the 3' UTR of these integrin α subunit gene transcripts, led to posttranscriptional repression. This resulted in a reduction of cell spreading, reducing tumour metastasis [373]. In line with this, miR-93 induced down-regulation of β1 integrin subunit promoted tumour growth and angiogenesis, whilst overexpression of miR-93 promoted angiogenesis and tumour growth [374]. Fibronectin is heavily involved in the regulation of migration due to its integrin binding properties. In line with this, upregulation of miR-143 downregulated the fibronectin type III domain containing 3B (FNDC3B) resulting in enhanced hepatitis B virus-related hepatocarcinoma metastasis [375]. Clearly, the role of miRNA expression in regulation of the ECM is a new field in which the role is becoming appreciated. The exact molecular mechanisms are still to be elucidated, however miRNAs are promising therapeutic targets.

In inflammation, the ECM has been linked to miRNA expression [360]. The SLRP decorin has been demonstrated as a regulator of miR-21 abundance in LPS-induced sepsis and sterile inflammation during tumour growth [289]. This occurs as decorin acts as an endogenous ligand for ***TLR2*** and TLR4, resulting in the production of the pro-inflammatory molecule programmed cell death 4 (PDCD4).

Decorin simultaneously inhibits translational repression of PDCD4 by decreasing the abundance of TGF-β and, subsequently, the translational inhibitor of PDCD4, miR-21.

1.6.2 miR-155 regulation by Tenascin- C

In the mouse, the ECM glycoprotein TN-C regulates miR-155 production thereby mediating inflammatory gene expression at the posttranscriptional level [188]. miR-155 is an exceptional miRNA in that it can enhance gene expression of TNF-α. One proposed mechanism for this enhancement of gene expression is that miR-155 stabilises mRNA transcripts [376]. Another potential mechanism involves the self-inhibitory action of the TNF-α 3' UTR as suggested by Tili *et al*.[377], but binding of the miRNA to the 3' UTR has not been demonstrated [378].

Piccinini *et al*. [188] established a key role for TN-C in the regulation of miR-155. *TNC*-/- mice were found to exhibit an impaired inflammatory response when injected with LPS, with significantly less expression of miR-155 and concurrent lower circulating levels of the pro-inflammatory cytokine TNF-α compared to *TNC*+/+ mice. This impaired phenotype was rescued by engraftment of TN-C expressing bone marrow cells. Furthermore, *in vitro*, LPS-induced TNF-α expression in mouse bone marrow-derived macrophages (BMDMs) lacking *TNC-/-* could be rescued by overexpression of miR-155. However, the same effect was not seen when *TNC-/-* BMDMs were exposed to soluble TN-C, which represents TN-C secreted by the liver in biological fluids during disease that is not incorporated into the ECM. Addition of TN-C at concentrations comparable to those measured in the conditioned media of *TNC*+/+ BMDMs after LPS stimulation was unable to rescue the phenotype, whilst higher concentrations inhibited cytokine synthesis. Both recombinant TN-C [312] and conditioned medium from HEK293 cells overexpressing TN-C (unpublished data) were used to address potential protein folding and thus activity issues. This suggests that TN-C must probably be in its ECM-associated form to exert its regulatory function on miR-155, highlighting the need for an *in vitro* model which would allow the investigation of the function of TN-C in gene regulation in the context of the ECM. Such model would not only help to determine whether TN-C exerts its regulatory effect in its soluble or

ECM-associated form, but also address the question as to whether this regulatory axis occurs also in human cells.

The lack of understanding of the molecular processes surrounding gene regulation by the ECM is partially due to the use of traditional two-dimensional (2D) culture methods failing to recapitulate the native microenvironment. However, reconstituting the ECM of the *in vivo* extracellular microenvironment *in vitro* is challenging due to the highly dynamic and varied nature of the ECM [379].

1.7 CURRENT MODELS TO STUDY HOW THE EXTRACELLULAR MATRIX INFLUENCES CELL FUNCTION

Many models are currently available to study the effect of the ECM microenvironment on cell function. Traditional models rely on 2D cell culture on flat and rigid surfaces such as tissue culture dishes (Figure 10 A). The limitations of these are increasingly being recognised. For example, a comparison of 60 cell lines and tissue samples, highlighted differences in gene expression from the corresponding cell line to its matching tissue, demonstrating the pitfalls in using non physiologically relevant models [379]. The majority of clinical trials failing has been attributed to misleading data from 2D monolayer culture systems in which the cellular response is unnatural due to the synthetic microenvironment [35]. Thus, there has been a recent effort to establish 3D culture methods that mimic the *in vivo* environment.

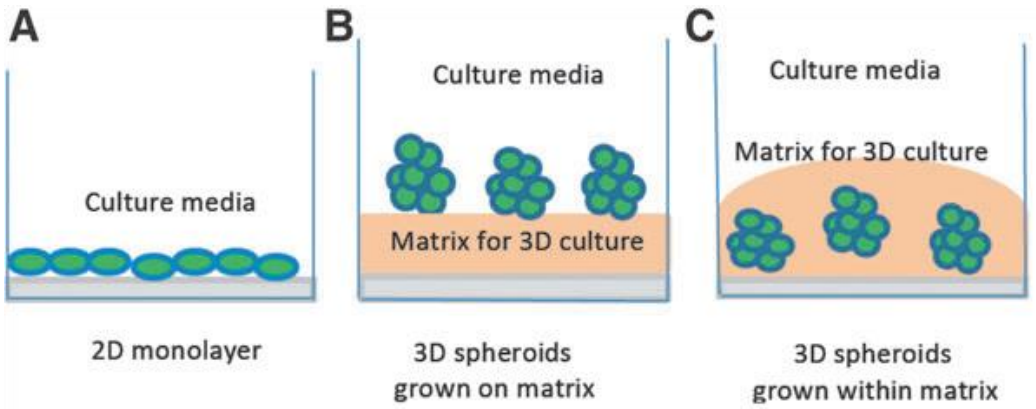


Figure 10. Schematic diagrams of culture methods.
A) Traditional 2D monolayer cell culture on a flat tissue culture dish that is not representative of any cells in vivo given that cells grow in a monolayer and are not in contact with any ECM proteins. **B)** Cells seeded on a cell derived matrix (CDM) – more relevant than traditional culture methods. **C)** Cells resuspended within a 3D CDM before it gelates – most appropriate to study how the ECM microenvironment influences cell function. The latter two methods result in cells making contacts with each other and ECM proteins in a 3D meshwork with the aim of recapitulating the ECM microenvironment. Adapted from Edmondson et al. [35].

1.7.1 Commercially available 3D culture methods

Many commercial models exist for studying the ECM microenvironment (Table 1). However, these are made of single or multiple components isolated from animals or tumours or are from non-biological sources and therefore do not closely mimic human tissue [380].

One currently available commercial model is Matrigel/Cultrex ™. This is a soluble extract of basement membrane proteins including laminin, type IV collagen, heparan sulfate, nidogen, unspecified proteases, but likely those usually expressed in cancers, growth factors TGF-β, FGF, EGF, and PDGF that is composed of natural components derived from the mouse Engelbreth-Holm-Swarm (EHS) tumour [381]. This substrate has been utilised with great success for the culture of stem cells, which, in conjunction with fibroblast conditioned media, can be used to replace the feeder layer of fibroblasts which usually generate an ECM substrate [382]. Furthermore, when these cells are injected into mice, all three germ layers successfully form. Matrigel has also been used to direct cell fate. For example, bone marrow derived stem cells cultured on Matrigel with FGF-4 and hepatocyte growth factor (HGF) exhibit characteristics functionally, morphologically and phenotypically similar to hepatocytes [383]. Further control of the microenvironment has been demonstrated in a study in which Matrigel can exhibit

an ‘aged’ microenvironment by deriving the components from tumours grown in aged mice or enriched for cartilage components, thus named ‘cartrigel’ [384].

However, a major drawback of this ECM model is its tumour origins; this ECM model is mainly suited for cancer studies and not for studies aiming to assess normal, healthy tissue. Further criticism comes from the fact that proteomic analysis revealed Matrigel™ contains an undefined amount of growth factors which may result in variability of results, a particular issue between batches [385]. This has been overcome in part by a further modified version which has been manipulated to be void of growth factors by the addition of a 20% NaCl precipitation step [386]. Even with the modification this partly undefined and tumour-derived murine ECM may not provide results predictive for healthy human tissue.

3D collagen gels have also been favoured as a model, particularly in studies of mechanotransduction [387]. As the name suggests, collagen gels consist of collagen only, usually of rat tail origin, and require a straight forward protocol of polymerisation and gelation. In such gels it is very difficult to modulate features of the microenvironment independently, for example modulating stiffness of collagen gels requires an alteration of density and pore size [388], which will ultimately affect both the stiffness and the structure. The polymerisation step of collagen gels also provides an opportunity for variation, with the temperature of polymerisation drastically affecting the organisation of fibrils [387]. Whilst this simplicity of collagen gels may aid in the convenience and ease of lab experiments, this simplistic model is not representative of the intricate and complex molecular structure and organisation of native *in vivo* ECM. Such models composed of one ECM component, including collagen, fibrin or hyaluronan, from synthetic or natural sources fail to recapitulate the organisational and molecular complexity of native ECM [389].

Table 1. Current available models to study the ECM microenvironment.
A limitation of many models is that components are not of human origin and thus elicit immunogenic responses. Other models are derived from tumorous origins and therefore not representative of healthy ECM. Adapted from Badylak *et al.* [390].

Brand	Manufacturer	Organism derived from	Natural or synthetic	Format
AlloDerm	Lifecell	Human skin	Natural	Dry sheet
AlloPatch®	Musculoskeletal Transplant Foundation	Human fascia lata	Natural	Dry sheet
Axis™ dermis	Mentor	Human dermis	Natural	Dry sheet
Bard® Dermal Allograft	Bard	Cadaveric human dermis	Natural	Dry sheet
CuffPatch™	Arthrotek	Porcine small intestinal submucosa (SIS)	Cross-linked	Hydrated sheet
DurADAPT™	Pegasus Biologicals	Horse pericardium	Cross-linked	Dry sheet
Dura-Guard®	Synovis Surgical	Bovine pericardium	Cross-linked	Hydrated sheet
Durasis®	Cook SIS	Porcine small intestinal submucosa	Natural	Dry sheet
Durepair®	TEI Biosciences	Fetal bovine skin	Natural	Dry sheet
FasLata®	Bard	Cadaveric fascia lata	Natural	Dry sheet
Graft Jacket®	Wright Medical Tech	Human skin	Natural	Dry sheet
Oasis®	Healthpoint	Porcine small intestinal submucosa	Natural	Dry sheet
OrthADAPT™	Pegasus Biologicals	Horse pericardium	Cross-linked	Dry sheet
Pelvicol®	Bard	Porcine dermis	Cross-linked	Hydrated sheet
Peri-Guard®	Synovis Surgical	Bovine pericardium	Cross-linked	Dry sheet
Permacol™	Tissue Science Laboratories	Porcine skin	Cross-linked	Hydrated sheet
PriMatrix™	TEI Biosciences	Fetal bovine skin	Natural	Dry sheet
Restore™	DePuy	Porcine small intestinal submucosa	Natural	Dry sheet
Stratasis®	Cook SIS	Porcine small intestinal submucosa	Natural	Dry sheet
SurgiMend™	TEI Biosciences	Fetal bovine skin	Natural	Dry sheet
Surgisis®	Cook SIS	Porcine small intestinal submucosa	Natural	Dry sheet
Suspend™	Mentor	Human fascia lata	Natural	Dry sheet
TissueMend®	TEI Biosciences	Fetal bovine skin	Natural	Dry sheet
Vascu-Guard®	Synovis Surgical	Bovine pericardium	Cross-linked	Dry sheet
Veritas®	Synovis Surgical	Bovine pericardium	Cross-linked	Hydrated sheet
Xelma™	Molnlycke	ECM protein, PGA, water	Cross-linked	Gel
Xenform™	TEI Biosciences	Fetal bovine skin	Natural	Dry sheet

Zimmer Collagen Patch®	Tissue Science Laboratories	Porcine dermis	Cross-linked	Hydrated sheet

These problems have led to the development of synthetic hydrogels. A popular choice is poly(ethylene glycol) (PEG) hydrogels which contain RGD integrin binding sites and thus confers capability to interact with many ECM proteins [388]. Utilising such synthetic models allows for greater control over modulation of the composition than natural counterparts and also may enable the identification of the specific components within the model that promote change in cell behaviours. For example, epithelial cell migration can be controlled by altering the susceptibility of the hydrogel to MMPs and by adding adhesion ligands [391]. This model was used to demonstrate angiogenesis using an *ex vivo* aortic arch explant assay, indicating that hydrogels are a functional model to study the microenvironment in cancer niches. However, this is still a synthetic model and therefore results should still be interpreted cautiously when being applied to human situations.

More recently, models are being generated in which the ECM components are derived from natural sources such as native tissue. In the field of regenerative medicine, and tissue engineering in particular, there is an increasing use of 3D scaffolds derived from decellularised organs and tissues, including rat heart [392], rat liver [393] and rat lungs [394]. These are decellularised by perfusion with detergents, which preserves architecture, and thus functionality of the ECM is maintained when successfully transplanted back into rats. Decellularisation is an important process as antigenic epitopes (antibody binding sites) remain on cell surfaces that would promote an immunologic response if the matrix is implanted in the same species (allogenic transplant) or a different species (xenogenic transplant) [390]. Abnormal immune activation by such epitopes could give misleading results during experimental studies. Again, modulation of these matrices here from their natural structure is difficult, thus lending its ultimate application to transplants as opposed to a model for cell biology research.

1.7.2 Cell derived matrices (CDMs)

Alternative, more physiologically relevant models are ‘cell derived matrices’ (CDMs) (Figure 10 B, C). Recapitulating the molecular complexity and organization of native tissue ECMs is difficult utilising

synthetic methods. This has motivated the use of native ECM itself as a biomaterial source. In this approach, matrices are derived from human cells such as fibroblasts, which deposit collagen rich matrices that are stable enough to allow for decellularisation and downstream applications. This represents a promising model, bypassing problems with modulation of specific components as CDMs facilitate customisation, including the choice of cell type generating the ECM, the presence or absence of a scaffold during culture and the option of genome editing of the cell generating the ECM [389]. However, generating cell derived matrices involves several important considerations which can affect the outcome of the ECM.

1.7.2.1 Cell types to deposit CDM

The first and arguably most important choice for generating CDM is the type of cell that will be exploited to produce matrix components. Cells express ECM components which would be present in their native environment or niche within the body. For example, mesenchymal stem cells deposit matrices more representative of either bone, cartilage or adipose tissue depending on their culture method, whilst fibroblasts produce a collagen rich matrix, more representative of vascular and dermal tissue. Cells can then be cultured as an adherent monolayer or as a monolayer on a scaffold which may aid a more 3D structure (Figure 10 B, C). Some scaffolds may be biodegradable thus enhancing biocompatibility [389].

Many inherent problems arise around cell choice. Primary cells are regarded as the best option as these cells most closely resemble the *in vivo* phenotype with regard to ECM production [389]. However, in some instances, acquiring large enough numbers of primary cells to deposit ECM can be difficult therefore requiring cell expansion *in vitro.* A primary problem of long-term expansion of cells is differentiation of the cells. Culturing cells *ex vivo* selects for faster expanding subpopulations, eventually leading to a genetic drift, cell differentiation and cells varying from their original phenotype. Whilst genetic drift in *in vitro* populations is debated, it has been demonstrated in serially passaged bovine chondrocytes which were allowed to deposit ECMs from passage (P) 0, 2 and 6. The ECMs from chondrocytes at P6 were vastly different to those deposited by P0 chondrocytes and composed of

much less total protein compared to P0 chondrocyte-derived matrices [395]. To circumvent this problem, immortalised cell lines may be used. However, this is not appropriate in all cases, as immortalised cell lines are usually tumour derived and differ substantially from the primary native phenotype of the cell [389].

A further aspect is that the cells utilised to derive matrix can be healthy or from disease origin. This would be particularly useful for rare or complicated disorders which are not easily modelled in mice models. The varied choice in cell type to generate CDMs allows the generation of many types of human ECM which will facilitate the generation of physiologically relevant ECM models to study disease pathogenesis.

1.7.2.2 Decellularisation of CDMs

Another important consideration when developing and utilising CDMs is the processing method for decellularisation (Table 2). A main drawback is that decellularisation can potentially damage the ECM. It has been demonstrated that the decellularisation process using ionic detergents can remove important ECM components such as GAGs and soluble collagen [396, 397]. Alternative methods such as mechanical action and snap freezing to remove cells can also damage the ECM, whilst enzymatic removal of cells can trigger an immunological response if the enzymes are not properly removed [397]. A further consideration is that denaturing agents such as sodium dodecyl sulphate are able to activate secreted pro-MMPs (without proteolysis) by disrupting their 'cysteine switch' [398]. This highlights the variation that may occur between cell derived matrices depending on the decellularisation process.

Table 2. Common methods of decellularisation methods and associated limitations.
Many physical and chemical methods exist, however triton removal is demonstrated as the most efficient. From Gilbert *et al.* [397].

	Mode of action	**Effects on ECM**
Physical decellularisation		
Snap freezing	Intracellular ice crystals disrupt cell membrane	ECM can be disrupted or fractured; remnant DNA
Mechanical force	Pressure can burst cells	Mechanical force can damage the ECM; remnant DNA

Mechanical agitation	Cell lysis, but more commonly used to facilitate chemical exposure and cellular material removal	Aggressive agitation or sonication can disrupt the ECM; remnant DNA
Chemical decellularisation		
Alkaline; acid	Solubilizes cytoplasmic components of cells; disrupts nucleic acids	Removes GAGs
Non-ionic detergents		
Triton X-100	Disrupts lipid–lipid and lipid–protein interactions, while leaving protein–protein interactions intact	Mixed results; efficiency dependent on tissue; removes GAGs
Ionic detergents		
Sodium dodecyl sulfate (SDS)	Solubilizes cytoplasmic and nuclear cellular membranes; tends to denature proteins	Removes nuclear remnants and cytoplasmic proteins; tends to disrupt native tissue structure, remove GAGs and damages collagen
Sodium deoxycholate	Solubilizes cytoplasmic and nuclear cellular membranes; tends to denature proteins; usually used with other reagents such as DNase or Triton	More disruptive to tissue structure than SDS
Triton X-200	Solubilizes cytoplasmic and nuclear cellular membranes; tends to denature proteins	Yields efficient cell removal when used with zwitterionic detergents
Zwitterionic detergents decellularisation		
CHAPS	Exhibit properties of non-ionic and ionic detergents	Efficient cell removal with ECM disruption similar to that of Triton X
Sulfobetaine-10 and -16 (SB-10, SB-16)	Exhibit properties of non-ionic and ionic detergents	Yield cell removal and mild ECM disruption when used with Triton X-200
Solvent decellularisation		
Tri(n-butyl)phosphate (TnBP)	Organic solvent that disrupts protein–protein interactions	Variable cell removal; loss of collagen content; minimal effect on mechanical properties
Enzymatic decellularisation		

Trypsin	Cleaves peptide bonds at the C-terminal side of Arg and Lys residues	Prolonged exposure can disrupt ECM structure; removes laminin, fibronectin, elastin, and GAGs
Endonucleases	Catalyse the hydrolysis of the interior bonds of ribonucleotide and deoxyribonucleotide chains	Difficult to remove from the tissue and could invoke an immune response
Exonucleases	Catalyse the hydrolysis of the terminal bonds of ribonucleotide and deoxyribonucleotide chains	Not reported, generally not used.
Non enzymatic agents		
EDTA, EGTA	Chelating agents that bind divalent metallic ions, thereby disrupting cell adhesion to ECM	Typically used with enzymatic methods (e.g., trypsin) which can degrade proteins
Hypotonic and hypertonic solutions	Efficient cell lysis by osmotic shock	Does not effectively remove the cellular remnants

1.7.3 Advantages of using CDMs over traditional culture methods to maintain primary cell phenotypes.

It has been suggested that tissue culture systems lack a vital component of cells native microenvironment [399]. Interestingly, CDMs have been indicated as promising substrates for restoring *in vitro* differentiated chondrocytes to their native differentiated state, more so than differentiated chondrocytes cultured on plastic [400]. Rat chondrocytes were cultured for 4 passages on plastic plates inducing dedifferentiation, characterised by the loss of collagen II. When the cells were then transferred to chondrocyte CDM and cultured, the cells exhibited redifferentiation including increased collagen type II deposition. The same study also demonstrated that chondrocytes grown on CDM had a higher proliferation rate than those grown on plastic [400].

A common problem in mesenchymal stem cell (MSC) culture is the loss of differentiation potential and senescence when expanded *in vitro*. However, MSCs serially passaged on CDM-coated plates proliferated considerably faster than uncoated plastic plates [399]. Similarly, *in vivo* transplantation assays demonstrated that MSCs cultured on their own matrix retained their capacity for skeletogenesis

unlike those cultured on plastic that lost their potential after 6-7 passages [399]. Again, in MSCs it has been shown that aged (18 months) MSCs can have their osteogenic and proliferation potential improved to levels comparable with young (3 months) MSCs via culturing on CDM from young (3 months) bone marrow stem cells (BMC) [401]. These experiments reinforce the notion that CDMs recapitulate the native microenvironment, allowing the propagation of signals causing cells to retain their *in vivo* phenotype, more so than traditional culture methods.

1.7.4 CDM customisation to study individual protein function.

In reverse genetics the gene of interest can be knocked out, causing a cessation of protein production from this gene, to ascertain the function of the gene [402]. This same approach can be employed to customise CDMs by knocking out the ECM gene encoding for the protein of interest, in the cell producing the CDM. This 'protein-of-interest null' (KO) CDM and the wildtype CDM can be utilised as a substrate for cells, or, depending on the protein, sans cells, to determine the effect that a lack of the protein causes. This is advantageous over synthetic models which come in a 'gelate and use' format which are only customisable to a certain extend i.e., exogenous addition of molecules [381, 385].

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) - CRISPR associated protein (Cas9) is a powerful tool that has recently revolutionised the gene editing field by facillilating targeted gene editing and knockouts [403-410]. CDMs are deposited by cells over a minimum period of 10-14 days which makes CRISPR the most suitable approch given that this method produces stable knockouts. This is precisely the reason that many other traditional methods such as small interfering (si) RNA are not suitable; siRNA induces transient knockdowns of approximately 2-4 days or with some methods up to 7 days, which would not persist for the full CDM deposition period [411]. Another reason alluded to – generally siRNAs do not result in complete knockouts, only knockdowns albeit up to 80 % in some cases [412]. This method would likely result in some variation of the production and incorportation of the target protein across different CDMs.

A further method of customisation, mentioned prevously, that has the benefit of being a quick and easy approach for the overexpression of proteins (sans CRISPR) is the exogenous addition of recombinant

proteins to already generated CDMs. In this way an 'overexpression' model is generated. By asking what the effect of a CDM 'overexpressing' the protein of interest is, we can assess a different side of the same coin that should complement the findings resulting from CDMs lacking a protein via clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR associated protein 9 (Cas9) gene editing.

1.8 CRISPR/CAS9 GENE EDITING

The RNA-guided CRISPR/Cas9 technology is a bacterial and archaeal adaptive immune system [405, 413, 414] that can be exploited to specifically modify or knockout genes.

The system was first demonstrated in *Streptococcus thermophilus* which acquired resistance to phages via incorporation of phage-derived sequences into its CRISPR locus [94]. In response to pathogen challenge, short nucleic acid fragments (protospacers) from invading pathogens are incorporated into the host chromosome at a CRISPR repeat element. The CRISPR loci thus comprise of host repeats and invader derived protospacers. This process is known as the adaption stage. Subsequently, during the transcription and expression stage, this results in the generation of short CRISPR-derived RNAs (crRNAs). This is followed by processing, which involves the cleavage of a long pre-crRNA into a 60 nt fragment [415]. Then, during the interference stage, the crRNAs are incorporated into a ribonucleoprotein complex with Cas endonucleases. The crRNA then acts as a guide by binding to complementary sequences of foreign nucleic acid, resulting in crRNA-directed Cas endonuclease-mediated cleavage of foreign nucleic acid and destroying the invader [413, 414]. These complexes can be likened to memory B cells, circulating in the host serving as memory for pathogens which have previously been encountered. Whilst the above summarises the basic mechanism for the CRISPR/Cas system, this can be divided into three types (I, II and III) and many subtypes which differ mainly/only at the cleavage mechanism of pre-crRNA [416] [417, 418].

1.8.1 CRISPR/Cas9 gene editing

The *Streptococcus pyogenes* Cas9 (SpCas9) system has been modified to allow targeted gene editing in mammalian cells. This is the gold standard CRISPR gene editing system and has been used to

successfully edit countless human, bacteria, mouse, fruit fly, rat pig and monkey cell lines [407]. Regardless of the Cas9 species, the molecular mechanism of targeting is the same. A 20 nt guide sequence (gRNA) is designed that is complementary to a specific locus within a gene of interest and immediately downstream of a protospacer adjacent motif (PAM). The PAM sequence varies between Cas9 species, however the 'traditional' Cas9 PAM sequence is 5' NGG 3', where N can be any nucleotide [413]. The PAM sequence is critical for guiding Cas9 to cleave the DNA strands inducing double strand breaks (DSBs) (Figure 11) [419] and PAMs are present every 8-12 bases in the mammalian genome [403], allowing flexibility in targeting. The gRNA is able to bind the complementary target region, whilst the adjacent PAM sequence enables cleavage and generation of a DSB at the target site by Cas9 to a high specificity. However, the widespread location of PAM sequences in the genome and tolerance for multiple mismatches between the gRNA and target region of the gene [403] can lead to off-targets.

Cas9 can be delivered into cells in many ways including viral, plasmid or ribonucleoprotein delivery. Concern is raised about viral delivery in whole organisms due to viral related issues such as carcinogenesis, insertional mutagenesis, targeted delivery and immunogenicity [420-422]. The choice of method depends on the cell type of interest or the organisms, with *in vivo* delivery posing more complications than *in vitro* delivery. Plasmid methods are widely used, in which the gRNA is cloned into an expression plasmid containing the expression cassette for Cas9, which, when inside the cell, results in the expression of a Cas9 and chimeric gRNA. However, this method also has drawbacks. Specifically, the plasmid DNA or parts of it can often be incorporated into the target genome at 'on' or 'off-target' sites [409]. Although transfection with plasmids is a 'transient' delivery method, the plasmid can also remain inside the cell for several days, posing further risks of 'off-target' edits [423].

1.8.2 CRISPR/Cas9n

To abrogate the risk of 'off targets' occurring, a mutant Cas9 (Cas9 nickase D10A) or 'Cas9n' has been developed that is not capable of inducing a DSB, due to the Cas9 only containing a single functional catalytic site instead of two. As such, Cas9n can only cleave one strand, resulting in a 'nick'. This system therefore requires two sgRNAs targeting loci in close proximity on each DNA strand (Figure

11). The requirement of two guides and Cas9 complexes binding simultaneously therefore results in higher specificity and subsequently less off targets [424].

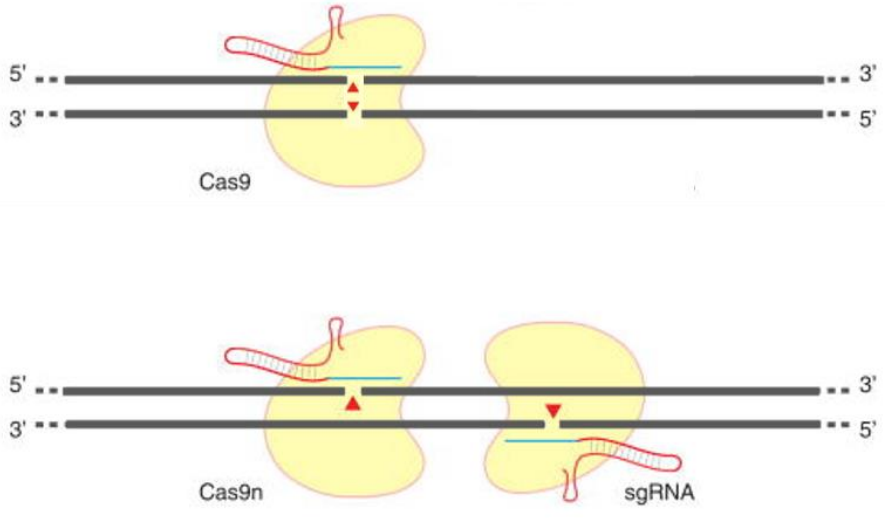


Figure 11. CRISPR/Cas9 (top) and CRISPR/Cas9n (bottom) gene editing schematic.
Red arrows indicate cleavage sites. Fused gRNA and crRNAs are shown as blue line (guide) and crRNA (red loop). As Cas9n mutants (bottom) can only cleave one DNA strand, due to having only one catalytic subunit compared to Cas9 (top), two sgRNAs complexed with Cas9n are required to induce a double strand break (DSB). This results in less off target cleavage due to both guides having to match to independent regions of the genome at once. Whilst Cas9n activity is more specific, it is much less efficient than Cas9, which only requires the sgRNA binding in one region of the genome to induce a DSB. Adapted from Ran *et al.* [424].

However, whilst solving specificity issues, the Cas9n is reported to lead to lower cleavage efficiencies due to the requirement of two Cas9n complexes binding simultaneously. The editing efficiencies in the different Cas9 types is debated. Many papers detail the editing specificities in their work. However, whilst not fully understood, it is clear that the efficiency is heavily reliant on a multitude of factors, including the cell type targeted, the genomic locus targeted, the gRNA design and sequence, the method of delivery, the concentrations of gRNA and Cas9 and their assembly into gRNA:Cas9 complexes [403, 409, 425]. For instance, some gRNA sequences have been shown to be less efficient or inactive even when perfectly complementary to their target and adjacent to a PAM sequence [426]. Also, in the case of plasmid delivery, it is well established that a 'G' is necessary at the 5' end of the gRNA for expression from a U6 promoter [427]. A 'G' is also favoured in the 1st or 2nd position closest to the PAM as this aids with Cas9 loading [427]. This is specific to Cas9 systems to 'knock out' protein production.

Plasmid methods of CRISPR have proved an impressive means that has revolutionised the field of gene editing, however the drawbacks of this method have led to the development of other CRISPR methods.

1.8.3 CRISPR/Cas9 ribonucleoprotein (RNP)

The latter two approaches discussed previously have traditionally relied on plasmid-based methods. This poses several problems, particularly with cells such as fibroblasts that are difficult to transfect. In such cells CRISPR/Cas9 edits the genome with low efficiency and 'off-targets' are not uncommon due to the relatively lengthy plasmid half-life. A newer method circumvents the use of plasmids. CRISPR/Cas9 ribonucleoprotein (RNP) uses recombinant Cas9 protein and a pre-synthesised gRNA, which are preassembled into an active RNP complex before being delivered into the cell, where the RNP is active immediately, unlike plasmid delivery methods which require transcription and translation of the Cas9 and gRNA [409]. This lack of delay makes this method more suitable for editing zygotes, where time is of the essence during the critical initial cell divisions [428]. Furthermore, this method has also demonstrated more efficient on-target cleavage and less off-target cleavage compared to plasmid-based methods [409, 410]. As well as being quicker to induce cleavages, this method is also incredibly straight forward and removes many limiting lab steps such as cloning or single cell sorting. The CRISPR-Cas9 experimental methods utilised in this project and gRNA design are discussed in great detail at the start of Chapter 2.

1.8.4 CRISPR and DNA repair by non-homologous end-joining (NHEJ) or homology-directed repair (HDR)

Following on from the CRISPR/Cas targeted DSBs the DNA damage repair pathway is activated. Mammalian cells repair DSBs by one of two endogenous DNA repair mechanisms: non-homologous end-joining (NHEJ) or homology-directed repair (HDR) (Figure 12). As the name suggests, the latter requires a homologous repair template and, as such, is restricted usually to the S or G_2 phase, when a sister chromatid is available to use as a repair template whilst NHEJ predominates throughout the cell cycle [429]. HDR is therefore much less efficient than NHEJ. However, HDR can be exploited to introduce precise genetic modifications of codons by recombination with a synthetic construct serving

as a repair template [430, 431]. The repair templates need to have regions (arms) homologous to the area targeted for modification. Repair templates can vary in format, including single-stranded DNA oligos up to 200 nt in length [432] and long double-stranded templates amongst others [433]. However, there are inherent limitations to repair templates. For example, linearised double stranded templates can be randomly inserted into the genome by NHEJ [431] and, as such, the low efficiency of HDR remains a hurdle.

NHEJ requires no repair template and can therefore function throughout the cell cycle as the main repair pathway. This process commonly results in insertion or deletion (Indels) mutations in the gene as NHEJ re-joins broken ends in an imprecise manner [434]. Such mutations can result in the construction of a knockout allele, especially if the number of mutated nucleotides is not a multiple of three, thereby introducing a frameshift, which can lead to a premature stop codon and thus gene disruption. Although it has been demonstrated that Cas9n is more specific in cleavage than Cas9, rates of NHEJ are lower following Cas9n-mediated DNA cleavage [404]. Nevertheless, it is clear that the NHEJ pathway can be exploited with CRISPR/Cas to result in targeted gene knockout, which could be utilised to investigate ECM gene function in CDMs as a model for the ECM microenvironment.

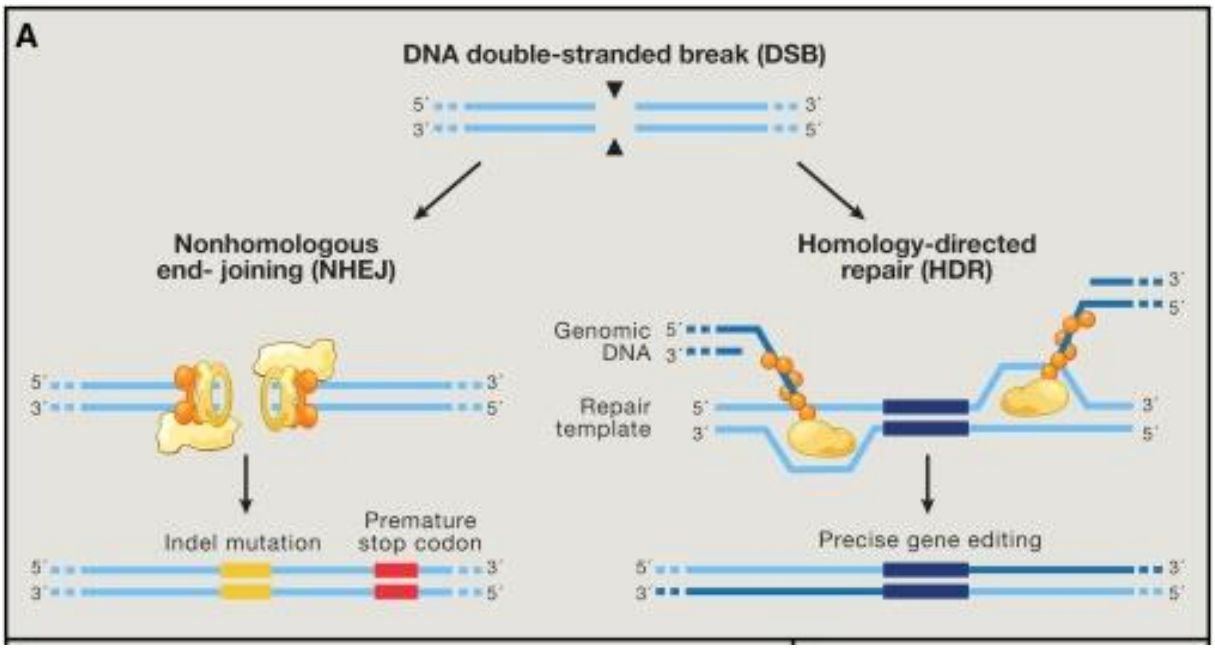


Figure 12. Repair of DSBs in DNA induced by Cas9 or Cas9n.
Two main repair pathways exist for repairing DNA breaks, non-homologous end-joining (NHEJ) and homology-directed repair (HDR). In the error prone NHEJ pathway, insertions or deletions (Indels) are introduced when the complementary strands are joined together by repair enzymes, regardless of sequence. This can eventually lead to frameshift mutations and gene knockout. During HDR accessory factors direct genomic recombination with homologous exogenous repair templates, usually in the form of sister chromatids during DNA replication, for precise DNA repair. From Hsu *et al.* [406].

1.9 AIMS AND OBJECTIVES OF THE PROJECT

The overarching aim of this project is to investigate how TN-C regulates inflammatory gene expression profiles in activated human macrophages by regulating miRNA levels. For this, two technologies will be combined: CRISPR/Cas9 gene editing and cell-derived matrices. Specifically, CRISPR/Cas9-based gene editing will be employed to stably knock-out *TNC* in human fibroblasts. These *TNC* knock-out cells will then be utilised to produce cell-derived matrices lacking TN-C. Wild type and *TNC* knock-out fibroblasts will be exploited as 'human extracellular matrix factories' for their innate ability to secrete and deposit abundant ECM molecules which are assembled into a three-dimensional ECM. Wild type matrices will be characterised at the protein level and assessed for compatibility with primary human monocytes/macrophages. On confirmation of compatibility, primary human monocytes will be cultured and differentiated into macrophages on these matrices and, following activation, their a) expression of candidate genes and b) complete coding and non-coding transcriptome will be analysed.

This will address many questions that have arisen from previous research. Namely, is the pro-inflammatory TN-C/miR-155/TNF-α axis specific to the mouse innate immune system, or does it occur also in humans? Is ECM-associated (insoluble) or plasma (soluble) TN-C responsible for the regulation of the miR-155/TNF-α axis in macrophages? Does TN-C regulate the expression of other genes?

This overarching aim will be investigated by addressing the following specific objectives:

1) To generate and characterize cell-derived matrices produced by human foreskin fibroblasts (WT CDM), which will be exploited as 'extracellular matrix factories'.
2) To engineer and characterize (genotype and phenotype) *TNC* knockout 'extracellular matrix factories' that produce matrices lacking TN-C (*TNC* KO CDM), by using CRISPR/Cas9 technologies.
3) To validate the generated matrices as a compatible substrate for the culture of primary human monocytes and differentiation into macrophages, and, following activation with LPS, analyse gene expression, including miR-155.

The combination of the CDM and CRISPR-Cas9 genome editing technologies will allow the development of a biological tool able to recapitulate the natural, human ECM microenvironment, more so than traditional culture methods, allowing physiologically relevant data to be generated. Beyond this specific application, this tool can further be applied to other cell types and functions (e.g., stem cell differentiation) and diseases (e.g. cancer). This approach could also be adopted to study many other ECM proteins in addition to TN-C. Furthermore, this tool will also allow for the replacement of animals to study the ECM, in line with the 3Rs.

Unfortunately, and unsurprisingly, as the Covid-19 pandemic occurred in my final year of study, progress on some of the objectives, principally objective number 3 utilising macrophages, was halted. Due to the university closure, and on subsequent reopening, the complex logistics of producing matrices, purifying monocytes from blood, culturing and differentiating macrophages while working under reduced hours and lab occupancy levels, part of the experiments planned for objective '3' were aborted. These included the validation of CDM compatibility with monocyte/macrophage culture using cells from at least two further donors and the transcriptome analysis of macrophages comparing their

inflammatory response when seeded on plastic, WT CDM and *TNC* KO CDM. To mitigate this, the project was adapted to follow up on a new avenue arising from interesting phenotypic observations of the *TNC* KO fibroblasts. Specifically, it was observed that *TNC* KO fibroblasts undergo morphological changes and exhibit a markedly reduced growth rate and suspected premature senescence, compared to similar age wild type fibroblasts. This will be discussed in detail in chapter 6.

2 MATERIALS AND METHODS

2.1 MOLECULAR BIOLOGY REAGENTS

Luria broth growth medium was purchased from OXOID Ltd. *Escherichia coli* (DH5a) cells were from Stratagene. Cloning reagents, including restriction digestion, phosphatase and ligase enzymes and buffers, DNA polymerases and their buffers, dNTPs and Quick-Load® Purple 1 kb and 100 bp DNA Ladder were from Roche or New England Biolabs (NEB). TRIzol® reagent was from Invitrogen. QuantiTect Reverse Transcription Kit was from Qiagen. GoTaq® qPCR Master mix, plasmid mini and midiprep kits and Reliaprep miRNA cell and tissue RNA extraction kit were from Promega. qPCR Taqman *TNC* primers and probes, and TaqMan™ MicroRNA Reverse Transcription Kit from Applied Biosystems. Agarose gel purification kits were from Monarch. The GenElute™ mammalian genomic DNA extraction kit was from Sigma-Aldrich. Lyophilised and desalted oligonucleotides were purchased from Sigma-Aldrich and resuspended in sterile water (100 µM) and stored at -20 °C.

2.2 PROTEIN CHEMISTRY REAGENTS

PageRuler™ Plus Pre-stained protein ladder (10 to 250 kDa) was from Thermo Scientific, SDS-PAGE electrophoresis apparatus, wet transfer apparatus and nitrocellulose membrane (0.45 µm) were from Bio-Rad. 30% (w/v) acrylamide/bis-acrylamide and N,N,N,N-tetramethylethylenediamine (TEMED) were from Sigma-Aldrich. The ECL Western blotting Detection Reagents were from GE Healthcare. Recombinant TN-C was from Merck Millipore (CC065). ELISA kits were from R&D Systems for TNF-α and PeptroTech for IL-6 and IL-8. Alcian blue was from Sigma (A5268).

2.3 CELL CULTURE

HBSS without calcium and magnesium, Eagle's Minimum Essential Medium (EMEM) without L-Glutamine, DMEM with L-Glutamine and Na-Pyruvate, RPMI 1640 with L-Glutamine, Trypsin-EDTA (200 mg/L Versene EDTA, 170.000 U Trypsin) were from Lonza. Fetal Bovine Serum (FBS) was from Gibco and M-CSF was from Peprotech. L-ascorbic acid, ammonium hydroxide (NH_4OH), Dulbecco's

phosphate buffered saline with calcium and magnesium, gelatin, penicillin-streptomycin (10,000 U/mL) and Trypan blue were from Sigma. Primocin™ was from Invivogen. LPS (*E. coli*, Serotype EH100 (Ra)) was from Enzo Life Sciences. DNase I was from Roche. Dimethyl sulfoxide (DMSO, ≥99.7 %) was from Fisher BioReagents. Cells were counted using a Neubauer haemocytometer.

2.4 SOLUTIONS AND BUFFERS

Table 3. Names and reagents necessary for composition of buffers and solutions

Buffer or solution	Reagents required
Ampicillin (1000X)	100 mg/ml Ampicillin in 50% ethanol and sterile diH_2O; stored at -20 °C
Vitamin C	50 mg/ml in PBS, sterile filtered; used at 1:1000
PBS (phosphate-buffered saline) (10X)	1.37 M NaCl, 27 mM KCl, 100 mM Na_2HPO_4, 18 mM, KH_2PO_4
PBST (1X)	0.05 % (v/v) Tween-20
TBS (10X)	50mM Tris-Cl pH 7.5, 150mM NaCl to plus water 1L
TBST 1X	5ml 10X TBS, 0.1% Tween-20, up to 50 mL with ddH_2O.
Ponceau S solution	0.1 % (w/v) in 5 % (v/v) acetic acid in ddH_2O
Western blot 5% milk blocking solution	5 % (w/v) Marvel milk powder in 1X PBST
Western blot 10 % milk blocking solution	10 % (w/v) Marvel milk powder in 1X PBST
Western blot 5 % BSA blocking solution	5 % (w/v) in 1X PBST
Western blot transfer buffer (10X)	15 g Tris base, 72 g glycine, up to 1 L with ddH_2O.
Western blot transfer buffer (1X)	100 ml of 10 X transfer buffer, 200 ml methanol, up to 1 L with ddH_2O; ice cold
Transfection 2x HBS solution	50 mM HEPES, 280 mM NaCl, 1.5 mM Na_2HPO_4, pH adjusted to 7.0 with 1 M HCl. Filter sterilised and stored at -20 °C
10% Ammonium persulphate (APS) solution	10 % (w/v) APS in ddH_2O. Stored at -20 °C

10% Sodium dodecyl sulphate (SDS) solution	10 % (w/v) SDS in ddH_2O
6 X Laemmli buffer	9 % β–mercaptoethanol, 375 mM Tris–HCl pH 6.8, 9 % (w/v) SDS, 50 % glycerol and 0.03 % bromophenol blue
Immunofluorescence blocking solution	5 % BSA (v/v) and 3% (v/v) goat serum in TBS
Immunofluorescence permeabilisation buffer	0.1% Triton X-100 in 1 x PBS
ELISA blocking and diluent buffer (R and D systems)	1 % BSA in 1 x PBS (sterile filtered)
ELISA blocking buffer (PeproTech)	1 % BSA in 1 x PBS (sterile filtered)
ELISA diluent (PeproTech)	0.05% Tween-20 (v/v), 0.1% BSA (w/v) in PBS (sterile filtered)
ELISA washing buffer	1 x PBST
ELISA stop solution	2N sulfuric acid in ddH_2O
3-(4,5-Dimenthylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT)	5 mg/ml in PBS, sterile filtered, 10 % (v/v) used in assay
MTT lysis buffer	10 % (w/v) SDS, 0.01M HCl2M
Adhesion assay staining buffer	3 % (v/v) paraformaldehyde, 1% (w/v) toluidine blue in PBS
Adhesion assay lysis buffer	2% (w/v) SDS in ddH_2O
Decellularisation buffer	20mM NH_4OH, 0.5% Triton X-100 (v/v) in PBS (sterile filtered)
TN-C coating buffer	50mM Tris-HCl pH 7.5, 2M NaCl, 5mM CaCl, 0.02% (v/v) Sodium Azide (sterile filtered)
Alcian blue buffers: 0.06 M 0.3 M 0.5 M 0.7 M 0.9 M	50 mg Alcian blue in 100 ml 0.2 M acetic acid pH 5.8 with: 1.22 g magnesium chloride ($MgCl_2$) 6.09 g $MgCl_2$ 10.15 g $MgCl_2$ 14.21 g $MgCl_2$ 18.27 g $MgCl_2$
1% Agarose gel	1 g agarose, 2 mL of 50X TAE or 10 ml of 10X TBE buffer in 100 mL H_2O.
TAE 50X	50mM EDTA (pH 8.0), 2.5M Tris-acetate
TBE 10X	1M Tris Base, 1M Boric acid, 20mM EDTA (pH 8.0)

2.5 MOLECULAR BIOLOGY TECHNIQUES

2.5.1 CRISPR/Cas9

2.5.1.1 CRISPR/Cas9 plasmid (Approach 1 and 2) sgRNA design and synthesis

20 bp long guide (g)RNA oligonucleotides were designed to target the *TNC* gene (gene ID: 3371, NC_000009.12) using MIT CRISPR design software (http://crispr.mit.edu; Table 4) (Hsu et al., 2013) and synthesised by Sigma Aldrich. Complementary overhang sequences necessary for cloning into mammalian expression vectors pSpCas9n(BB)-2A-GFP (PX461 and PX458 D10 mutant) were added. Where necessary, an additional G (underlined) was also added at the 5' end of each guide to initiate RNA polymerase III-dependent U6 promoter transcription of the gRNA, Cas9 and GFP reporter gene. For the schematic detailing the gRNA binding locations in the *TNC* gene see Figure 33 and Figure 34 in Chapter 4.

2.5.1.2 CRISPR/Cas9 RNP (approach 3) sgRNA design and synthesis

20 bp long guide (g)RNA oligonucleotides (Table 2) were designed, as previously in 2.5.1.1, to target the *TNC* gene (gene ID: 3371, NC_000009.12). No additional nucleotides to the guide sequence were necessary in this instance, as these guides were not cloned into an expression vector. For the schematic detailing the gRNA binding locations in the *TNC* gene see Figure 35 in Chapter 4

2.5.1.3 Assembly of the ribonuclear protein (RNP) complex

1 µl of gRNA (100 µM) and Alt-R® CRISPR-Cas9 tracrRNA (100 µM) were mixed together and diluted in nuclease-free duplex buffer to 25 µM. The RNA mix was boiled at 95˚C for 5 minutes and cooled at room temperature for 10 minutes before 2.9 µl of the annealed RNA (72.5 pmol) was complexed with 1 µl of Alt-R® S.p. Cas9 Nuclease V3 Cas9 (60 pmol/µl) to a ratio of 1:2:1 to form the RNP. After a 10 minutes period to allow complexing, 0.6 µl of Alt-R® electroporator enhancer (100µM; IDT) was added and incubated at room temperature with the RNP complex for 5 minutes.

Table 4. CRISPR-Cas9n sgRNA sequences.
Details of ~20bp guides complimentary to *TNC* target regions in exon 2 or exon 7, with complementary mammalian overhangs necessary for cloning shown in bold and additional G necessary to initiate U6 dependant promoter transcription underlined.

ID	Sequence
Guides for cloning into CRISPR/Cas9n PX458 (Approach 1)	
GB *TNC* A1 5'-3' top	**CACC**GCTGAGTCATGGCCCCCATGG
GB *TNC* A1 5'-3' bottom	**AAAC**ACCTTCGGTAGCGAGGGCAAC
GB *TNC* A2 5'-3' top	**CACC**GTTGCCCTCGCTACCGAAGGT
GB *TNC* A2 5'-3' bottom	**AAAC**GACACTGAGTACGAGGTGTC
GB *TNC* B1 5'-3' top	**CACC**GGACCCCACCTTCGGTAGCG
GB *TNC* B1 5'-3' bottom	**AAAC**CGCTACCGAAGGTGGGGTCC
GB *TNC* B2 5'-3' top	**CACC**GATCCGGCACAAGCGACAGAG
GB *TNC* B2 5'-3' bottom	**AAAC**CTCTGTCGCTTGTGCCGGATC
Guides for cloning into CRISPR/Cas9 PX461 (Approach 2)	
GB *TNC* 2A3 5'-3' top	**CACC**GCCATGACTCAGCTGTTGGC
GB *TNC* 2A3 5'-3' bottom	**AAAC**GCCAACAGCTGAGTCATGGC
GB *TNC* 7A 5'-3' top	**CACC**GTGGCACGTCTTTGATGCCGT
GB *TNC* 7A 5'-3' bottom	**AAAC**ACGGCATCAAAGACGTGCCAC
GB *TNC* 7B 5'-3' top	**CACC**GACACCTCGTACTCAGTGTC
GB *TNC* 7B 5'-3' bottom	**AAAC**GACACTGAGTACGAGGTGTC
Guides for CRISPR/Cas9 RNP (Approach 3)	
A2	TTGCCCTCGCTACCGAAGGT
7A	TGGCACGTCTTTGATGCCGT

2.5.1.4 CRISPR/Cas9 sgRNA phosphorylation and annealing

2 µl of 100 µM CRISPR-Cas9 sgRNAs (Table 4) were phosphorylated with 2 µl of 10 x T4 DNA ligase buffer, 10 units of polynucleotide kinase in a final volume of 20 µl with ddH_2O and incubated at 37 °C for 60 minutes, followed by enzyme inactivation incubation at 65 °C for 20 minutes.

To anneal the complementary phosphorylated top and bottom sgRNAs, 5 µl of each were mixed, made up to 50 µl with ddH_2O and cycled as in Table 5:

Table 5. Thermocycling protocol to anneal complementary phosphorylated oligos.

Temperature (°C)	**Time**
95	3 min
Reduce by 1 °C per min until 26 °C is reached	
25	5 min

2.5.1.5 Restriction digestion of mammalian expression vector

The mammalian expression vectors pSpCas9n(BB)-2A-GFP (PX461 and PX458 D10 mutant) were digested with *Bbs1* for 2 hours at 37°C as in Table 6:

Table 6. Components of a CRISPR vector digestion reaction.

10 x NEBuffer 2.1 Buffer	2.0 µl (to make 1 x)
PX461 vector	1 µg
BbsI enzyme	10 Units (1 µl)
ddH_2O	To a total volume of 20 µl

Vectors were then dephosphorylated to prevent re-ligation of the vector by adding 1 unit of TSAP for 1 hour at 37°C followed by 20 minutes at 74°C to inactivate TSAP. The vector was analysed and purified by agarose gel electrophoresis and quantified by Nanodrop.

2.5.2 Ligation of DNA fragments with vectors

gRNAs were ligated into CRISPR plasmids after restriction *bsbI* enzyme digestion to produce new CRISPR plasmids targeting *TNC*. 2 µl of annealed and phosphorylated oligos were added to 100 ng of BSb1 digested vector, 2 µl of 10 x T4 DNA ligase buffer, 3U of T4 DNA ligase plus sterile ddH_2O to a final volume of 20 µl. For each set of ligations reactions a negative control consisting of all reagents minus gRNA insert was included to indicate the level of background undigested/re-ligating vector. Ligation reactions were incubated overnight at 4°C before transformation of *Escherichia coli* DH5a competent cells.

2.5.3 Agarose gel electrophoresis

DNA fragments were analysed on 0.6, 1, 1.5 or 2 % (w/v) agarose gels in 1 x TBE buffer and stained with 1 µl ethidium bromide per 50 ml of melted agarose. PCR or digestion products were mixed with 5 µl loading dye per 1 µl of DNA sample alongside either 3 ul of Quick Load® 1 kb (NEB) or Thermo Scientific™ *MassRuler™ Low Range* ladder. DNA bands were resolved by electrophoresis in BioRad tanks and 0.5 x TBE buffer at 90 V for a length of time dependent upon agarose gel concentration and PCR product size. Bands were visualised using a UV transilluminator. DNA fragments were excised from gels and purified using the Monarch™ DNA extraction kit following manufacturers protocol.

2.5.4 Quantification of nucleic acids

DNA quantity and quality were measured using a NanoDrop 1000 spectrophotometer (Thermo Scientific). The purity was determined by comparing the absorbance values at 260nm and 280nm wavelength (A260/280), and 260nm and 230nm (A260/230). Purity ratios at OD260/280nm for RNA or DNA were 2.0 and 1.8 respectively.

2.5.5 PCR amplification for genotyping of gene-edited cells

Primers 20 nt long flanking the target region to be edited and in other non-target regions of the *TNC* gene were designed for each CRISPR/Cas9 gRNA pair (Table 7 and 8) using NCBI primer-BLAST https://www.ncbi.nlm.nih.gov/tools/primer-blast/. A 200 bp region of *COL1A1* was amplified as a control using Fw and Rv primers 5' -3' CACGTCTCGGTCATGGTACCT and GTCGAGGGCCAAGACGAA.

Table 7. Primers for genotyping CRISPR/Cas9 and CRISPR/Cas9n targeted regions of *TNC* gene. Primers denoted on PCR gels and schematic diagrams by another label are identified (red).

Primer	Sequence 5' – 3'	Amplicon length (bp)	Tm (°C)	GC content %
Primers for screening exon 2				
A 1 Fw	GATACCCTAGAGCCCTAGAG	180	56.1	55
A 1 Rv	CTGGCAGGGTGGCGTTCACC	180	74.4	70
A 2 Fw	CTCTGTGCTTCTAAATCCCC	266	60.3	50
A 2 Rv	CTGGCTGGTTCTCTTCTGGC	266	66.3	60
B 1 Fw	GGGCCATGACTCAGCTGTTG	177	68.2	60
B 1 Rv	CCACTGGCAGCTTGATGTTG	177	66.9	55
B 2 Rv	CCACTGGCTGACTCCAGATC	214	65	60
B 3 Fw (F6 Fw)	CCAAGGGACCAGAGCTACAG	821	63.8	60
B 3 Rv (F6 Rv)	CCTCCAGTCTGCTCAGCAGC	821	67.8	65
Primers for screening exon 7				
7A Fw (F7 Fw)	CTCTGCTAAGCCTCAGGCAA	**352**	**64.6**	**55**
7A Rv (F7 Rv)	ACAGGGGGAGGCCTTTTACC	352	67.2	60
7B Fw	TACGGCATCAAAGACGTGCC	185	68.2	55

7B Rv	ACAGGGGGAGGCCTTTTACC	185	67.2	60

Table 8 . Primers for genotyping regions of the *TNC* gene not targeted by CRISPR/Cas9.
Primers denoted on PCR gels and schematic diagrams by another label are identified (red).

Primer	**Sequence 5' – 3'**	**Amplicon length (bp)**	**Tm (°C)**	**GC content %**
Primers for screening intron 1				
1 Fw	GTCTGTTTTGCCCTCATCACT	773	63.1	47.62
1 Rv	GGGTGAGAAGAACCCCTGTG	773	65.8	60
1.1 Fw (F1)	TGCCCCACCTTTCTCTTGAC	528	61.8	55
1.1 Rv	TCCACACCATCTGCTGGTTC	528	61.7	55
Primers for screening intron 2				
2 Fw (F2)	GCTGCCTGTGATAGTGCTGA	651	61.2	55
2 Rv	GCATGTGGAAACAGGGCATC	651	61.4	55
Primers for screening intron 4				
4 Fw (F3)	GTATAACCCTCCCTCGCCAC	1049	64.4	60
4 Rv	TGCACACAGTGCTCTAAGCC	1049	64.3	55
Primers for screening intron 6				
7 Fw (F4)	GCCCACTTCCTGTCACTTGT	535	62	55
7 Rv	CTTGCCTGTTTTTGCAGGCT	535	61.5	50
Primers for screening exon 8-9				
8-9 Fw	TCCCAGGAATCTTCGACGTG	595	67.4	55
8-9 Rv	TGACTCTGTCTCATCTGCCCA	595	66.2	52.38

For a schematic detailing the primer binding regions within the *TNC* gene and the approach taken to genotype the CRISPR/Cas9 *TNC* KO cell line see Figure 52 and Figure 53. Briefly, as amplification

of *TNC* over the CRISPR target sites was not possible, regions of the gene, proceeding and preceding the break sites were amplified to confirm the DNA was not degraded (Figure 52, page 160-161). After confirmation of this, pairs of forward pairs or reverse pairs of primers were utilised to determine whether a hypothesised inversion of the fragment resulting from CRISPR/Cas9 cleavage at the 2 target sites in exon 2 and exon 7, had occurred (Figure 53, page 163). This work is detailed thoroughly in Chapter 4.

2.5.5.1 Protocol 1: PCR BIO taq PCR

Unless otherwise stated, PCR was performed in a 25 µl reaction volume consisting of: 50-100 ng of template or dH_2O (non-template control), 1 X PCRBIO buffer (containing 1mM dNTPs and 3mM $MgCl_2$), 0.2 µM of 10mM forward and reverse primers and 0.625 U of Q5 high fidelity DNA polymerase and cycled (Table 9). Extension time was modified depending on the size of the expected product.

Table 9. Standard PCR thermocyler protocol to amplify *TNC* genomic DNA.

Temperature °C	Time	Cycles
95	1 min	1
95	1 min	35
55-60	15 s	
72	15 s/kb	

2.5.5.2 Touchdown PCR protocol

Where necessary to remove non-specific amplification a touchdown protocol was used (Table 10)

Table 10. Touchdown PCR thermocycler protocol to amplify *TNC* genomic DNA.
The annealing temperature drops by 2 °C every 2 cycles until 60 °C is reached.

Temperature (°C)	**Time**	**Cycles**
95	1 min	1
95	1 min	20
78	30 s	
72	30 s	
Annealing temperature drops by 2 °C every 2 cycles until 60 (°C).		
95	1 min	20
60	1 min	
72	30 sec	
72	5 min	1

2.5.5.3 Protocol 2: PCR using Q5 Taq for long range amplicons

Unless otherwise stated, PCR was performed in a 25 µl reaction volume consisting of: 50-100 ng of template or dH_2O (non-template control), 1 X PCR Q5 buffer (containing 2 mM Mg^{++}, 200uM dNTPs, 0.5 µM of forward and reverse primer and 0.5 µM of Q5 high fidelity DNA polymerase and cycled (Table 11). Extension time was modified depending on the size of the expected product.

Table 11. PCR thermocycler protocol to amplify *TNC* genomic DNA using Q5 Taq polymerase.

Temperature (°C)	**Time**	**Cycles**
98	30 min	1
98	10 s	35
67	30 s	
72	30 s/kb	
72	2 min	1

2.5.6 Off-target analysis of CRISPR/Cas9 guides targeting *TNC*

Benchling software [403, 435](http://benchling.co.uk) was utilised to identify predicted off-target regions for each gRNA. Primers were designed (Table 12) to amplify regions spanning the predicted off-target regions and synthesised by Sigma.

Table 12. Primers for screening predicted off-target regions in CRISPR/Cas9 nucleofected cells.
All predicted off-targets residing adjacent to the classical Cas9 'NGG' PAM sequence were screened. One off-target region per guide with the 'NAG' PAM sequence were screened and are indicated with '*'.

Primer	Sequence 5' – 3'	Amplicon length (bp)	Tm (°C)	GC content %
Primers for screening exon 2 predicted off-targets				
Adenylate Cyclase 4 (ADCY4)				
Fw	CCCAGGAACTCAATCTCTGGC	315	67.1	57.1
Rv	GTTCTAGGAGCCATGGGGGT	315	66.2	60
Calcium Modulating Ligand (CAMLG)				
Fw	AGTCATCGCCCTCGCAGC	312	69.2	66.6
Rv	GGGAGAGGGAAGAATGACCCT	312	66.6	57.1
Transmembrane protein 42 (TMEN 42) *				
Fw	TGGGGCGTATTCAACTGTCT	304	64.4	50
Rv	TGGCGTGATTAGAGGCGAAA	304	67.7	50
Primers for screening exon 7 predicted off-targets				
Von Willebrand Factor A Domain Containing 1 (VWA1) (guide 7B)				
Fw	CTAGTGGGGCCTCCAATCTC	495	64.8	60
Rv	AGATGACGATGCGCTCTGG	495	66.7	57.8
Adenosine Deaminase (ADA) * (guide 7B)				
	CAGGATGGGCTGGATTCTGG	470	69	60
	GGTTGGGCTTGTCTTGGACT	470	65.3	55
Protein Phosphatase 4 Catalytic Subunit (PPP4C) (guide 7A)				
Fw	CGCCCTCATCTCCTATCGTG	316	66.9	60
Rv	CACGACCCCTCCAGAGAATG	316	67	60
protein tyrosine phosphatase (PTPN14) * (guide 7A)				
Fw	GGAGGTTTTTCCATCCTTGGG	479	67.4	52.4
Rv	CAGATGCCACCGAAAGGACT	479	66.4	55

2.5.7 Ethanol precipitation of DNA

To concentrate DNA, 100% ethanol and 3 M sodium acetate pH 5.2 were added in volumes of 200% and 10%, respectively, of the original sample volume. Samples were incubated at – 20 °C for 30 minutes, followed by centrifugation at 13,000 RPM for 15 minutes at 4 °C. The pellet was washed in 70 % ethanol and centrifugation as before, and the pellet air-dried before resuspension in ddH_2O.

2.5.8 Sequencing

Sequencing reactions were performed with 500 ng of plasmid or 50 ng of PCR product by Source Bioscience Ltd using reverse primers for the corresponding amplicon as in Table 5.

2.5.9 Genomic DNA purification from mammalian cells

Genomic DNA was purified from mammalian cells using the GenElute™ Mammalian Genomic DNA Miniprep Kit, following manufacturer's instructions. Cells were trypsinised as in **2. 5. 2.** and centrifuged at room temperature at 1500 RPM for 5 minutes. The genomic DNA was purified and stored at -20°C.

2.5.10 RNA extraction and purification

For total RNA extraction, medium was removed and 1 mL of TRIzol® Reagent (Invitrogen) per 10 cm^2 was added. Cells were lysed following manufacturer's instructions. Total RNA for RNAseq was extracted using ReliaPrep™ miRNA Cell and Tissue Kit (Promega) as per manufacturers recommendations. Approximately 1 x 10^6 fibroblasts were harvested per column. RNA concentration and purity was quantified using the NanoDrop and for RNAseq samples RNA concentration and integrity were assessed using a TapeStation (Agilent; University of Nottingham, Deep Sequencing facility) and a Bioanalyser by Novogene (Cambridge, UK). Samples were stored at – 80°C.

2.5.11 Whole transcriptome RNA sequencing

Whole transcriptome sequencing (miRNAs, lncRNA, miRNAs, mRNAs and circRNA) and analysis was carried out by Novogene (Cambridge, UK).

2.5.11.1 RNA preparation and sequencing

Total RNA was extracted from WT and *TN-C* KO primary human foreskin fibroblasts using the ReliaPrep™ miRNA Cell and Tissue Kit (Promega) according to the manufacturer's recommendations. Total RNA (2ug) stored in RNase-free water was used for library preparation. mRNA and lncRNA was prepared for sequencing using Illumina Ribo-Zero Plus rRNA Depletion Kit and NEBNext® Ultra™ Directional RNA Library Prep Kit. Small RNA was prepared using NEB Next® Multiplex Small RNA Library Prep Set for Illumina®. These methods were used to construct 250-300 bp insert libraries and 10-50 bp insert libraries respectively following the manufacturer's recommendations. Libraries were sequenced by Novogene on an Illumina Novaseq 6000 platform.

2.5.11.2 Data analysis by Novogene

Data analysis was carried out by Novogen, Cambridge. The quality of raw reads was first assessed using an in-house script. Clean reads were obtained for future analysis after removing reads with low quality, adapter sequences and those sequences containing multiple 'N' bases. Reads were aligned to the human genome (*hg38*) with STAR (v2.6.1d). The read numbers mapped to each gene was quantified with FeatureCounts (v1.5.0-p3). Fragments Per Kilobase of transcript per Million mapped reads (FPKM) of each gene was calculated based on gene length and mapped read counts. Differential gene expression analysis was performed using the DESeq2 package (v1.20.0) for samples with bio-replicate. Resulting *P* values were adjusted using the Benjamini and Hochberg approach to control the false discovery rate (FDR, adjusted *P* value). Genes with an adjusted $P < 0.05$ were deemed differentially expressed. For samples without biological replicates, read counts for each sample were firstly adjusted by Trimmed Mean of M-values (TMM), followed by differential expression analysis in EdgeR package (v3.24.3). The *P* values were adjusted using the Benjamini and Hochberg methods. Adjusted *P* value of 0.005 and log2Fc (Fold Change) of 1 were set as the threshold for significantly differential expression. Hierarchical clustering of union differentially expressed genes was performed using $\log_2$(FPKM+1) values through clustering rows and columns in R (cluster_rows=TRUE & cluster_ cols=TRUE). This generated several .csv files with this information.

Further bioinformatic analysis was carried out by myself utilising the differentially expressed gene (DEG) .csv files resulting from Novogene's bioinformatics processing for small RNA (mRNA and miRNA) transcripts.

2.5.11.3 mRNA transcript analysis

Initially, a broad approach was taken utilising the list of all differentially expressed mRNA transcripts, both 'TRUE' and 'FALSE' for significance. A catalogue of genes of interest were retrieved from 3 online databases [436-438] of the senescence associated secretory phenotype (SASP) or the differentially expressed transcript list was imported into the Cell Age database [439]. The data (gene name, transcript expression, adjusted P value) was compiled from the aforementioned transcript .csv files provided by Novogene. This information was imported into the online software multi experiment viewer *(*MeV) to plot heatmaps using log2 transformed expression data, in order to assess the relative mRNA abundance. The adjusted P value (as calculated by Novogene) was utilised for determining if differentially expression was significant. Next, a more targeted approach was taken, in which significantly differentially expressed genes (as determined by the adjusted P value) were analysed. Genes with more than -2 or less than 2 fold changes or infinity log changes were filtered out from the list. The remaining list was imported into MeV, as previously, to generate heatmaps of mRNA abundance on a $\log_2$ scale. *In sillico* GSEA with Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathway enrichment analyses on Webgestalt [440] was utilised to determine which biological pathways may be altered. Webgestalt generated P values for this output and FDR values which are reported where relevant.

2.5.11.4 miRNA transcript analysis

Significantly differentially expressed miRs (determined using adjusted P value as calculated by Novogene's analysis) were filtered to remove miRs with more than -2 or less than 2 fold changes. This list was next input into miRBase (www.mirbase.org) for *in silico* screening of validated human targets genes. Next GSEA and KEGG analysis of this generated miR target list on Webgestalt was utilised to determine which biological pathways may be affected. This analysis generated P values.

2.5.12 TaqMan® quantitative polymerase chain reaction for miRNA gene expression quantification

Relative quantification of miRNA expression was measured by quantitative real-time PCR (qPCR). TaqMan uses a sequence specific probe for the target genes (has-miR-155) flanked by a forward and reverse primer. Probes contain a 5' fluorophore (FAM) and a 3' quencher (NFQ). As Taq polymerase extends the cDNA sequence the 3' end of the quencher is degraded, separating the fluorophore from the quencher, thus allowing emission and measurement of a fluorescent signal. qPCRs were performed with a Qiagen Rotor-Gene Q machine followed by data analysis (ΔΔCt method).

2.5.12.1 Reverse transcription of miRNA RNA

cDNA was reverse transcribed from 20 ng total RNA containing small RNAs using specific stem-loop RT primers that stabilise miRNAs. The components of the RT mastermix are listed below (table 13).

Table 13. RT master mix for qPCR of miRNAs.

Component	Initial Concentration	Volume
dNTPs (with dTTP)	100 mM	0.15 µl
Multiscribe™ Reverse Transcriptase	50 U/µl	1.00 µl
Reverse Transcription Buffer	10 X	1.50 µl
RNase inhibitor	20 U/µl	0.19 µl
TaqMan® MicroRNA RT Primer	5 X	3.0 µl
TaqMan® U6 Primer	5 X	3.0 µl
Total RNA sample	20 ng/µl	1.0 µl
Nuclease-free water		Up to 15 µl
Total		15.0 µl

The RT mastermix with RNA added was cycled as in Table 14.

Table 14. Taqman reverse transcription incubation protocol to synthesise cDNA from RNA.

Temperature (°C)	**Time**
16	30 min
42	30 min
85	4 min
4	hold

2.5.12.2 miRNA Taqman® qPCR

Diluted cDNA (1:3) was amplified in triplicate using specific TaqMan® primers and probes (Hs01115665_m1 and Hs02800695_m1) in a final volume of 10 µl consisting of 5 µl TaqMan 2X universal PCR master mix (No AMPErase UNG), 2 µl Nuclease-free water, 0.5 µl TaqMan MicroRNA target specific primers 20 X, and 2 µl of diluted cDNA. Samples were cycled as in Table 15. miRNA expression was normalized to the endogenous invariant control small nuclear RNA U6 using the ΔΔCt method.

Table 15. Thermocycling protocol for Taqman qPCR.

	Temperature (°C)	Time	Cycles
Enzyme inactivation	95	10 min	
Denaturation	95	15 sec	40
Annealing extension	60	1 min	

2.5.13 SYBR green qPCR for mRNA expression quantification

TNC and *HPRT1 mRNA* expression was quantified using SYBR green qPCR.

2.5.13.1 Reverse transcription of mRNA

400 ng of RNA was reverse transcribed using the QIAGEN QuanitiTect Reverse Transcription Kit. Samples were incubated 2 minutes with genomic DNA elimination buffer followed by incubation at 42 °C for 15 minutes with the reverse transcription mix. The Reverse Transcriptase enzyme was inactivated by incubation at 95 °C for 3 minutes.

2.5.13.2 SYBR mRNA qPCR

mRNA expression was quantified by SYBR green qPCR using GoTaq® qPCR Master mix and specific primers for TNC and HPRT1. Expression of the gene of interest was normalized to the housekeeping gene HPRT1 and untreated control (ΔΔCt method). The PCR reaction was assembled as in Table 16 and cycled as in Table 17. Melt curve analysis was carried out at the end of the run to determine if there was a single product. Products were also ran on a gel to determine a single product was present.

Table 16. SYBR green qPCR reaction mix.

Reagent	Volume (µl)
GoTaq® qPCR Master Mix 2X	5.0
0.4 µM Forward primer	0.4
0.4 µM Reverse primer	0.4
cDNA (diluted 1:3)	3.0
Nuclease free water up to 10 µl	1.2

Table 17. SYBR green qPCR cycle details

Cycle Step	Temperature ° C	Duration	Cycles
Enzyme inactivation	95	2 min	Hold
Denaturation	95	5 sec	40
Annealing	62	30 sec	
Extension	68	15 sec	
Melting			

2.6 PROTEIN ANALYSIS

2.6.1 Tris-glycine Sodium Dodecyl Sulphate Polyacrylamide Gel electrophoresis (SDS-PAGE).

Lysates were resolved together with 5 µl of PAGE- ruler plus pre-stained marker on 8% SDS-PAGE gel (Table 18) for FLAG tagged Cas9, and 10 % SDS-PAGE gels for TN-C and fibronectin. Proteins were extracted in 300 µl Laemmli buffer containing 5 % β-mercaptoethanol per 6 well plate and boiled

for 15 minutes at 95 °C. HEK-293 protein lysate was diluted 10-fold and prepared with 1 X loading buffer in a total volume of 30 µl and boiled for 15 minutes at 95 °C. Lysate from BJ fibroblasts were resolved in 1 X Tris-glycine running buffer at 20 mA for 30 minutes then at 30 mA for 1 h 40 minutes, using a BioRad tank and BioRad PowerPac Universal.

Table 18. Composition of solutions and buffers used to prepare different percentage SDS-PAGE gels (resolving and stacking) which varied depending on the size of the protein of interest.

Component	**Resolving gel**		**Stacking gel**
	8 %	10%	3.9 %
30 % acrylamide/ 0.8 % bisacrylamide	2 ml	2.5 ml	0.32 ml
4X Tris-Cl/SDS, pH 8.8	1.875 ml	1.875 ml	-
4X Tris-Cl/SDS, pH 6.8	-	-	0.62 ml
H_2O	3.625 ml	3.125 ml	1.525 ml
10 % (w/v) ammonium persulfate (APS)	25 µl	25 ul	12.5 µl
N,N,N',N'-Tetramethylethylenediamine (TEMED)	5 µl	5 ul	2.5 µl

2.6.2 Western blotting

Protein samples separated on SDS-PAGE gels were electroblotted onto nitrocellulose membranes. Membranes were first hydrated with water and then equilibrated in 1X, ice-cold transfer buffer for 10 minutes with shaking. Wet protein transfer was carried out using a BIORAD PowerPac Universal, at 350 mA for 60 – 90 minutes in ice-cold 1 X transfer buffer or at 30 V overnight in ice-cold transfer buffer in the cold room at 4°C. Generally larger proteins were transferred overnight to ensure successful transfer. Protein transfer was confirmed by visualization of the pre-stained protein ladder and staining

with 1 % Ponceau S solution. Membranes were destained with ddH_2O and then blocked with 10 mL of blocking buffer for 1 h at room temperature or overnight at 4 degrees before incubating with primary antibody overnight at 4°C as in Table 19. After 3 washes with 1 X PBST for 10 minutes with shaking, membranes were incubated in 5 mL of secondary antibody as in Table 19 for 1 hour at room temperature. After 3 washes with 1 X PBST for 10 minutes with shaking, bands were visualised on ImageQuant^{tm} LAS-4000 using the enhanced chemiluminescence (ECL) method. After detection, antibodies were stripped from the membranes with Re-blot plus stripping solution (Brand) and stored dry in a sealed plastic bag.

Table 19. Antibodies and buffers used for western blotting.

Antibody against target protein	**Size (kDa)**	**Blocking buffer**	**Primary antibody**		**Secondary antibody**
			Dilution	Source	Dilution/ raised against
Mouse- anti- flag (F108)	131 (FLAG-tagged Cas9)	5 % milk/ PBST	1:10,000 in 5 % milk/PBST	Sigma	1:5000 Anti-mouse IgG-HRP
Mouse-anti- TN-C (N- terminus MAB1908)	320	5 % BSA/PBST	1:1000 in 2% BSA/PBST	Millipore	1:5000 Anti–mouse IgG-HRP
Rabbit- anti- Fibronectin (F3648)	220	10 % milk/PBST	1:1000 in 10% milk/PBST	Sigma	1:5000 Anti-rabbit IgG-HRP
Rabbit – anti- Alpha tubulin (Ab52866)	50	2 % milk/PBST	1:5000 in 2% milk/PBST	Abcam	1:5000 Anti-rabbit IgG-HRP

2.6.3 Immunocytochemistry

Cell-derived matrices were grown in IBIDI chambers or on 18mm glass coverslips and fixed with 4 % paraformaldehyde in PBS at 4˚C for 15 minutes. Samples were analysed for the presence of collagen type I and II, and fibronectin. Matrices were incubated with blocking solution at room temperature for 1 hour followed by primary antibody in blocking solution at room temperature for 1 hour (Table 20). After 3 washes with 1 X TBS, matrices were incubated with the secondary antibody (Table 20) in

blocking buffer at room temperature for 1 hour. After 4 washes with TBS, samples were mounted in prolong gold (when coverslips were used) or had PBS added in plates. Samples were analysed by confocal microscopy (Zeiss LSM 510 confocal microscope) by Hilary Collins or by fluorescent microscopy (EVOS fluorescent microscope). For staining of phalloidin the above was carried out with the addition of a permeabilization step by incubation of samples in permeabilisation buffer for 15 minutes, as per manufacturers protocol.

Table 20. Primary and secondary antibodies for immunocytochemistry.

Primary antibody and probes	Dilution	Source	Secondary antibody (1:500)
Anti-fibronectin F3648	1: 200	Sigma	Anti-Rabbit Alexa-fluor 546/488/594
Anti-collagen I NB600408.1	1: 200	Novusbio	
Anti-collagen II NB600.594	1: 200	Novusbio	
Phalloidin Fluor™ - 488	1:200	ThermoFisher	-
DAPI staining	1:100	Invitrogen	-
Draq5	1:5000	Thermofisher	-

2.6.4 Senescence – associated β – galactosidase (SA-βG) staining of fibroblasts

Fibroblasts at PDLs indicated in figures were seeded at a low densities of 1 x 10^4 per 12 well of a 12 well plate. Positive control fibroblasts were treated with 150 µM of H_2O_2 for 2 hours and cultured for a further 72 hours. Cellular senescence was determined using a Senescence β-galactosidase staining kit (Cell Signalling technologies, #9860). Cells were fixed in 1 X fixing solution diluted in ddH_2O for 10 minutes at room temperature and then sealed plates were placed inside a sealed bag and incubated in the X-gal staining solution at pH 6.0 for 48 hours at 37°C in a dry incubator without CO_2. Per ml, the staining solution was comprised of 930 µl of 1 X staining solution (diluted in ddH_2O), 1 X solution A, 1 X solution B and 1 mg of X – gal solution.

2.6.5 ELISA of Tenascin- C, Interleukin -6 and -8

TNF-α production was measured using the TNF-α Duo Set ELISA kit from R&D Systems. For IL-6 and IL-8 ELISA kits were purchased from PeproTech. 96-well ELISA plates were coated with 50 µl of capture antibody diluted in PBS as indicated in Table 21 and incubated overnight at 4 °C. Plates were washed 3 times with washing buffer ensuring all liquid was removed. Wells were blocked with 50 µl of blocking buffer for 1 hour at room temperature and washed 3 times with washing buffer.

For assessment of human TNF-α, conditioned media of macrophages stimulated with LPS was centrifuged at 1500 g for 5 mins to remove any cell debris and supernatants were diluted 1:4 in blocking buffer. For human IL-6 and IL-8, supernatant was diluted in diluent at 1:2 and 1:4 respectively. A 7-point standard curve was prepared for all ELISAs by 1:3 dilution series of the protein standard at the starting concentration indicated in Table 21, in blocking buffer for TNF-α or diluent for IL–6 and –8. The samples and standards (50 µl) were incubated in capture antibody–coated plates at room temperature for 2 hours. Blank samples consisted of all reagents with complete media treated and stored as per samples, instead of cell supernatant. Plates were washed 3 times with washing buffer and 50 µl detection antibody diluted in diluent added at the concentration indicated in Table 21 before incubation at room temperature for 2 hours. Plates were washed 3 times in washing buffer. A 1:40 dilution of streptavidin-HRP was added to each well and incubated at room temperature for 20 minutes in the dark, before 50 µl of stop solution was added to each well. Absorption was read on a BioTek Synergy ™ HTX multimode microplate reader at 450 nm and blank corrected using the 'blank' sample.

Table 21. Concentrations of antibodies and standards for use in ELISAs.

Target	Capture antibody	Standard	Detection antibody
TNF-a	4 µg/ml	10 ng/ml	50 ng/ml
IL-6	0.50 µg/ml	2000 pg/ml	0.1 µg/ml
IL-8	1.00 µg/ml	150 pg/ml	0.25 µg/ml

2.7 BACTERIAL TECHNIQUES

2.7.1 Transformation of bacterial cells with CRISPR ligation products

Escherichia coli DH5a cells were utilised in all bacterial transformations. 50µl of heat-competent cells was thawed on ice and 50 ng of vector/plasmid and 10 µl of ligation products (CRISPR plasmid and gRNA) were added. The mixture was incubated on ice for 30 minutes, heat shocked at 42°C for 2 minutes, before incubating on ice for 5 minutes. 950 µl of LB medium was added and cultured at 37°C for 1 hour. Cultures were centrifuged for 2 minutes at 7,000 RPM. 920 µl of the supernatant was removed and the pellet was resuspended before plating on LB/Ampicillin agar plates and incubating overnight at 37°C. For each transformation heat-competent *E.coli* cells were subjected to the procedure without the addition of any plasmid DNA, which should act as a negative control due to being incapable of expressing ampicillin resistance. A positive control of undigested vector was also included. The negative ligation control, consisting of digested vector with no insert was also transformed.

2.7.2 Screening of transformed colonies harbouring CRISPR/plasmids

Single bacterial colonies were picked from transformation plates and used to inoculate a 5 ml starter culture of LB with ampicillin and incubated at 37°C in a shaking incubator for 16 - 18 hours until the culture was visibly turbid. These cultures were then miniprepped and sequenced to confirm the correct identity of the gRNA insert. On confirmation, the remaining starter culture was transferred to a 96 ml volume of LB/Ampicillin broth and grown overnight at 37°C in a shaking incubator and conical flasks. These cultures were then midiprepped and sequenced a final time to confirm the correct CRISPR plasmid sequence. A small aliquot of culture was stored at 4°C for use in creating a glycerol stock following sequencing results.

2.8 CELL CULTURE TECHNIQUES

2.8.1 Cell lines

Human BJ skin fibroblasts were from ATCC. Human embryonic kidney 293 (HEK293) and human breast adenocarcinoma MDA-MB-231s were provided by Hilary Collins. Peripheral human blood monocytes were isolated by Anna Piccinini.

2.8.2 Cell culture

BJ cells were cultured in EMEM medium, supplemented with 10 % (v/v) FBS, L- Glutamine (1 X) and Penicillin/Streptomycin (1 X) or Primocin™ (100 ug/ml). HEK293 and MDA-MB-231 cells were cultured in DMEM medium supplemented with 10 % (v/v) FBS and Penicillin/Streptomycin (1 X). Cells were cultured to 70-80 % confluence at 37°C with 5 % CO_2. Complete growth medium was changed every 2–3 days. While HEK293 cells were split by gentle pipetting, BJ and MDA-MB-231 cells were washed twice with 1 X PBS before adding 1 X trypsin/EDTA and incubating at 37°C until cells detached. FBS containing media was added to neutralise the trypsin/EDTA activity and the suspension centrifuged at 1500 RPM for 5 minutes before resuspending the pellet in complete media. HEK293 cells were subcultivated at ratios of 1:2 to 1:10, MDA-MB-231 at ratios of 1:3 to 1:6 and BJ cells at ratios of 1:2 to 1:9. For aged fibroblasts the split ratio gradually decreased to no more than 1:2 or 1:3. Primary human monocytes were seeded at a density of 1 x 10^6 / ml and differentiated to monocytes in RPMI1640 supplemented with 5% FBS and Penicillin/Streptomycin (1 X) and M-CSF (100 ng/ml) for 5 days. After this, macrophages were cultured without M-CSF.

2.8.3 LPS stimulation

BJ fibroblasts, HEK293 and MDA-MB-231 cells were stimulated with 100 ng/ml LPS in complete media, whilst macrophages were stimulated with 1 ng/ml LPS in complete media.

2.8.4 Cell cycle synchronisation

BJ cells were cultured to 80% confluence in 10 cm dishes, serum starved for 24 hours, before adding complete growth medium with 15% FBS and 20 mM Aphidicolin for 48 hours. Cells were washed twice with 1 X PBS and released from the G_1 hold with the addition of complete growth media with 10 % FBS. Cells were harvested as described below for flow cytometry analysis.

2.8.5 Flow cytometry

2.8.5.1 Analysis of synchronised cells using Propidium Iodine staining

Harvested BJ fibroblasts were washed with 1 x PBS and centrifugation at 1500 RPM for 5 minutes, before being resuspended in 500 µl 1 x PBS per 10 cm dish. Cell suspensions were gently vortexed for 5 seconds, then fixed by the addition of 4.5 ml of ice cold 70% ethanol. Samples were centrifuged at 1500 RPM for 5 minutes and the ethanol wash removed. Cell pellets were washed twice with 500 µl of 1 x PBS before resuspension in 20 µl of Sigma RNase A solution (catalogue #R6148) and incubation for 2 minutes at room temperature. Finally, 400 µl of PI staining solution was added to each sample and thoroughly mixed. Samples were transferred to 5 ml FACS tubes and allowed to sit for 4 hours in the dark at 4°C. The fluorescence of each cell suspension was quantified using a Beckman Coulter FC500 flow cytometer (University of Nottingham) with excitation at 536 nm and emission measured at 617 nm wavelengths. An unsynchronised and unstained sample was also analysed.

2.8.5.2 Analysis of macrophage maturation

Monocytes seeded at 8 x 10^4 /96-well and differentiated with M-CSF (100 ng/ml) in the presence or in the absence of CDM were with or without BJ-derived matrix. Following permeabilization of cells, CD-68-PE expression was assessed by Anna Piccinini by FACS on an FC500 flow cytometer (University of Nottingham) and data analysed using Weasel.

2.8.6 Total cell lysates for western blotting

Total cell lysates from HEK293, MDA-MB-231 and BJ cells and their CDMs were prepared using 1 x Laemmli buffer with 5 % β–mercaptoethanol and resolved on 8% or 10% SDS-PAGE gels and analysed by western blotting as per section 2.6.2.

2.8.7 Gelatin coating of plates and coverslips

Plates and coverslips were coated with 0.2 % sterile gelatin for 1 hour at 37°C, before washing with PBS and cross-linking under UV light for one cycle of the UV light filter in a TC hood. Plates/coverslips were equilibrated with serum-free media at 37°C for 30 minutes and cells seeded immediately or within 24 hours (plates stored at 4°C).

2.8.8 Transfection of cells

2.8.8.1 Calcium phosphate co-precipitation

HEK293 cells were transfected by calcium phosphate co-precipitation. Cells were seeded at a density of 4 x 10^5 cells per well in 6-well plate 24 hours prior to transfection. 2 hours prior to transfection, DMEM was replaced with fresh media. For each sample, 200 µl of 2 x HEPES Buffered Saline (HBS) solution and a separate calcium chloride reaction mix consisting of 25 µl of 2 M calcium chloride solution, 1 µg of each vector, was made up to 100 µl with sterile diH_2O. Calcium chloride was added dropwise to the HBS solution, whilst gently vortexing the HBS. The mixture was allowed to precipitate for 20 minutes, before adding it dropwise to the cells. After 16 hours the media was removed and the cells were washed with PBS, before adding fresh growth media. The transfection efficiency was evaluated by light and fluorescence microscopy 24 hours later. Genomic DNA or proteins were extracted 48-72 hours after transfection.

MDA-MB-231 cells were transfected as above, but were stimulated with LPS 4 hours before transfection, and transfected with 6 µg of each vector. A 2 minutes 15 % glycerol shock step was also performed before replacing with fresh media 16 hours after transfection.

2.8.8.2 Electroporation of fibroblasts for CRISPR/Cas9 plasmid (approach 1)

On the day of transfection, cells were trypsinised as in **2. 5. 2** and 2.5 x 10^5 cells per reaction were counted before centrifuging at 90 g for 10 minutes. Cells were resuspended in 20 µl of SF cell line solution (82% buffer, 18% supplement: Lonza) per reaction, before adding 1 µg of each vector or pmaxGFP control. The reactions were transferred to a 16-well strip cuvette and pulsed with programme FF-120 (4D Nucleofector, Lonza). After incubation at room temperature for 10 minutes, 80 µl of pre-warmed media without antibiotics was added to each cuvette before transferring to a pre-warmed 6-well plate with complete EMEM medium without antibiotics. Cell viability and transfection efficiency were analysed by light and fluorescence microscopy 24 hours later. Positive cells, expressing the green fluorescent protein (GFP) reporter gene were selected by Fluorescence Activated Cell Sorting (FACS) 48 hours after transfection.

2.8.8.3 Electroporation of fibroblasts for CRISPR/Cas9 RNP (approach 3)

Experiments carried out as in section 2.5.7 with the exception that 15.5 µl of P3 solution from a P3 Primary Cell 4D-Nucleofector® X Kit (82% buffer, 18% supplement; Lonza) was mixed together with 4.5 µl of RNP complex for a final volume of 20 µl per reaction. Reactions were pulsed with programme CM-138 (4D nucleofector, Lonza). Cells were expanded without selection.

2.8.9 Selection of transfected cells (Approach 1 and 2) using fluorescence activated cell sorting (FACS)

Cells transfected with CRISPR/Cas9 or CRISPR/Cas9n vectors were sorted 48 hours after transfection based on their GFP expression using the Beckman Coulter MoFlo XDP (Flow Cytometry Facility, University of Nottingham). For sorting, cells were detached as in **2. 5. 2.** 48- 72 hours after transfection. Cells were resuspended in 500 µl of complete growth medium and transferred into a 5 ml FACS tube and kept on ice. Cells were double sorted; the first sort to separate GFP^+ cells from non-fluorescent cells (bulk sorting) and then single GFP^+ sorted cells were each seeded into a single well of 96-well plates containing 100 µl/well of complete growth medium. The remaining bulk GFP^+ cells were seeded in a

6-well plate and grown to 70 % confluence before genotyping. An additional 100 µl/well of complete media was added to plates on return to the lab before placing in a 5 % CO_2 incubator. Flow cytometry data was analysed using Weasel (https://frankbattye.com.au/Weasel/).

2.8.10 Preparation of cell-derived matrices

BJ cells were resuspended at a cell density of 1 x 10^5 cells/mL and plated as follows: 12.5 ml in a 10 cm dish ,1 ml in a 12-well plate, 1 ml in a IBIDI 35ml round dish, 2 ml in a 6 well plate. IBIDI 8 well chambers cells were seeded at 2.7 x 10^4/300 µl, 96 well plate cells were resuspended at 1.25 x 10^4 and 100 µl seeded. Upon cells reaching 100 % confluence, fresh growth medium was supplemented with 50 µg/ml L-ascorbic acid (L-AA; Sigma). The L-AA-containing medium was changed every day for 14 days to promote collagen rich matrix deposition. Matrices were analysed by immunocytochemistry after growing on 18 mm coverslips in 12-well plates or in IBIDI 8–well chambers.

2.8.11 Matrix decellularisation

BJ cells were washed twice with 1 x PBS before pre-warmed extraction buffer was added (6 mL/10cm^2). Cells were lysed for 2 minutes until no intact cells were visible by phase microscopy. The extraction buffer was removed, and matrices washed 3 times with 1 x PBS containing Ca^{2+} and Mg^{2+}. Residual DNA was digested with 10 µg/ml DNase I in 1 x PBS containing Ca^{2+} and Mg^{2+} for 30 minutes at 37˚C. DNase I was removed, and the matrix washed twice as above. Matrices were fixed with 4 % paraformaldehyde or used immediately.

2.8.12 TN-C coating of WT CDMs/ Generation of CDMs overexpressing TN-C

Decellularised CDMs in 12-well plates were coated with full-length human TN-C (100 µg/ml Sigma-Aldrich CC065) diluted in *TNC*-Azide coating buffer at concentrations of 1.5 µg/ml, 3 µg/ml and 6 µg/ml, in a total of 1 ml/well. Plates were wrapped tightly in parafilm and incubated overnight at 4ºC. Buffer was aspirated and wells washed twice with 1 x PBS before blocking wells with 100 µl of sterile-

filtered 1% BSA for 30 minutes at room temperature. Wells were washed twice with 1 x PBS and cells fixed and stained for the analysis of TN-C or macrophages seeded immediately.

2.8.13 MTT assay

Cells were seeded at 3600 cells per well in a 96-well plate and incubated overnight before LPS stimulation. Media was aspirated and 10% (v/v) MTT solution was added to each well for 4 hours and incubated at 37˚C, 5% CO_2 until purple formazan crystal formation was visible. MTT stop solution was added to solubilise formazan and the absorbance read at 590 nm using a BioTek Synergy HTX multimode microplate reader. Unstimulated cells acted as controls. A CDM matrix only was also included as a control.

2.8.14 Adhesion assay

Monocytes seeded at 8 x 10^4 /96-well and differentiated with M-CSF (100 ng/ml) in the presence or in the absence of CDMs were tested for their ability to adhere to the substrate. Briefly, mature macrophages were removed from the CDM by inverting the plate and washing gently twice with 1 x PBS. Cells were fixed for 20 minutes with 100 µl of staining buffer. Wells were washed twice with 1 x PBS and then lysed in lysis buffer for 5 minutes. Absorbance was read at 590nm.

2.8.15 Alcian blue staining of BJ-derived CDMs

Matrices were generated as described above in 12-well plates, decellularised, fixed with 4% PFA and incubated overnight at 4ºC with 300 µl of the Alcian blue/$MgCl_2$ staining solutions at the differing $MgCl_2$ concentrations as indicated. Plates were washed four times with 1 x PBS and imaged by phase microscopy. When necessary, after decellularisation, wells were treated with trypsin for 10 minutes to degrade the proteoglycan core of the GAGs. This was to act as a heparinase substitute, allowing removal of GAGs during the wash step, to assess staining specificity.

2.8.16 Optical profiling

CDMs were generated in 60 mm plates before decellularisation. CDMs were unfixed and thickness and roughness analysed immediately by optical profiling using a Zeta-20 optical profiler (KLA). The focus finder setting was applied to enable focussing on the matrix in order to set the 'top' and 'bottom' positions for measuring. Plates were sectioned into 5 sample areas with 44 sample points per area.

2.8.17 Environmental Scanning Electron Microscopy (SEM)

Decellularised matrices deposited on 18 mm coverslips with macrophages, or macrophages on coverslips were fixed and washed ten times with sterile ddH_2O to remove salt crystals resulting from storage at 4 °C in PBS. Samples were mounted onto the pre-cooled Peltier stage at 20°C and analysed using FEI Quanta650 ESEM. The water vapour pressure in the chamber was controlled to maintain a humidity of approximately 90%. For matrices plus macrophages the temperature was lowered to -50°C and humidity reduced to 78 - 80%. For macrophages on coverslips the temperature was maintained at 20°C and 78 – 80% humidity, the voltages were also changed to optimise imaging of the cells. Samples were analysed by Ms Nicola J. Weston at the Nanoscale and Microscale Research Centre (nmRC).

3 GENERATION AND CHARACTERISATION OF FIBROBLAST CDMS

3.1 INTRODUCTION

The view of the ECM as a passive scaffold is now outdated - it is becoming increasingly clear that there is a reciprocal and dynamic exchange of information between the ECM and cells to regulate all aspects of cell behaviour [1]. Integrins are the most widely studied ECM component with regard to their role in mediating interactions between cells and ECM molecules. Integrins allow ECM linkage to the cytoskeleton, resulting in changes to cell morphology and function [441]. Components such as TN-C have been implicated in posttranscriptional regulation of inflammatory genes [312]. The active role and interaction of the ECM with the innate immune response and pathogens during infection is also being recognised [442]. However, studies on the exact role are difficult due to the lack of available physiological models. Many commercial models of ECM exist, however these are subject to limitations: mainly derived from animal and/or tumour origins [380, 443], the latter in which TN-C expression is commonly upregulated; and in a similar vein such animals derived models contain, immunogenic and pathogenic component contamination, prohibiting the application of these CDMs in immune cell culture. Synthetic ECM models, such as hydrogels, including Matrigel [385] still fail to fully recapitulate the molecular, biochemical, and structural complexity and intricacy of naturally occurring ECM. Synthetic poly-ethylene-glycol (PEG) hydrogel models usually emulate either the three-dimensionality or stiffness [444]. Synthetic models also lack the complete catalogue of ECM components; composed of only a select few components only, and absent of the numerous versions of ECM proteins resulting from posttranslational modifications, including glycosylation, phosphorylation etc, and isoforms resulting from alternative splicing. Matrigel is also victim to variable compositional and structural characteristics between batches [385, 445]. This makes it hard to carry out well controlled studied with these types of models and also highlights problems which may occur if such models are to be used for generating ECM for human therapies. This means current ECM models are not suitable when exploring

the physiological role of healthy, human ECM in gene regulation. Therefore, the first objective of this project was to develop and characterise a physiologically relevant, healthy human ECM model.

Fibroblasts are one of the most commonly occurring cell types in connective tissues and the principal producers of the ECM and TN-C. Among normal human fibroblast cell lines, skin fibroblasts such as the BJ cell line are a particularly suitable cell model as they grow into a confluent monolayer, have a relatively long lifespan in comparison with other normal fibroblast cell lines and, due to their intrinsic ability to secrete and assemble a functional ECM, they allow for the production of cell-derived matrices (CDMs; Figure 13) which are used here as 3D physiological matrices that mimic the cellular microenvironment. Once cells have assembled enough ECM, they can be removed by decellularisation to leave a cell-free CDM, which can be used as a substrate for the culture of other cells of interest.

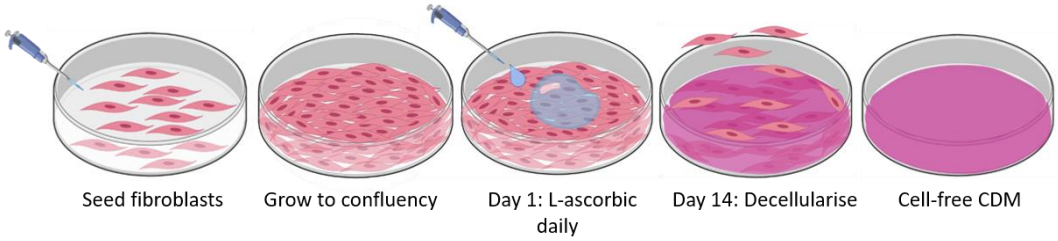


Figure 13. Schematic diagram of the process used to create cell-derived matrices (CDM).
Cells are seeded and cultured to 100 % confluency, before adding L-ascorbic acid for 10-14 days. Cells are removed with decellularisation buffer composed of 20mM ammonium hydroxide (NH_4OH) and 0.5 % Triton X-100, followed by treatment with DNase. BJ-derived- CDM is then left coating the tissue culture dish.

3.2 BJ-DERIVED CDM MODEL: PRELIMINARY ASSESSMENT

3.2.1 Biochemical analyses

The first objective was to investigate the BJ-derived CDM with specific focus on the components synthesised and if the resulting CDM was robust enough to handle to allow decellularisation and, further down the line, subsequent application as a cell culture substrate. To do this BJ fibroblasts were cultured on tissue culture plastic and allowed to deposit CDM for 14-days in the presence of L-ascorbic acid (50 µg /ml). This work was carried out by Dr Anna Piccinini and Er Pey Ling. BJ fibroblasts grow in a confluent, characteristically skin-like monolayer in which alignments of the spindle shaped fibroblasts can be seen (Figure 14 A). This organisation and alignment was reflected in the decellularised BJ-

derived matrix, which exhibited an organised meshwork and striated pattern, a documented occurrence [446].

Next, the presence of the ECM components Collagen 1 (COL1) was verified by immunostaining. COL1 was assessed as this is one of the most abundant ECM components (along with COL3 and FN) in skin connective tissue, in which fibroblasts reside *and* synthesise and therefore should be very prevalent in BJ-derived CDM. Collagens in particular account for 30 % of total protein mass in the body [116]. Immunostaining of the non -decellularised BJ-derived CDM demonstrated a rich COL1 network, interspersed with ubiquitous DAPI staining of the nucleus and actin cytoskeleton staining of the resident fibroblasts.

This prevalent DAPI staining is in marked contrast to the decellularised BJ-derived CDM in which little/no DAPI staining can be seen (Figure 14 A). Removal of fibroblasts and residual DNA was achieved following the protocol by Kaukonen *et al.* [447] by incubation with a Triton X-100 and NH_4OH containing buffer followed by a DNAase I treatment. The lack of DAPI staining indicates that fibroblasts residual DNA was efficiently removed. Together, this indicates that BJ-fibroblasts can secret a COL1 rich CDM that can successfully be decellularised.

As previously mentioned, COL1, 3 and FN are the most abundant ECM components in skin connective tissue. Therefore, the presence and deposition of these proteins into a BJ-derived CDM was investigated. To determine the content of BJ-derived CDMs, decellularised BJ-CDM was separated into soluble and insoluble fractions using a modified RIPA buffer and urea lysis buffer. Dialysis and subsequent protein quantification by Bradford staining revealed that insoluble proteins accounted for the majority of proteins in the CDM (0.329 ± 0.01 mg/ml) with far less insoluble proteins (0.055 ± 0.01 mg/ml) (Figure 14 B). Proteins incorporated into the meshwork of the ECM, such as FN and COL are insoluble, whilst ECM-associated molecules such as growth factors are soluble [448]. Therefore, the finding that the majority of proteins are insoluble is consistent with the assembly of proteins into an ECM.

Taken further, after solubilisation, SDS-PAGE electrophoresis, colloidal blue staining of the insoluble and soluble extracts revealed two distinct bands in the insoluble fractions (Figure 14 C). These bands

correspond to the large structural proteins collagen (130 kDa) and FN (250 kDa). This is expected and indicates collagen cross-linking and assembly into insoluble fibers, and FN assembly into insoluble fibrils. If not formed into fibers and fibrils, these proteins would be found in the soluble fraction. The presence of these bands can also be seen in the 'cells and matrix' samples, the intensity of which looks similar, indicating that the decellularisation process does not result in perturbation of these structural components. The above demonstrates the assembly of large structural proteins into a BJ-derived ECM.

Confirmation of the assembly of an ECM was completed by immunofluorescence analysis of collagen 1 (COL1), collagen 3 (COL3) and fibronectin (FN) in decellularised BJ-derived CDM, which indicated a dense meshwork of these ECM molecules (Figure 14 D). Interestingly, COL3 particularly mirrors the pattern of cells in the non-decellularised matrix (Figure 14 A top left). It has been demonstrated that COL1 can orientate cells [449], however to what extent this occurs with fibroblasts that are in situ before and during depositing collagen remains unclear. Though it does seem likely that fibroblasts direct assembly in alignment with themselves given that many ECM components require cell-mediated integrin-dependent assembly (i.e. fibronectin) [450]. However, it is clear that FN and COL are assembled into a highly organised meshwork. COL1 and COL3 have a similar molecular structure but a distinctively different localisation patterns with COL1 density appearing increased compared to COL3. This is consistent with COL1 comprising the main structural protein in the ECM followed by COL3 [1]. FN can be seen as a very consistent network. The FN network appears to have no patterns that mirror the fibroblasts alignment, in comparison to COL1 and 3 which may be consistent with this protein binding to many ECM proteins, including collagens. This confirms the generation of BJ-derived ECM rich in the main components of skin connective tissue that can successfully be decellularised.

As this research project focuses on the ECM with specific reference to the glycoprotein TN-C and its role in the LPS- induced inflammatory response [188], the final preliminary work assessed if TN-C is synthesised and assembled into an ECM by BJ fibroblasts. After separation of BJ-derived matrices into insoluble and soluble fractions, as previously and following SDS-PAGE electrophoresis, western blotting demonstrated TN-C is found largely in the insoluble fraction and, in smaller amounts, in the soluble fraction. It is known that TN-C has a large number of binding partners, further enhanced by

alternative splicing of TN-C FN type III like repeats [200], including fibronectin [182], collagen [451] and proteoglycans such as perlecan and aggrecan [183], that is consistent with the finding that TN-C is found in the insoluble fraction. The location of TN-C in the soluble fraction can be explained given that TN-C can also bind soluble factors such as PDGF/VEGF, phosphatidylinositoL-glycan biosynthesis class F protein (PIGF)-2, -3, TGF-β and insulin-like growth factor (IGF-1/-2) amongst others [200, 328]. Together, these experiments demonstrated the assembly of a *TNC* rich BJ- derived ECM.

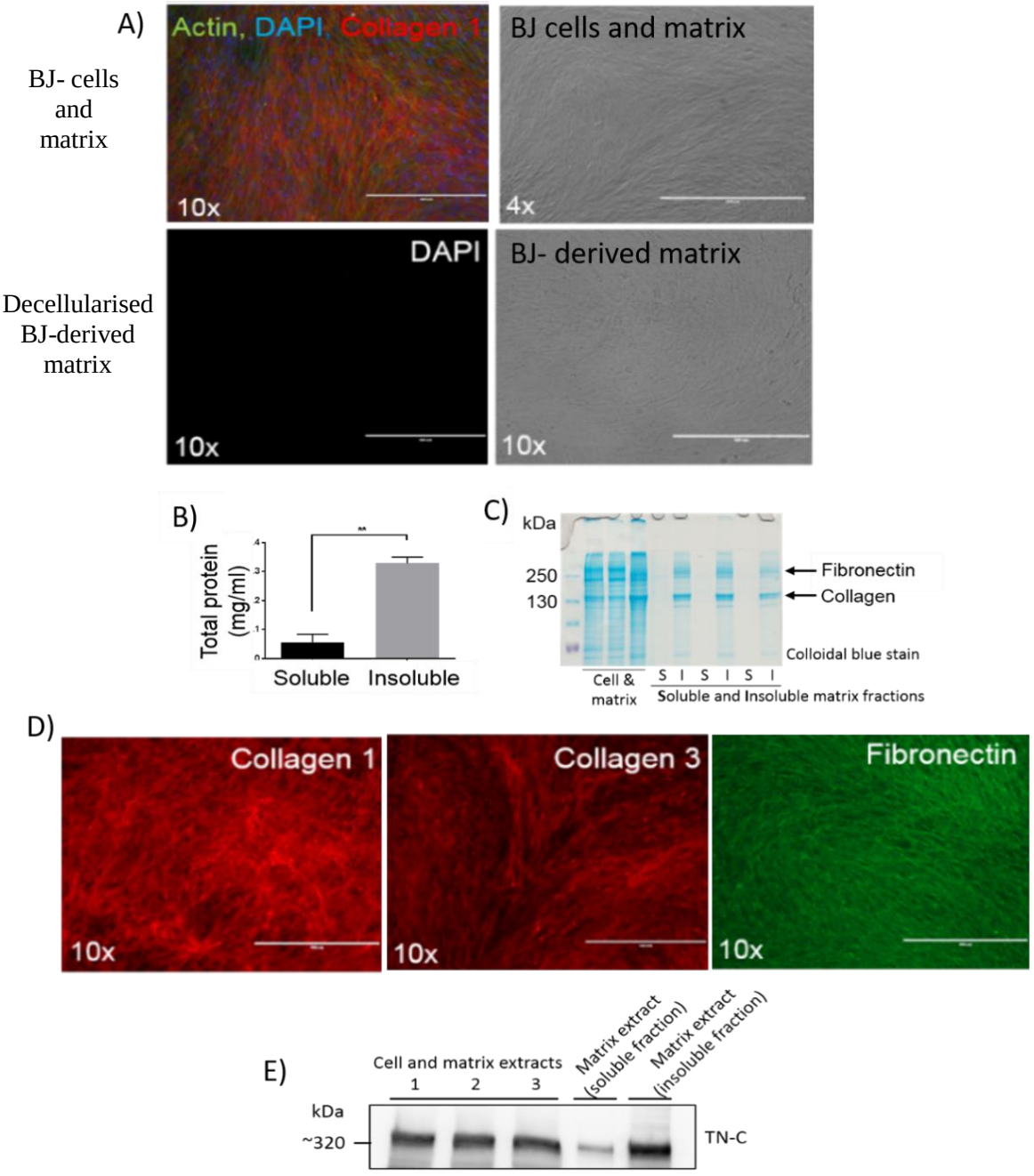


Figure 14. BJ-derived CDMs secrete a COL1, COL3 and FN rich matrix which is assembled into a 3D ECM and can be successfully decellularised.
A) BJ derived matrices deposited in the presence of L-ascorbic acid, without decelluarisation were stained for actin, DAPI, COL1 (top left) and the corresponding light microscopy image (top right). BJ derived matrices deposited in the presence of L-ascorbic acid, followed by decelluarisation and DNase treatment were stained for DAPI (bottom left) and the corresponding light microscopy image of BJ-derived CDM (bottom right). Scale bar 400 µM. **B)** Total proteins (mg/ml) in soluble and insoluble matrix extracts after solubilisation in RIPA buffer or urea lysis buffer, following dialysis and quantification by a Bradford assay (mean ±SD n = 3). **C)** Colloidal blue staining of soluble and insoluble extracts from BJ-derived matrix and cells and decellularised BJ-derived CDM on a 5 % polyacrylamide gel following fractionation into soluble and insoluble fractions and SDS-page electrophoresis. N =3. **D)** BJ derived matrices deposited in the presence of L-ascorbic acid, followed by decellularisation and DNase treatment were stained for COL1, COL3, and FN. Scale bar 400 µm. **E)** Immunoblot analysis of TN-C in BJ-derived CDM and cell extracts, fractionated into insoluble and soluble fractions. N = 3. Work carried out by Dr Anna Piccinini and Er Pey Ling.

3.3 BJ-DERIVED CDM: OPTIMIZATION OF A STANDARDISED AND REPRODUCIBLE ECM MODEL

Following on from previous work, BJ fibroblasts were cultured to 100 % confluence and supplemented with L-ascorbic acid (50 μg/ml) every day for 14 days to generate CDMs (Figure 13) as per the protocol from Kakounen *et al* [447]. Early work demonstrated L-ascorbic acid was necessary for collagen formation in tissue [452] and later was shown to promote the transcription of collagen genes and collagen mRNA stability in skin fibroblasts [453, 454]. L-ascorbic acid is now well established as a cofactor for LOX, which contributes to tissue and organ homeostasis and regeneration via collagen cross-linking [121]. By altering the amount of L-ascorbic acid supplemented to fibroblasts (50 ug/ml daily or 2 x per week for 14 days), we confirmed its activity in stimulating collagen biosynthesis and cross-linking without affecting the deposition of other ECM molecules, specifically TN-C and FN (Figure 15). Given that collagen 1 is the main structural ECM protein and contributes to mechanical attributes such as stiffness [455] this indicated that ECM architecture and stiffness could be modulated in this way. Collagen expression appeared the densest with the most consistent network in the daily supplementation condition, therefore this method was utilised to generate stable matrices for the duration of the project.

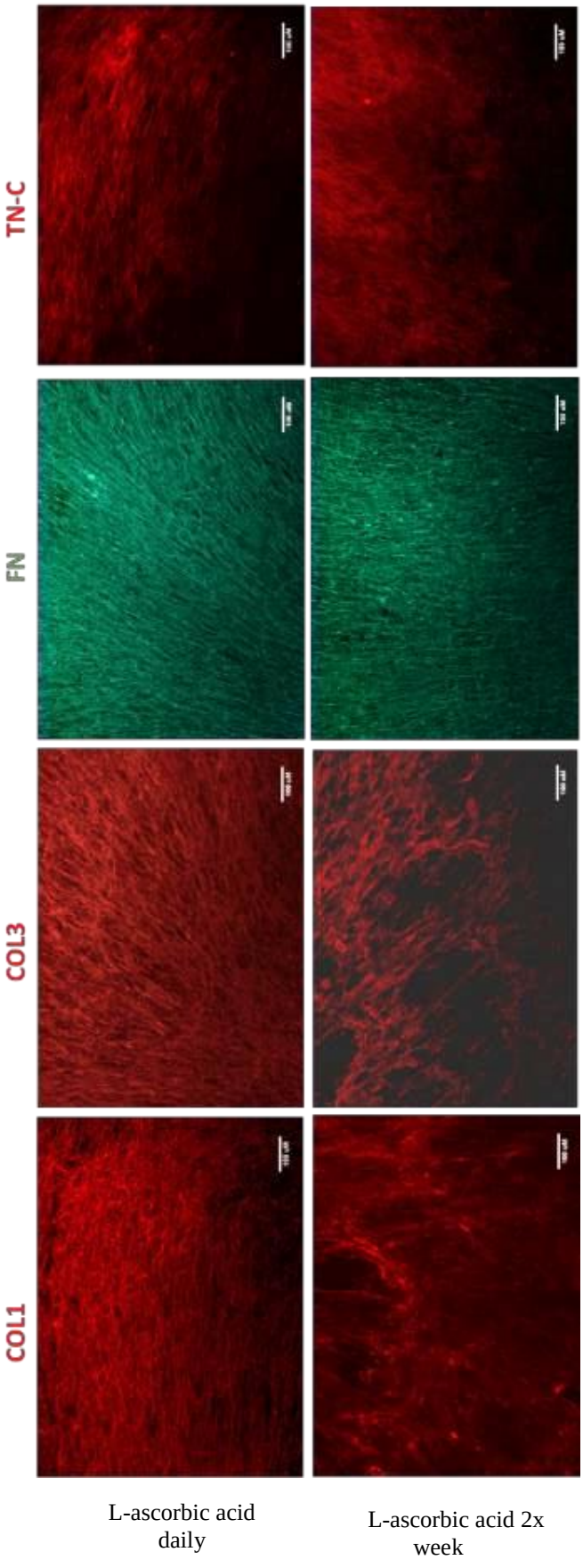


Figure 15. Modulation of L-ascorbic acid results in differences in collagen deposition and cross-linking.
Immunostaining of COL1, COL3, FN and TN-C after fibroblasts were cultured in the presence of L-ascorbic acid (50 µg/ml) daily or twice a week, for 14 days. DAPI staining of genomic DNA (blue) was used to stain the nuclei. Scale bar 100 uM.

Following on from the preliminary work by Anna Piccinini and Er Pey Lin, after the generation of BJ-derived CDM with L-ascorbic acid daily, CDM proteins and fibroblasts within the CDM were assessed via immunofluorescence by Anna Piccinini. Consistent with the preliminary findings, light microscopy and DAPI staining demonstrated fibroblasts aligned confluently and nematically in a monolayer and produced an ECM with a highly linearized pattern, both locally and globally, and both before and after decellularisation (discussed later) (Figure 16 B, C). This has been demonstrated before in fibroblast-derived matrices [456]. This linearized pattern may have particular physiological relevance in vitro and interestingly this pattern is not replicated in solubilised commercial models, such as Matrigel (Figure 16 D from [457]). As mentioned previously, this pattern may be due to fibroblasts directing ECM molecule assembly via integrins and therefore the difference in architecture could be due to the cell-independent method of thermal polymerisation. The actin cytoskeleton could be seen throughout the CDM, with its deposition almost masking COL3, compared to that of COL1 in which COL1 and actin staining appeared very similar. The consistent actin staining is as expected due to the large numbers of fibroblasts residing within the ECM. This analysis of CDMs indeed confirmed that BJ fibroblasts secrete large amounts of ECM proteins that appear in a dense meshwork and allowed us to generate a reproducible and effective method for producing matrices containing collagens, FN and TN-C (Figure 16 B, C).

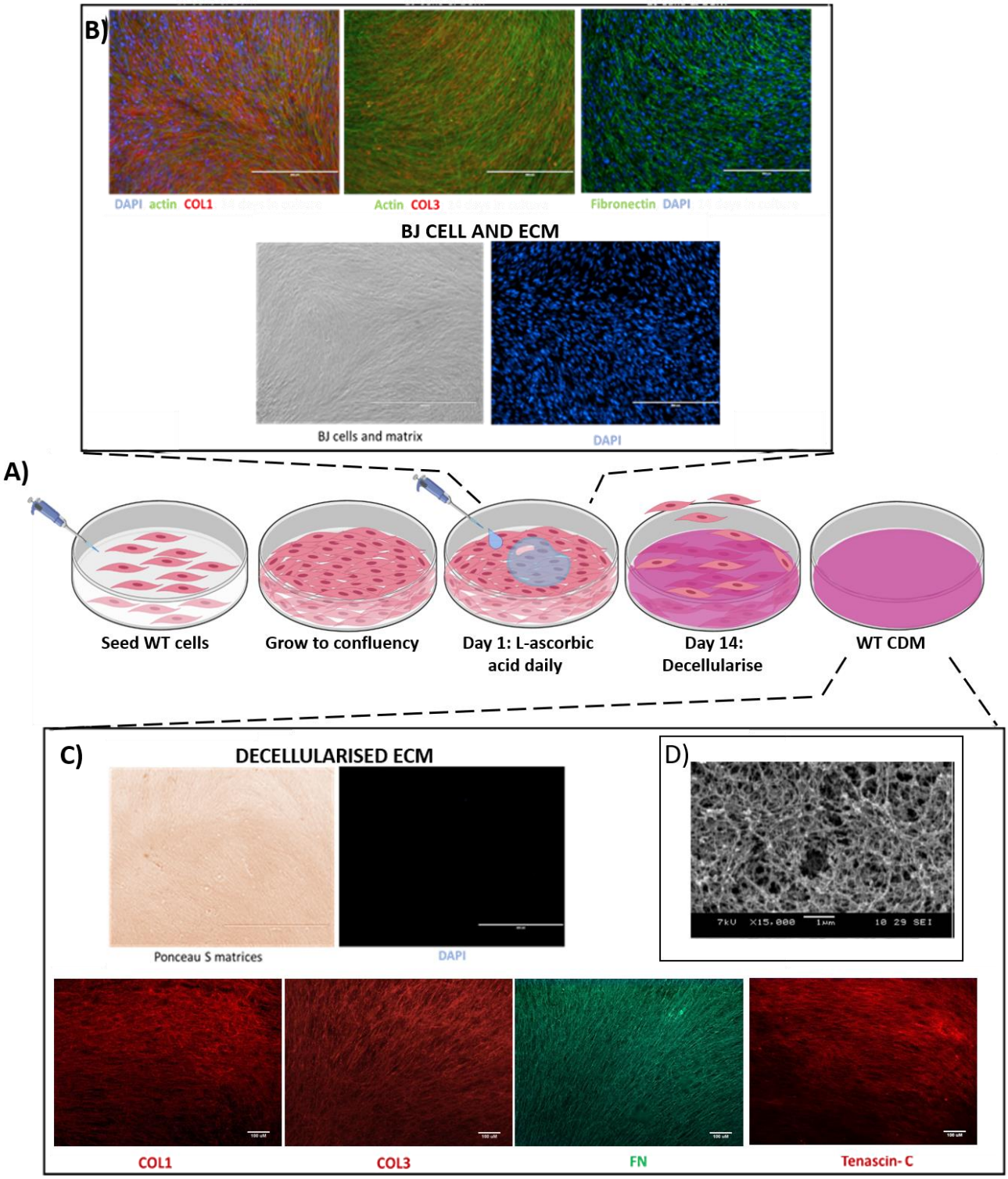
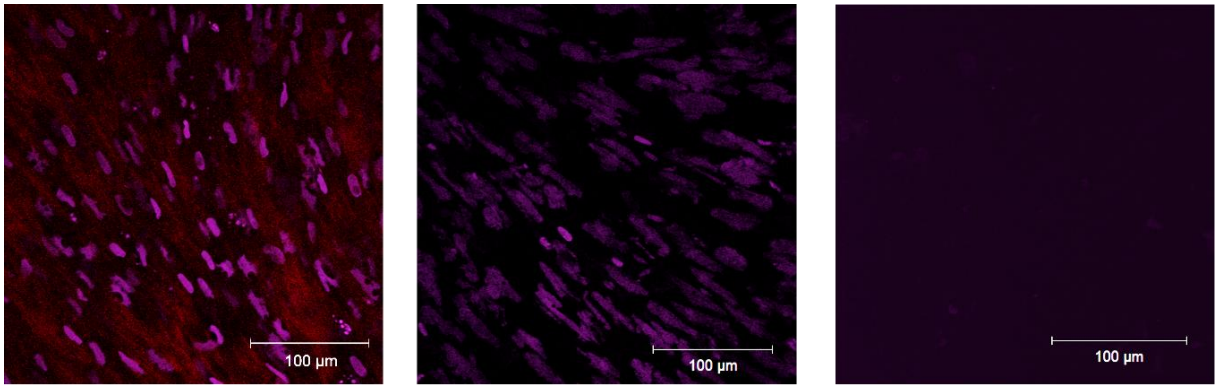


Figure 16. Generation and decellularisation of CDMs by BJ fibroblasts.
A) Workflow of CDM generation. BJ fibroblasts were seeded onto tissue culture dishes for 100% confluence 24 hours after seeding and cultured for a further 14 days with complete media supplemented with 50 µg/ml L-ascorbic acid. Matrices were decellularised using a buffer containing 0.5% Triton X-100 and 20mM ammonium hydroxide, followed by DNase I treatment. **B)** Top panels: immunofluorescence staining of BJ cells and their ECM molecules; actin (green), COL1 and COL3 (red) and FN (green). Bottom panels: bright field microscopy image (left) and DAPI staining (right) of BJ cells and their ECM before decellularisation. Scale bar 400um. All images in panel B by Anna Piccinini. **C)** Top panels: Ponceau S staining of decellularised matrix (left) and DAPI staining (right); scale bar 400um. Bottom panels: Immunofluorescence of the ECM molecules COL1 and COL3 (red), FN (green) and Tenascin-C (red) after decellularisation. Scale bar 100um. **D)** Scanning electron microscopy images of Matrigel scale bar 1 µM, taken from **[457]**.

3.4 DECELLULARISATION OF CDMS

To utilise CDMs as a substrate for cells, including macrophages, an efficient decellularisation process is needed in order to avoid an immunological response caused by remaining immunogenic components, such as fibroblasts residual DNA (14). This issue has been demonstrated in porcine-derived orthopaedic scaffolds which have caused inflammatory response in patients, resulting from residual porcine DNA [458]. However, the decellularisation method must not be so aggressive that it damages the ECM. Damages poses two problems, firstly to investigate the role of the healthy ECM an undamaged and intact ECM recapitulating the native ECM environment is required. Secondly, many ECM fragments can act as bioactive matrikines or DAMPs that activate the innate immune response, in a process of sterile inflammation [459]. For example the TLR2 signalling pathway can be activated by the fibronectin extra domain A fragment [460], TLR4 can be activated by fibronectin fragments [285, 458] and TN-C FBG domain [188]. Hyaluronan and biglycan can also activate TLR2 and 4 [288, 461, 462]. ECM molecules can also enhance the inflammatory response by modulating cytokine expression, for example collagen fragments augmenting or supressing IL-1β [463]. Ponceau S staining of total proteins, immunofluorescence analysis of the ECM proteins fibronectin, COL1, COL3 and TN-C, and staining of genomic DNA with DAPI confirm the decellularisation process effectively removes cells and any remaining genomic DNA. Decellularised matrices appear to be intact with a large visible number of interconnected fibers arranged into a dense sheet (Figure 16 C). This demonstrates that BJ-derived CDMs can effectively be decellularised whilst maintaining ECM architecture.

Many papers have studied the damage to matrix proteins in whole lungs resulting from decellularisation, including collagen, laminin, elastin and fibronectin [464]. There appears to be trade-off between effective cell and DNA removal and maintenance of ECM architecture; however Triton X- 100 containing buffers appear to be the most effective at removing the majority of cells, whilst leaving ECM components unperturbed [465]. Treatment with DNase I after incubation with the decellularisation buffer efficiently removed any residual DNA (Figure 17).



No decellularisation | Treated with decellularisation buffer only | Treated with decellularisation buffer and DNase

Figure 17. Treatment with Triton X-100 containing buffer followed by DNase I treatment successfully removes cells and residual DNA.
Confocal images of Draq5 stained DNA in BJ-derived CDMs following either; no decelluarisation (left), decellularisation for 2 mins with Triton X-100 and NH_40H containing decellularisation buffer (middle) or decellularisation buffer followed by DNase I treatment (right). Scale bar 100 µm.

Triton-X-100-containing decellularisation buffers have been shown to strip glycosaminoglycans (GAGs) of deposited proteoglycans, however this method has been demonstrated as superior in retaining collagen and GAGs within deposited ECM compared to other decellularisation methods [466]. The removal of GAGs is problematic due to their role in both the structure and function of the ECM (22). Alcian blue staining of cells with their matrix and decellularised matrices (CDMs) was used to discern different sulfated GAGs. Alcian Blue stains with increasing selectivity as increasing amounts of magnesium chloride are incorporated into the dye solution [467]. This indicated that decellularisation does not remove GAGs; an increase in GAG staining is seen after decellularisation. This was particularly evident for more strongly sulfated GAG types such as heparan sulphate and keratin sulfate (Figure 18 A). The increase in staining is at first surprising, however as proteoglycans interact with numerous ECM components, it could be that an 'unmasking' effect of bound GAGs occurs after decellularisation. Staining specificity was tested by very limited trypsin treatment of CDMs after performing alcian blue staining; trypsin allows the removal of GAGs by cleaving proteoglycans at their protein core (Figure 18 B). A reduction in staining is therefore indicative of specific staining; this can be seen particularly at the stronger concentration of $MgCl_2$. In summary, we have developed an efficient, reproducible method for both generating CDMs rich in fibril forming proteins, and decellularising the CDMs effectively to remove potentially immunologic components without damaging the matrix.

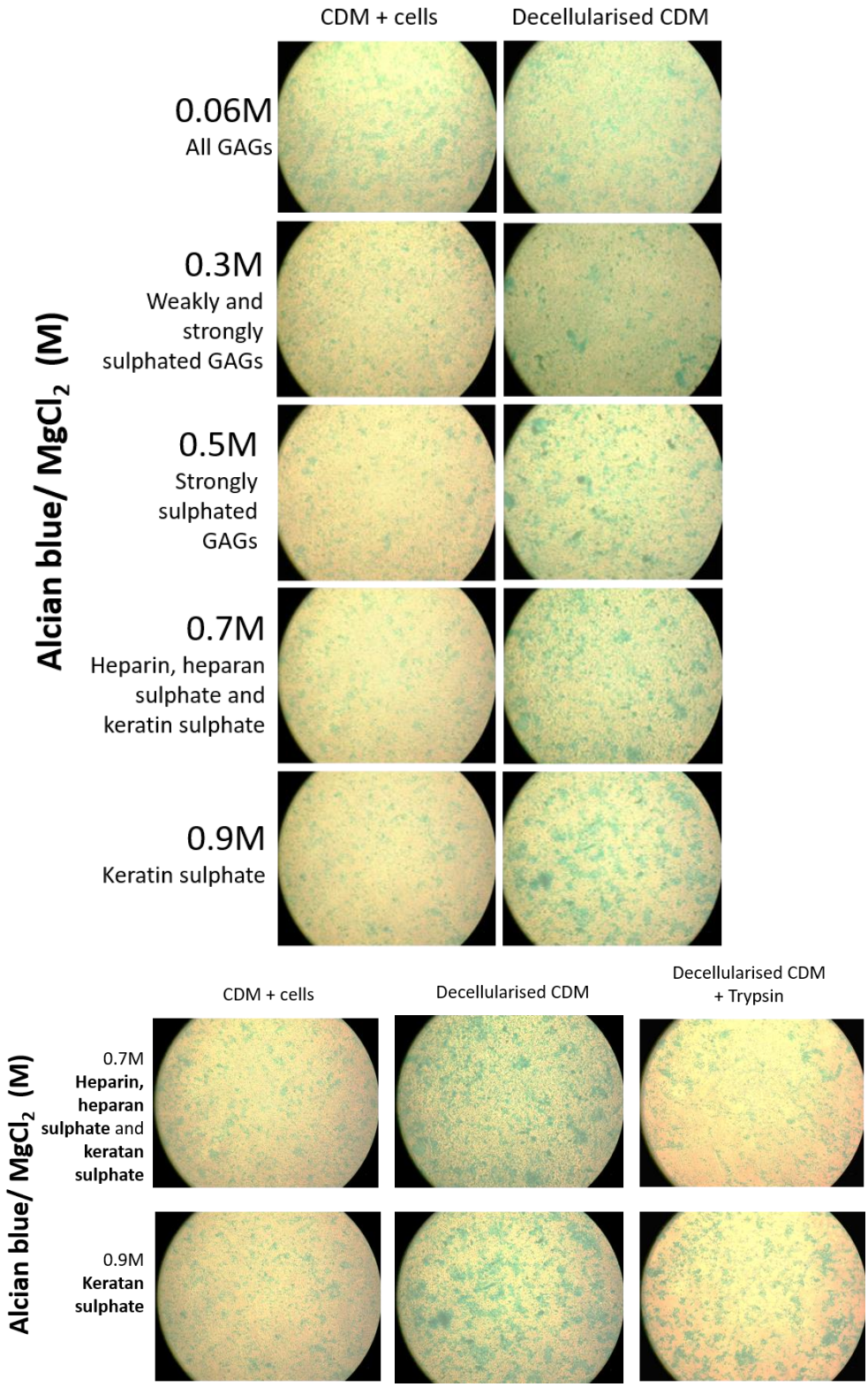


Figure 18. Alcian blue/$MgCl_2$ staining (M) of GAGs in BJ-CDM demonstrates GAGs are not removed during the decellularisation process.
A) BJ-CDM **(**left column**)** and decellularised BJ-CDM **(**right column**).** 4X phase contrast microscopy of ECM before or after decellularisation and alcian blue staining of GAGs. **B)** Stain specificity was tested by short trypsin treatment of stained CDMs **(**right column**)** to degrade the core proteoglycan, thereby detaching the stained GAGs from the matrices. Representative images from 3 experiments.

3.5 CHARACTERISATION OF THE THREE-DIMENSIONALITY OF FIBROBLAST-DERIVED CDMS

Current culture methods generally rely on 2-dimentional (2D) monolayer culture of cells on flat and rigid surfaces. Whilst this has proven to be an invaluable method for cell-based experiments, the limitations of this are increasingly being recognised. Mainly, that this culture method is not physiologically relevant to *in vivo* conditions, where the majority of cells are surrounded by and embedded within a 3D matrix composed of numerous ECM proteins and numerous other cells types. The most striking difference between 2D and 3D culture systems is the mechanical environment; stiff glass or tissue-culture plastic versus the soft and dynamic consistency of the ECM. Cells in 2D cultures are also limited to 2D interactions. All these factors regulate gene expression in cells. This problem has been illustrated by studies mainly focussing on cancers, in which 2D culture data has been non-predictive for subsequent *in vivo* experiments or 3D culture experiments [379, 468, 469]. Very few cancer drugs pass successfully to clinical use, approximately 10% [470] with the failure attributed to non-physiological relevant data from 2D culture models [35]. Whilst on a slight tangent, a particularly good example is in solid cancers, in which many have a hypoxic region that subsequently affects drug outcomes [471]. The development of such hypoxic regions is not possible to the same extent in a 2D model compared to a 3D model, hence leading to non- predictive data based on the lack of structural relevance to *in vivo* tumour structure [472]. The generation of 3D matrices would therefore move one step closer to recapitulate *in vivo* environmental conditions of ECM and generate results closer to those generated from *in vivo* models. The three dimensionality of the BJ-derived ECM was therefore assessed.

3.5.1 Measuring CDM thickness by confocal microscopy

Confocal microscopy is a powerful tool for recording 3D images with increasing resolution over conventional light microscopy. Z-stacking is a processing method which takes multiple images at different focal distances (slices) which can then be combined to form a composite 3D image with a greater depth of field than individual images (Figure 19). This is achievable as most biological samples have a thickness of several microns. In this way the thickness is semi-quantified by measuring immunostained ECM components throughout the sample from the first in-focus plane to the last. This

method of immunostaining (normally COL or FN) followed by microscopy is the favoured method used to assess fibroblast-CDM thickness [447, 473-477].

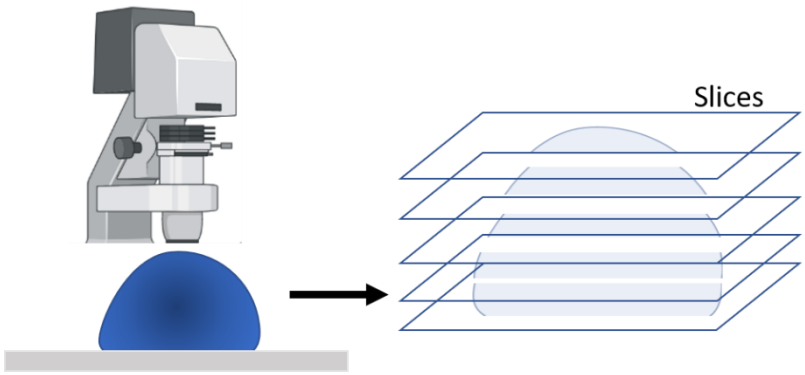


Figure 19. Schematic of Z-stack analysis.
Images of each 'slice' of samples are taken that can then be assembled into a composite image, allowing visualisation of 3D structures.

To assess the optimum thickness of BJ-derived CDM, 'mono' and 'double' layer BJ-derived CDMs were generated. The single and double nomenclature refers to how many layers of cells were seeded to generate the CDM; the 'single' layer was generated as per the protocol (Figure 13) in which BJ fibroblasts are grown to confluency and supplemented with L-ascorbic acid for 14 days, before decellularising. The 'double' layer CDM is generated by seeding fibroblasts on a pre-existing 14-day decellularised ECM. This second layer of cells was allowed to deposit matrix for a further 14 days, before decellularisation. ECM components were then stained for immunofluorescence analysis of thickness by confocal microscopy Z-stacks (imaging carried out with Hilary Collins). It was hypothesised that the second layer of fibroblasts would secrete ECM components onto the pre-existing CDM and generate a thicker matrix.

After immunostaining of collagen within the decellularised 'mono' BJ-derived ECM, the thickness of the CDM was estimated by taking thickness measurements calculated from confocal microscopy Z-stacks from different areas of the sample. Z-stack slices were taken at µm intervals. The 'mono' layer was 18.38 µm thick (mean ± 2.387, N =8; Figure 20). It was then tested whether the thickness of the CDM could be increased by seeding BJ cells on top of the first decellularised CDM and culturing for a further 14 days followed by decellularisation, immunostaining and confocal microscopy analysis. The

thickness of the 'double layer' CDM was 20.5µm (mean ± 1.852; n=8; Figure 20) and much less dense, with distinct holes in the matrix (Figure 20). Whilst this is surprising that the matrix was not thicker, the holes were most likely the result of the second layer of fibroblasts migrating through the first layer of matrix, as opposed to adhering on the top of the first CDM and laying down new matrix. Only a few cell types, such as neutrophils and lymphocytes can migrate through the matrix without altering the ECM architecture [478], whereas it is well known of fibroblasts remodelling capabilities.

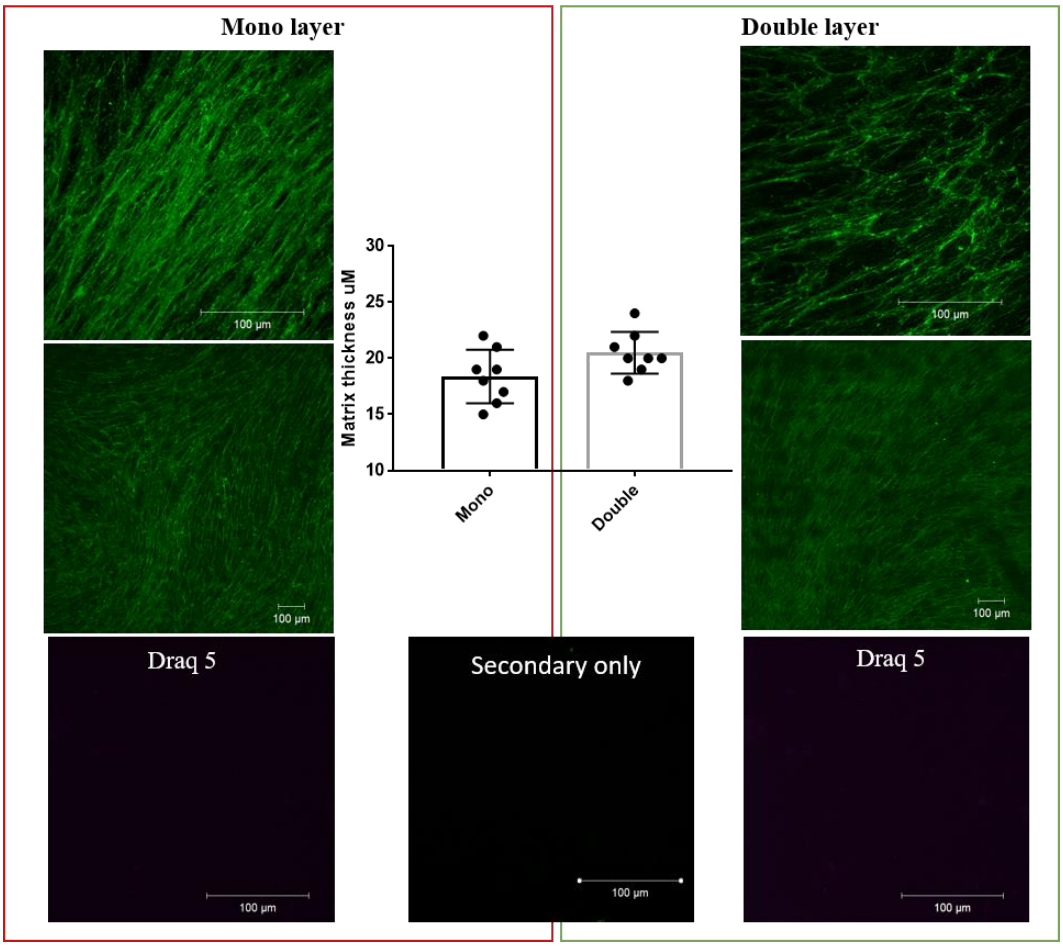


Figure 20. Thickness analysis by confocal microscopy Z-stacking demonstrated that a thicker CDM could not be generated by seeding fibroblasts on pre-existing CDM.
Immunostaining and confocal analysis of decellularised CDMs deposited for 14 days by one layer of cells (red box) or 28 days in total by two layers of cells (green box) in complete media with 50 µg/ml L-Ascorbic acid. The 28-day matrices required decellularisation after 14 days and reseeding of new cells on the first layer of matrix. Thickness was calculated and shown as a mean ± SD; N = 8. Scale bar 100 µm.

This migration was possible as confocal analysis of the single layer CDM stained for FN and COL1 showed that it is a three-dimensional substrate (Figure 21), not a one-dimensional coating. Furthermore,

the decellularisation was once again confirmed to be effective as staining of genomic DNA with Draq5 indicated no BJ fibroblast DNA visible during confocal analysis.

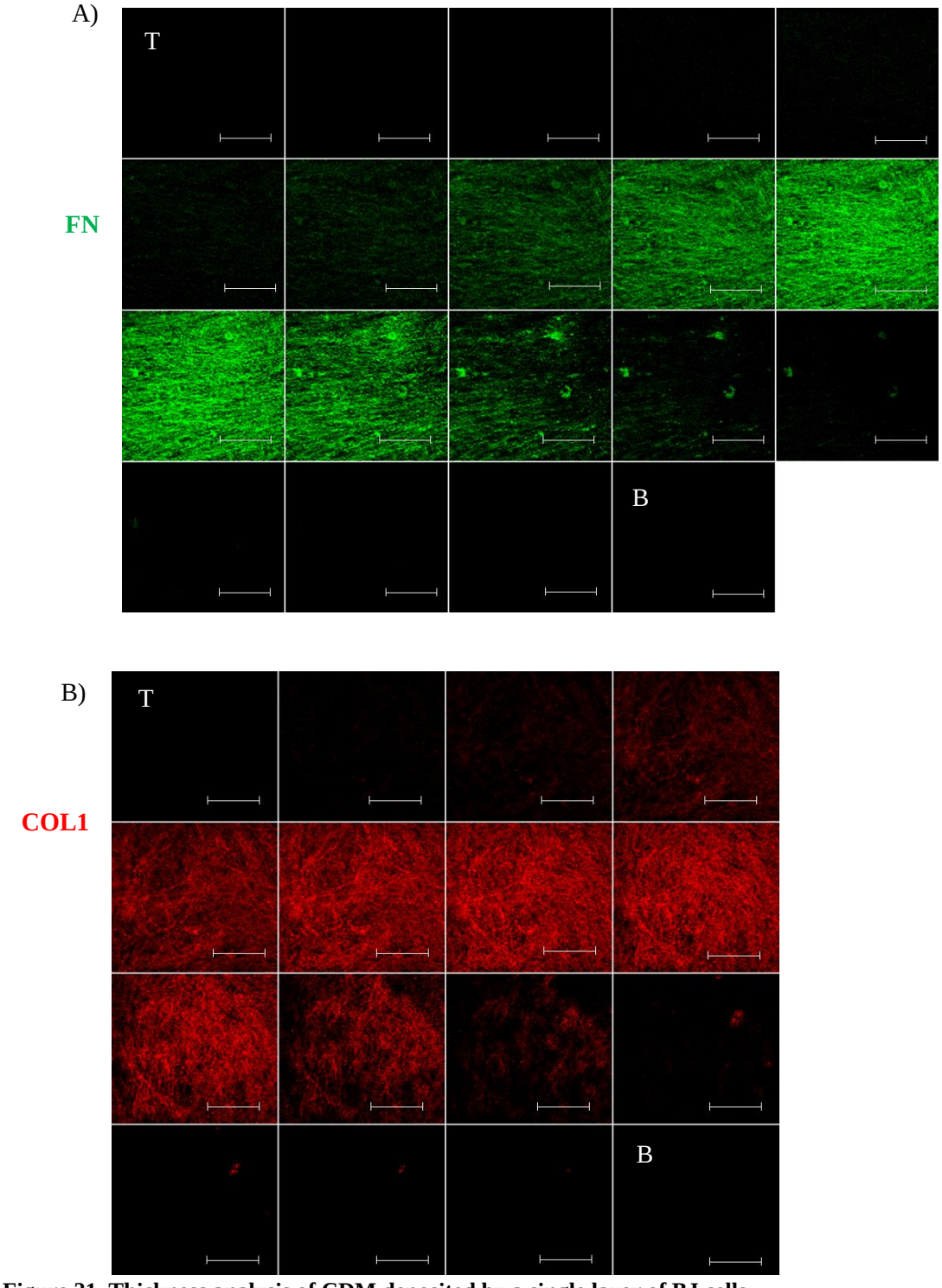
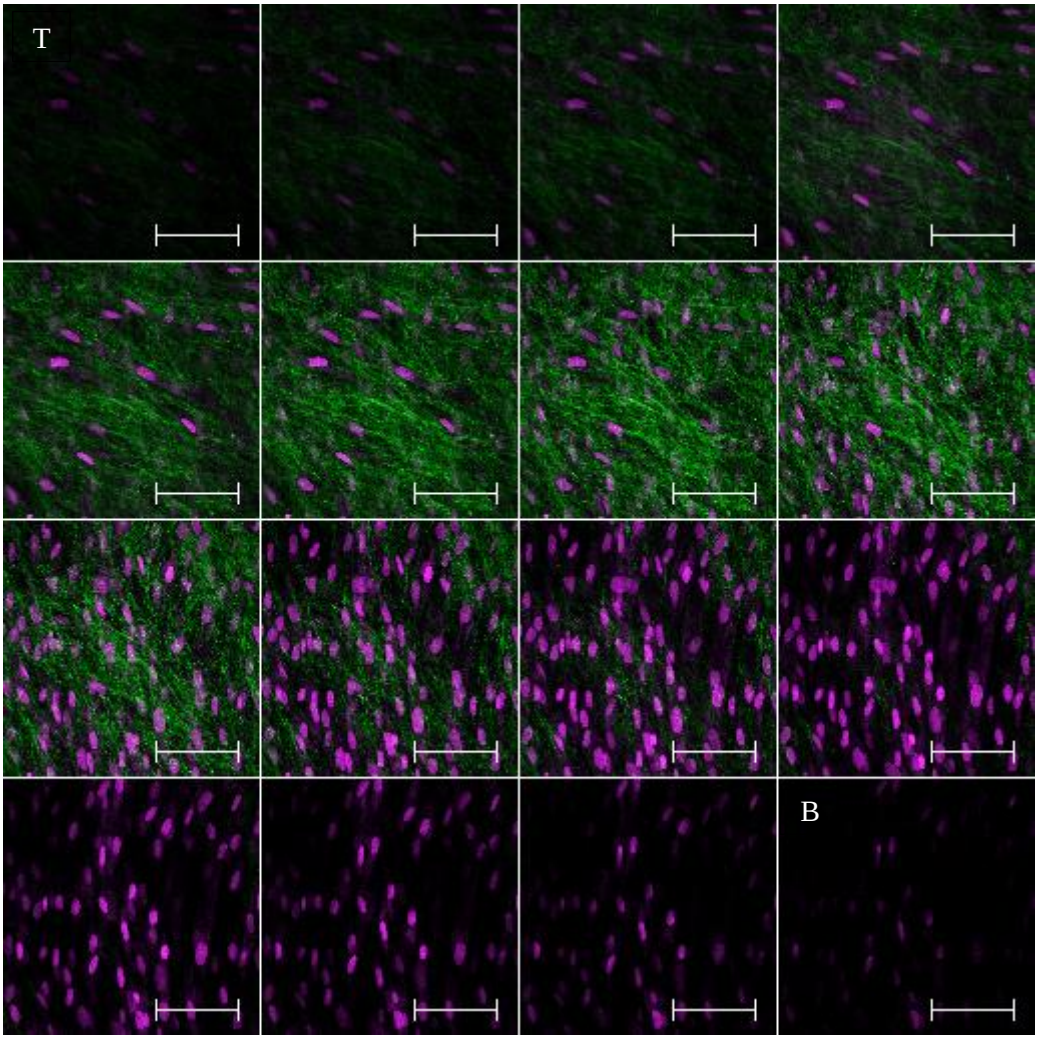


Figure 21. Thickness analysis of CDM deposited by a single layer of BJ cells.
Confocal Z-stack analysis of ECM proteins deposited in 14-day BJ-derived CDM following decellularisation and immunofluorescence staining for **(A)** FN (green) and **(B)** COL1 (red) Top (T) and bottom (B) of Z-stacks are indicated. One representative image per sample. Each Z-stack slice corresponds to 1 µm. Scale bar 100 µm.

As the ‘double’ layer matrix had been decellularised it was not possible to visualise cells and their presence throughout the matrix via Draq5 staining. To visually discern if the holes were indeed due to fibroblasts migrating through the matrix, as opposed to depositing new ECM material, the above process to generate ‘mono’ and ‘double’ layer matrices was repeated. After generation of the first layer of CDM by allowing fibroblasts to deposit CMD for 14 days (Figure 22 A), the first matrix layer was decellularised, fixed and stained with primary and secondary antibodies for detection of COL1 (green), followed by seeding of a new layer of cells. After a further 14-day deposition, the matrix with cells was fixed and stained for COL1 again (red) (Figure 22 B). It was clear to see cells present throughout the z-slices, in both the ‘single’ and ‘double’ layer CDMs, along with a network of COL1 and FN. As the z -stacks image from top to bottom it appears many cells stay oriented above the CDM, generating a layer of CDM beneath.

In the double layer matrix (Figure 22 B) from top to bottom, newly synthesised collagen can be seen (red), which on progression through z-stacks decreases and collagen synthesised by the first layer of fibroblasts is present (green). Some staining overlapping occurred. Whilst it may appear on first inspection that this was due to staining the same ECM molecule, this is unlikely; after staining the first layer, saturation of the COL1 (green) antibody sites should have occurred. It seems more likely that this mixed staining is due to the second layer of fibroblast cells migration and remodelling throughout the first layer along with some deposition of new COL1 (red). This would explain the simultaneous staining, representing new and old collagen localised together. Given that the mono and double layer matrices were similar in thickness, and that cells were present throughout both single and double 3D matrices and the mono takes 14 days to generate compared to 28 for the double layer, mono layer matrices were utilised for the rest of the project.

A)



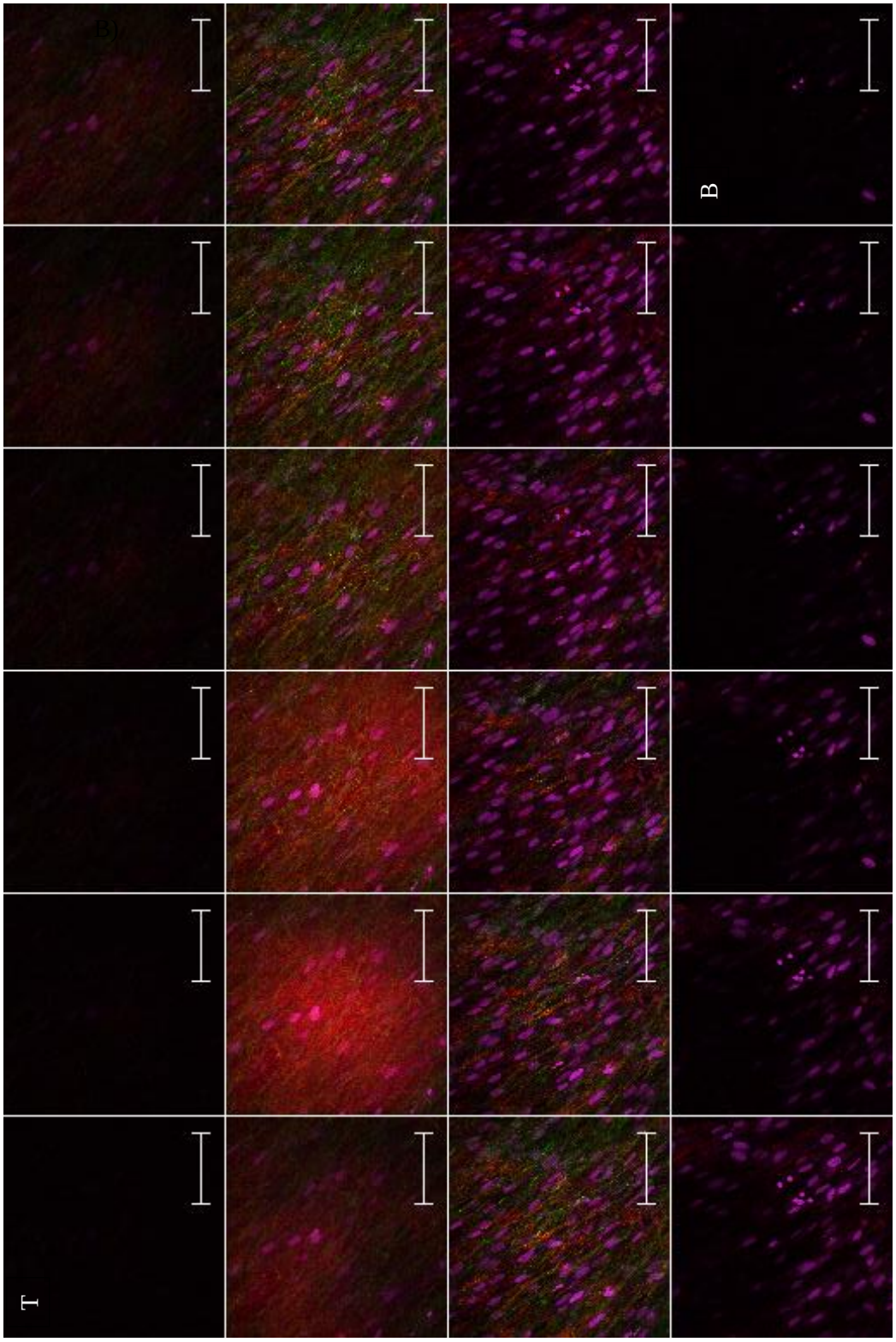


Figure 22. Staining of mono and double BJ-derived matrix layers to discern migrating fibroblasts.
Confocal microscope Z-stacks of 'single layer' **(A)** and 'double layer' **(B)** CDM. Single layer CDMs were generated by seeding fibroblasts in IBIDI chambers and growing to 100% confluence before fixing, staining COL1 (green) and cells (Draq5 purple). 'Double layer' CDM was generated by seeding a layer of BJ fibroblasts onto pre-existing 14-day CDM that had been decellularised and stained for COL1, for a further 14 days before fixation and staining of COL1 and cells (DAPI in blue). Each Z stack slice corresponds to 1 µm. Top (T) and bottom (B) of z-stacks are indicated. Scale bar 100 µm.

It is known that fixation of fibroblast CDMs can cause shrinkage of samples [479]. On utilisation of the BJ-derived CDM as a substrate for cells further down the line, fixation would not be required. The thickness measurements may therefore not be representative of the CDM utilised in experiments. Therefore, as further assessment of the 3D structure and the effect of fixation on the matrices, COL1 and FN were stained in BJ-derived CDM either without (unfixed) or following fixation with PFA (fixed). In fixed CDMs, an intact network of fibers can be seen, however both COL1 and FN staining is present in less slices, compared to unfixed CDM. As each slice represents 1 μm it is clear that there are less slices with stained ECM molecules in fixed COL1 and FN samples (Figure 23). It also appears that the intensity and density of the fixed samples appears markedly increased compared to the unfixed samples; this is particularly evident in comparing the first few images in each condition for both COL1 and FN. This may be indicative of a similar amount of ECM material present in a smaller, more condensed structure. This appears to demonstrate that PFA causes some shrinkage of samples. Whilst the unfixed measurement is likely closer to the 'native' thickness of the CDM, this measurement may still be inaccurate given that confocal imaging only measures the thickness by distribution of one molecule within the ECM (FN or COL1). A different method which could measure the distribution of all the molecules within the ECM, including GAGs, was therefore required.

However, taken together, the above results demonstrate the generation of a consistent, reproducible, 3D-ECM, rich in GAGs, COL1, COL3, FN and TN-C. An efficient decellularisation process to remove cells without damaging or removing vital components from the ECM has also been validated.

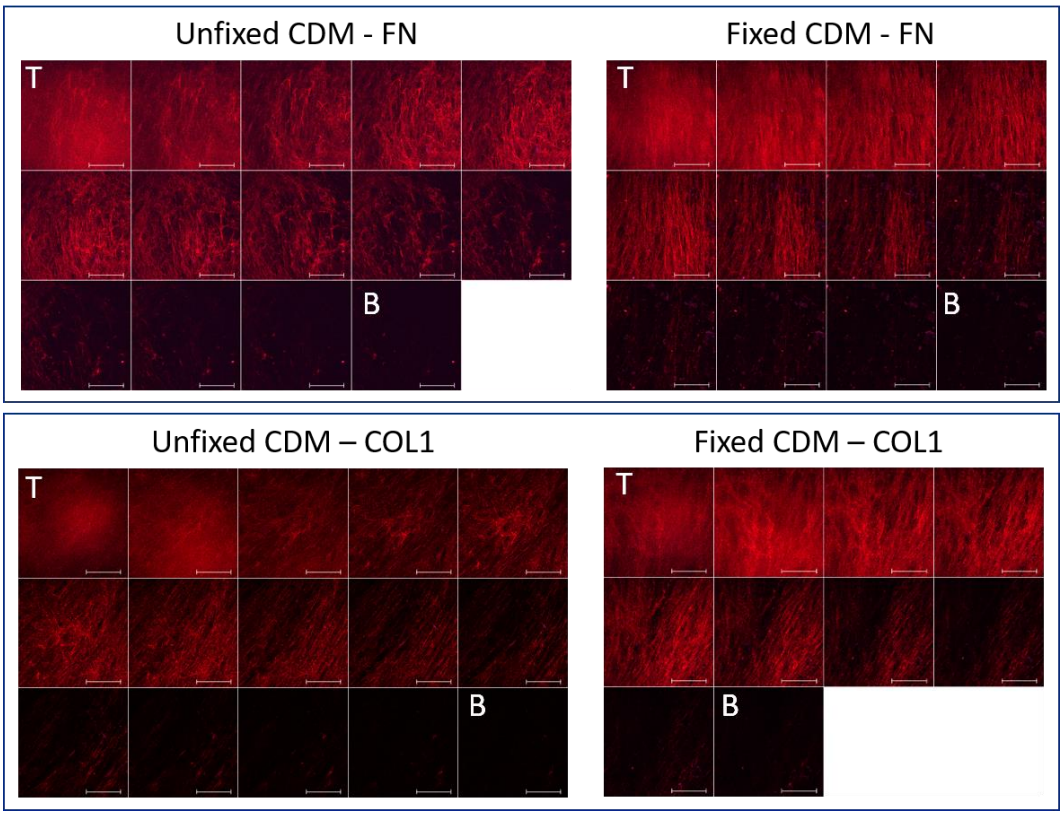


Figure 23. PFA causes shrinking of decellularised BJ-CDM.
BJ-derived CDM was decellularised and stained for COL1 and FN following either fixation by paraformaldehyde (PFA) or without fixation and immediately imaged by confocal microscopy. Each image represents one Z-stack slice at 1 µm. Top (T) and bottom (B) of Z-stacks indicated. Scale bar 100 µm.

3.5.2 Measuring CDM thickness by optical profiling

Optical profilers utilise wavelengths of light as rulers (Figure 24). They operate by comparing the optical difference between a test surface and a reference surface. Specifically, a beam of light from a single source is split, directed and reflected from a sample and reference point, and differences between these beams are calculated as the height of the sample. These differences are known as constructive and destructive interferences, depending on whether the distance is increased or decreased, compared to the reference, and result in light and dark banded interference fringes, respectively. This method can also be used to visualise surface roughness without damaging the sample as this is a non-contact method. Optical profiling has a nm range, thus resulting in more precise measurements of the matrix than the confocal microscope, which measures in 1 µm slices.

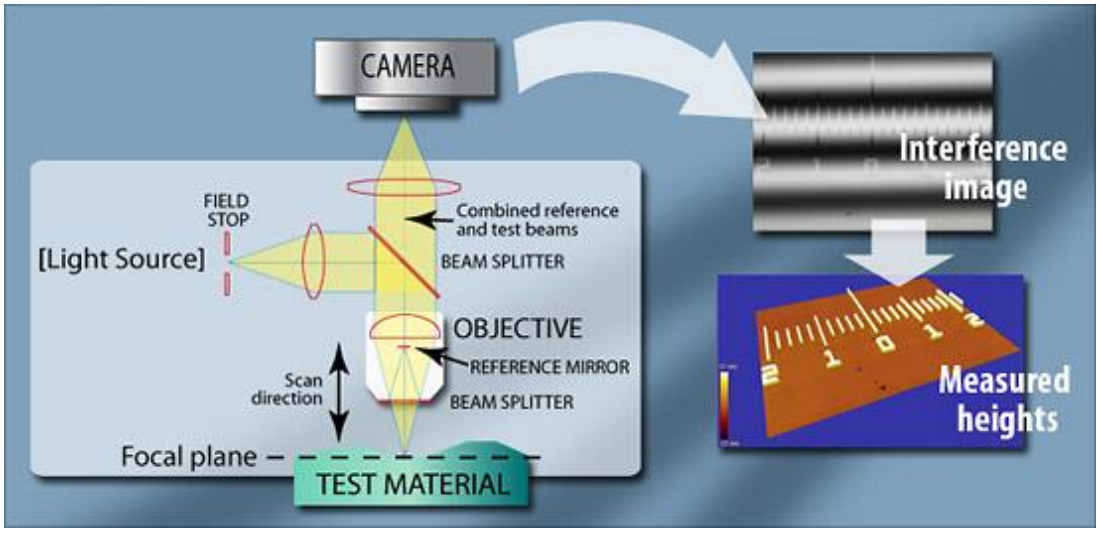


Figure 24. Schematic of how optical profiling works using the wavelength of light as a ruler.
A wavelength of light is split, directed and reflected from a sample and a reference point (known reference wavelength). This results in wavelength differences when the sample height differs from the reference wavelength. This creates interference fringes; dark and light bands depending on whether the wavelength difference is increased (constructive) or decreased (destructive) from the known reference wavelength. This then allows the conversion of the bands into a measurement of height differences.
From https://www.zygo.com/?/met/profilers/opticalprofilersabout.htm.

As optical profiling relies on contrast to correctly manually focus the sample, unfixed, decellularised BJ-derived CDM was assessed after Ponceau S staining. Ponceau S staining stains all proteins within the sample, thus allowing thickness measurements based on the total protein content of the CDM, as opposed to a specific molecule as was the case with confocal. Thickness measurements were taken over a 1740µm cross section (Figure 25 D) in each of the 5 sample areas and were varied but relatively similar between plates (Figure 25, A, B, C). Optical profiling revealed an average CDM thickness of 45.9 µm ± 6.48 (mean ± SEM) (Figure 25 D). The variations in this method are expected due to cells naturally aligning differently within the plate and each cell secreting ECM components with naturally small deviations. In plate 1 it appears that samples 1-3 are very consistent with measurements, in comparison to sample area 5 and 6. As sample area 5 corresponds to the centre of the plate it could be that more deposition naturally occurred here. However, this does not correspond to the centre area sample points in plate 2 and 3 in which the thickest matrix is detected in sample area 4. Sample areas 1, 2, 3 and 4 all occur on the same region of the periphery of the plate. Sample areas with no thickness measurements of 0 may also indicate matrix lifting from the plate. However, if this was true then almost the entirety of plate 3 matrix would have been detached and floating in the plate. This would have been

visible by eye, especially as the matrices were visibly red from the Ponceau S staining, and this was not the case. Some different shades can be seen- this resulted from removal of PBS . Some shadows can also be seen. The optical profiler is set up as a top-down microscope and light from the surrounding environment casts shadows onto the plate. This method also allowed surface topography visualization, indicating a rough surface (Figure 25 B). This rough surface seen more clearly in the magnified topography image is expected, due to the natural meshwork deposition of the ECM which is emphasised after the removal of cells. The area may have appeared smoother if cells were still present. The roughness could result from 'pockets' in which the cells were embedded throughout the matrix. Optical profiling allowed a thickness measurement established using the entirety of the deposited ECM proteins.

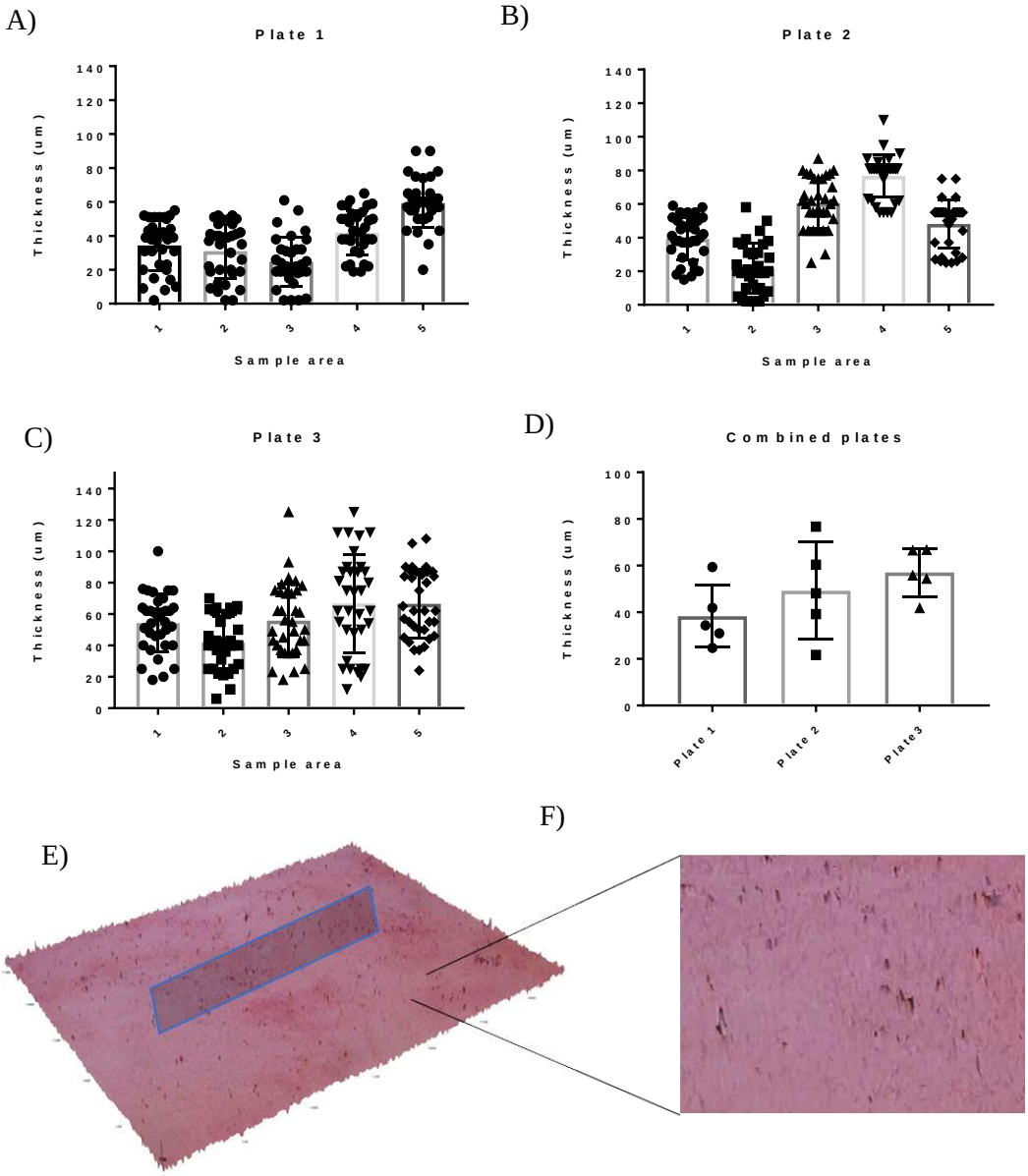


Figure 25. Optical profiling of the thickness (µm) of 14-day BJ-CDMs grown in 60mm dishes, after decellularisation and Ponceau S staining.
Data shown as mean ± SD in **A-C**. **D)** Combined plates data represented as mean ± SEM; 1-way ANOVA, p= 0.246, N =3). **E)** Surface topography of one representative sample area of 14-day BJ CDM obtained by optical profiling. The sample region 1740µm in length (indicated by the blue outlined transparent slice) indicates where sample points were taken along for each area.

3.5.3 Fibroblast-derived matrices fibers show local and global alignment

OrientationJ, an ImageJ plugin, was utilised to determine the alignment of individual matrix components, COL1 and FN. This software calculates the orientation of each pixel within an image and assigns a hue for each angle of local fiber orientation, which ranges from -90 ° to + 90 ° relative to

horizontal. In this way, aligned fibers are displayed in the same colour. This software has previously been used to assess collagen orientation [480, 481]. Following this angle calculation and hue assignment, a distribution histogram can be generated, allowing differently orientated fiber clusters to be identified.

OrientationJ analysis of three representative confocal images of immunostained COL1 14-day fibroblasts-CDMs revealed small regions of local alignment within matrices, but not global alignment (Figure 26 A-C). I have focussed on COL1 mainly as this is the main structural component of the ECM. It is important to note that we were not assessing whether the COL1 or FN in each matrix was aligned in the same orientation, as we did not try to control fibroblast alignment and therefore matrix alignment, hence I will not go into too much detail in describing the specific orientation angles but rather the general orientation of the fibers within the images on the whole. Further analysis at higher magnification for two COL1 images indeed demonstrated strong local alignment (Figure 26 D-F).

Next, the alignment of FN fibers was assessed, in the same manner as above (Figure 27). FN appeared to have a stronger global alignment than collagen, likely arising from this molecules role as a provisional matrix. Unfortunately due to Covid it was not possible to take further images.

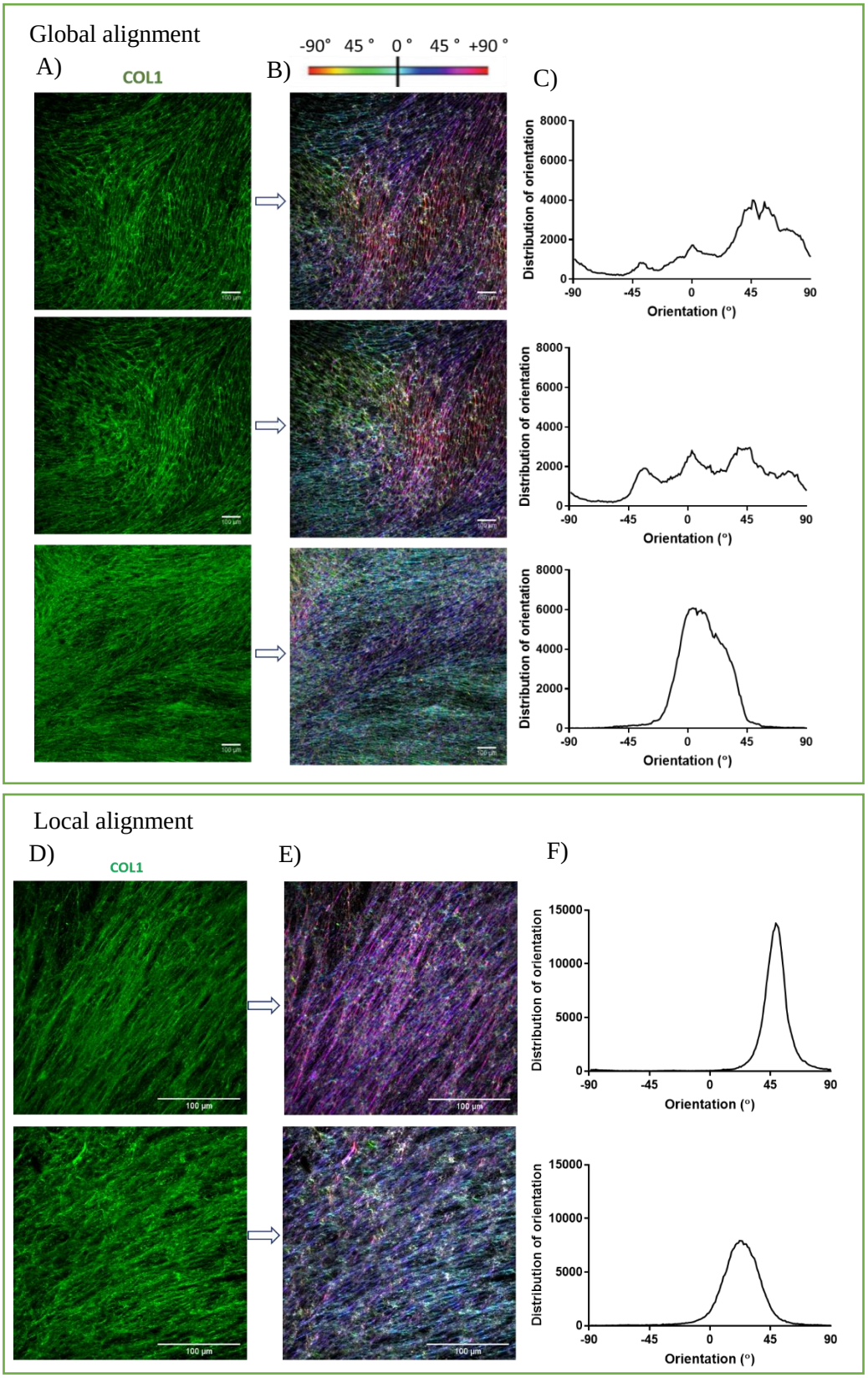


Figure 26. COL1 fibers do not display a global alignment but rather smaller regions of local alignment. Confocal microscopy images of immunostained COL1 fibers in 14-day BJ-CDMs **(A)** and processing using OrientationJ after hue assignment for each pixel and the determined angle of fiber orientation **(B)**. Quantification of fiber orientations from each image **(C).** Increased magnification confocal images of immunostained COL1 fibers in 14-day BJ-CDMs **(D)** and after hue assignment **(E)** and quantification **(F)** using Orientation J. Scale bar 100 um.

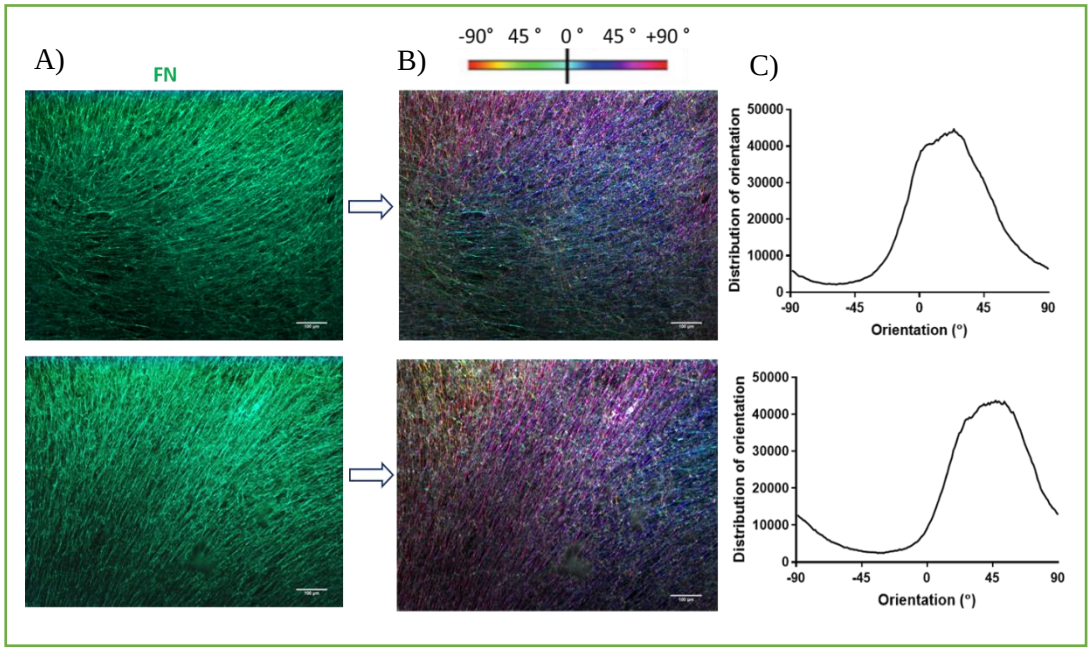


Figure 27. FN fibers display a generally global alignment.
Microscopy images of immunostained FN fibers in 14-day BJ fibroblast decellularised matrices **(A)** and processing using OrientationJ after hue assignment for each pixel and the determined angle of fiber orientation **(B).** Quantification of fiber orientations from each image **(C)**. Scalae bar 100µm.

3.6 SUMMARY

Due to the lack of physiologically relevant currently available ECM models to study gene regulation, we have generated an efficient and reproducible protocol for the creation of a BJ-derived ECM (CDM) that can be utilised as a human ECM model. We have characterised the CDM at the protein level, confirming ECM molecules deposition, and, more importantly assembly into a 3D ECM that is robust enough to allow handling. We have optimised a protocol for efficient decellularisation of the CDM without perturbing the dense meshwork of deposited structural ECM proteins or GAGs.

3.7 DISCUSSION

Gene expression is a highly organised and finely tuned process that is finely controlled to orchestrate a variety of complex processes allowing organism homeostasis. This is especially true for the generation and maintenance of the ECM. The ECM itself is also crucial in the generation and propagation of these complex signalling pathways. However, its role is unclear, which is partially due to a lack of physiological relevant models. To overcome the limitations of pre-existing synthetic models, which

usually rely on tumorous origins/animals sources [389], human BJ fibroblasts were utilised to generate a healthy, naturally-derived human ECM model, applicable for the culture of inflammatory cells and subsequent study of inflammation.

Human fibroblasts are a popular cell type for CDM generation due to their inherent ability to secrete and assemble large amounts of ECM molecules and their differing origins around the body which can be used to generate CDMs from many tissues [389, 456, 479]. Characterisation of 14-day decellularised BJ CDMs at the protein level revealed a CDM rich in collagen 1, collagen 3, fibronectin. Tenascin-C was also present, at a much lesser extent, consistent with the knowledge that this protein is conspicuously absent in adult tissues, with some exceptions, until disease or injury occurs [183]. This is also in line with the native composition of skin tissue.

The method to generate CDM as detailed by *Kaukonen et al.* [447] has been utilised by many studies. We have optimised many aspects, including the decellularisation process, to generate a robust method to facilitate the generation of reproducible BJ-derived CDMs, free from BJ fibroblasts and their DNA that would elicit an immune response. Supplementation of media with ascorbic acid, a cofactor for LOX and therefore crucial in collagen cross-linking, has previously shown to results in the generation of CDM rings (mimicking aorta) with a much higher tensile strength than CDM rings deposited in unsupplemented media [482]. Furthermore, *in vivo*, aligned fibrous matrices result in a stiffer and stronger matrix, when compared with an unorganised matrix [483]. In organised and stronger CDM rings mimicking aorta, cells were found more aligned with matrix fibers and evenly distributed throughout the CDM compared to fibrin and collagen rings, which displayed stratified cell layers [482]. In line with this, BJ fibroblasts displayed a highly aligned cell layer that resulted in a highly organised, dense BJ-derived CDMs that were stable enough for downstream applications. This highlights a role for cell-mediated ECM alignment and indicates that this alignment is a crucial aspect for structural stability and mechanical strength.

It is known that cultured fibroblasts show a preference for a parallel arrangement and our study demonstrates that fibroblasts shown an inherent preference for alignment with other fibroblasts as well

as collagen and fibronectin within their deposited CDM [446]. One proposal is that steric interactions result in nematic fibroblast alignment, but this does not fully explain the lack of global collagen orientation [446, 484]. Although neither of these methods were used, it has been demonstrated that most connective tissue cells align in response to stretching or by providing topological cues utilising a grooved or patterned scaffold [485] [486, 487]. *In vitro*, experiments have demonstrated that monocytes, mesenchymal stem cells, muscle cells, and fibroblasts orient in parallel to uniaxial stretch via contact guidance [449, 488-490] and demonstrates that mechanical signals play a significant role in tissue morphogenesis [491]. Collagen fibers in collagen gels also align in response to uniaxial stretching [492]. Furthermore, fibroblasts under strain on (strained) collagen gels, the simplest tissue-equivalents, both adjust their orientation to collagen fibers that are in alignment with the strain and further align collagen fibers in the same direction [449, 493]. This also provides a method by which cells not directly in contact with other cells can still communicate, via collagen fiber tension [493]. These studies suggest that uniaxial stretching of the entire CDM may be required to facilitate a global alignment. Although it has been demonstrated that fibroblasts on strained collagen gels can align in response to stretch without concurrent collagen fiber alignment [494], the differences here are likely due to the stiffness of the substrate – this was a 'soft' gel and cells are more resistant to forces on elastic gels due to the compliancy of the gel resulting in less tension exerted over a smaller area that other cells may not be in contact with. Together, these experiments suggest that cells can align independently of matrix fibers, but promote matrix alignment and align in response to an appropriate mechanical signal.

In a similar vein, cell contractility and motility, facilitated by culture of cells on stiff tissue culture plastic/glass coverslips could result in local regions of strain [495]. This would explain the regions of locally aligned collagen resulting from clusters of aligned cells, given that there was no singular strain to orient the cells and CDM as a whole.

In contrast, fibronectin fibers appeared largely orientated. Unfortunately, the CDMs were not stained for both collagen and fibronectin simultaneously so only speculations can be made as to why these differences have occurred. The previous experiments, whilst not directly assessing fibronectin tension on gels, can likely be applied to fibronectin given that this is also a fibrous, structural protein.

Fibronectin is present in developing tissues prior to collagen deposition; fibronectin acts as a scaffold for collagen deposition. It has been suggest that fibronectin may orient cells in alignment with collagen fibrils and therefore facilitates the generation of a CDM with parallel collagen fibers [496]. This may explain the global alignment; as fibronectin is present before many other ECM components in tissue in the developing organism, fibronectin would likely be able to self-regulate its assembly in a highly-organised manner which is necessary to provide a successful base for the deposition of other ECM components.

Inter- and intra-matrix differences in alignment may also be explained by the experimental set up. Matrix deposition was carried out in IBIDI chambers for collagen, but on coverslips in 6 well-plates for the matrices stained for fibronectin. The IBIDI chambers are square whilst the 6-well plates are round. Whilst no specific and intentional stretching forces were applied to the matrices, culture in these dishes would result in tensions applied to the deposited matrices and cells within. The fibroblasts seeded on coverslips for deposition are also subject to an additional strain – not only where the matrix adjoins the side of the well, but also where the coverslip and matrix connects to the base of the well. The strain could be speculated to be equal and therefore stretching the matrices in every angle from 0 to 360 degrees. However, this would potentially result in less aligned fibronectin, assuming the highly–ordered self-regulatory mechanism is not true. In the CDM deposited in the square IBIDI chamber, these forces are likely to be less equal and result in biaxial strain perpendicular to the IBIDI well edges which should result in a more globally aligned collagen network, unless the two axis of strain result in cells attempting to align with multiple axis of strain. Furthermore, coverslips can be rotated at any angle from 0-360 degrees when mounting and imaging, and the removal of coverslips from wells can cause stretching and damage to the deposited CDM particularly around the edges of the coverslip. The former factor specifically does not allow the direct comparison of the orientations of matrices seeded in different wells; however, this is less of an issue in IBIDI chambers in which matrices are not disturbed by rotation during immunostaining and imaging. Further experiments, controlling the orientation of coverslips and specifically assessing tension and co-staining for both fibronectin and collagen are required to ascertain

the role of experimental set up in determining alignment of cells and the deposited ECM molecules, although it is clear that cells are intrinsically able to orient themselves and their deposited ECM proteins.

The thickness of the BJ-derived matrices was assessed by two methods, which displayed different results, explainable by the fixation method utilised for immunostaining analysis, however both demonstrated a substantially thick matrix that was visible by eye. Whilst many studies utilise CDMs as substrates, particularly those from fibroblasts, very few studies assess this characteristic. Matrices generated from primary human foreskin fibroblasts over an 8-day period with L-ascorbic acid and WI-38 fibroblasts for an unspecified period 'until confluency' (and measured by fibronectin immunostaining) have been demonstrated as 9 μm thick, and 8 μm thick, although this had shrank from 10 μm before decellularisation, respectively [497]. We also saw shrinkage after fixation for staining and it is likely that these unfixed matrices would be thicker if assessed by optical profiling which can assess the thickness by measurement of all proteins, as opposed to one component, such as fibronectin. These matrices were also deposited over a shorter 8-day time period; visual inspection of 8 -day matrices from BJ fibroblasts generally looked poor. In an attempt to increase the thickness of BJ-derived CDMs, 'double layer matrices' were generated. This was unsuccessful as it appeared fibroblasts deposited on an already present BJ-derived CDM migrated and remodelled more so than depositing new matrix. Generally, cells are embedded in a degradable scaffold to increase thickness [389]. The larger thickness measurement acquired by optical profiling results from measuring all the proteins (Ponceau S stained) within the CDM and as the CDMs were not fixed, no shrinkage resulting from PFA fixation occurred. It may also be that, as the CDMs were unfixed and hydrated, longer proteins (i.e. PGs with GAG chains extending outwards from the ECM surface) could have resulted in the thicker measurement. However, given that the ECM would be unfixed and hydrated *in vivo* this likely reflects a more physiological thickness of the CDM.

Another group was able to generate an average 395 μm thick CDM from dermal fibroblasts by growing in serum-free, high glucose DMEM supplemented with many other factors including insulin and selenious acid (lipid precursor), ethanolamine, L-ascorbic acid phosphate magnesium salt *n*- hydrate, EGF in human serum albumin, basic FGF, dexamethasone, glutamax and glutathione [479].

Furthermore, the CDM was assembled in highly organised and distinguishable layers. EGF, together with ascorbate promotes growth and protein synthesis by fibroblasts [498]; basic FGF regulates proliferation of ECM production by fibroblasts [499]; in serum-free media dexamethasone significantly increases fibroblast proliferation by upregulating PDGF-*a* receptor and enhancing fibronectin production that acts as a scaffold for the organisation of other ECM components [500-502], although newer studies have demonstrated that dexamethasone inhibits fibroblast proliferation whilst increasing fibronectin and pro-collagen synthesis [503] and finally insulin upregulates TGF-β1 synthesis and subsequent ECM protein synthesis amongst many other actions [504]. Ahlors *et al.* [479] suggested serum may contain inhibitory factors given that fibroblasts grown in the same conditions with serum were thinner and had a lower ultimate tensile strain. The addition of these components to the protocol for CDM generation optimised in this work may aid in the generation of a thicker CDM.

The production of CDMs by human cells has the potential for many useful applications. The mechanical robustness and defined composition of the ECM means it could be possible for use as highly stable and biocompatible tissue graph or scaffold to aid in healing of injuries or wounds [479]. The ultimate tensile strain of dermal fibroblast-derived CDM has been found to be higher than collagen gels and fibrin gels, but slightly less than skin [479]. Interestingly, in this study, although the collagen gels had the greatest collagen density and fibril diameters, they were significantly weaker than CDMs grown under the same conditions, indicating that not only the composition but also the assembly gives the ECM its mechanical structure.

Whilst the stiffness of the fibroblast CDMs was not assessed, this is an important mechanical feature of ECM that functions to modulate gene expression and ultimately dictate development, tissue maintenance and disease progression [505]. For example, matrix stiffness has been shown to modulate cell functions including morphogenesis, differentiation, migration and proliferation [506-509]. Cancer cells preferentially metastasize along stiff collagen fibers residing in stiffer ECM than normal breast cancer tissue [510]. Whilst the elastic modulus (stiffness) of synthetic substrates has been assessed, very little is known in comparison about the stiffness of naturally produced CDM [511]. The stiffness of CDMs produced from dermal fibroblasts, WI-38 lung fibroblasts and primary fibroblasts have been

demonstrated as similar to the *in vivo* range, even with production on tissue culture plastic that has a stiffness of gigapascals [473, 479, 512, 513]. Elastic modulus ranges from several 100 pascals in the brain to several gigapascals in cortical bone [514]. However, it could be that cells seeded on lab plasticware interpret the plastic substrate as a bony substrate and hence produce CDM with stiffness that would usually be in contact with bone. Indeed, if we take the aforementioned examples of brain tissue and bone, a CDM of approximately 100 pascals produced from primary and lung fibroblasts [473, 513] and 300 pascals from dermal fibroblasts [479] produced on tissue culture plastic possibly mirrors the stiffness of brain tissue in contact with skull bone. In addition, cancer-associated fibroblasts, isolated from tumours, which usually produce ECM with increased stiffness compared to normal ECM, produced a CDM stiffer than those from a normal fibroblast [513]. This not only validates CDMs an *in vitro* model for health and disease states but reinforces the role of stiffness in general ECM development and homeostasis.

Whilst it was planned to carry out atomic force microscopy (AFM), the gold standard for measuring stiffness, time constraints and the complexity of the fibroblast CDM sample did not allow for this. However, future work should assess this characteristic of BJ-derived CDM.

4 ENGINEERING *TNC* KNOCKOUT FIBROBLASTS USING CRISPR/CAS9 GENE-EDITING TO PRODUCE CDM DEVOID OF TN-C

4.1 INTRODUCTION

After the generation and validation of a reproducible BJ fibroblast-derived human extracellular matrix (ECM) model in chapter 1, the next objective was to generate BJ fibroblast cell-derived matrices (CDMs) lacking TN-C (KO CDMs) and utilise cell-free CDM as a substrate for the culture of relevant cells (Figure 28). Specifically, this would allow the investigation of the impact of ECM-associated TN-C on primary human macrophages during the LPS-induced inflammatory response. Macrophages are 'first responders' in the inflammatory response. They characteristically monitor the extracellular environment, constantly sampling and responding to encountered pathogens and their components such as lipopolysaccharide (LPS) found in the outer membrane of Gram-negative bacteria. Upon activation of macrophages by LPS through toll-like-receptor 4 (TLR4), drastic changes in gene expression occur to induce and propagate a proper inflammatory response that results in the clearance of the pathogen and repair of any damage to the tissue [291, 515]. Macrophages can therefore be seeded on TN-C containing CDMs (WT CDMs) or TN-C $^{-}$ (KO CDMs) matrices and stimulated with LPS, and the resultant TN-C-dependent differences in gene expression assessed. Targeted gene editing using CRISPR/Cas9 was utilised to knockout the *TNC* gene in human BJ fibroblasts, with three approaches designed.

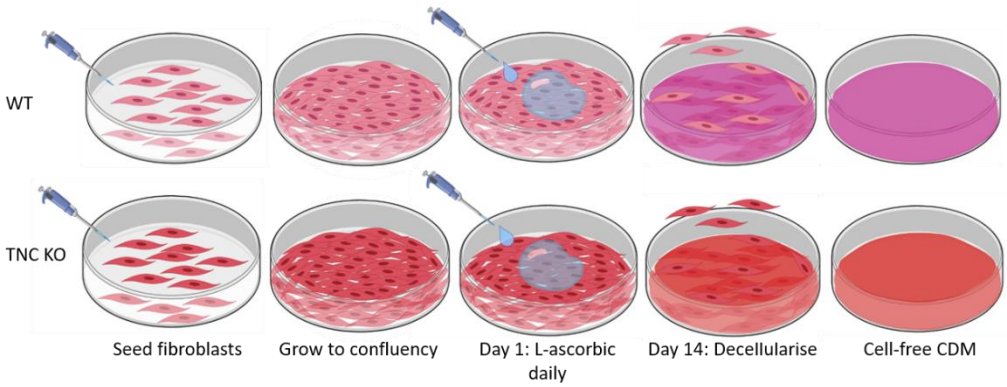


Figure 28. Schematic of the generation of WT CDMs (top) or *TNC* KO CDMs (bottom) following CRISPR/Cas9 gene editing.
BJ fibroblasts are seeded to confluence before culture for a further 14-days in the presence of L-ascorbic acid. Cells are removed by decellularisation leaving cell-free matrix.

CRISPR/Cas9 can be targeted to a specific genomic locus by designing 20bp guides (guide RNAs or gRNAs) complementary to the target region sequence. Cas9 endonuclease activity is reliant on a 5' Protospacer Adjacent Motif (PAM) sequence directly adjacent to a gRNA [424]. After a double strand break (DSB) has been induced, the cell will repair the damage by one of several pathways. Large insertions or deletions (INDELS) occur when the error-prone repair pathway non-homologous end joining (NHEJ) is utilised to fix the DBSs, in which ends are ligated together with no repair template. By targeting the DSB to the ATG translation start codon, or a gene region vital for function, INDELS can lead to frameshifts and a premature stop codon which should result in premature termination of translation of the gene.

4.2 PLASMID-BASED CRISPR APPROACHES 1 AND 2

We utilised two different plasmid-based strategies to target the *TNC* gene in BJ fibroblasts. Approach 1 (CRISPR/Cas9nickase) and 2 (CRISPR/Cas9) utilise an expression plasmid through which nucleofected cells can express the CRISPR/Cas9 or CRISPR/Cas9n complex (Figure 29). The CRISPR/Cas9n system, in which one of the two endonuclease activities is not functional, has the advantage of reduced off-target effects. This results from the paired guide format which requires two guides and their associated Cas9 binding at the correct locus to induce a DSB. Any binding at an off-target region will only induce a single strand break (SSB) only. SSBs are usually repaired by the high-

fidelity base-excision repair pathway (BER). Approach 2 uses CRISPR/Cas9 (full nuclease), which is more efficient than the nickase system, requiring only a singl guide to target Cas9 to induce a DSB. This approach targeted exon 2 and 7 simultaneously.

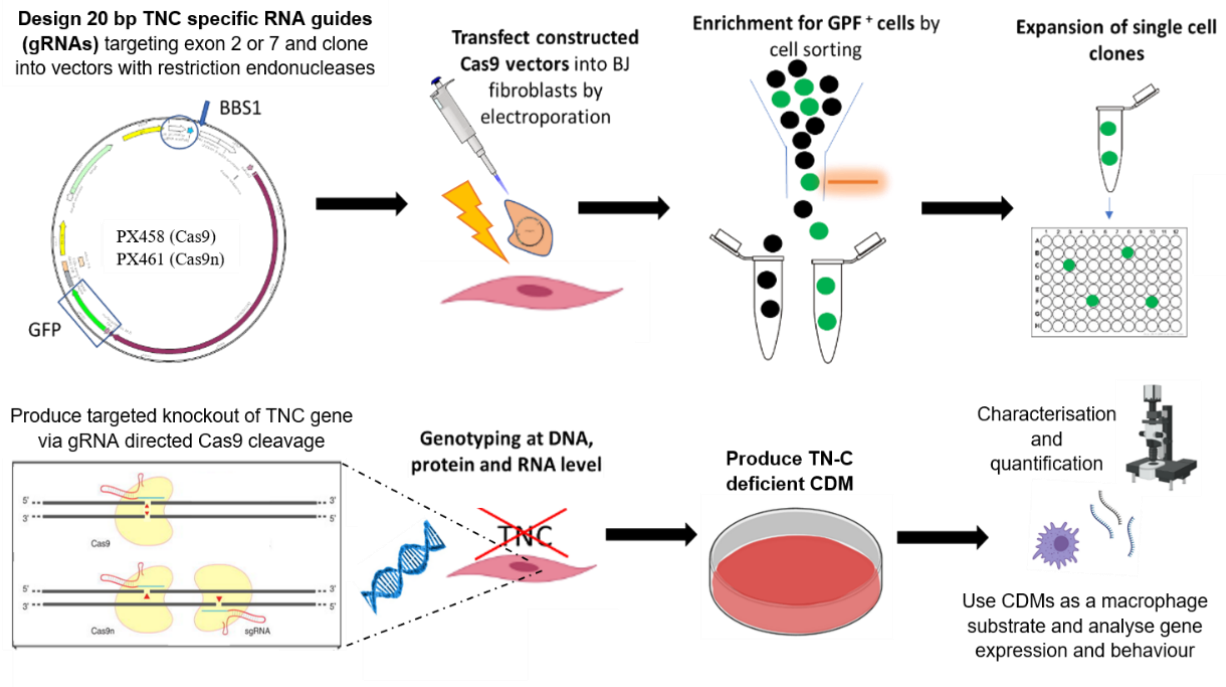


Figure 29. Schematic overview of the generation of *TNC* knockout fibroblasts and matrices using CRISPR/Cas9 and CRISPR/Cas9n plasmid approaches 1 and 2.
First, guides are designed targeting the *TNC* gene at exon 2 or exon 7 and cloned into mammalian expression vectors PX458 or PX461 containing a GFP reporter gene. Fibroblasts are transfected with the CRISPR vectors, selected by GFP expression into single cells and expanded into clonal populations. Cell-derived matrices (CDMs) are created from these *TNC* knockout fibroblasts and used as a substrate for the culture of cells of interest.

4.3 RIBONUCLEOPROTEIN-CRISPR APPROACH 3

A third approach using CRISPR/Cas9 ribonucleoprotein (Cas9 RNP) (Figure 30) was targeted to exon 2 or 7 or both exons simultaneously. In this, the ribonucleoprotein complex is produced in vitro and then transfected into the cells where it is immediately active. The latter results in this method being more effective over both plasmid approaches. Cas9 protein has a half-life significantly shorter than a Cas9 expressing plasmid. Previous research in the lab has shown that Cas9 expression (inferred by GFP expression) can be seen up to 9 days after transfection. This method therefore reduces the possibility of

off-target effects due to the more transient presence of Cas9, however it does not completely abrogate the chance of off-target effects.

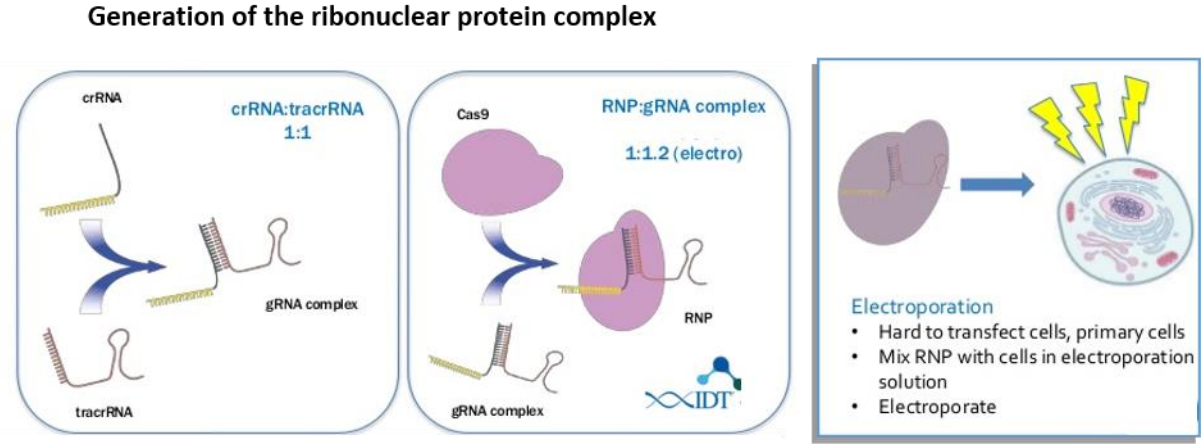


Figure 30. Generation of the CRISPR ribonucleoprotein (RNP) complex.
crRNA and tracrRNA are annealed together to form the gRNA complex, before complexing with Cas9 to form the RNP gRNA complex at a 1:1.2 ratio of Cas9 to gRNA and nucleofecting BJ fibroblasts. Adapted from IDT (https://eu.idtdna.com/)

4.4 TARGET REGIONS

Two different regions important for both the structure and function of TN-C were targeted by all three approaches: exon 2, containing the ATG translation start codon and encoding for the assembly domain necessary for TN-C oligomerisation and exon 7, encoding the integrin-binding RGD motif. Approach 1 using Cas9n was targeted to either exon 2 or 7. Approach 2 and 3 with Cas9 or Cas9 RNP targeted both exon 2 and 7 separately and exon 2 and 7 simultaneously. The latter was done as *TNC* is very large with a multidomain structure and targeting only the ATG may generate a truncated protein which could still retain some activity. Thus, targeting a functional site in exon 7, in addition to the ATG codon in exon 2, increases the chance of a complete knockout or a non-functional protein. All of the targeted sites are expressed by every *TNC* isoform following alternative splicing.

The edited *TNC* knockout fibroblasts can then be utilised to produce an ECM devoid of TN-C in the form of a cell-derived-matrix (*TNC* KO CDM). Macrophages can then be cultured on the WT CDMs or *TNC* KO CDMs to investigate if human TN-C regulates the expression of miR-155 and other inflammatory genes as previously shown in mice.

4.5 DESIGN OF CRISPR GUIDE RNAS

CRISPR/Cas9n and CRISPR/Cas9 guide design was carried out using the CRISPR design software (http://crispr.mit.edu/). The region submitted was either a 250 bp TNC coding sequence from exon 2, including the ATG (Figure 31), or from exon 7, including the RGD domain with highly conserved amino acids shown in bold (Figure 32). The software generates a list of gRNAs located next to the 3' NGG PAM sequence. Guides are scored from 0-100 with higher scores indicating less off-targets and therefore increased guide specificity: guides with a score of >50 and 0 off-targets were selected (table 22). In the case of Cas9 guides, guides were also selected by the criteria that no predicted off-targets were in ECM or ECM-associated genes. For subsequent analysis of the predicted off-targets of the guides Benchling software (http://www.benchling.com) was utilised due to the MIT software being closed down (table 22). In this, the software again generates a list of guides adjacent to the PAM sequence. These are scored utilising the same algorithm for off-targets as the MIT software [403] with a score from 0 to 100. However, these guides are also given an on-target score (0-100) [435]. The software then generates the predicted off-target sequences in the off-target gene. This allows PCR primers to be designed to amplify the predicted off-target region; the resultant PCRs are discussed later in this chapter.

Exon 2 probe sequence (250bp)

```
361 ATACCCTAGA GCCCTAGAGC CCCAGCAGCA CCCAGCCAAA CCCACCTCCA CCATGGGGGC
421 CATGACTCAG CTGTTGGCAG GTGTCTTTCT TGCTTTCCTT GCCCTCGCTA CCGAAGGTGG
481 GGTCCTCAAG AAAGTCATCC GGCACAAGCG ACAGAGTGGG GTGAACGCCA CCCTGCCAGA
541 AGAGAACCAG CCAGTGGTGT TTAACCACGT TTACAACATC AAGCTGCCAG TGGGATCCCA
601 GTGTTCGGTG
```

Figure 31. The 250bp coding sequence submitted (http://crispr.mit.edu/) for exon 2 for guide design. The ATG start codon is shown in bold.

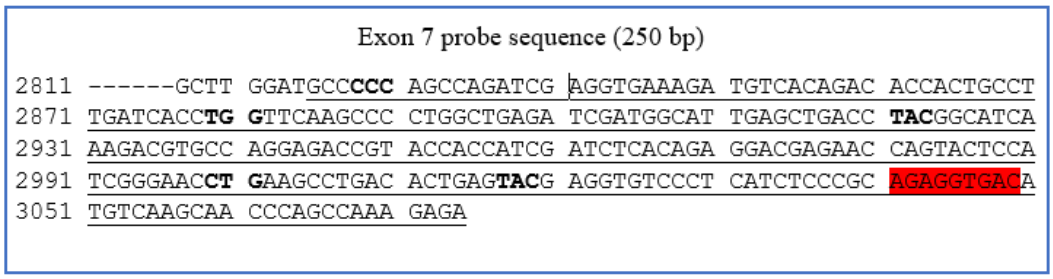


Figure 32. The 250bp coding sequence (underlined) submitted to (http://crispr.mit.edu/) for exon 7 guide design.
The RGD domain is shown in red whilst codons for highly conserved amino acids are shown in bold.

Table 22. Guide sequences for all approaches of CRISPR.
Target exons and the specific chromosomal location of guide binding sites are indicated. The on-target and off-target scores from the two softwares used to generate guides are indicated with scores from 0-100. Higher scores indicate more specific guides.

			MIT software [403]	**Benchling software**	
				Doench [435]	**MIT [403]**
Guide name	**Sequence**	**Locus**	**Score (0-100)**	**On-target score (0-100)**	**Off-target score (0-100)**
Exon 2					
A1 (Approach 1)	CTGAGTCATGGCCCCCATGG	chr9:+115091018	60	57.5	57.5
A2	TTGCCCTCGCTACCGAAGGT	chr9:-115090956		45.2	
B1 (Approach 1)	GGACCCCACCTTCGGTAGCG	chr9:+115090963	81	59.2	79
B2	ATCCGGCACAAGCGACAGAG	chr9:-115090918		61.7	
A2 (Approach 2 and 3)	TTGCCCTCGCTACCGAAGGT	chr9:-115090956	96	45.2	96
2A3 (Approach 2)	GCCATGACTCAGCTGTTGGC	chr9:-115090996	71	47	68.4
Exon 7					
7A (Approach 3)	TGGCACGTCTTTGATGCCGT	chr9:+115078105	94	65.8	93.8
7B (Approach 2)	GACACCTCGTACTCAGTGTC	chr9:+115078019	86	53.2	84.5

For approach 1 (CRISPR/Cas9n) targeting exon 2, the A pair of guides, consisting of gRNA A1 and A2 with an offset of 28 bp, have an aggregate score of 60. The B pair of gRNAs consisting of the gRNAs B1 and B2, which have an offset of 11 bp, have an aggregate score of 81 and 0 predicted off-targets (Figure 33). For approach 2 (CRISPR/Cas9) targeting exon 2, the A2 guide had a score of 96 with 2 predicted off-targets whilst guide 2A3 had a score of 71 and 2 off-targets in pseudogenes (Figure 34). For approach 2 targeting exon 7, guide 7A had a score of 94 and 1 predicted off-target whilst guide 7B had a score of 86 and 1 predicted off-target (Figure 34). For approach 3 (CRISPR RNP) targeting exons 2 and 7, the guides were identical to A2 and 7A and as such had the same score and off-targets as previously (Figure 35, table 13).

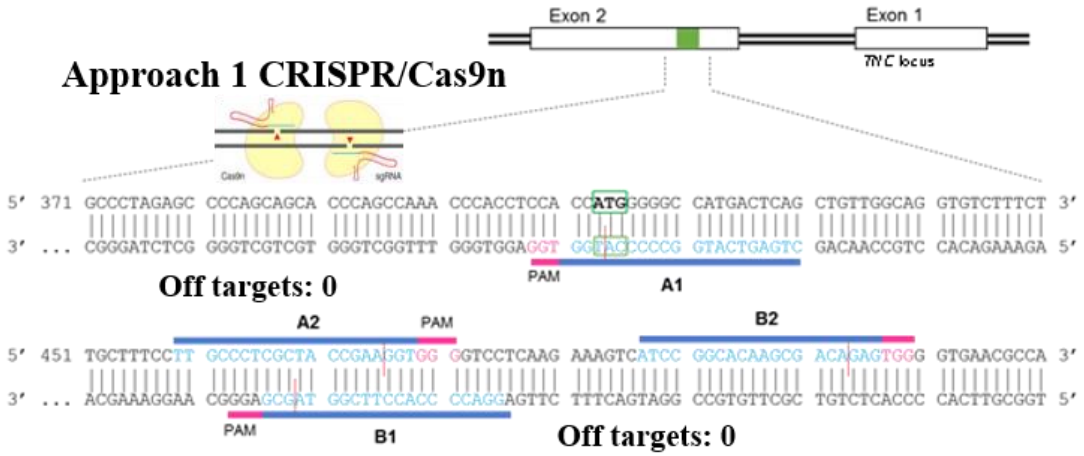


Figure 33. Schematic of approach 1 with CRISPR/Cas9n (plasmid) guides targeting exon 2 to knockout the *TNC* gene.
The translation start codon ATG is boxed in green, PAM sequences located adjacent to 20 nt guides are highlighted in pink and blue respectively. Two guides are required to induce a double strand break (indicated by red line, cutting 3 bases upstream of the PAM sequence). Predicted 'Off-targets' for pairs are indicated.

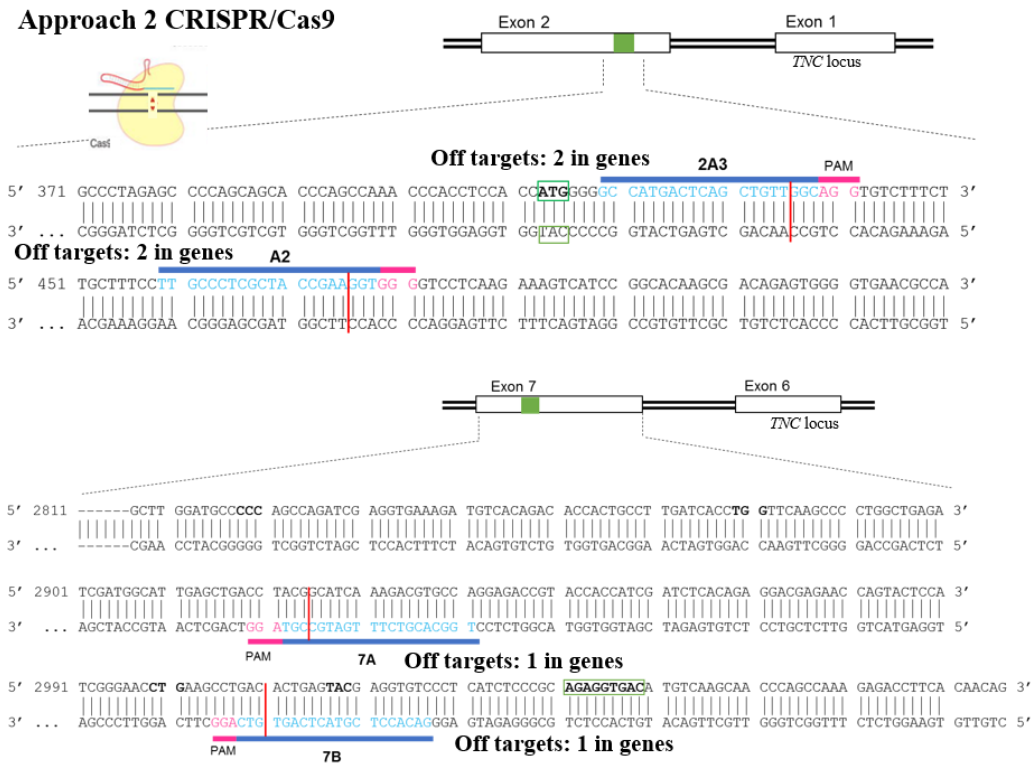


Figure 34. Schematic of approach 2 with CRISPR/Cas9 (plasmid) guides targeting exon 2 and exon 7 to knockout *TNC* gene.

The translation start codon ATG is boxed in green (top panel), PAM sequences located adjacent to 20 nt guides are highlighted in pink and blue, respectively. The RGD domain is boxed in green (bottom panel). One guide is sufficient to induce a double strand break (indicated by red line, cutting 3 bases upstream of the PAM sequence). 'Number of predicted off-targets' for these guides are indicated, however these guides were chosen as they do not have off-targets in ECM or ECM-associated protein coding genes.

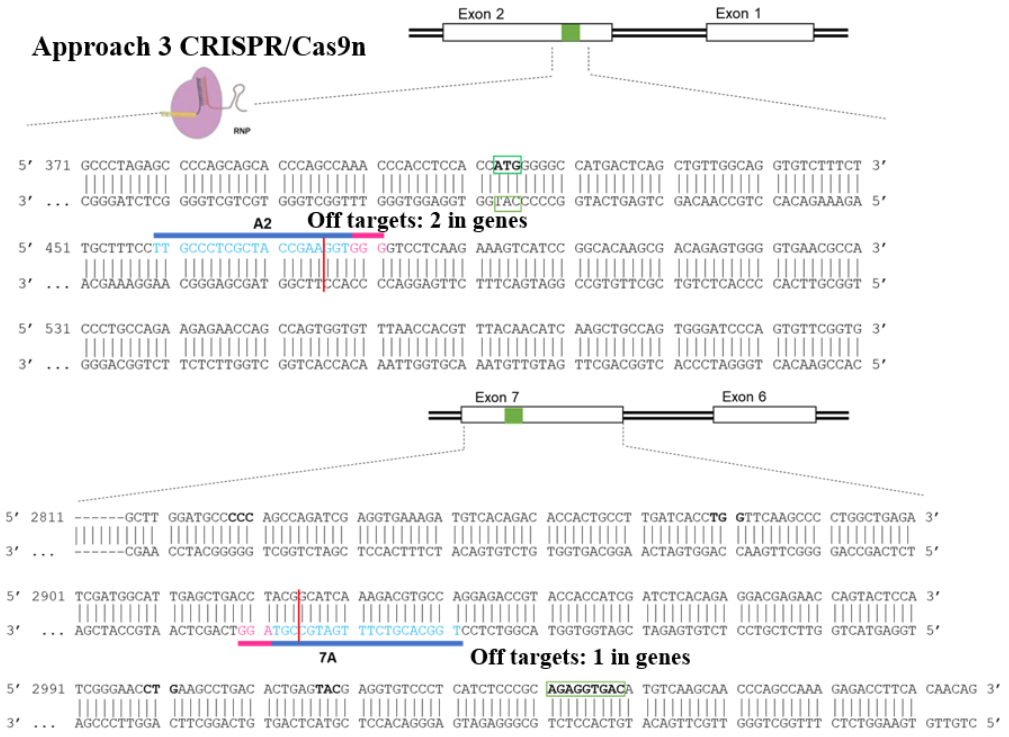


Figure 35. CRISPR/Cas9 ribonuclear protein (RNP) approach.
Schematic of approach 3 with CRISPR/Cas9 RNP guides targeting exon 2 and exon 7 to knockout *TNC* gene. The translation start codon ATG (top panel) is boxed in green, PAM sequences located adjacent to 20 nt guides are highlighted in pink and blue, respectively. The RGD domain is boxed in green (bottom panel). One guide is sufficient to induce a double strand break (indicated by red line, cutting 3 bases upstream of the PAM sequence). 'Number of predicted off-targets' for these guides are indicated, however these guides were chosen as they do not have off-targets in ECM or ECM-associated protein coding genes.

4.6 Cloning of CRISPR guide RNAs

Following on from guide design and synthesis by Sigma-Aldrich, top and bottom guides targeting exon 2 or exon 7 were annealed and cloned into the CRISPR/Cas9 or CRISPR/Cas9n vectors described above. For this, vectors were digested with BbsI restriction enzyme followed by gel purification to remove the resulting 22 nucleotide fragment (Figure 36).

The guides oligonucleotides were annealed, phosphorylated with T4 polynucleotide kinase and ligated into the expression vectors. *E. coli* DH5a cells were transformed by heat shock and selected on agar plates containing ampicillin, followed by purification of plasmid DNA. Agarose gel electrophoresis and extraction confirmed the gRNA inserts had ligated in the correct position and orientation in the vector, with no undesired mutations. Sequencing primers spanned the insert and vector joining region.

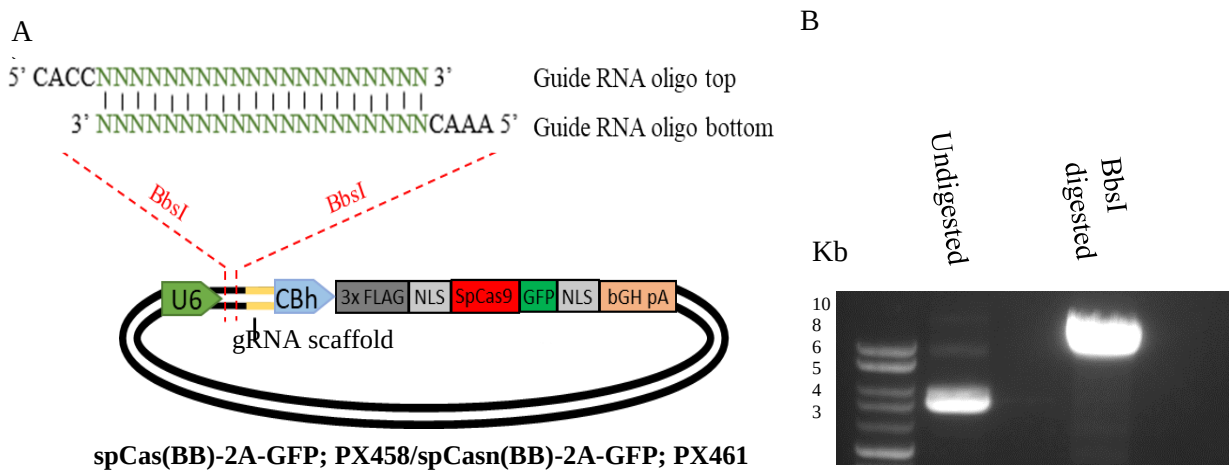


Figure 36. Plasmid map and restriction digest of plasmid.
A) Schematic of spCas(BB)-2A-GFP; PX458 (CRISPR/Cas9), spCasn(BB)-2A-GFP; PX461 (CRISPR/Cas9n) and BbsI restriction enzyme target sites for cloning of the 20nt guide (green 'N') into the vector, sgRNA scaffold is shown in yellow. **B)** Agarose gel electrophoresis of 2 µg of BbsI digested PX458 vector and 500 ng undigested vector control on a 0.6 % agarose gel.

4.7 VALIDATION OF CRISPR-CAS9 AND -CAS9N VECTORS

4.7.1 Transfection into mammalian cells

Due to BJ fibroblasts being difficult to transfect cells and slow growing, HEK-293 cells were used to validate the constructed vectors. RNA-Seq data from the Human Protein Atlas confirmed HEK-293s express TNC (Figure 37). Previous work from the lab confirmed this by qPCR. HEK-293s were transfected by the calcium-phosphate method with 1 µg of each vector and the gRNAs for CRISPR/Cas9n targeting exon 2 (A1 and A2; B1 and B2).

Figure 37. HEK-293s express *TNC*.
Data from The Human Protein Atlas.

The A pair and B pair nickase vectors demonstrated a good transfection efficiency of ~ 70 % (Figure 38). Similarly, the CRISPR/Cas9 vectors demonstrated a good transfection efficiency, however HEK293s transfected with 7B displayed a higher transfection efficiency than those transfected with vector 7A (Figure 39). Non-transfected and mock-transfected cells showed no GFP fluorescence whilst cells transfected with the positive control pmaxGFP plasmid showed GFP fluorescence. As Cas9 and GFP are expressed as a fusion ORF on the same transcript, GFP expression parallels Cas9 expression and can therefore be used to extrapolate Cas9 expression. This data indicates that the constructed vectors can be transfected into mammalian cells and express Cas9 required for *TNC* gene editing at the target locus in *TNC*.

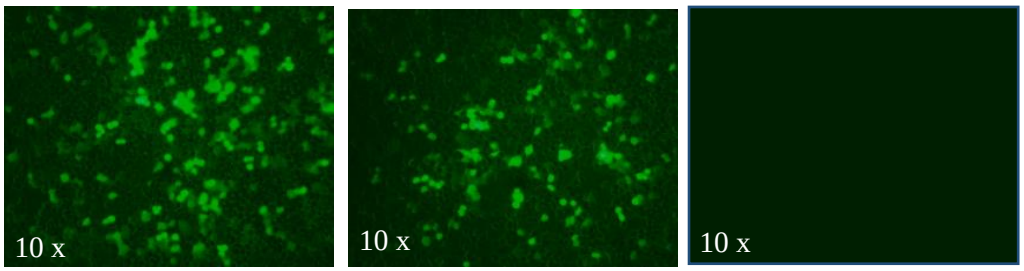


Figure 38. Calcium phosphate transfection of HEK-293s with CRISPR/Cas9n vectors.
Expression of reporter green fluorescent protein (GFP) in HEK293s 48 hrs after transfection with 2 µg total of CRISPR/Cas9n plasmids targeting exon 2. One representative image for each condition per transfection is shown.

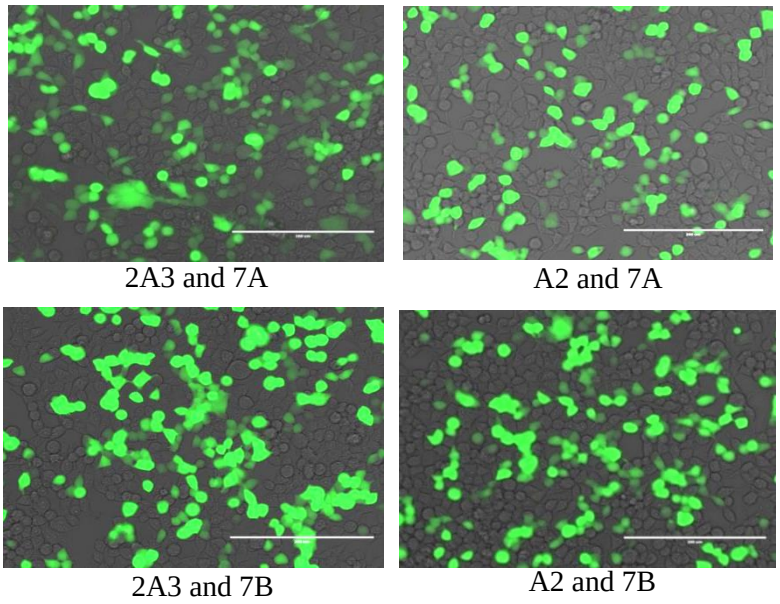


Figure 39. Calcium phosphate transfection of HEK-293s with CRISPR/Cas9 vectors.
Expression of reporter green fluorescent protein (GFP) in HEK-293s 48 hours after transfection with 2 µg total of CRISPR/cas9 plasmids; with one plasmid targeting exon 2 and one plasmid exon 7. One representative image for each condition per transfection is shown. Scale bar: 200 µm.

4.7.2 Validation of Cas9n expression

To verify Cas9n expression by the cloned vectors, HEK293 cells were transfected with 2 µg of each CRISPR/Ca9n vector (A1, A2, B1, B2) and after 24 hours total proteins were extracted. The expression of Cas9n was quantified by western blotting (Figure 40). The Cas9n enzyme can be seen at the expected size of 130 kDa. The other faint bands are most likely FLAG-tagged cleavage products as Cas9n is expressed as a fusion ORF with GFP and a self-cleaving 2A linker. Western blotting of extracted proteins from transfected HEK293s indicates nickase vectors express Cas9.

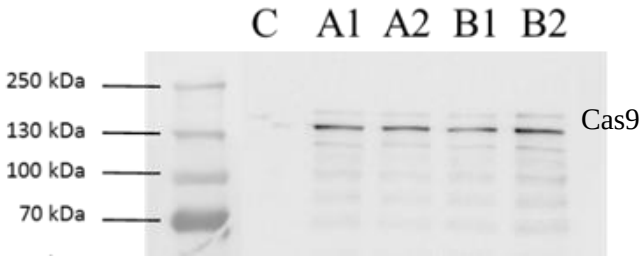


Figure 40. Western blot of FLAG-tagged Cas9n proteins from HEK293s transfected with CRISPR/Cas9n vectors.
HEK-293s were transfected with 1 µg of either A1, A2, B1, B2 or mock transfected control (C). Proteins were harvested 48 hours after transfection into Laemmli buffer and resolved by SDS-PAGE on an 8% polyacrylamide gel.

4.7.3 Validation of *TNC* gene edit using CRISPR/Cas9n or CRISPR/Cas9 in HEK-293 cells

To determine if the transfected vectors cause a CRISPR/Cas9-mediated *TNC* edit, genomic DNA (gDNA) was extracted from bulk populations 48 hours after transfection and subjected to PCR amplification and sequencing. Single nickase vectors were included as negative controls for their inability to induce DSBs. Controls should therefore be wild type DNA when sequenced. Several sets of PCR primers were designed that flanked the targeted regions that were ~30 to 500bpt upstream or downstream of the PAM sequence directly adjacent to the guide, depending on whether the guide targets the top or bottom strand of DNA. Amplicon sizes were also designed to be distinguishable from each primer set. Target regions were amplified with a large 781 bp primer set, and further PCRs with smaller sets could be carried out if mutant products could be seen. This was done to avoid PCR failure due to primer proximity to the targeted region. In some experiments the condition '+ AZT (azidothymidine)' can be seen; this small molecule was reported to decrease homology directed repair and thus increase error-prone repair NHEJ [516]. However, this treatment was abandoned early on due to lack of success and extra stress on cells during the challenging process of generating knockout cells.

Based on the size of the PCR products no obvious INDELs could be detected in HEK293s transfected with either CRISPR/Cas9 vectors (Figure 41 A) or CRISPR/Cas9n (Figure 41 B) vectors after touchdown PCR with primer set B 3 (Table 8). Products sizes were all the expected 781bp; *TNC* gene

mutants are expected to be shorter or longer, depending on whether insertions or deletions (INDELS) have occurred.

To confirm that a very small base pair change had not been detected, a standard PCR to generate a smaller amplicon was performed. To screen exon 2, targeted by Cas9 vectors 2A3 and A2, primer set 2 that generated a 232bp fragment was used. Exon 7 targeted by Cas9 vectors 7A and 7B was screened using primer set 7A and primer set 7B to generate 352 bp fragments and 185 bp fragments, respectively. Regions targeted by Cas9n pair A were amplified using primer set A2 to generate a 232 bp fragment whilst B pair regions targeted by Cas9n pair B were amplified with primers B 1 Fw and B 2 Rv to generate a 271 bp fragment. Whilst some band shifts were observed, band extraction and sequencing revealed no mutations were present in the HEK293s transfected with Cas9 vectors or Cas9n vectors. This could be because gDNA was extracted from the entire cell population, which contained a mixture of wild type and mutant cells. Thus, the mutants may not have been prevalent enough in the population to be detected by PCR. Additionally, the expression of *TNC* may also be very low (34 Ct) resulting in the gene being condensed into heterochromatin and subsequently lead to a physical obstacle in which Cas9 is not able to bind the target region for cleavage.

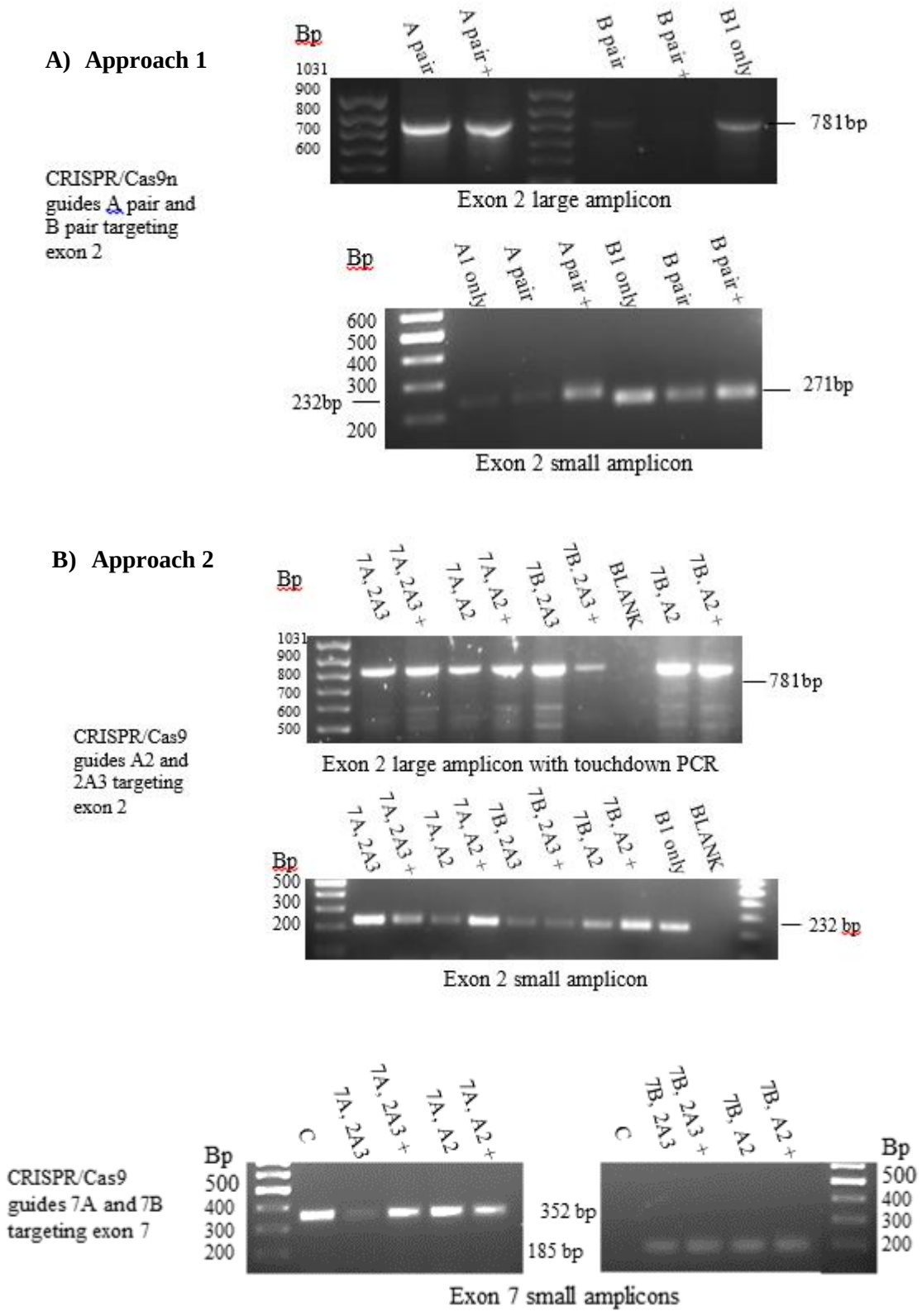


Figure 41. Genotyping by PCR of transfected HEK293s demonstrated no CRISPR/Cas9 edits had occurred. PCR products generated from the bulk GFP$^+$ population of HEK293s after genomic extraction following transfection with: **(A)** 2 µg total of 'A pair' or 'B pair' CRISPR/Cas9n vectors targeting exon 2 or **(B)** CRISPR/Cas9 vectors targeting exon 2 and 7. Controls are one nickase vector alone. + indicated a treatment with azidothymidine (AZT) during transfection that was later abandoned. 25 µl of a 25µl reaction was ran on a 1.5 % w/v agarose gel. Control **(C)** indicates cells transfected with single guide only, which should therefore be the unedited and exhibit a WT sequence.

4.7.4 Validation of *TNC* gene edit via CRISPR/Cas9n in MDA-MB-231 cells

To verify our hypothesis that the lack of a gene edit by CRISPR/Cas9n in HEK293s was due to low expression levels of *TNC* that may not be sufficient enough to achieve and detect a gene edit, MDA-MB-231 cells, which are derived from breast adenocarcinoma and are known to express high levels of *TNC* [517], were used to validate a *TNC* edit. First, to ensure that this cell line was suitable to validate CRISPR/Cas9 gene editing of *TNC*, expression of *TNC* was confirmed by RTqPCR. As cell stimulation with LPS is known to induce *TNC* expression [224] and MDA-MB-231s express the LPS receptor TLR4 [518], cells were stimulated with LPS. After LPS stimulation (100 ng/ml), RNA extraction and reverse transcription, TNC mRNA was quantified by qPCR (Figure 42). MDA-MB-231s *TNC* expression peaked 4 hours after stimulation (22 Ct), 15-fold more than unstimulated cells (25 Ct).

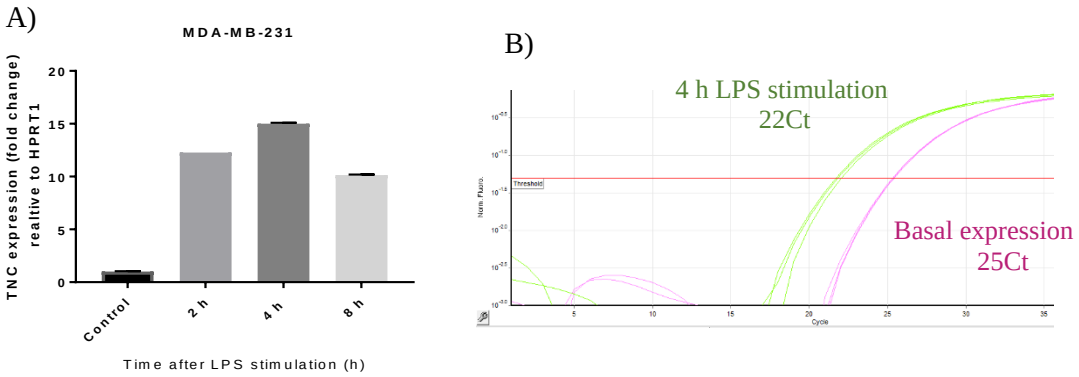


Figure 42. MDA-MB-231 express *TNC* and are responsive to LPS stimulation.
A) Quantitative RT-PCR analysis of *TNC* mRNA in MDA-MB-231s after LPS (100 ng/ml) stimulation. Gene expression (fold-change) normalised to *HPRT1* housekeeping gene and relative to unstimulated control. **B)** Amplification plot showing peak stimulation at 4 hours compared to unstimulated basal expression. Results from 1 experiment conducted in triplicate, mean +/- SD.

MDA-MB-231s were stimulated with LPS (4 hours before transfection (+); or left in LPS supplemented media overnight (++)) and transfected by the calcium phosphate method, before sorting GFP^+ cells into a bulk population by FACS as previously described. GFP^+ cells were diluted manually to single cells and seeded into 96-well plates, with the remainder into 6-well plates for bulk expansion. gDNA extraction of the bulk sample and PCR amplification using primer set B 1Fw and B 1Rv and B 1Fw and B 2Rv, generating products 177 bp and 277 bp fragments spanning the CRISPR/Cas9n targeted region, was carried out. PCR resulted in expected product sizes indicating no obvious indels (Figure 43 A).

After expansion of manually diluted single cells into 6-well plates, gDNA was extracted and amplified using the large primer set B 3Fw and B 3Rv to generate a 781 bp fragment (Figure 43 B). Multiple bands of different sizes were seen (B pair lane, Figure 43 B) that were not the expected WT product size, indicating an insertion had occurred. These bands were extracted and sequenced demonstrating three products. An intermediate product (green box) was determined to be mixed template by sequencing (Figure 43 C). A premature stop codon was found in the largest product (blue box Figure 43 B, lane 'B Pair'). The lowest band boxed in yellow was wild type sequence and translated in silico to compare with the mutant product sequence. The three bands present were not unexpected due to the manual dilution in which the resulting cell lines may not be clonal, the MDA-MB-231 cell line is also near triploid. This experiment indicated the CRISPR/Cas9n B pair of guides were functional and could produce a targeted *TNC* gene edit providing proof of principle. However, due to repeated failures in expanding fibroblasts clones transfected with the nickase plasmids, and the knowledge that Cas9 is more efficient than Cas9n, ultimately the nickase system was not used to generate a *TNC* KO cell line.

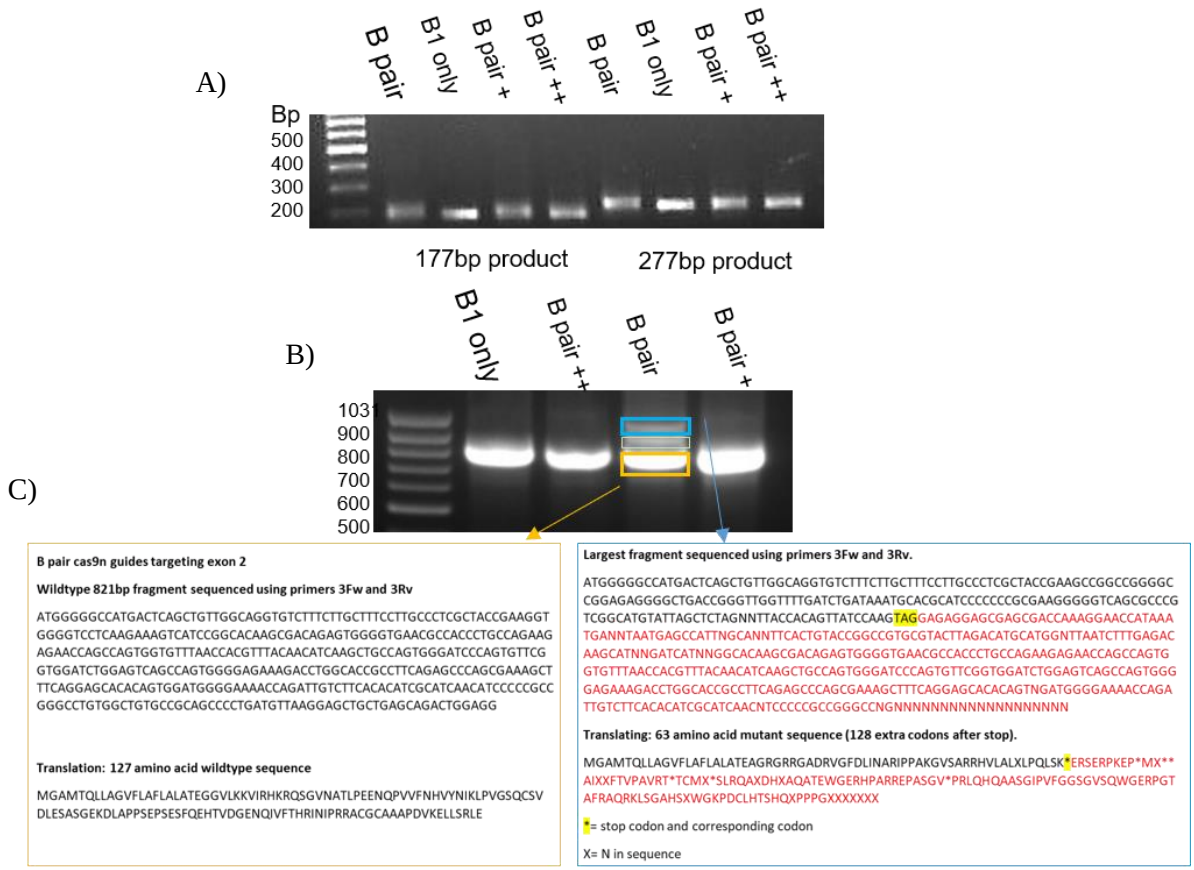


Figure 43. Genotyping of mutant alleles by PCR of MDA-MB-231s cells transfected with CRISPR/Cas9n 'B pair' stimulated with LPS.
PCR of gDNA from MDA-MB-231s after transfection with 2 µg in total of CRISPR/Cas9n 'B pair' vectors targeting exon 2, in conjunction with 4 hours LPS treatment (+) or overnight LPS treatment (++) following cell sorting into bulk **(A)** or single cells **(B).** In single sorted cells, mutant alleles, boxed in blue and green, can be seen in lane 'B pair'. The wild type product is boxed in yellow. **C)** Wild type (yellow) and mutant band (blue) sequence translation indicating a premature stop codon in the mutant allele. 25 µl of 25 µl reaction was run on a 1.5 % agarose gel.

4.8 Optimisation of transfection and recovery of BJ fibroblasts

4.8.1 Cell synchronisation

Low proliferation rates of BJ cells pose a number of challenges; low transfection efficiencies and, as a result, low expression of Cas9 and a low frequency of NHEJ repair events. Cell cycle synchronisation has been shown to increase transfection efficiency [519] and also total CRISPR editing [429]. Cell cycle synchronisation was achieved using aphidicolin and successfully halted the cell cycle at G_1 (Figure 44), as demonstrated by propidium iodide (PI) staining of cells and subsequent FACS analysis. Upon aphidicolin removal, cell progressed through the cell cycle, with 56.6% of cells entering S phase 2 hours after release.

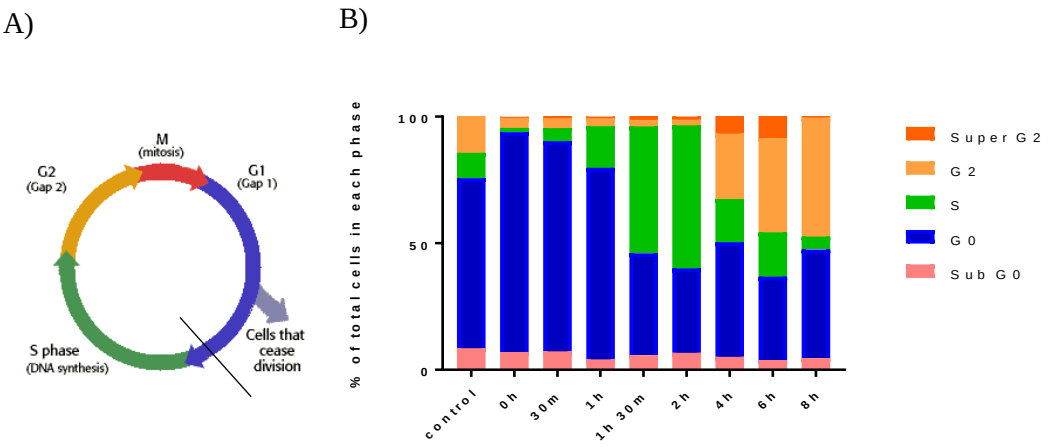


Figure 44. Cell cycle synchronisation achieved with aphidicolin successfully halted the cell cycle at G_1.
A) Schematic showing the stages of the cell cycle, with aphidicolin halting the cycle at G1 (blue). **B)** Cells were harvested and stained with PI before analysis by FACS showed as percentage of cells progressing through the cycle after aphidicolin removal. Colours correspond to the stages of the cycle in **(A).**

Transfection of BJ cells with 2 µg in total of CRISPR/Cas9n B pair vectors, after cell cycle synchronisation release at 30 minutes and 1 and 2 hours resulted in high cell death (Figure 45) and a reduced transfection efficiency compared to unsynchronised transfected cells.

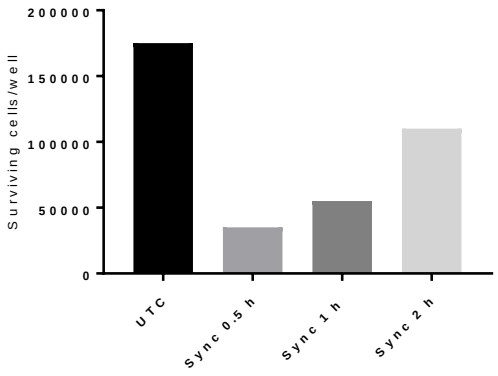


Figure 45. Cell survival following cell cycle synchronisation and transfection with 'B pair' guides.
Cells were synchronised at G_1 with aphidicolin (20mM) followed by release for 30 mins, 1 h or 2 hrs before transfection with 2 µg total of CRISPR/Cas9n B pair guides. Cells were harvested and counted and the surviving cells from the 200,000 transfected calculated. Data are from one experiment with one measurement taken at each condition.

Transfection after 2 hours release resulted in the least cell death, however a drastic reduction in transfection efficiency was still observed with far less cells expressing GFP (Figure 46). Cell synchronisation was therefore abandoned.

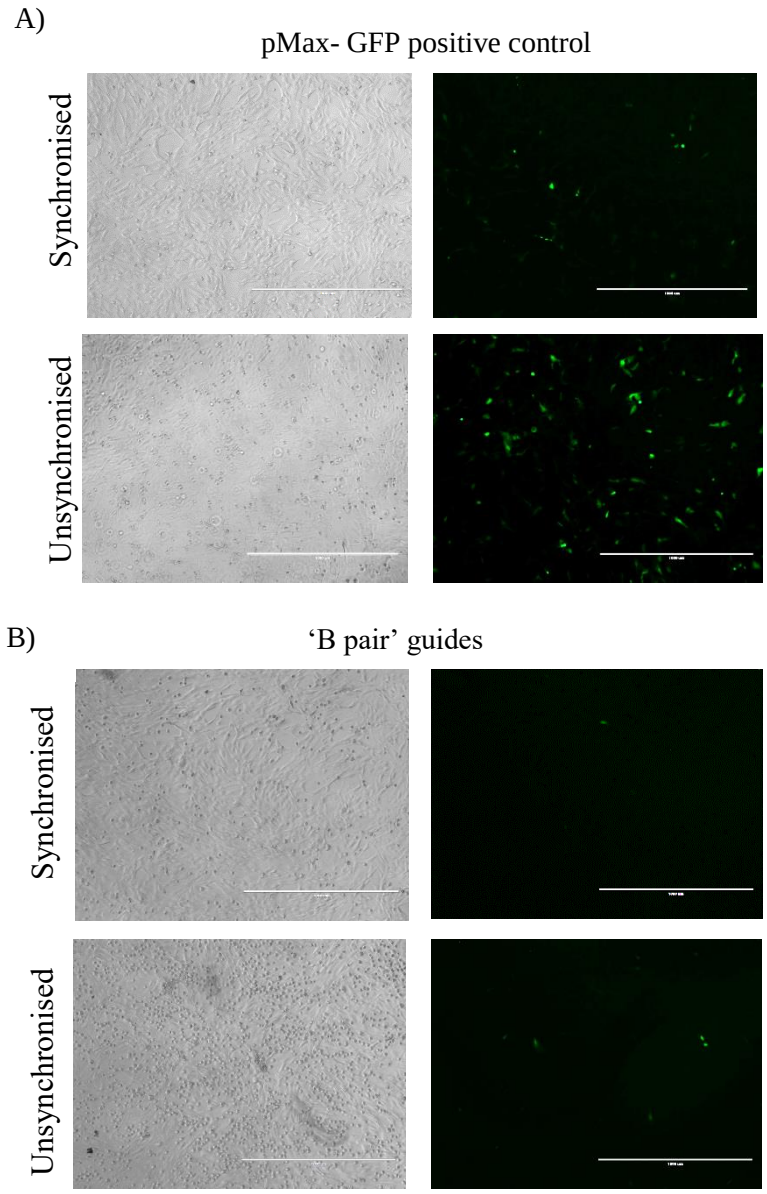


Figure 46. Transfection efficiency after cell synchronisation with aphidicolin and release for 2 hrs.
BJ cells expressing GFP reporter gene 24 hrs after synchronisation with aphidicolin and released for 2 hrs or unsynchronised followed by transfection with positive control GFP: A) or 2 µg total of 'B pair' CRISPR/Cas9n: B). Scale bar: 1000 µm

4.8.2 Does LPS stimulation of fibroblasts increase transfection efficiency and gene editing efficiency?

Genes that are not actively expressed can be within condensed heterochromatin; thereby blocking access and cleavage of the DNA strands by the Cas9 enzymes. Correspondingly, genes that are being actively expressed are not condensed, thereby allowing access and cleavage by Cas9/Cas9n. Whilst BJ cells do express basal levels of *TNC* (23 Ct), stimulation with LPS further increases it with peak expression at 4 hours (20 Ct), 3 fold more than unstimulated cells (23 Ct) (Figure 47). Furthermore,

LPS stimulation induces the expression of cell survival and proliferation genes [520]. Thus LPS stimulation was tested for its ability to enhance the transfection efficiency and *TNC* gene editing efficiency.

A) B)

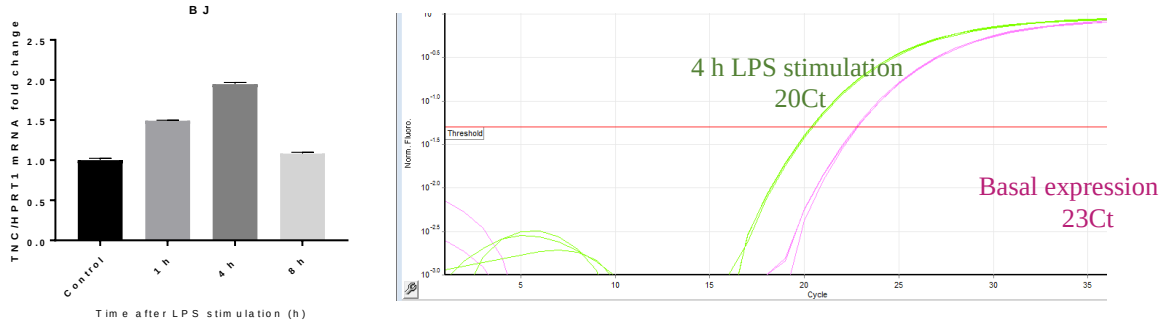


Figure 47. BJ fibroblasts express *TNC* and are responsive to LPS stimulation.
A) Quantitative RT-PCR analysis of *TNC* mRNA in BJ fibroblasts after LPS (100 ng/ml) stimulation, normalised to *HPRT1* housekeeping gene and relative to unstimulated control. **B)** Amplification plot showing peak stimulation at 4 hours compared to unstimulated basal expression. Results from one experiment conducted in triplicate, mean +/- SD.

BJ fibroblasts were stimulated with 100 ng/ml of LPS for 18 hours after transfection as Cas9n expression, inferred by GFP expression and GFP + cells, is induced approximately 24 hours after transfection, while *TNC* expression peaks at 4 hours after stimulation. Stimulation at 18 hours should therefore result in parallel peak *TNC* and Cas9n expression. Stimulation increased transfection efficiency as verified by FACS analysis of the % of GFP^+ cells (table 14) during cell sorting, compared to unstimulated transfected cells.

Table 14. FACs analysis of GFP^+ fibroblasts following transfection with 'B pair' CRISPR/Cas9n guides alone or in conjunction with an LPS stimulation 18 h after transfection.
Single guide control (B1 only) also shown.

Sample	Total cells sorted	GFP^+ cells	GFP^+ cells (%)
B pair	85000	4950	5.82
B pair + LPS	68000	6290	9.25
B1 only	93000	5520	5.9

4.8.3 Optimisation of BJ single cell culture and expansion

Cells containing the CRISPR plasmids after transfection are sorted into single GFP^+ cells by FACS. Recovery and expansion of BJ single cells is difficult as they naturally do not grow well in isolation therefore a series of plating conditions were tested to ensure the highest possible survival rate of potentially gene-edited single GFP^+ cells. 96-well plates were either left uncoated (plastic only) or coated with i) 0.2% gelatin on the day of seeding, ii) 0.2% gelatin the day before seeding (stored in PBS/Pen-Strep or media overnight) or iii) BJ cell-derived matrices (CDM). Cell were fed with either complete growth media or 50% BJ conditioned media. ~ 5 cells were seeded per well, and upon the first well reaching confluence, live cells in all wells were counted (Figure 48). ~5 cells were seeded as opposed to a single cell to ensure that a lack of observed growth was due to a suboptimum plating condition and not an absence of a cell in the well at seeding. Gelatin-coated wells with complete growth media resulted in the best growth of cells (Figure 48). 96-well plates were therefore coated with 0.2% gelatin to recover CRISPR/Cas9n transfected single cell clones.

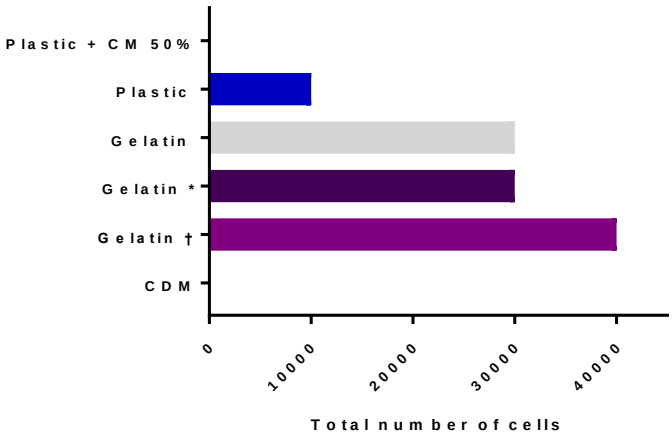


Figure 48. Fibroblast cell recovery upon seeding on different plating conditions.
Number of cells after ~5 cells were seeded onto either uncoated (plastic), uncoated with BJ conditioned media (CM), gelatin-coated and stored overnight in PBS*, gelatin-coated on the day of seeding †, or BJ cell derived matrix (CDM) and cultured until a well reached confluency. Count shown is total cells from four 96 wells pooled to enable counting of small cell numbers.

4.9 GENERATION OF *TNC* KNOCKOUT FIBROBLASTS USING CRISPR/CAS9 (APPROACH 2)

After the validation of Cas9 targeting vectors and optimisation of transfection and clonal expansion of GFP+ cells, BJ fibroblasts were nucleofected with two CRISPR/Cas9 vectors targeting *TNC* exon 2 and exon 7, and GFP +ve cells selected by FACS (Figure 49). Two rounds of enrichment occurred, the first to separate GFP expressing and thus CRISPR plasmid harbouring cells from non-fluorescent cells (bulk sort) and then to select the cells with the highest expression of GFP. Single GFP + cells were plated into individual gelatin-coated wells of a 96 well plate by FACS and expanded into clonal populations.

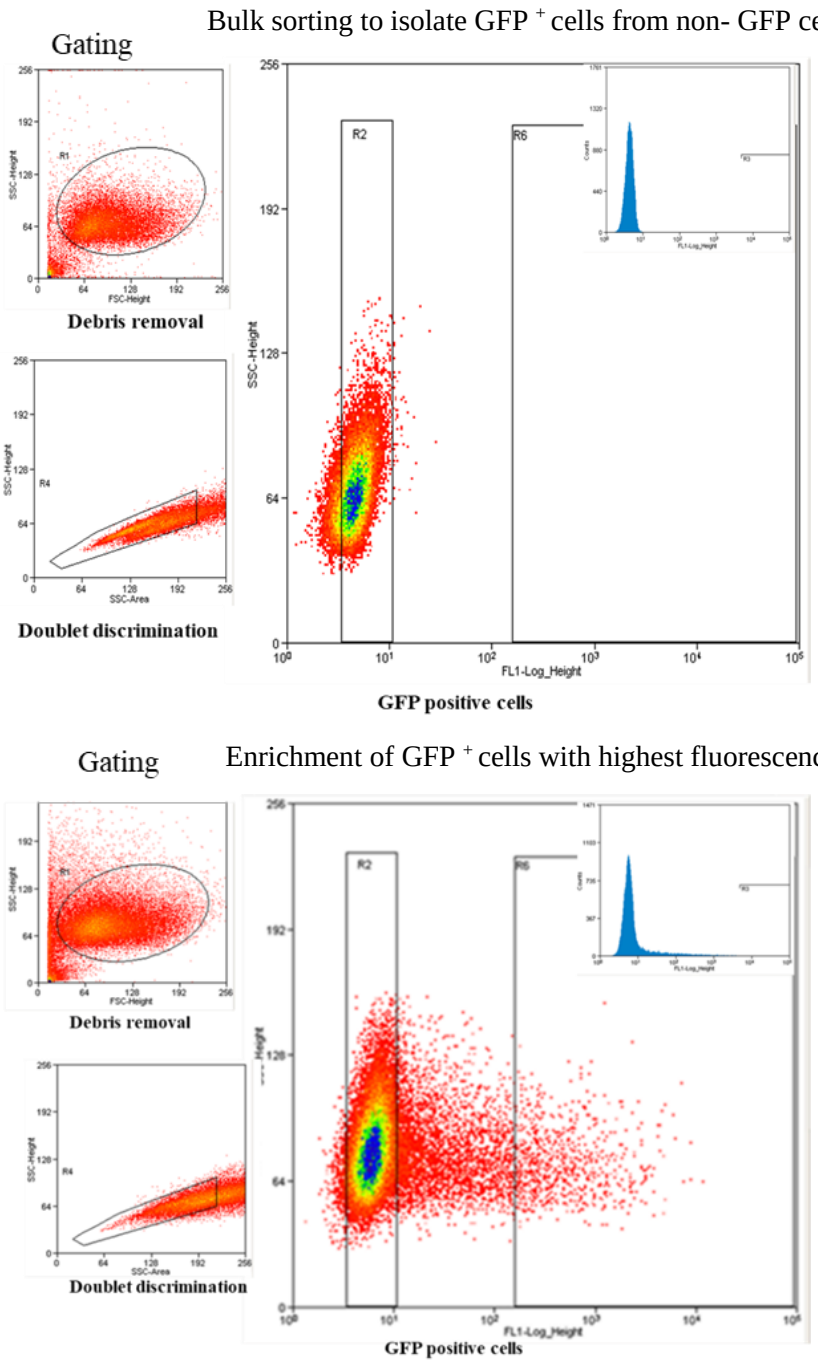


Figure 49. FACS analysis of BJ fibroblasts transfected with CRISPR/Cas9 plasmids encoding *TNC* targeting guides 7B and A2 to enrich for the fibroblasts expressing the highest levels of GFP.
Dot plots and the corresponding histograms of BJ fibroblasts undergoing FACS analysis after transfection with CRISPR/Cas9 plasmids encoding guides targeting exon 7 and exon 2. Firstly, cells underwent a bulk sort to separate GFP expressing fibroblasts from non-fluorescent fibroblasts. Next, single cells expressing the highest levels of GFP (R6) in the population were plated after gating to exclude debris/dead cells (R1) and doublets (R4). As GFP is utilised to infer Cas9 expression, the fibroblasts with the highest chance of a gene edit were seeded when selecting for the highest GFP expression.

4.9.1 Assessment of potential *TNC* KO fibroblasts at the protein level

4.9.1.1 Western blotting and immunofluorescence revealed a complete knockout at the protein level, whilst other key structural ECM proteins were produced and deposited

After successful expansion of the one surviving clone from several hundred plates, potential knockout fibroblasts at PDL 54 were allowed to deposit matrices for 14 days on 10 cm dishes for western blotting or on 18 mm coverslips for immunofluorescence analysis. Western blotting of crude lysates of cells and matrix revealed robust expression of TN-C in WT controls, however TN-C was not detected in the GFP+ clone indicating successful knockout of TNC expression (Figure 50 A). Note the TN-C antibody, detects an epitope in the N- terminus region, which is able to bind to all TNC transcripts, thus confirming a complete knockout. As a control, a second antibody confirmed equal levels of FN protein in both WT and targeted cells. Neither TN-C nor FN were detected in cell supernatents.

Absence of TNC expression in the targeted cells was further confirmed by immunofluorescence, using the same antibody (Figure 50 B). Future work should utilise an antibody that detects an epitope in the C-terminus for a more comprehensive assessment of protein production. Immunofluorescence analysis of COL1, COL3 and FN confirmed that these proteins were still deposited and incorporated into the ECM unperturbed by TNC knockout (Figure 50 C). Together these data indicated a full *TNC* gene knockout at the protein level. Some areas of FN appeared less dense. Handling of the coverslips can cause damage around the edges of the coverslips and some differences attributable to removal of the coverslips from the tissue culture well can cause this phenotype.

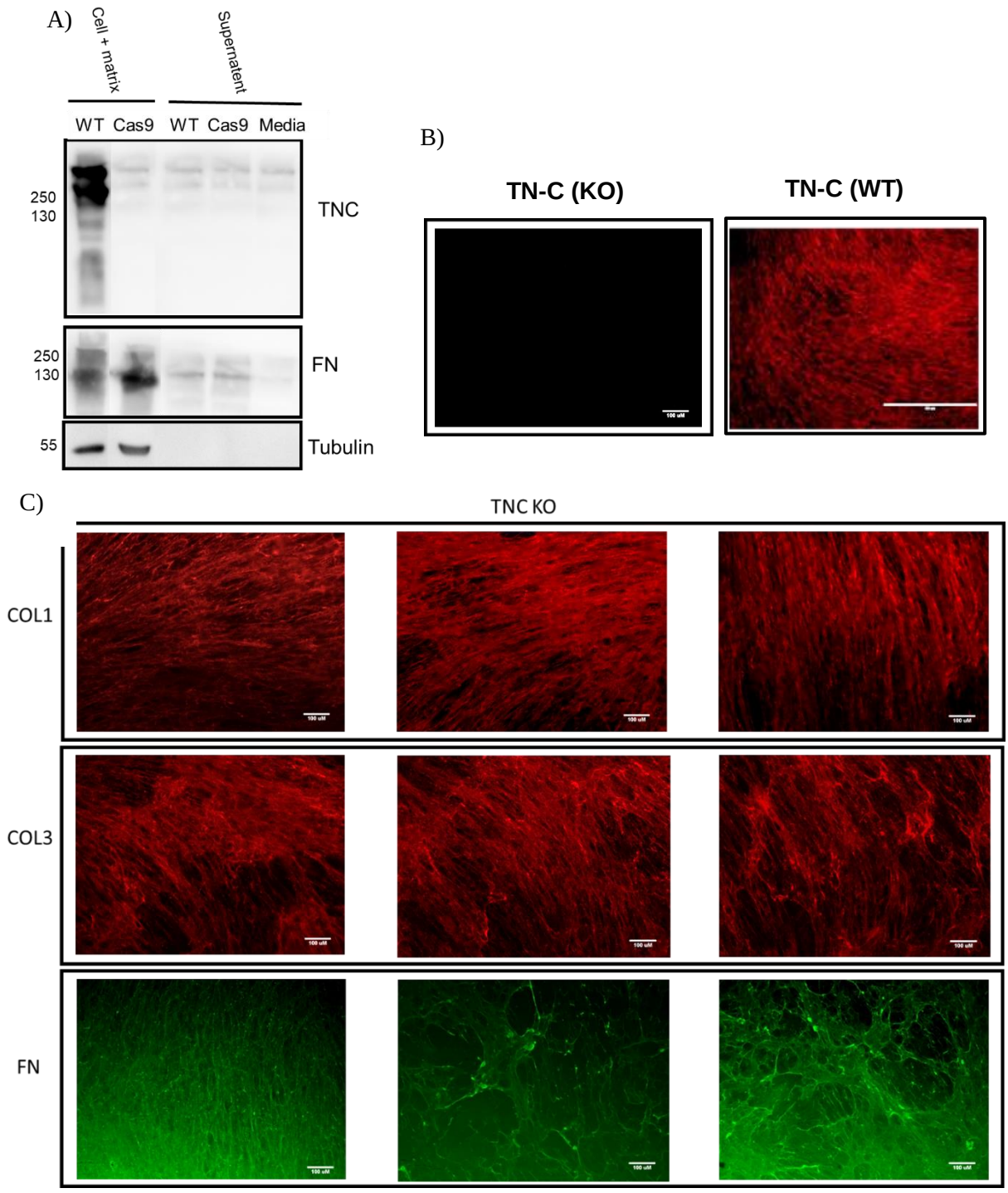


Figure 50. CRISPR/Cas9 genome editing to knock out *TNC* in BJ fibroblasts.
A) Total protein was extracted from WT BJ cells and their CDM (WT) and CRISPR transfected BJ cells and their CDM (Cas9) and their supernatant. After resolving on a 10% SDS-PAGE, TN-C, fibronectin (FN) and tubulin protein levels were assessed by western blotting. A complete media control (Media) was also loaded to assess non-specific antibody binding to FBS components. B) Immunofluorescence analysis of TN-C in WT CDM and *TNC* KO CDM, the latter at PDL 54. C) Immunofluorescence analysis of Collagen-1, collagen-3 and fibronectin in 14-day decellularised *TNC* KO matrices. DAPI staining of gDNA (not shown) confirmed cells had been effectively removed. Scale bar 100 μm.

4.9.1.2 RNA-Seq demonstrated TNC KO does not affect the transcription of ECM molecules

To assess whether depletion of *TNC* in BJ cells impacts on gene expression, RNA-Seq analysis was performed using total RNA isolated from the WT and *TNC* KO clone. The transcriptome of *TNC* KO and WT fibroblasts at PDL 55 was analysed by Novogene, Cambridge. Differentially expressed genes (adjusted P value >0.05 as determined by Novogene) were filtered to remove TCONS transcripts and this list of 123 genes was entered into MatrisomeDB which is a catalogue of more than 300 core ECM proteins and many more ECM-associated proteins [521]. The core matrisome was focused on here and revealed 12 ECM proteins that were significantly differentially expressed (Figure 51). The largest changes were exhibited by laminin subunit beta 4 (*LAMB2;* 4.4 log2Fc, adjusted P value = 0.015), tubulointerstitial nephritis antigen like 1 (*TINGAL1;* -4 log2Fc, adjusted P value = 0.045), insulin like growth factor binding protein 2 (*IGFBP2;* -3.6 log2Fc, adjusted P value = 0.020) and fraser extracellular matrix complex subunit 1 *(FRAS1;* 1.06 log2Fc, adjusted P value = 0.003) in KO fibroblasts compared to WT. The changes in expression of the other genes were small; less than 1 log2Fc or greater than -1 log2Fc. Unfortunately, there was insufficient time to validate these DEGs by RTqPCR, which is planned as future work. This preliminary RNA-Seq analysis supports the results of the immunofluorescence and confirms that genetic ablation of *TNC* does not affect the production and deposition of COL1/3 and FN, and (if we assume transcription changes relate to protein levels) that other ECM molecules are largely unaffected. Therefore, both ECM models, WT and KO, that will be generated by these the WT and *TNC* KO fibroblasts will be very similar.

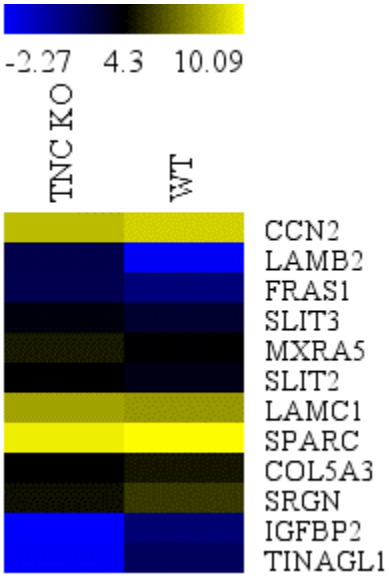


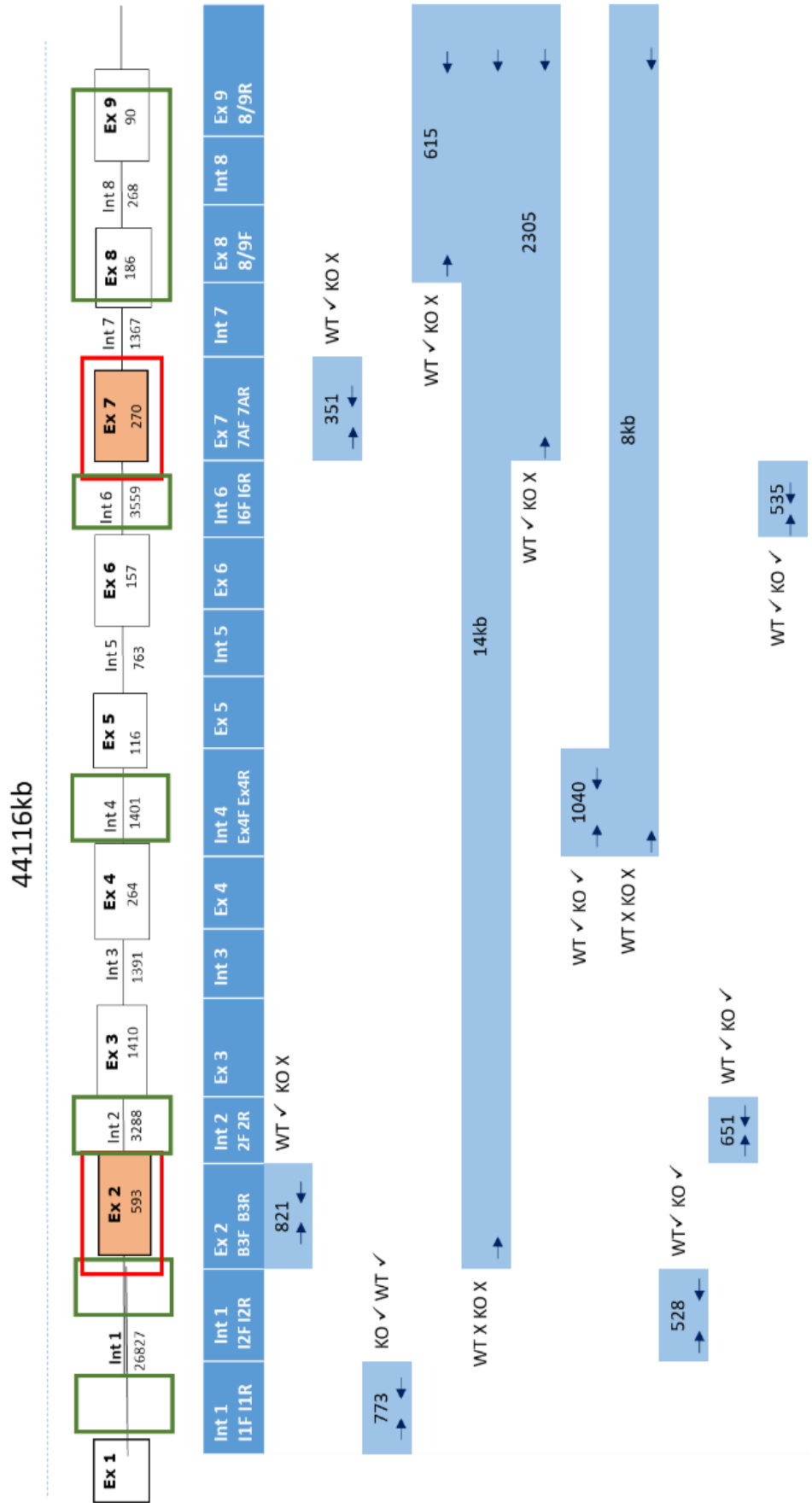
Figure 51. ECM protein transcription is unperturbed by genetic ablation of *TNC*.
The transcriptome of WT and '*TNC*' KO fibroblasts was assessed for changes in core ECM protein gene expression. A list of 123 significantly differentially expressed genes (adjusted P value > 0.05 as determined by Novogene) from WT and *TNC* KO fibroblasts was entered into MatrisomeDB [521] and the resulting hits for ECM proteins within this list displayed as a heatmap; high expression (yellow) or low expression (blue) shown in mRNA abundance on a log2 scale. Heatmaps were made using multi experiment viewer (MeV). Data from three independent experiments. Abbreviations: cellular communication network factor 2 (*CCN2)*, slit guidance ligand 3 (*SLIT3*), matrix remodelling associated 5 (*MXRA5*), slit guidance ligand 3 (*SLIT2)*, laminin subunit gamma 1, secreted protein acidic and cysteine rich (*SPARC*), collagen type V alpha 3 chain (*COL5A3)*, serglycin (*SRGN*).

4.9.2 Genotyping of potential *TNC* KO cell line

Concurrently, whilst generating CDMs to assess the potential *TNC* KO at the protein level, the genotype of this clone was also established. Genotyping of the clonal cell line was challenging due to targeting two sites of the gene simultaneously, thereby inducing two double strand breaks in the one gene. As PCR amplification across the double strand break site of an 800bp fragment in exon 2 and 351bp fragment in exon 7 was not possible using *TNC* knockout gDNA it was next assessed that *TNC* gDNA from the knockout cell line was present and not degraded. To do this, primers were designed that amplified regions throughout the TNC gene both preceding and proceeding both target sites (Figure 52). Regions from intron 1 to exon 9 could be amplified in both gDNA from the knockout and wild type cell line (Figure 52 and Figure 53). Specifically, regions either side of the break sites were indicated as present by successful amplification of a 500bp fragment from intron 1 (primer pair Int1.1Fw and Int1.1Rv); a 651bp fragment in intron 2 (primer pair Int2Fw and Int2Rv); a 535bp fragment in intron 6

(Int6fw and Int6Rv) and a 595bp fragment spanning exon 8 and 9 (primer pair Ex8 Fw and Ex8 Rv) (Figure 52 and Figure 53B).

As amplification either side of the break sites was possible, yet not over the break sites, it was hypothesized that a large chromosomal rearrangement had occurred, in which the entire fragment between the two break sites had moved to a different chromosome (translocation) or the fragment had been inverted (transversion). The latter was possible to identify by utilising pairs of primers that face that same orientation, with one primer binding site preceding the break site, and one primer binding site proceeding the break site. In wild type gDNA no amplification of product is expected as both primers will bind the same DNA strand. In contrast, in gDNA with an inversion, the primer bound on the DNA preceding the break site will be oppositely oriented to the primer bound to the inverted DNA fragment, thus leading to amplification (Figure 53 C, primers identified in methods table 8).


44116kb
Ex 1
Int 1
26827
Ex 2
593
Int 2
3288
Ex 3
1410
Int 3
1391
Ex 4
264
Int 4
1401
Ex 5
116
Int 5
763
Ex 6
157
Int 6
3559
Ex 7
270
Int 7
1367
Ex 8
186
Int 8
268
Ex 9
90
Int 1 I1F I1R
Ex 2 B3F B3R
Int 2 2F 2R
Ex 3
Int 3
Ex 4
Int 4 Ex4F Ex4R
Ex 5
Int 5
Ex 6
Int 6 I6F I6R
Ex 7 7AF 7AR
Int 7
Ex 8 8/9F
Int 8
Ex 9 8/9R
773
KO ✓ WT ✓
821
WT ✓ KO X
351
WT ✓ KO X
615
WT ✓ KO X
14kb
WT X KO X
2305
WT ✓ KO X
1040
WT ✓ KO ✓
8kb
WT X KO X
528
WT✓ KO ✓
651
WT ✓ KO ✓
535
WT ✓ KO ✓

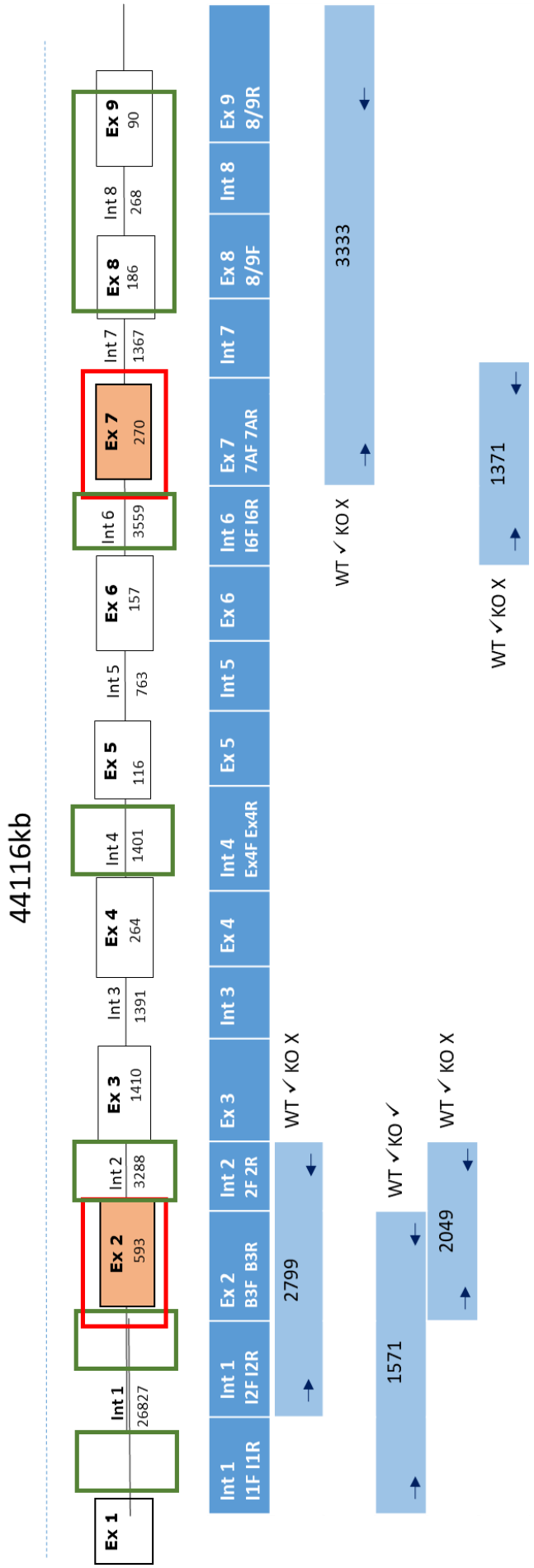


Figure 52. Schematic of the *TNC* gene with detailed introns (black line) and exons (black boxes) sizes.
PCRs were carried out to genotype CRISPR/Cas9 targeted *TNC* gDNA. Cas9 targeted exons 2 and 7, filled in orange whilst primers corresponding to each exon are labelled in line with schematic. Amplicon sizes detailed in the table. gDNA amplification was possible throughout the *TNC* gene (green boxes) but not across the Cas9 targeted exons (red boxes).

To test this hypothesis pairs of forward primers (F1 and F2) were utilised in a PCR (Figure 53 C) which successfully generated a 12Kb fragment from *TNC* KO gDNA, but not wild type gDNA. Correspondingly, 'primer walking' using forward primers (in reverse orientation on the inverted DNA) closer to the break site successfully generated smaller 10Kb (primer set F1 and F3) and 3Kb fragment (Primer set F1 and F4).

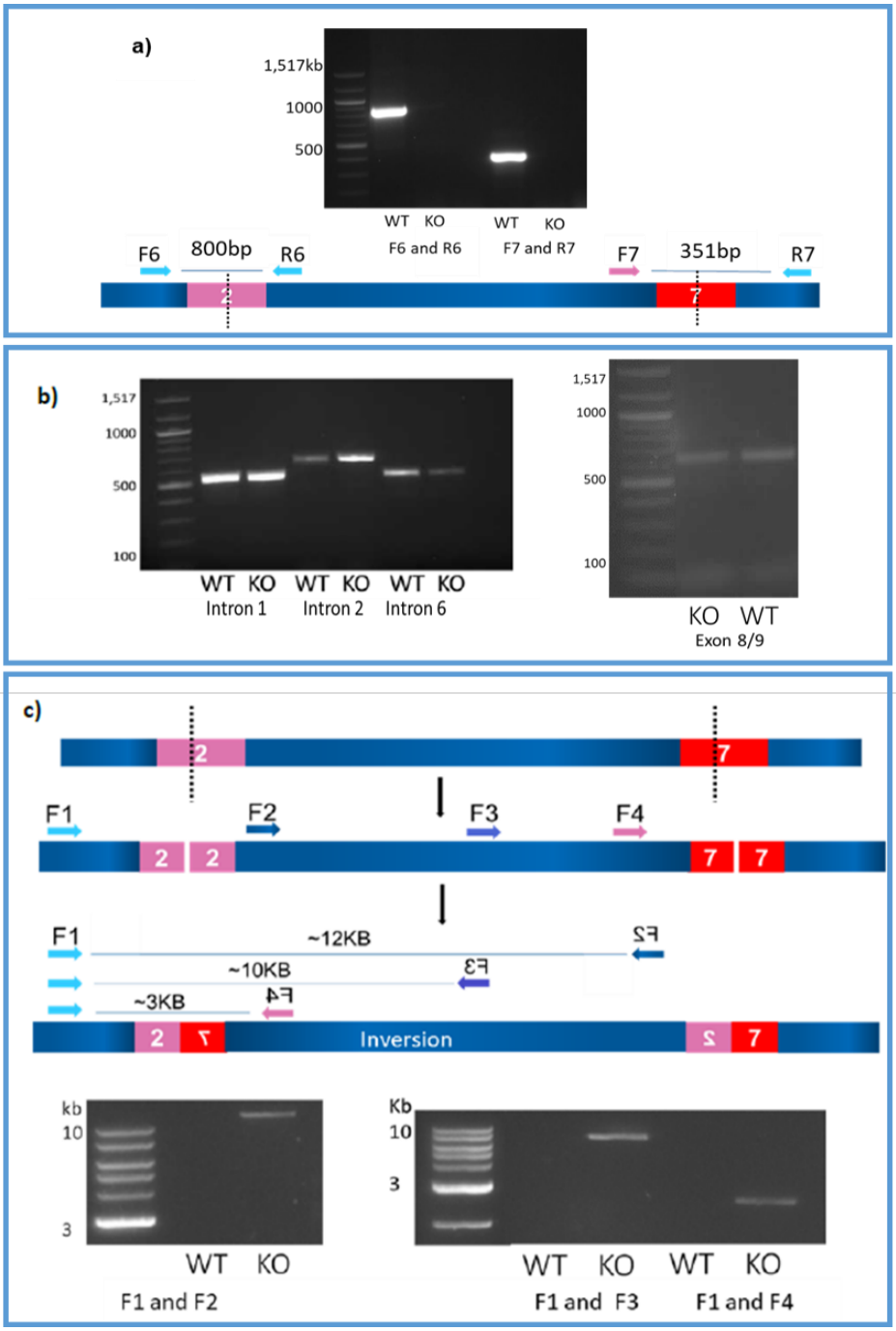


Figure 53. Genotyping of the gDNA from wildtype BJ cells (WT) or BJ cells transfected with CRISPR/Cas9 plasmids targeting exon 2 and 7 (KO) indicated a large rearrangement had occurred in the KO cell line. **A)** PCRs using primers spanning the CRISPR target sites in exon 2 and exon 7 were successful in amplifying WT DNA (F6 and R6; 800bp and F7 and R7; 351bp), however failed in KO gDNA. **B)** PCRs amplifying *TNC* regions not targeted by CRISPR/Cas9 guides, including intron 1, intron 2 and intron 6 were successful in both WT and KO gDNA. **C)** Due to the presence of *TNC* DNA, yet failure of PCRs spanning the CRISPR-targeted break site and the knowledge that CRISPR/Cas9 cleavage is very efficient, it was hypothesised that an inversion had occurred. Pairs of forward primers in the inventory were utilised to investigate this. Forward primer pairs targeting regions preceding the cut site in exon 1 (F1), and downstream of the break site to bind either intron 2 (F2; 12kb product), intron 4 (F3; 10kb product) or intron 6 (F4;3kb product) indicate a large inversion has occurred, possibly spanning the break site from exon 2 to exon 7. No products were seen in WT gDNA using these forward primer pairs indicating genuine binding of primers to inverted KO DNA. 25 µl of 25 µl reaction was loaded on 1% or 0.6% agarose gel.

A further PCR with a primer even closer to the break site produced a fragment short enough to sequence. This confirmed an inversion had occurred, with the 5' end of exon 2 joined with the 5' end of exon 7 (Figure 54 A). This approach was utilised to genotype the exon 7 Cas9 target site; this time utilising pairs of reverse primers. A fragment was successfully amplified from TNC KO gDNA, but not wild type gDNA (Figure 54 A). Sequencing revealed exons 2 and 7 were joined at both 5' and 3' break sites (Figure 54 B, C). Together, this confirms specific targeting of exons 2 and 7 by CRISPR/Cas9, generating a 12Kb fragment that was then inverted and re-joined into original break sites in the *TNC* gene. This large genetic rearrangement resulted in a complete KO of *TNC* production.

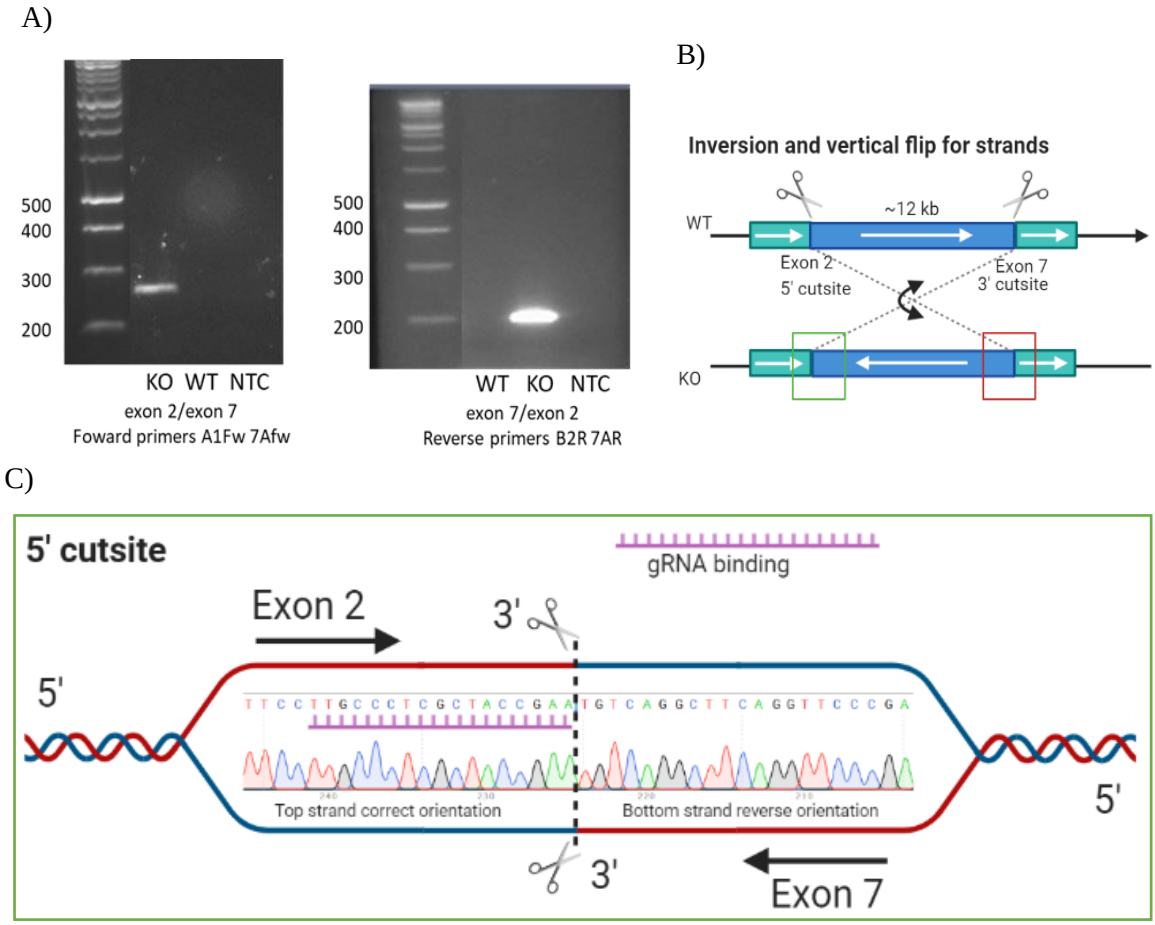


Figure 54. Genotyping of gDNA from wildtype BJ cells (WT) or *TNC*-KO BJ cells (KO) indicates a CRISPR/Cas9- induced ~12 Kb inversion in the KO BJ fibroblasts.
A) Inversion PCRs of WT or KO gDNA using pairs of forward primers (left) or reverse primers (right) spanning the CRISPR target sites in exon 2 and exon 7 of *TNC*. 25ul was loaded on a 1.5% agarose gel. **B)** Schematic demonstrating the genetic rearrangement in *TNC* KO gDNA resulting from Cas9 cleavage at exons 2 (5 ‘ cut site green) and 7 (3’ cut site red). Of note is that the top and bottom strands have swapped as well as an inversion of sequence. **C)** Sequencing of band extracted CRISPR/Cas9 targeted *TNC* gDNA (from A) demonstrating an inversion of a fragment spanning 12Kb from the 5’ cut site in exon 2 (green) to the 3’ cut site in exon 7 (red). Fragment ends correspond to Cas9 cleavage at the target site, 3 nucleotides upstream of the PAM sequence (remaining guide binding region highlighted by pink oligo).

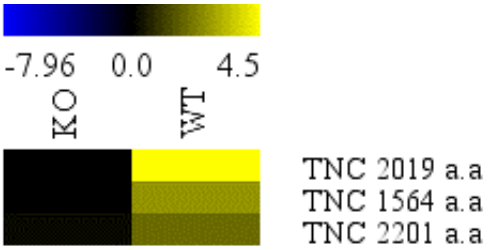
4.9.3 Genotyping at the RNA level revealed no functional transcripts were produced from *TNC* KO fibroblasts.

Additionally, RNA-Seq of WT and *TNC* KO fibroblasts also demonstrated a loss of *TNC* transcripts. A TCONS transcript was produced from *TNC* KO fibroblasts, but not in wild type, however this transcript does not produce a protein (determined using Ensemble), and TCONS transcripts were removed from the analysis, as detailed further in chapter 6. In wild type fibroblasts three transcripts were detected corresponding to the full length isoform ENST00000350763 (2201 amino acids; a.a), a

slightly smaller isoform ENST00000341037 (2019 a.a) and 1 small isoform ENST00000537320 (1564 a.a). The expression of these transcripts was nearly '0' in TNC KO fibroblasts with a log2Fc change of -12.5, and -3.4, -5.7, respectively (Figure 55 A). This supports a previous study which assessed the expression of TNC isoforms from fibroblasts by western blot, however only the largest 2201 a.a isoform and 1546 a.a isoform was detected, with the predominant form the large isoform [522]. This study also demonstrated that the full length 2201 a.a isoform was more predominant than the smaller 1546 a.a isoform. Here, the expression of the 2201 a.a and 2019 a.a isoform transcripts are greater than the small isoform, and the 2019 a.a transcript had the highest expression. However, transcripts may not be equally translated and it could be that in the previous study the resolution of the gel did not allow discrimination between the two large isoforms.

Furthermore, mapping of the *TNC* reads to the reference genome using IGV (Figure 55 B) indicated that *TNC* transcripts starting from exon 2 to exon 7 (the two loci where the Cas9 cleavage occurred) were absent in the *TNC* KO fibroblasts compared to the WT fibroblasts. Exon 1 was present in all samples. WT fibroblasts displayed good coverage and reads throughout the *TNC* gene. However, reads mapped to the 3' region of the gene from exon 7 onwards demonstrated stark changes between WT and *TNC* KO samples; in the WT fibroblasts the reads profile matched at each exon to the 5' end of the gene, as expected, and each WT sample had a similar transcript profile. In the *TNC* KO fibroblasts, reads were present from exon 7 onwards, however these were present throughout introns and exons and did not match the transcript profile of the WT fibroblasts. A sashimi plot of splice junctions in a representative sample from the *TNC* KO and WT fibroblasts further indicated severe gene disruption resulting from CRISPR/Cas9 (Figure 56). In the *TNC* KO fibroblasts exon 1 and exon 8 are joined excluding exons 2-7, compared to WT fibroblasts in which exon junctions are clearly aligned with the expected transcript and the inclusions of exons 1-8 occur sequentially. This indicates the CRISPR/Cas9 inversion resulted in aberrant splicing Together, this supports the results from Sanger sequencing and western blotting, and confirms that the *TNC* KO fibroblasts are indeed knockouts.

A)



B)

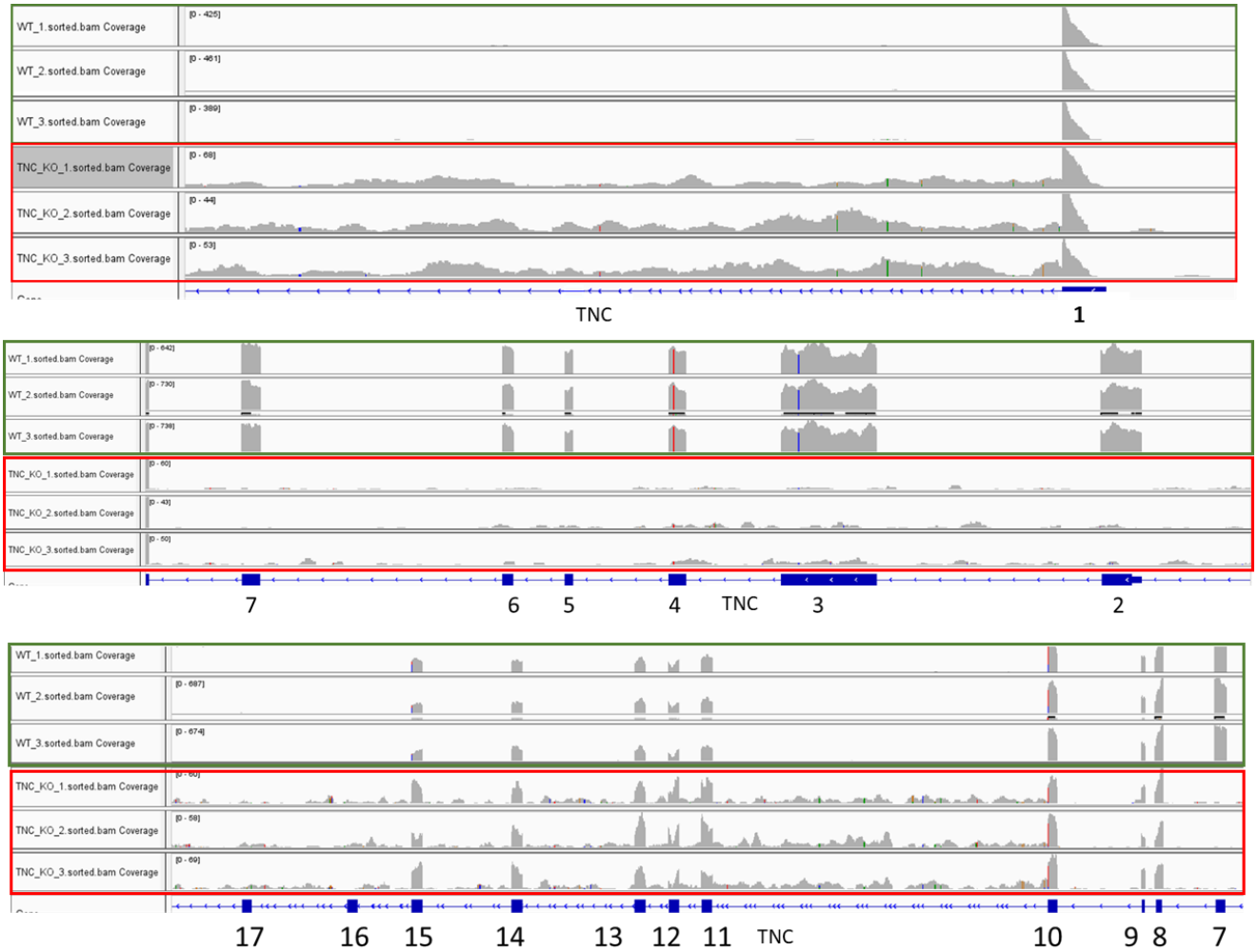


Figure 55. RNA-Seq and read mapping demonstrates a complete KO of *TNC* in the *TNC* KO fibroblasts compared to WT fibroblasts.
A) Levels of *TNC* protein coding transcripts were zero in the *TNC* KO fibroblasts relative to WT fibroblasts, indicating complete KO of *TNC* by CRISPR/Cas9 gene editing. Three isoforms of *TNC* were detected, these are denoted by the size of the protein produced from these transcripts in amino acids (a.a). The 2201 a.a. isoform is the full length protein. **B)** RNA-Seq .BAM files are indexed, sorted and then the reads are mapped across the reference human genome GRCH38. Transcripts indicated an absence of transcripts spanning from the two CRISPR target regions in exons 2 and 7 (exons indicated by numbers) and subsequent abnormal transcript alignment in the *TNC* KO fibroblasts (red box) compared to WT fibroblasts (green box) to the end of the gene (representative image shown from exon 7-17). Data from three independent experiments.

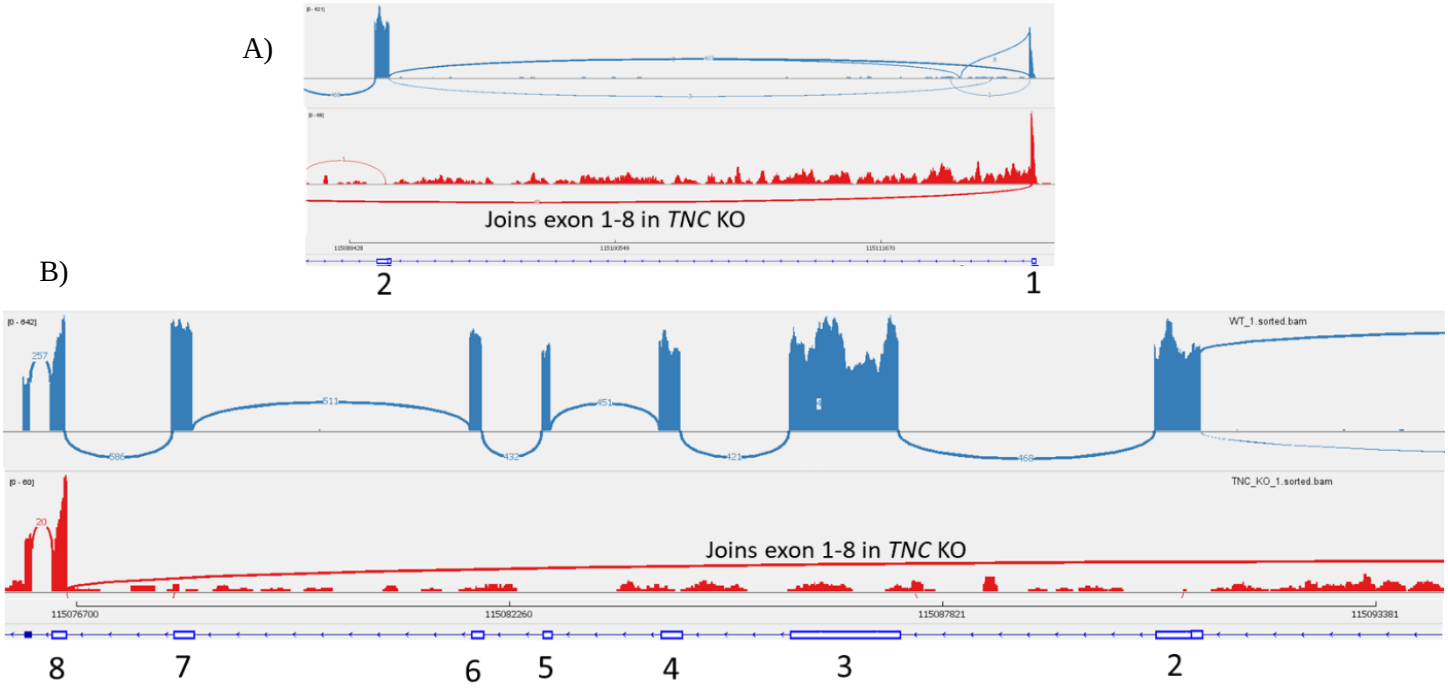


Figure 56. Sashimi plots from one representative WT fibroblast and *TNC* KO fibroblast sample indicate aberrant exon splicing in the *TNC* KO fibroblasts that would likely produce a non-functional protein. Due to the very large intron (26827 bp) separating exon 1 and exon 2 it was not possible to plot one sashimi plot that covered from exon 1 to exon 8, hence two plots are shown for the WT fibroblast (blue) and *TNC* KO fibroblasts (red) with exon 1-2 in panel A) and exon 2-8 in panel B). This resulted in the splicing line 'joins exon 1-8 in *TNC* KO fibroblasts' automatically changing position however in both cases indicates that exon 1 is joined with exon 8. The expected transcript is shown at the bottom of the plot and exons indicated with numbers.

4.10 Generation of *TNC* knockout fibroblasts using CRISPR/Cas9 RNP (Approach 3)

Due to the lengthy and inefficient process to generate clonal cell populations using the CRISPR/Cas9 plasmid method, BJ fibroblasts were nucleofected with CRISPR/Cas9 RNP with guides targeting exon 2 (A2) or 7 (7A), or both exons (++) simultaneously. Note the exon 7 guide is different to the guide originally chosen for the CRISPR/Cas9 plasmid method as analysis of both the 7A and 7B guides with new software from the reagent supplier indicated the 7A guide would be slightly more efficient (Integrated Technologies Alt-R CRISPR custom guide RNAs design tool). Nucleofected cells were expanded on uncoated plates, without any sorting, and allowed to deposit matrix in 10cm dishes for 14 days, as previously. Western blotting of crude cell and matrix extract confirmed the three cell

populations had reduced levels of the TN-C protein, whilst fibronectin and tubulin were not perturbed (Figure 57 A). Densitometric analysis of TN-C relative to WT fibroblasts and normalised to tubulin revealed a reduction of approximately 79% in the 7A and ++ cell populations and 66% in the A2 cell population (Figure 57 B). Together these results demonstrate the generation of three cell populations in which TN-C production is markedly reduced compared to WT BJ fibroblasts.

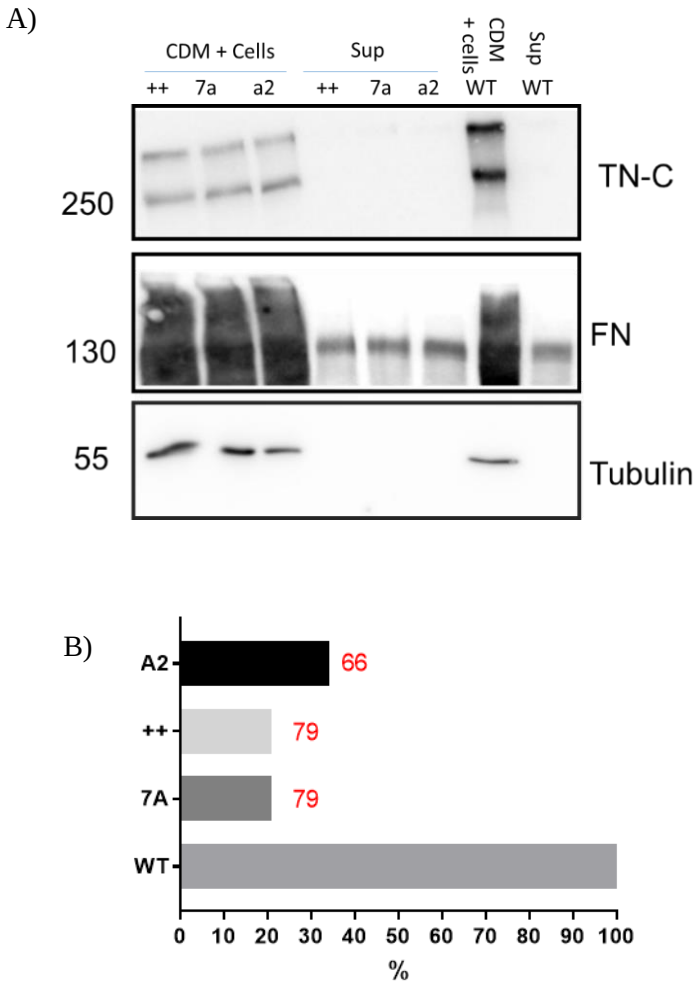


Figure 57. Fibroblasts transfected with CRISPR/Cas9 RNP exhibit a decreased production of *TNC* at the protein level.
A) Total protein and supernatant (sup) was extracted and collected from WT BJ cells and their CDM (WT) and BJ cells nucleofected with CRISPR RNP 7A, A2 or 7A and A2 (++) and their CDM. After resolving on a 6% SDS-PAGE gel, TN-C, fibronectin (FN) and tubulin protein levels were assessed by western blotting. **B)** Densitometric analysis using Bio-Rad Image lab software to determine TN-C protein production (%) compared to WT TN-C protein production and normalised against tubulin loading. The decrease of TN-C production is indicated (red lettering %).

4.11 OFF -TARGET ANALYSIS OF CRISPR/CAS9 AND CRISPR/CAS9 RNP EDITED CELLS

4.11.1 Off-target prediction

As off-target effects are a major criticism of CRISPR/Cas9, primers were designed (Table 12, materials and methods) to amplify predicted off-target regions of the genes that were identified utilising the Benchling software (Table 23). A prerequisite for Cas9 activity and cleavage is the presence of the species-specific 'NGG' PAM sequence, however Cas9 is also capable, with less specificity, of cleaving targets with an 'NAG' PAM [403, 523]. Predicted off-targets with these PAMs were therefore included. No-off targets were predicted for approach 1, Cas9n, due to the paired format and therefore high specificity of this method. The predicted off-target regions with 'NGG' PAMs were Von Willebrand Factor A Domain Containing 1 (*VWA1*) for guide '7B' and, Calcium modulating ligand (*CAMLG*) and Adenylate cyclase 4 (*ADCY4*) for guide 'A2'. Predicted off-targets harbouring the 'NAG' PAM sequence for guide '7B' were adenosine deaminase (*ADA*), GEM Interacting Protein *(GMIP),* SKI Family Transcriptional Corepressor 1 *(SKOR1), SMPDL38,* 5-phosphohydroxy-L-lysine phospho-lyase *(PHYKPL),* Bromodomain Adjacent To Zinc Finger Domain 1A *(BAZ1A),* Testis Expressed 2 *(TEX2)* and Transmembrane Protein 204 *(TMEM204)*. The predicted 'NAG' off- targets for guide 'A2' were transmembrane protein 42 (*TMEM42),* Kynurenine-oxoglutarate transaminase 1 (*CCLB1),* WD Repeat Domain 90 *(WDR90),* Adaptor Related Protein Complex 3 *(AP3D1)* and Ciliogenesis Associated TTC17 Interacting Protein *(C2orf62)*. These guides were utilised to generate the CRISPR/Cas9 *TNC*-KO cell line (approach 2). As the 'A2' guide sequence was identical to generate the populations of cells utilising CRISPR/Cas9 RNP (approach 3), the predicted 'NGG' and 'NAG' off-targets were identical. However, guide '7A' was utilised in approach 3 and generated one predicted 'NGG' off-target in serine/threonine-protein phosphatase 4 catalytic subunit C (*PPP4C*). The 'NAG' predicted-off targets for '7A' were tyrosine-protein phosphatase non-receptor type 14 (*PTPN14), RP5-1014d13,* nuclear receptor coactivator 4 *(NCOA4)* and RP11-557J10.0 a lncRNA. Off-targets were not predicted in genes coding for ECM or ECM-associated proteins with the exception *VWA1*, a blood glycoprotein. Generally

the scores for the predicted off-targets, in particular the 'NAG' off-targets are very low; >1, compared to the score of *TNC* at 100.

Table 23. Predicted off-targets from the Benchling software for all CRISPR approaches.
Guides and their specific chromosomal binding region in the off-target genes are indicated. Nucleotides that differ from the guide sequence shown are shown in red. Predicted off-targets with a 'NGG' PAM sequence are highlighted in blue, those with an 'NAG' are highlighted in green.

Guide name	Off-target gene	Sequence 5'- 3'	Locus	Score /100
Exon 2				
A2 (Approach 2 and 3†)	*ADCY4*	CTCCCCTAGCTCCCGAAGGT	chr14:+24329445	0.82
	CAMLG	TGGCCGTCGCTACCGACGGC	chr5:+134738648	0.22
	TMEM42	TCGACCTCGCTGCCGAAGGC	Chr3:-44862105	0.45
	CCBL1	TTGCCGGCGCTGCCGATGGT	Chr9: 1.29E +08	0.13
	WDR90	TTGCCTTCTCCACCGATGGT	Chr16:666124	0,13
	AP3D1	TCGCCATCGCCACAGAAGGT	Chr19:-2116208	0.09
	C2orf62	TTGCCCTCGGTGTCCAAGGT	Chr2:-2.18E+08	0.07
Exon 7				
7A (Approach 3†)	*PPP4C*	TGGCAGGTCATTGATGCCAC	chr16:+30083083	0.14
	PTPN14	TGGTGCGTCTTTGAGGTCGT	chr1:-2.14E+08	0.19
	RP5-1014d13	TGTCAGCTCTTTGATGCTGT	Chr22: 37881832	0.14
	NCOA4	TGGAATGTCTTAGAAGCCGT	Chr10:-46015213	0.13
	RP11-557J10.3	TGGAACGTCTTAGAAGCCAT	Chr4:-1.46E+08	0.08
7B (Approach 2†)	*VWA1*	GCCACGTCGTAGTCCGTGTC	chr1:-1439278	0.14
	ADA	GTCATCTGGTAATCAGTGTC	Chr20:44621078	0.8
	GMIP	CCCACCCCGTACACAGTGTC	Chr19:-19630336	0.46
	SKOR1	GTCACCTCAAACTCAGTGAC	Chr15: 67826585	0.35
	SMPDL38	GATACCTTGTAGTCAGGGTC	Chr1:-27945267	0.34
	PHYKPL	GAGACCTCTTGCTCAGGGTC	Chr5:1.78E+08	0.23
	BAZ1A	AACACCTGTTACTCACTGTC	Chr14:-34758611	0.19
	TEX2	TCCACCTCGTCCTCACTGTC	Chr17;64212847	0.19
	TMEM204	GACTCCACGTAGTCATTGTC	Chr16:-1555000	0.1

† Approach 1: CRISPR/Cas9n, Approach 2: CRISPR/Cas9, Approach 3: CRISPR/Cas9 RNP

4.12 EXPERIMENTAL VALIDATION OF PREDICTED OFF-TARGETS

4.12.1.1 PCR amplification of predicted off- target region indicated no off-target editing had occurred.

PCR was carried out to assess if any guides had bound and induced Cas9 cleavage at an off-target site. To identify if guide 'A2' had bound and resulted in Cas9 cleavage at a predicted off-target region in genes *CAMLG* and *ADCY4*, PCR amplification of gDNA extracted from wild type BJ fibroblasts and *TNC* KO BJ fibroblasts was carried out utilising primer pairs *ADCY4* Fw and Rv and *CAMLG* Fw and Rv (Figure 58 A). These primers were expected to generate a product size of 315bp and 312bp, respectively. Following PCR amplification and gel DNA extraction and purification, the resulting products were sequenced using the reverse primers (Figure 58 B). Target regions from both genes were successfully amplified from both WT and KO gDNA and sequencing and multiple alignments using Clustal Omega revealed no change from WT sequence. This indicates no 'A2' guide off-target effects. One point to note is that sequencing was always carried out with the reverse primer, however sequences were reverse complemented, where necessary, so that sequence matched identically with the predicted off-target sequence in 5' – 3' orientation in table 15 to make it explicitly clear where the predicted guide binding site occurred.

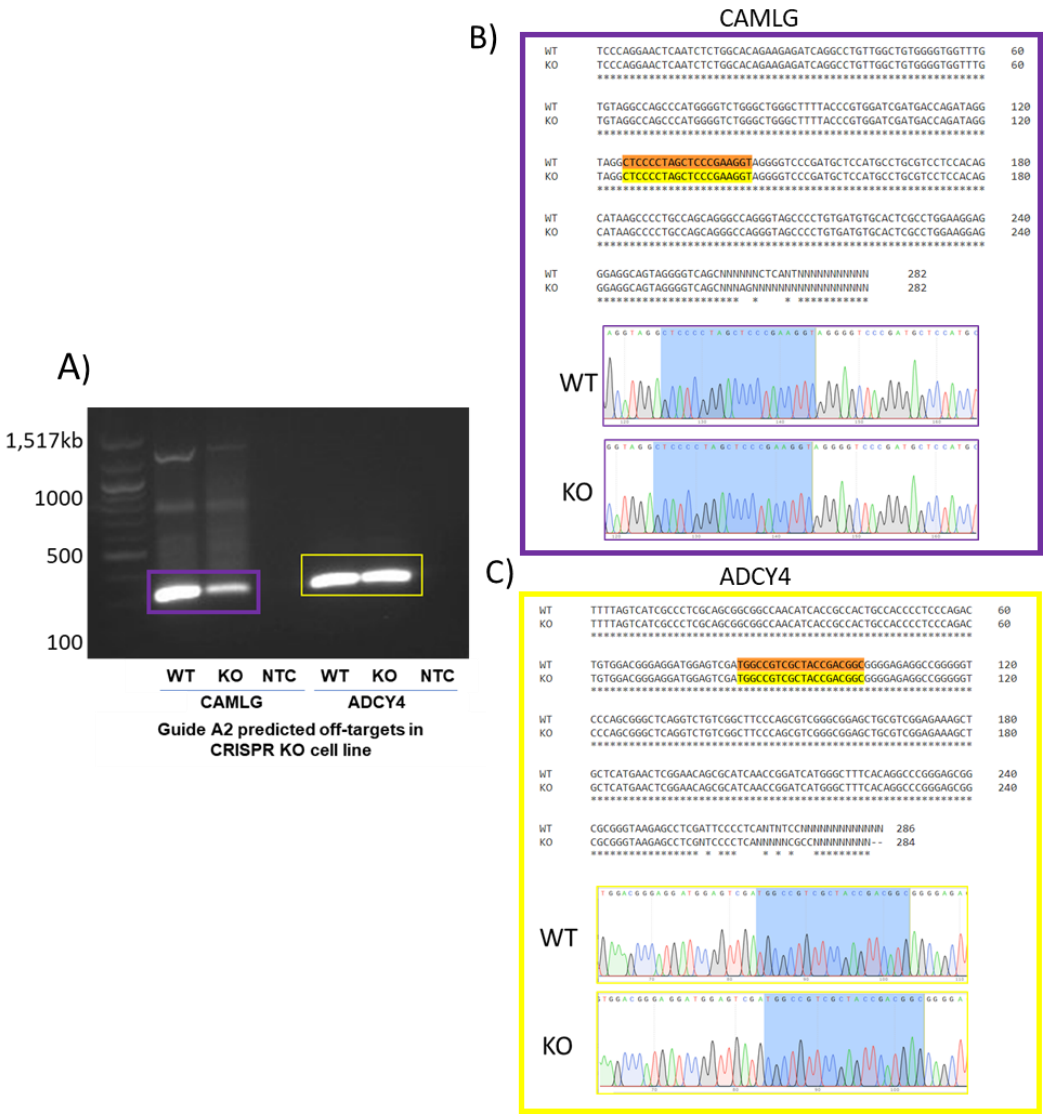


Figure 58. Analysis of predicted-off target regions resulting from guide 'A2' in CRISPR/Cas9 *TNC* KO fibroblasts.
A) PCRs from gDNA extracted from WT and KO BJ fibroblasts were carried out utilising primers to amplify a region spanning the predicted-off target regions ((https://benchling.com) in *ADCY4* and *CAMLG*. 25µl of PCR produce was ran on a 1% agarose gel. **B)** and **C)** Multiple alignment of the band extracted and sequenced WT and KO PCR products. Asterix's indicate identical sequences whilst the predicted-off target sequence region is highlighted in orange (WT) and yellow (KO). Electropherograms for each product are shown with the predicted off-target region identified (blue).

Next, it was assessed whether guide '7B' had induced any mutations at the predicted off-target site in the *VWA1* gene by PCR of gDNA from WT and *TNC* KO BJ fibroblasts (Figure 59). Primers *VWA1* Fw and Rv were used. Multiple bands were seen from KO gDNA including a product at the expected size of 495bp (light blue box). However, amplification of WT gDNA was unsuccessful after repeated attempts (Figure 59 A). Sequencing of the largest and second largest product (blue boxes) revealed products were identical, with the latter being a shorter fragment. The shortest product (red box) was

mixed sequence. Multiple alignment of the largest fragment with the WT *VWA1* sequence generated from an *In-Silico* PCR (https://genome.ucsc.edu/cgi-bin/hgPcr) revealed no difference in sequence from WT (Figure 59 B). This indicated that guide '7B' had no off-target effects. Taken together, these findings indicated no off-target effects had occurred in the *TNC* KO cell line at the predicted sites.

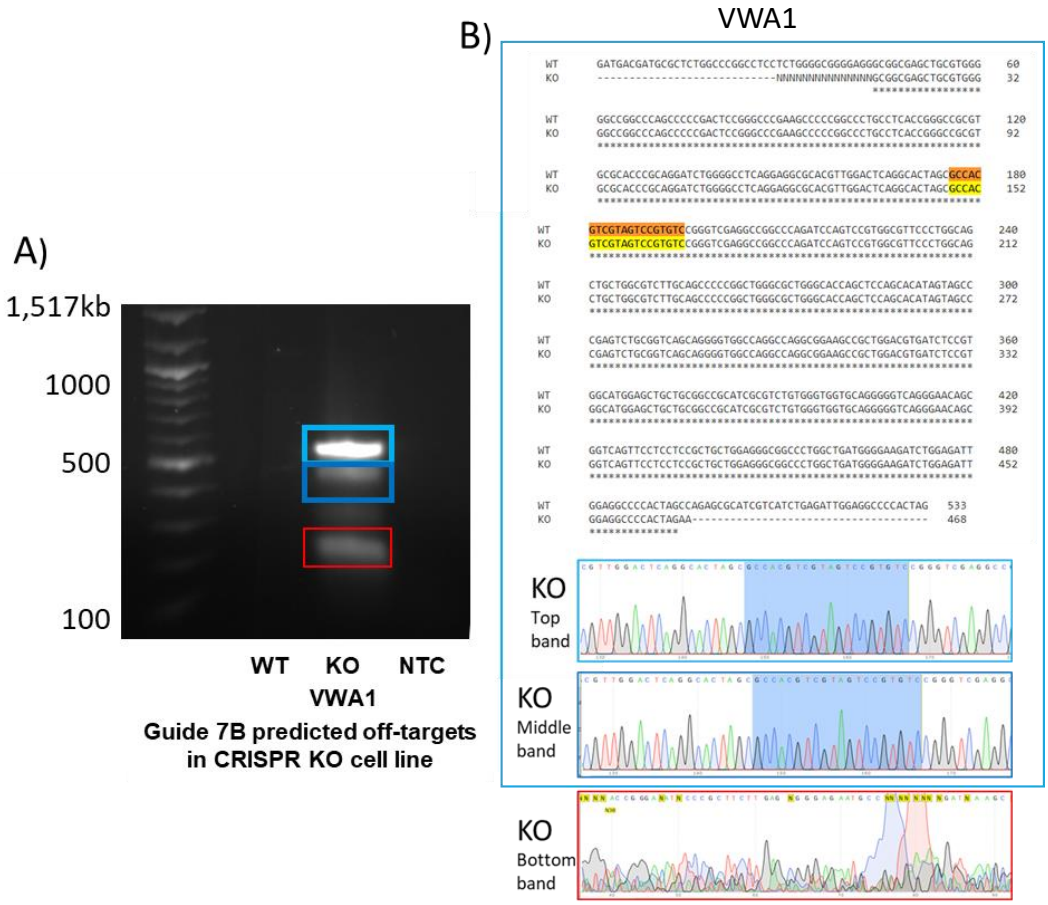


Figure 59. Analysis of predicted off-target regions resulting from guide '7B' in CRISPR/Cas9 *TNC* KO BJ fibroblasts established no off target editing had occurred.
A) PCRs from gDNA extracted from WT and KO BJ fibroblasts were carried out utilising primers to amplify a region spanning the predicted-off target regions (https://benchling.com) in *PPPC4*. 25µl of product was loaded on a 1% agarose gel. **B)** Multiple alignment of the band extracted and sequenced KO PCR products. WT sequence for the multiple alignment was generated using In-Silico PCR (https://genome.ucsc.edu/cgi-bin/hgPcr). Asterix's indicate identical sequences whilst the predicted-off target sequence region is highlighted in orange (WT) and yellow (KO). Electropherograms for each product are shown with the predicted off-target region identified (blue).

The same procedure was utilised to assess whether off-target effects had occurred in the three cell populations, termed 'A2', '7A' and '7A and A2 (++)' derived from the guides they were nucleofected

with, utilising CRISPR/Cas9 RNP. First, it was assessed whether guide 'A2' had induced any mutation in *CAMLG* and *ADCY4* utilising primer pairs as previously. Products at the expected size, for both gene regions were produced from gDNA from WT (Cas9 only), 'A2' and '++' cells (Figure 60 A). Sequencing of CRISPR/Cas9 RNP products revealed they were identical to WT, indicating no off-target effects had occurred (Figure 60 B). Next, it was assessed whether there were any off-target effects resulting from guide '7A'. gDNA was amplified from WT, '7A' or '++' fibroblasts utilising primers *PPP4C* Fw and Rv to generate an expected product of 316bp. Products were amplified successfully from all gDNA (Figure 61 A). Band extraction and sequencing, and multiple alignments of the sequences demonstrated identical sequence to that of WT, in '7A' and '++' cell populations, indicating no off-target effects have occurred (Figure 61 B).

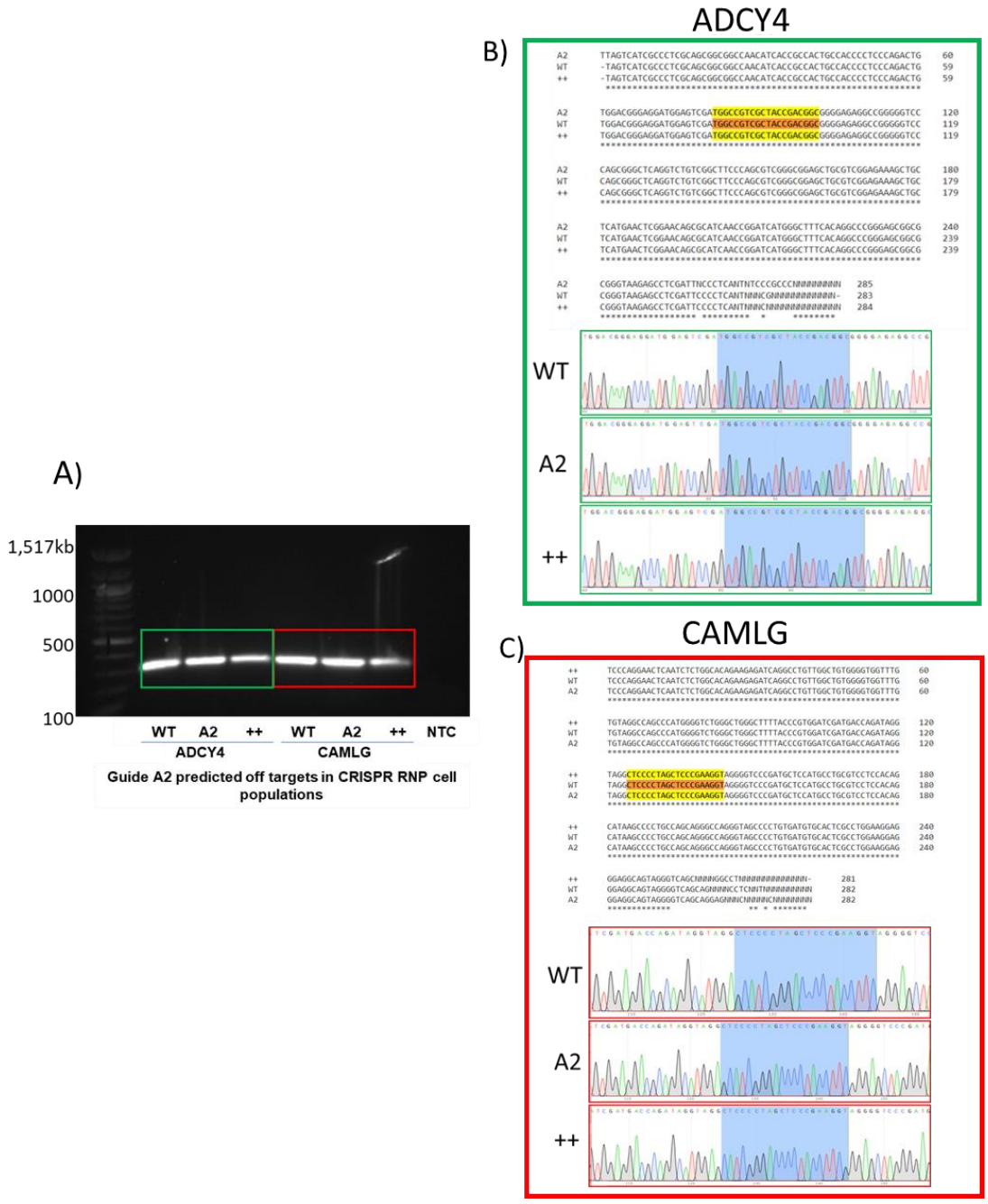


Figure 60. Analysis of predicted-off target regions resulting from guide 'A2' in CRISPR/Cas9 *TNC* RNP BJ fibroblasts established no off-target editing had occurred.
A) PCRs from gDNA extracted from WT and CRISPR/Cas9 RNP nucleofected fibroblasts were carried out utilising primers to amplify a region spanning the predicted-off target regions (https://benchling.com) in *ADCY4* and *CAMLG*. 25 µl of PCR product was loaded on a 1% gel. **B)** and **C)** Multiple alignment of the band extracted and sequenced WT and CRISPR/Cas9 RNP PCR products. Asterix's indicate identical sequences whilst the predicted-off target sequence region is highlighted in orange (WT) and yellow (KO). Electropherograms for each product are shown with the predicted off-target region identified (blue).

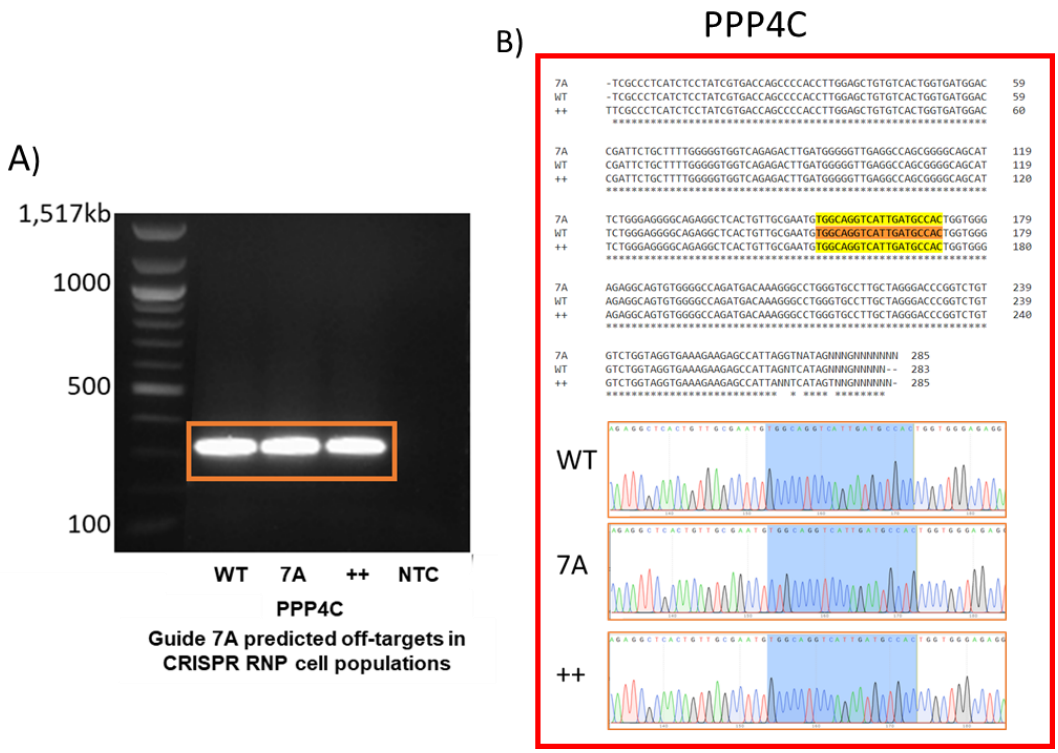


Figure 61. Analysis of predicted-off target regions resulting from guide '7A' in CRISPR/Cas9 *TNC* RNP BJ fibroblasts established no off-target editing had occurred.
A) PCRs from gDNA extracted from WT and CRISPR/Cas9 RNP nucleofected BJ fibroblasts were carried out, utilising primers to amplify a region spanning the predicted-off target regions (https://benchling.com) in *PPP4C*. 25µl of product was loaded on a 1% agarose gel. **B)** Multiple alignment of the band extracted and sequenced WT and CRISPR/Cas9 RNP PCR products. Asterix's indicate identical sequences whilst the predicted-off target sequence region is highlighted in orange (WT) and yellow (KO). Electropherograms for each product are shown with the predicted off-target region identified (blue).

Given that Cas9 is capable of binding 'NAG' sites, albeit with less efficiency than 'NGG' sites [403, 523], the highest scoring 'NAG' predicted off-target site, per guide, was screened by PCR. Successful amplification of all three genes occurred; a region of the *ADA* gene was amplified with primers *ADA* Fw and Rv, a region of *TMEN42* was amplified with *TMEN42* Fw and Rv and a genomic region of *PTPN14* was amplified with primers PTPN4 Fw and Rv (Figure 62 A, C, E). Sequencing confirmed amplicons from WT, *TNC* KO, A2, 7A, 7A and A2 (++) gDNA were all at the expected sizes of 495bp, 304 bp and 479 bp for *ADA*, *TMEN42* and *PTPN14*, respectively (Figure 62 B, D, F). A shorter product in *ADA* was also present from both gDNA samples that was identical to the larger intended product. A shorter product could also be seen in the 7A and A2 lane for the *PTPN14* gene, however this could not be sequenced due to the small amount of PCR product recovered from the band extraction. Given that this faint band was not present in the WT lane, it suggests this may be a different product, however it

could be a shorter yet identical product to the larger predominant band. Given the low score for this guide and this off-target area, an off-target edit is highly unlikely to have occurred here. The poor quality and large size of the bands in the *TMEN42* gel resulted from technical issues with the LAS imager, however under a UV light box the bands were unsmeared and narrow, thus could be excised accurately. The fainter bands at the bottom were not excised due to the faintness under the UV light box. As the faint bands were identical in all lanes, this suggested a short but identical product to the larger band, as previously demonstrated in the *ADA* gel.

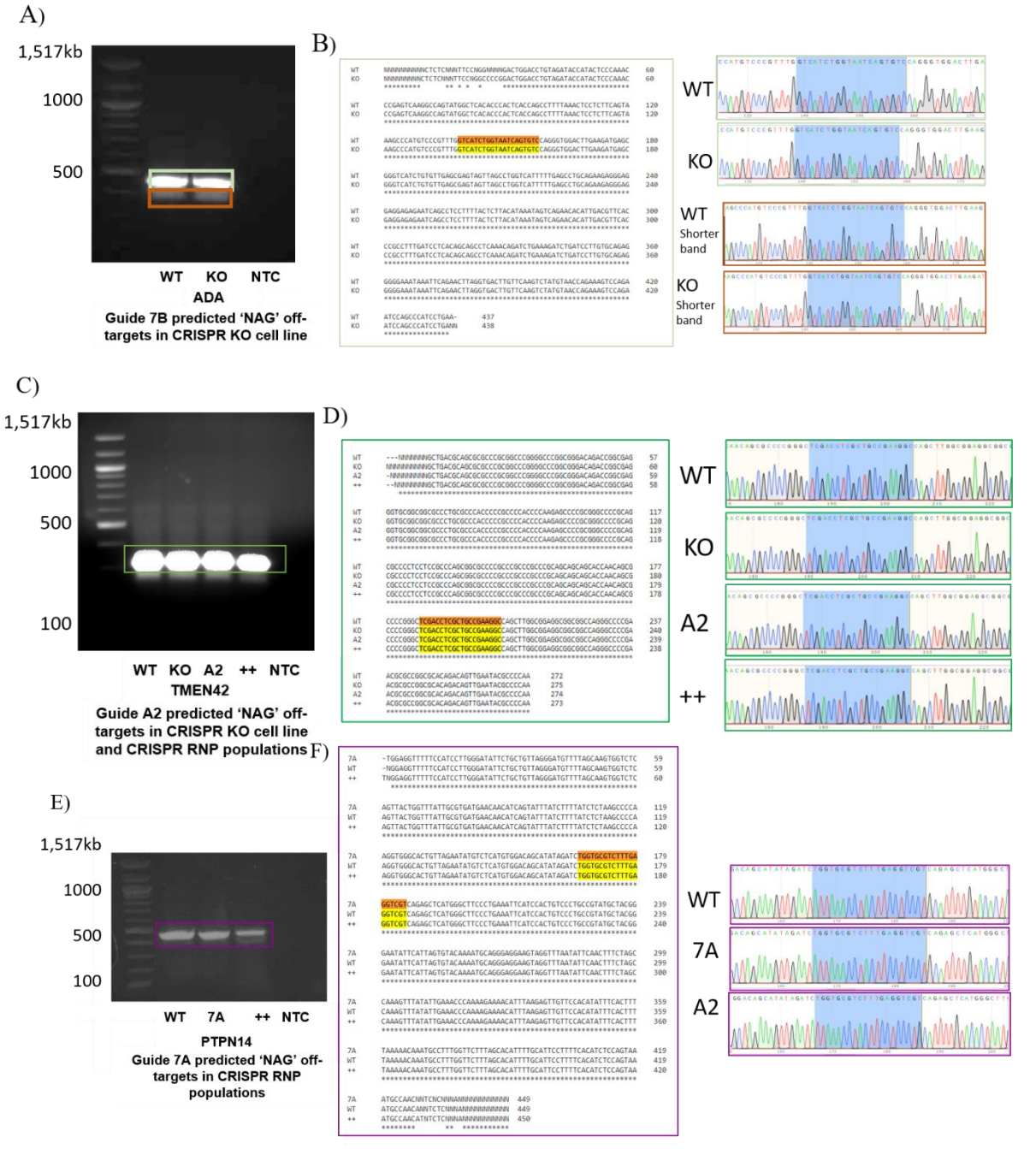


Figure 62. Analysis of the highest scoring predicted 'NAG' off target region per guide in complete or partial *TNC* KO BJ fibroblasts demonstrated no off-target editing had occurred.
PCRs from gDNA extracted from WT, *TNC* KO, or partial *TNC* KO BJ fibroblasts nucleofected with either CRISPR/Cas9 or CRISPR/Cas9 RNP were carried out, utilising primers to amplify a region spanning the predicted-off target regions (https://benchling.com) in the *ADA* gene **(A),** the *TMEN42* gene **(C)** and *PTPN14* gene **(E).** Multiple alignment of the band extracted and sequenced products for *ADA* **(B),** *TMEN42* **(D),** and *PTPN14* **(F).** * indicates identical sequences whilst the predicted-off target sequence region is highlighted in orange (WT) and yellow (KO). The corresponding electropherograms for each product are shown with the predicted off-target region identified (blue).

To test the robustness of our off-target prediction and validation approach, an additional random gDNA site which was not predicted as off-target was analysed. For this, *COL1A1* was chosen as this gene is very relevant given the ECM aspect of the project. Amplification of gDNA from WT, *TNC* KO and partial *TNC* KOs using primers *COL1A1* Fw and Rv was successful, however the band was not at the expected size (Figure 63). A repeat PCR returned the same results with band sizes far larger than the expected 200 bp, approximately 1500 bp. The primers were originally designed for qPCR, hence the small amplification product. Sequencing revealed the sequences were highly similar for at least 660 bp across all cell populations (Figure 63 B), this was lost after this point, when base calling ended for some populations or the peaks became of poor quality and miscalled bases. Blast analysis revealed approximately 98 % sequence homology to the *COL1A1* gene on chromosome 17; this was without removing any 'N' base calls. The *COL1A1* gene is a highly repetitive gene, with the protein characteristically composed of repeating repeats. It may be that the primers bound to a similar region further away from the intended binding region resulting in the much larger product. This seems likely and is supported by the fact that some of the lanes appear to have smears and multiple bands. The image quality is poor, again due to a malfunctioning LAS imager. There were predominant bands at the same size in each lane that were accurately excised for sequencing after visualisation on a UV light box (Figure 63 C). It appears no off-target editing occurred. Generally, for all the sequencing products, some unmatched base calls occurred, as indicated by the lack of an asterix in the sequence alignments. Manual inspection revealed these were miscalls, which commonly occurred with repeat nucleotides and mainly at the start of the sequencing which generally suffers inaccuracy until the chemistry stabilises. Collectively, these results demonstrate no-off target effects in these specific genes resulting from off-target CRISPR/Cas9 cleavage.

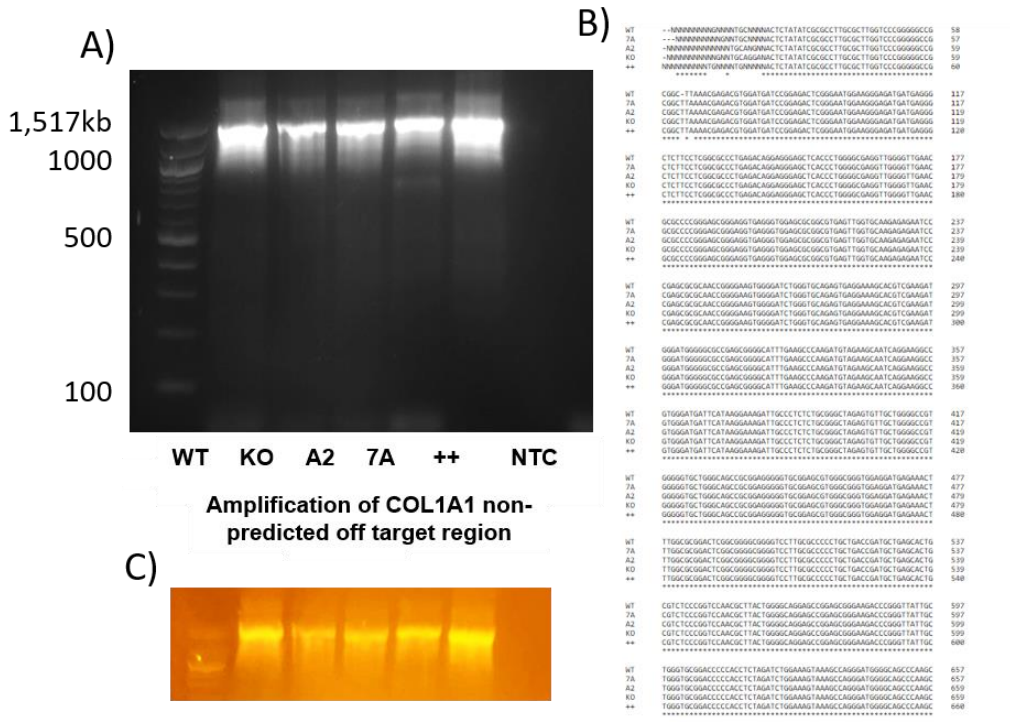


Figure 63. PCR analysis of a region of the *COL1A1* genome indicated no off-target editing by Cas9 had occurred in a non-predicted region in *TNC* KO fibroblasts.
A) PCRs from gDNA extracted from WT, *TNC* KO and partial *TNC* KO fibroblasts using primers to amplify a genomic region of *COL1A1*. **B)** A Clustal Omega multiple alignment of sequenced PCR products. C) Gel image acquired using a UV lightbox.

4.12.1.2 Off-targeted validation utilised RNA-Seq of WT and TNC KO fibroblasts at PDL 55 indicated no off-target editing had taken place

RNA-Seq of WT and *TNC* KO fibroblasts at PDL 55 was employed to ascertain if any off-target edits had occurred in any of the bioinformatically predicted off-targets, including those screened by PCR, as well as those that had not been screened by PCR (Figure 64). Many genes hosting predicted off-targets were not detected in either WT or *TNC* KO fibroblasts: *VWA1, CAMLG, ADCY4, ADA GMIP, SKOR1, SMPDL38, PHYKPL, BAZ1A, TMEM42, CCLB1, WDR90*, and *C2orf62*. Expression was not significantly different in any of the detected genes (adjusted P value > 0.05, as determined by Novogene), as demonstrated by the similar colour denoting expression level (log_2 scale, Figure 64), indicating no off-target edits had occurred in the additional genes to those from the section above assessed by PCR.

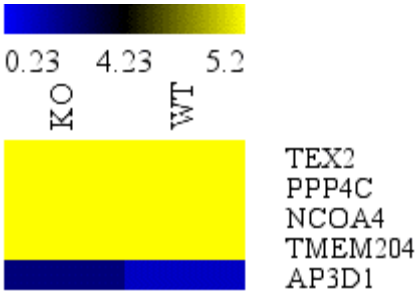


Figure 64. RNA-Seq of WT and *TNC* KO fibroblasts indicated no off-target edits had occurred in additional predicted off-targets.
Unfortunately, not all genes hosting predicted off-targets were expressed in either *TNC* KO or WT fibroblasts, however a few loci of predicted off-targets were detected in addition to those screened by PCR. Increased expression (mRNA abundance on a $\log_2$ scale) is denoted by yellow whilst decreased expression is denoted by blue.

4.13 Summary

A stable and complete *TNC* KO ECM model was engineered using CRISPR/Cas9 genome editing of human BJ foreskin fibroblasts that generated a pair of double-strand breaks at two targeted gene loci. This led to a ~12 Kb inversion and a flipping of the top and bottom strands of the entire *TNC* genomic fragment, from the target site near the ATG start codon in exon 2 to the target site in exon 7, created by CRISPR/Cas9 cleavage and subsequently a complete KO of *TNC*. This ECM model is representative of healthy human ECM and has been demonstrated to deposit a rich network of collagen-1, collagen-3 and fibronectin, as in wild type fibroblast-derived ECM. Screening of bioinformatically predicted off-targets indicated precise on-target editing of our designed gRNAs, with no off-target effects in predicted off-target regions.

4.14 Discussion

TN-C is a well-studied matricellular protein, conforming to the major matricellular characteristics of a highly restricted expression pattern in adults, with the exception of disease or injury and that it does not have a structural role, unlike other core matrisome proteins [183]. Whilst it was first documented as an

anti-adhesion protein with early *in vitro* cell studies [182, 239], its function remained unclear due the fact that *TNC* KO mice exhibited a normal phenotype [231, 232]. Now, it is established that TN-C has a role in wound healing and tissue repair, and is upregulated in many cancers and inflammatory conditions [183]. This is partially due to its complex domain structure which allows interaction with a wide range of ECM proteins and receptors, including fibronectins, perlecans, glypicans, heparins, integrins and syndecans, controlling many aspects of cell behaviour, including adhesion, proliferation and migration [184]. More recently, *TNC* expression was demonstrated necessary to generate an appropriate inflammatory response to bacterial lipopolysaccharide in mice, by regulating macrophage miR-155 expression and subsequent TNF-α release [188]. In order to investigate the relevance of this finding to humans, a healthy human ECM model was developed. Just as knockout mice were used to establish the role of *TNC* in mice, the next step was to generate a *TNC* knockout ECM model, utilising the ECM model validated in chapter 1.

CRISPR is a powerful tool that has revolutionised genomic editing, allowing relatively quick and incredibly specific editing of any genomic loci, providing the locus meets the criteria for species specific Cas9 cleavage. This approach has been utilised to edit, knockout and overexpress many proteins in different organisms, and to generate cell lines with the genetic identity of known disease-causing mutations. We chose to use CRISPR/Cas9 to produce a stable *TNC* knockout cell line that would then be utilised to generate a *TNC* knockout 3D ECM model. As opposed to 2D culture models, 3D models allow the elucidation of how ECM molecules, such as *TNC*, influence the inflammatory response in humans. This was done by utilising the ECM models as a substrate for the culture and differentiation of primary human monocytes (chapter 3).

We successfully generated a complete *TNC* knockout BJ fibroblast cell line which we characterised at the DNA level. We found that targeted CRISPR/Cas9 gene editing resulted in the introduction of a ~12 kb inversion of *TNC* itself into *TNC* and thus a paracentric inversion. Flipping of the top and bottom strands also occurred. This inverted segment was the ~12 kb complete fragment generated from cleavage at the target sites in exon 2 and exon 7 of *TNC*. We confirmed a complete KO at the protein level, supporting the genotyping results. At the transcript level we determined that the transcripts

encoding the main isoforms of TN-C were also negligible in the *TN*-C KO cell line, compared to WT fibroblasts in which expression was far greater. Aberrant splicing was also indicated in the *TNC* KO fibroblasts cell line by using IGV viewer, which demonstrated exon 1- 8 were joined with the exclusion of exon 2-7, compared to the WT fibroblasts. In WT fibroblasts, the full length isoform (2201 amino acids), a slightly smaller isoform (2019 amino acids) and a small isoform (1564 amino acids) were detected. This is in line with a previous study which identified expression of the small and largest isoforms from fibroblasts [522]. Whilst the generation of a Cas9-induced DSB at the target locus is well established, with the resulting sequence following the error-prone NHEJ repair pathway varying from hundreds to thousands of base pairs, the inversion at first appeared surprising. However, this is a well-documented event induced by utilising pairs of sgRNAs during CRISPR/Cas9 genomic editing. Large chromosomal rearrangements, including inversions, have been demonstrated in the body-cavity-based lymphoma cell line (BCBL-1) infected with Kaposi's sarcoma-associated herpesvirus (KSHV). For this, two sgRNAs and Cas9 expressed from a single plasmid resulted in complete knockout of the viral ORF-57 via an inversion of approximately 1.5 kb. This inverted fragment sequence composition was deduced by inverse PCRs and sequencing and found to be the fragment resulting from cleavage at the two target sites [524]. In lung adenocarcinoma, translocations and inversions can be 'driver events' [525]. One study chose to model two common inversion driver events, including a paracentric inversion, analogous to the inversion that occurred in the *TNC* KO fibroblasts, using CRISPR/Cas9 with a pair of sgRNAs targeting 2 target loci, identified from patient samples. This introduction of two simultaneous DSBs in HEK-293Ts successfully introduced a specific 12 Mb inversion. Furthermore, this inversion was not present in control cells expressing only one sgRNA [408]. Similarly, a study using HEK293Ts investigated whether there was a size limitation to such CRISPR/Cas9-mediated inversions by designing multiple pairs of sgRNAs, for several chromosomes. Here, they successfully induced inversions ranging from 35 bp to 807.5 kb and found that the inversions occurred 3 bp upstream from the PAM sequence, in line with Cas9 cleavage. This study also demonstrated that micro insertions and deletions (scarring) occurred, suggesting a mechanism of NHEJ [526]. A study prior to this, which screened the resulting murine erythroleukemia (MEL) clones for different types of genetic rearrangements after transfection with pairs of CRISPR/Cas9 plasmids, also confirmed that inversions

commonly occur [527]. Again, it was demonstrated that these inversions can occur when the sgRNAs target sites are over 1 megabase apart. However, it was shown that deletions occur more frequently than inversions on an allele basis, in which deleted alleles accounted for 26.8% (149/558) whilst inversions only account for 12.9% (72/558). It was also demonstrated that scarring (INDEL formation) was the most common outcome at 60.3% (556) [527], although scarring did not occur here. *Kraft et al.* transfected G4 mouse embryonic stem cells with pairs of CRISPR plasmids with sgRNAs 232 kb apart and adjacent to the pitx1 promoter, and also demonstrated that deletions and inversions were commonly found (5%) [528]. Another study that transfected CD44+ erythroid cells with pairs of sgRNAs targeting the MCS-R2 α-globin enhancer also found a similar occurrence of inversions (6%), however deletions were again more common [529]. Similarly, in a study in which zygotes were transplanted into female mice, following cytoplasmic injection of pairs of sgRNAs and Cas9 mRNA into zygotes, 4/41 offspring contained inversions at the target locus (the entirety of the *Cntn6* gene) whilst 7 offspring contained deletions [530]. Unfortunately, it cannot be deduced whether the proportions of different mutations types in the *TNC* genome would be in line with the previous literature; only one clone survived expansion. The results throughout these studies show very similar probabilities of developing an inversion, however the small differences are likely due to the different CRISPR/plasmids and transfection methods used, and the different cell lines and target loci. The inverse relationship between deletion size and deletion frequency [527, 528] and the repeated demonstration that utilising pairs of sgRNAs can successfully induce large chromosomal rearrangements may aid in designing strategies for the knockout of other large genes.

The most prevalent criticism of genome editing with CRISPR/Cas9 is the question of unintended alteration to the genome; off-target effects. This has been the focus of many studies, however not all studies utilising CRISPR assess the potential off-target effects. Off-targets which contained 3-4 mismatches were predicted utilising the Benchling software and subsequently validated by PCR amplification and Sanger sequencing and RNA-Seq. No off-targets were demonstrated in the *TNC* KO cell line at any of the predicted-off target sites. There are large amounts of conflicting literature on this topic. Schaefer *et al*. [531] claimed that CRISPR-Cas9 resulted in unexpected off-targets in mice, as

confirmed by whole genome sequencing. However, this paper has been retracted due to more recent research challenging the conclusions of the work and the methodologies; the main critique was that the proper controls of a mouse sibling was missing. More recently, three independent studies utilising embryos have demonstrated the dangers of editing with CRISPR. One study identified off-target edits ranging from 4 to 20 kb in regions beyond the target gene as well as within the target gene *POU5F1* itself, resulting in loss of heterozygosity [532]. Segmental and complete loss of the paternal chromosome 6 following targeting of the *EYS* loci was a common outcome in a further study attempting to restore the blindness-causing disrupted reading frame [533]. A final independent study on embryos also demonstrated loss of heterozygosity after CRISPR/Cas9 targeting at not only *MYBPC3* target sites, but regions adjacent to the target site. This resulted from a repair mechanism known as 'gene conversion' in which one intact homolog is utilised as a repair template, leading to a copy of that chromosome [534]. Loss of heterozygosity presents a serious safety concern in itself but also indicates that gross rearrangements can occur resulting from CRISPR/Cas9 genome editing.

Whilst these studies focussed on the much wider context of the chromosome after genome editing, studies probing the specific outcomes of the sequences of sgRNAs have identified that sgRNAs are capable of introducing mutations with 3-6 mismatches from the target sequence in both U20S GFP reporter cells and in HEK293 cells [425]. Given this and the fact that mutations can cause large chromosomal rearrangements both on- and off-target, it seems surprising that no bioinformatically predicted off-targets were experimentally found in the *TNC* KO cell lines in the present study. This is most likely explained by the fact that the majority of the predicted off-targets had 4 mismatches to the guide gRNA/target site, with at least 1 occurring within the last 10-12 nucleotides at the critical 3' end of the gRNA proximal to the PAM site. Indeed, mismatches in the 3' end of gRNAs are less well tolerated than those in the 5' end with regard to Cas9 activity [425], confirming previous findings that a single mismatch up to 11 nucleotides 5' of the PAM sequence is enough to abolish Cas9 cleavage, whilst those further upstream are tolerated in HEK293T [535]. However, there are exceptions - not all 5' mismatches are well tolerated and some 3' mismatches still result in Cas9 cleavage [425]. These studies and others [413, 536] indicate that perfect base-pairing of 10-12 bp seed sequence directly 5′ of

the PAM (PAM-proximal) determines Cas9 specificity. A larger study assessing >700 sgRNA variants and INDELS at >100 off-target genomic loci in HEK293T and HEK293FT cells confirmed the previous findings that single mismatches are tolerated more so in PAM-distal locations than PAM-proximal locations, but also found that nearly all bases in the guide provide varying degrees of Cas9 specificity [403].

Other factors affecting off-target cleavage have been identified. Hsu *et al.* demonstrated that Cas9 cleavage specificity (ratio of off-target to on-target mutations) was reduced and INDEL mutation occurrence increased with increasing concentrations of Cas9 and sgRNA (10 ng to 400 ng of Cas9-sgRNA plasmid). Correspondingly, the reduction in concentration led to increased specificity, but a less on-target cleavage [403]. An *in vitro* study utilising high-throughput DNA-sequencing and computational analysis revealed that excess concentrations of Cas9 to sgRNA ratio decreases cleavage specificity [537]. Unsurprisingly, these studies establish a complex role for both sgRNA sequence, Cas9 concentration and target locus. Whilst the *TNC* KO clone received a total of 1 μg CRISPR/Cas9 + sgRNA plasmids, which is a higher concentration than these studies, it could be that in this cell line the enzyme was not in excess to allow tolerance of the mismatch and subsequent cleavage. These studies utilised different expression plasmids and largely relied on HEK-293s, a genetically abnormal, polyploid cell population. This is in contrast to BJ fibroblasts, which are a primary diploid cell line. It could be that off-target mutations are tolerated better by HEK-293s and that fibroblasts with off-target effects simply died; this is purely speculation since only one clone survived. If larger numbers of cells had survived the expansion, it may be that off-target effects were found. Furthermore, whereas specific predicted off-targets [435] were assessed in the *TNC* KO clone, the studies reported in the literature took a much broader approach, utilising deep sequencing, the TE17 endonuclease or surveyor assays, which have a higher and/or broader detection rate and therefore may pick up more off-target mutations including smaller SNPs or substitutions. Predictive software is only as good as the information compiled to generate the predictions and it may simply be that off-target editing occurred in regions that were not screened. The utilisation of primers targeting the *COL1A* gene was used to try address this concern, as this was not a predicted target and therefore should not, and did not exhibit a mutation in the region

screened. However, in line with our findings, a recent study on human articular chondrocytes screened 17 putative predicted off-targets generated from software by Sanger sequencing and found no off-target editing had occurred [538]. The presence of the species-specific trinucleotide PAM sequence is essential for Cas9 activity. Cas9 also has a tolerance for NAG PAM sequences, but with much less efficiency [403, 523]. Initial work indicated no off-target edits had occurred in the highest scoring 'NAG' predicted off-target region per guide. However, only deep sequencing would allow for a full and unequivocal assessment of whether any off-target effects had occurred. Overall, our results demonstrate that Cas9-induced DNA breaks promote efficient rearrangement between pairs of targeted loci, and that there were no-off targets at predicted-off target sites.

Transfection of fibroblasts with Cas9 ribonuclear protein (RNP) removed many limiting factors of generating a *TNC* KO cell line via CRISPR with plasmids, including transfection efficiency and single-cell expansion; bacterial contamination frequently occurred in the low numbers of cells recovered after single-cell sorting and explains why only a single clone survived. Whilst the mixed population of edited CRISPR RNP fibroblasts did not allow genotyping as in the clonal KO cell line, western blotting revealed drastically reduced TN-C expression. The risk of off-target effects are mitigated with CRISPR/Cas9 RNP due to the shortened half-life of approximately 12 hours compared to expression of Cas9 from a vector [409]. Our lab was able to detect GFP expression for up to 9 days in CRISPR/Cas9 transfected fibroblasts. Indeed off-targets were also not detected in any of the predicted off-targets in these cell populations, however this conclusion of no off-targets is subject to the same caveats discussed previously.

Prior to this and because of the limited off-target effects associated with CRISPR/Cas9n, we attempted to validate the Cas9n vectors in 2 other cell lines; MDA-MB-231 and HEK-293. These cells are more robust and amenable to transfection, albeit a transformed cell line. Mutants were not detected in HEKs, which could be attributable to low expression and therefore condensed chromatin (heterochromatin) blocking Cas9 access. The current effect of high-order chromatin configurations on RNA-guided Cas9 cleavage is ill-defined. However, a clever model system was developed in which chromatin confirmation (condensed heterochromatin or relaxed euchromatin) of reporter constructs was controlled

by treatment with doxycycline. This treatment released an epigenetic silencing KRAB fusion protein from a recognition element in the gene which resulted in the relaxation of heterochromatin to euchromatin. Following transfection with CRISPR/Cas9 plasmids, reporter gene expression of GFP and mCherry was then measured. As mCherry was cloned out of frame, only ORF correcting Cas9-induced INDELs (in the GFP target site) would result in mCherry gene expression and a corresponding lack of GFP expression. This study demonstrated that condensed heterochromatin significantly hindered the formation of Cas9-induced INDELs for several Cas9 variants [539]. It was also demonstrated that TALENs, the other key designer nuclease platform, were affected more strongly by chromatin configuration [539], indicating the effectiveness of CRISPR/Cas9 over TALENs. A large scale genetic screen of 18,380 sgRNAs targeting 1539 genes and an *in vitro* binding and cleavage assay with Cas9 or dCas9 confirmed these findings, demonstrating that nucleosomes impede Cas9 access to DNA [540]. A further study of dCas9 using CHIP-Seq to analyse the genome wide binding sites of 12 sgRNAs in HEK-293s found that dCas9 binding sites were enriched in open chromatin regions [541]. Whilst these studies did not investigate Cas9n, it is logical that the chromatin structure would affect Cas9n access similarly since dCas9 differs only by a mutation in the catalytic site and is therefore structurally identical in shape and size. This may partially explain why no genomic edits were detected in HEK293s.

To address this issue of condensed chromatin structure, Cas9n vectors were tested in MDA-MB-231s, which are known to express high levels of TN-C due to their tumorous origin [517]. Cells were treated with LPS, which is known to induce *TNC* transcription and thus euchromatin structure enabling transcription. We successfully detected an INDEL mutation that would result in a premature stop codon *in vitro* in one allele; this gDNA was harvested from a cell not subjected to LPS before or during the transfection, indicating that LPS did not enhance INDEL generation in this instance. Most recently, Doudna and her team have discovered a hyper compact species of Cas9 – Cas9 Φ from bacteriophage that is half the size of Cas9 and still able to cleave DNA in a targeted fashion in animal, bacteria and plant cells. Future work utilising this species should be more efficient for delivery into the cell and may also enable targeting of condensed regions of chromatin [542].

Many other reasons as to why more INDELs were not detected in either cell line exists. Cas9 associates with the DSB for up to 6 hours after a DSB [543] which may prevent repair machinery access to the break [543] and explain variable and slow (hours) repair rates following Cas9-induced DSBs [544]. Again this has not been tested in Cas9n, and may well be different since only a nick occurs. The lack of detected mutants may be explained by the intuitive thought that the efficiency of Cas9n is lower than Cas9, since it requires two guides to bind as a pair to induce a break [424, 536]. Conversely, Cas9n pairs have been shown, in specific instances, to be more efficient [545]. However, single nicks in DNA are preferentially repaired by HDR rather than the error prone NHEJ [546], and it has been shown that Cas9n facilitates HDR with minimal INDEL formation [535]. Transfection following cell synchronisation was tested in fibroblasts as transfection and CRISPR editing efficiency were shown to be increased during S-phase [429, 519]. However, this resulted in high cell death compared to controls, and would likely have resulted in even more HDR events as a necessary copy of the *TNC* gene would have been available to utilise as a repair template. NHEJ events could in theory have been forced by transfecting fibroblasts in G1 phase, however high cell death resulted from this. Cas9n is favourably utilised in studies not wishing to knockout protein production and rather produce targeted edits. Together, it is not surprising that a lack of mutants were detected in HEK-293s and MDA-MB-231s. It is clear that concentrations of Cas9 and sgRNA, and sgRNA sequences can drastically alter specificity. In this instance, it does appear that Cas9 is more efficient than Cas9n, since one surviving cell contained the desired targeted gene edit in *TNC*, whilst this was not seen in either clones or bulk populations of HEK-293s or MDA-MB-231s.

Another reason why so few fibroblasts, HEK-293s or MDA-MB-231s with edits were detected may be due to the very recent finding that Cas9-induced DSBs activate a p53 dependent cell-cycle arrest pathway, which subsequently leads to death. Although a high efficiency of INDELs has been achieved in human induced pluripotent stem cells (iPSCs), edited cells viability was significantly decreased [547]. This finding was reinforced by Haapaniemi *et al*. [548] who identified that CRISPR/Cas9 induced a p53-mediated damage response in RPE1 cells. Correspondingly, inhibition of p53 led to slight increase of HDR following Cas9 cleavage. The findings that many fibroblast single cell clones survived

the initial expansion after transfection with Cas9 but subsequently died before genotyping could occur is in line with these previous studies. Furthermore, given that several cell populations were recovered for HEK-293s and MDA-MB-231s, yet none were a clonal *TNC* KO, also supports this theory. Together, these studies indicate that the DSB required to introduce the desired edits may actually result in selection against the cells containing the desired DSB and subsequent edit [549]. It does therefore not seem surprising that so few edits were detected.

This work has shown that it is possible to knockout an ECM gene in BJ foreskin fibroblasts, in a targeted approach using CRISPR/Cas9 and that knockout of *TNC* does not affect BJ cells innate ability to deposit collagen-1, collagen-3 and fibronectin. However, future approaches with this cell line should be focussed on utilising the ribonuclear Cas9 in order to remove the limiting factor of cell sorting, which was a lengthy and inefficient process for generating a stable cell line; clonal populations are not possible without cell sorting, but rounds of RNP transfection or more than one guide at once may facilitate even higher gene knockdown percentages.

Assessment of the *TNC* KO BJ-derived CDM demonstrated a fibronectin matrix that appeared less dense than the WT BJ derived matrix from chapter 1. Images were taken in regions around the edge of the coverslip and some towards the middle. The apparently decreased fibronectin matrix density may be the result of damage to the CDM from handling of the coverslips in which removal causes ripping of the matrix, due to the tension generated by attachment to the tissue culture dish. This damage was seen more so in matrices deposited by *TNC* KO fibroblasts, suggesting that the deposition and assembly may be less in these matrices, creating a weaker matrix more susceptible to damage.

5 ANALYSIS OF INFLAMMATORY GENE EXPRESSION FROM *TNC* KO FIBROBLASTS

5.1 UTILISING CDMS AS SUBSTRATE FOR MACROPHAGES TO INVESTIGATE HOW THE ECM REGULATES GENE EXPRESSION

During the inflammatory response, the initial detection of the pathogen is mediated by several classes of evolutionary conserved pattern recognition receptors (PRRs) including TLRs. PRRs recognise conserved microbial structures or pathogen-associated molecular patterns (PAMPs). Cell wall components of bacteria can function as PAMPs such as LPS from gram-negative bacteria which solely activates TLR4, on the macrophage cell surface [550, 551] with mutations resulting in increased susceptibility to infection [552]. Molecules released from damaged cells, such as TN-C, can also function as endogenous ligands for PRRs, known as damage-associated molecular patterns (DAMPs) [326]. PPR activation triggers a cascade of intracellular signalling pathways that co-ordinately lead to the expression of inflammatory mediators, including cytokines such as TNF-α and transcription factors like NF-kB, necessary to eradicate the pathogens.

Previous research by Piccinini *et al*. has demonstrated the role of TN-C in the miR-155/TNF-α axis during the LPS-induced inflammatory response in which TN-C is necessary for an appropriate immune response [188]. . In $TNC^{+/+}$ mice LPS stimulation results in the expression of miR-155 and subsequent production of TNF-α by macrophages. In contrast, $TNC^{-/-}$ knockout mice have a reduction in both miR-155 expression and TNF-α production and as such demonstrate an impaired inflammatory response. miR-155 is an exceptional miRNA in that it enhances TNF-α production by stabilising TNF-α mRNA transcripts [553], whilst usually miRNAs work to silence gene expression by targeting the transcript for degradation or repressing translation of the transcript [356]. Several questions arose from the Piccinini study: Does this regulatory axis occur in human cells? Is soluble (secreted in bodily fluid) or insoluble (ECM-associated) TN-C responsible for this effect? Does TN-C regulate the expression of other genes?

The aim was therefore to validate the BJ-derived CDM as a substrate for human macrophages. BJ derived TN-C containing CDM (WT) and BJ-derived CDM lacking TN-C (KO CDM) were then used as healthy human ECM models to investigate the role of human ECM-associated TN-C on inflammatory gene expression during an LPS-induced inflammatory response.

5.2 ECM MODELS

Model 1 utilised WT CDM and CRISPR/Cas9-generated TN-C KO CDM for the culture and differentiation of monocytes to macrophages (macrophages on WT CDM and macrophages on KO CDM, respectively), which were then stimulated with LPS to simulate infection and produce an inflammatory response. Expression of miR-155 was analysed by RNA extraction and RT-qPCR. TNF-α expression was quantified by an ELISA using supernatant (Figure 65 A, C).

Model 2 was developed to assess the corresponding role of TN-C by 'overexpression' of TN-C. Due to time constraints, CRISPRa (discussed later) for the overexpression of *TNC* was not feasible. Instead, WT CDMs were generated as previously, followed by the addition of full- length purified human TN-C protein to the CDMs. TN-C was allowed to incorporate into the ECM before thorough washing of unbound TN-C. This model was utilised as described above for differentiation and culture of macrophages (macrophages on TN-C$^+$ CDM) and the analysis of miR-155 and TNF-α (Figure 65 B, C). In both models, WT CDM, or uncoated plastic (macrophages on plastic) could be utilised as controls (Figure 65 C).

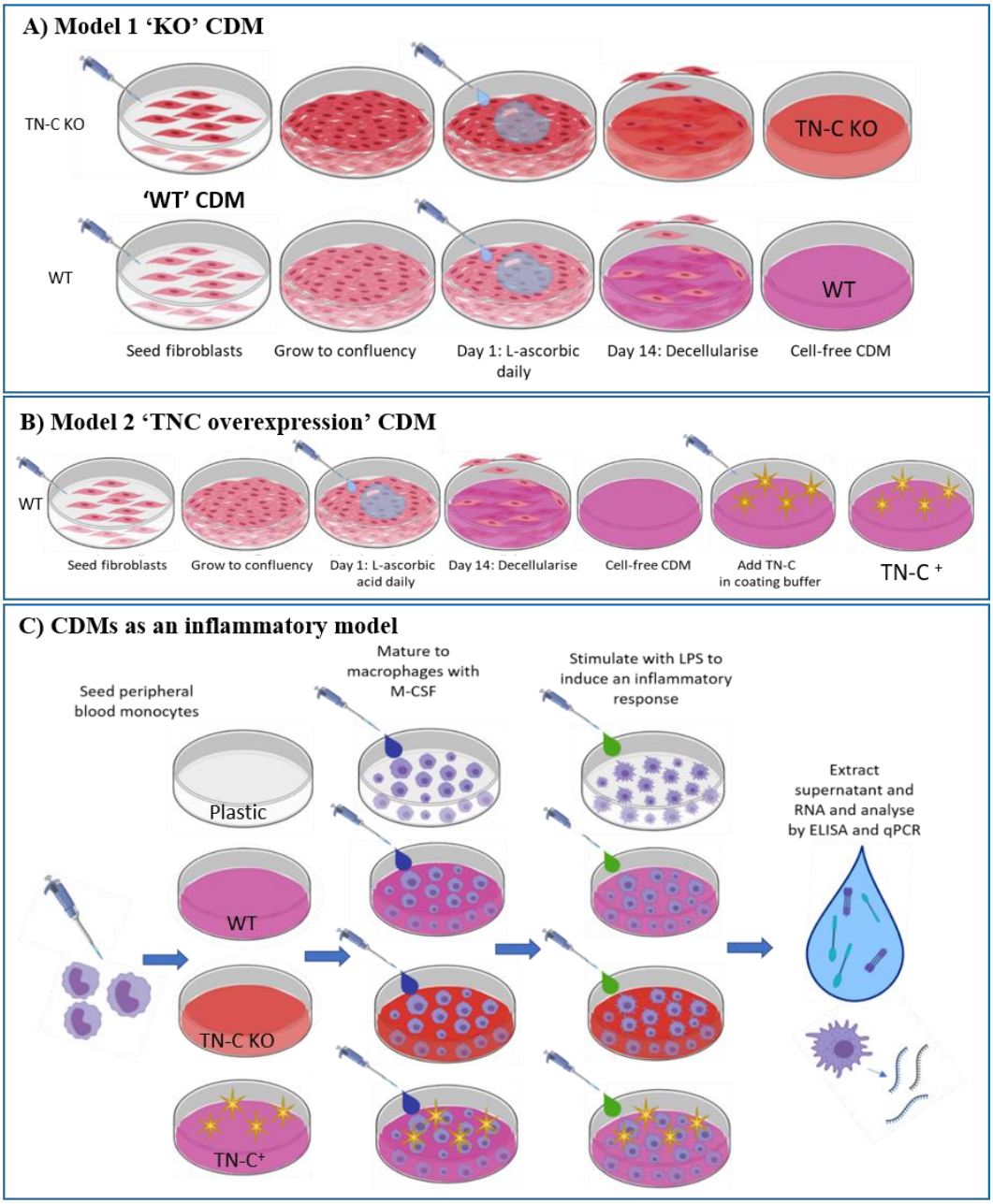


Figure 65. The generation and utilisation of WT, KO and TN-C + (*TNC* overexpression) CDMs as models of healthy human ECM.

A) Schematic of model 1 in which WT fibroblasts or TN-C KO fibroblasts (generated by CRISPR/Cas9) were seeded and grown to confluence before culturing for 10-14 days in the presence of L-ascorbic acid. During this time fibroblasts secrete ECM proteins that are assembled into a 3D ECM. CDMs were decellularised using decellularisation buffer composed of 20 mM ammonium hydroxide (NH_4OH) and 0.5 % Triton X-100 before DNase I treatment to remove residual DNA, which resulted in cell free WT and TN-C KO CDMs. **B)** In model 2, CDMs are generated with WT cells as previously, but recombinant TN-C was added after decellularisation and allowed to incorporate into the already deposited ECM. The resulting ECM was therefore used as a 'TN-C overexpression' model. **C)** Schematic of the approach taken to generate an inflammatory response. Peripheral blood monocytes (PBMs) were seeded, differentiated to mature macrophages with M-CSF (100 ng/ml) before stimulating with LPS (1 ng/ml) to induce an inflammatory response. Inflammatory cytokines in the supernatant were analysed by ELISA, whilst the expression of the pro-inflammatory miR-155 and the anti-inflammatory miR–132 were analysed by RT-qPCR.

5.3 ANALYSIS OF CDM COMPATIBILITY AND DIFFERENTIATION OF PRIMARY HUMAN MONOCYTES

Firstly, to determine whether this model is a suitable substrate for the culture of human macrophages, monocytes were seeded on either uncoated plastic or WT CDMs in the presence of 100 ng/ml of macrophage-colony stimulating factor (M-CSF) for 5 days. The maturation of macrophages on WT CDM and on plastic was determined by measuring the expression of the maturation marker CD68 by FACS analysis. CD68 expression was nearly identical with a negligible increase from macrophages on WT-CDM (0.1 %) indicating cells differentiate and mature successfully and similarly on both substrates (Figure 66 A).

In order to ascertain that macrophages were viable on the different substrates following stimulation with LPS, macrophages were cultured on WT CDM or plastic and viability assessed by an MTT assay. Cell viability was altered over the time course; the most striking difference was a sharp increase in viability at 8 hours. This increase is not surprising as LPS results in the induction of many genes, including cell survival and proliferation genes [520], which by 8 hours would confer their effects. This was followed by a dip in viability by 24 hours that was still marginally increased compared to unstimulated controls. It could be that the effect of LPS-induced genes was still present at this time point. The variation throughout the time course was mainly in a temporal manner; cell viability did not differ between macrophages on each substrate, indicating that the CDM substrate was not detrimental to macrophage survival (Figure 66 B).

Next, the adhesion of macrophages differentiated on plastic or CDM substrates, was tested. *In vivo*, these immune cells exhibit a highly migratory phenotype, which is necessary for responding to pathogens and accumulating at sites of infection, as well as to patrol and sample the cell microenvironment. In contrast, it is well known that macrophage adhesion to plastic is extremely strong [554] and that, logically, the adhesion of a cell is linked to its migratory properties [555]. Reflecting this, macrophages on CDM were 50 % less adhesive than macrophages on plastic (Figure 66 C). This suggests that macrophages may adhere in a more physiological manner on CDMs compared to plastic.

It should be noted that these experiments only consist of one biological replicate and that 4-5 donors are usually necessary to draw conclusions. However, even with one biological replicate these experiments do suggest that CDMs are a suitable and promising substrate for the culture of macrophages.

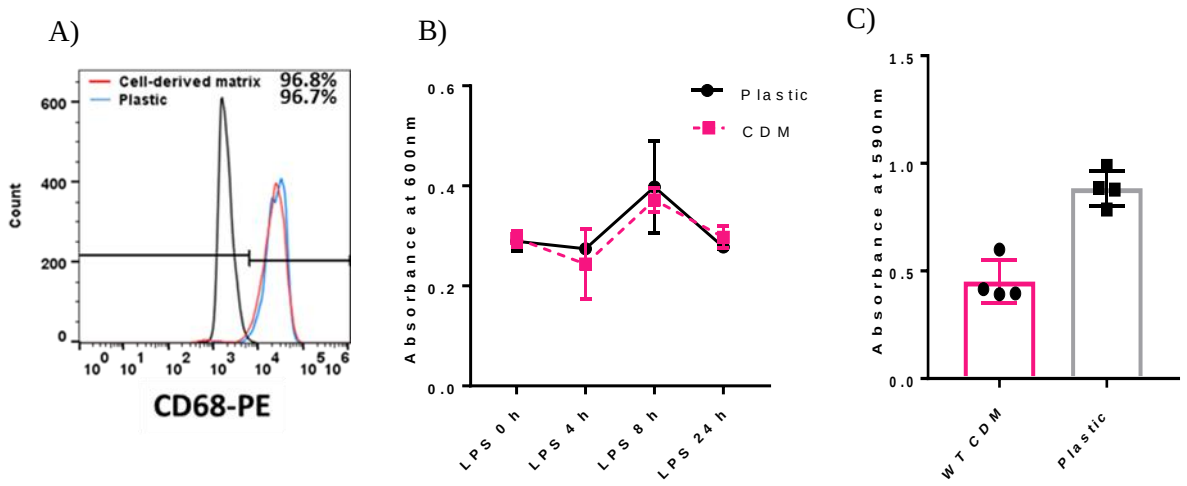


Figure 66. BJ-CDMs are a suitable substrate for the differentiation and culture of human macrophages. **A)** FACS analysis of the macrophage maturation marker CD68 from macrophages on plastic or macrophages on WT CDM after differentiation of monocytes with M-CSF (100 ng/ml) for 5 days. **B)** MTT analysis to assess the viability of macrophages on WT CDM or macrophages on plastic, untreated (0) and stimulated with LPS for 4, 8 and 24 hours. Cells were treated with MTT and the absorbance read after complete cell lysis. Mean ± SD from 3 technical replicates from 1 human donor. **C)** Adhesion assay of macrophages on WT CDM or plastic. Macrophages were differentiated on the two substrates as described previously. Following removal of unbound cells, adherent cells were stained with Toluidine blue. Cells were fixed, lysed and the absorbance read at 590 nm. Mean ± SD from 3 technical replicates from 1 human donor.

5.4 INVESTIGATION OF MACROPHAGE ADHERENCE ON GLASS COVERSLIPS OR WT CDM

To investigate the adhesion of monocytes further and visibly, a relatively new method of Environmental Scanning Electron Microscopy (ESEM) was carried out by Nicole Weston. Conventional SEM (CSEM) cannot observe biological samples with significant amounts of water for several reasons: during scanning the water evaporates causing sample distortion and collapse, the water vapours interfere decreasing the vacuum and interfere with the sample omitted electrons [556]. ESEMs main advantage over CSEM is that samples do not need to be fixed, dried or coated with metals before imaging, lending

itself to the imaging of naturally hydrophilic samples, such as the CDM [557]. In this way ESEM allows for imaging of cells in a state as close as possible to their physiological undried state. For example, samples do not appear stereotypically shrunken; artefacts of the drying process [556]. CSEM works by scanning an electron beam over the sample and then collects sample-emitted electrons (secondary electrons) under a high vacuum to prevent scattering of the beam or electrons. The secondary electrons are collected by a detector to form an image. In principle, ESEM operates in the same manner except requires no vacuum but instead a high-pressure gaseous environment which controls electron scatter and a specialised detector collects the secondary electrons to form an image. The temperature and humidity can be controlled to maintain the sample in its natural state. Due to the change in experimental setup, there is a slight degradation in image quality and resolution compared with CSEM, however this is outweighed by the possibility and ease of analysing many new types of samples [557].

For ESEM analysis, human monocytes were seeded on glass coverslips or WT CDM and differentiated to macrophages with M-CSF, as previously, before imaging by Nicole Weston. Macrophages on glass exhibited a much more uniform appearance, and appeared to adhere very strongly which resulted in a very flattened appearance with the cell nucleus clearly visible (Figure 67 A, C). The cytoplasm surrounding the nucleus is also clearly visible and appears in a relatively consistent diameter around the cells; almost a “fried egg” phenotype. Whilst some cells appear elongated, as if migrating, this is quite rare. This is in stark contrast to macrophages on CDM, in which the cell cytoplasm is more difficult to distinguish as the cells exhibit a more rounded phenotype. This cell population appears much less uniform in shape and the cells exhibit a preference to make contacts with other cells by long protrusions. The CDM is evident as a meshwork surrounding the cells (Figure 67 B, D). Salt crystals can be seen (green box) in both conditions – samples require thorough washing before imaging to remove any residual salt from the PBS storage buffer. Salt crystals are more prevalent in the CDM condition, indicating that the ECM network may naturally retain more water and therefore more salt, and that extra washing is required for these samples. Some partially dehydrated cells can be seen in both substrate conditions (blue arrows) in the higher magnification images. This results from decreased temperature in conjunction with increased time in the imaging chamber; eventually samples will dehydrate in the

chamber if imaged for long enough. Nevertheless, these images support the previous indication that macrophages are much less adhesive on WT CDM than glass or uncoated coverslips, thus supporting CDMs as a physiological substrate in which macrophages are able to migrate and patrol, as they would *in vivo*.

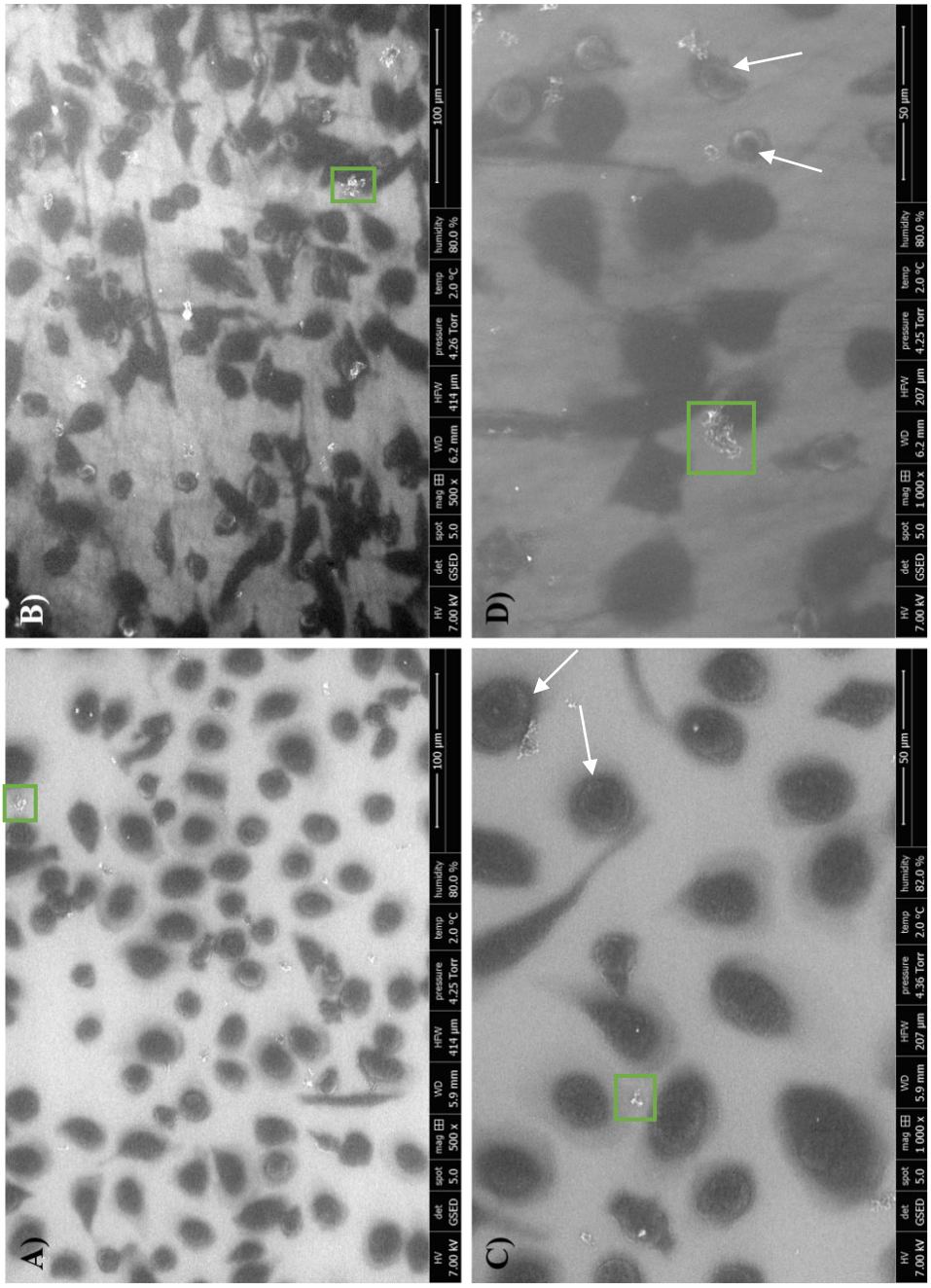


Figure 67. Macrophage behaviour is regulated differently on glass or WT CDM. ESEM of monocytes matured to macrophages in the presence of MCS-F for 5 days on glass coverslips **(A and C)** or on CDM **(B and D).** Scale bar 100 µM (top) and 50 µM (bottom). White arrows indicate partially dehydrated cells. Green box indicates a salt crystal example.

5.5 PRELIMINARY INVESTIGATION OF MACROPHAGE GENE REGULATION BY THE ECM

To determine how CDM substrates impact macrophage miRNA expression, monocytes were differentiated into macrophages in 10 cm dishes and stimulated with LPS before RNA extraction, reverse transcription and qPCR analysis of miR-155 and miR-132; two well documented, early LPS-responsive miRNAs involved in the inflammatory response [558]. Macrophages differentiated and

cultured on WT CDM displayed peak expression of miR-155 at 8 hours after LPS stimulation before rapidly decreasing at 24 hours. This is in stark contrast to macrophages differentiated and cultured on standard tissue culture plastic, which exhibited a strong linear and exponential induction of miR-155 at 24 hours (Figure 68 B, C). This same pattern of induction is seen for miR-132 in which the induction peaks 8 hours after LPS stimulation and then decreases by 24 hours in macrophages on CDM. Similarly, a strong linear and exponential induction of miR-132 at 24 hours in macrophages on glass. *In vivo*, this oscillating pattern of induction of pro-inflammatory miRNAs followed by downregulation is common, to avoid a prolonged and damaging inflammatory response. As expected, a much stronger induction was seen for miR-155 than miR-132 although induction of both miRNAs was stronger in macrophages on plastic. The stable expression of the U6 housekeeping gene indicates that these differences in gene expression are not due to defects in transcription and represent genuine expression differences (Figure 68 A). Together, these results indicate that CDMs are a promising model for the culture of macrophages to study inflammation, and may produce more physiologically relevant results than traditional culture methods.

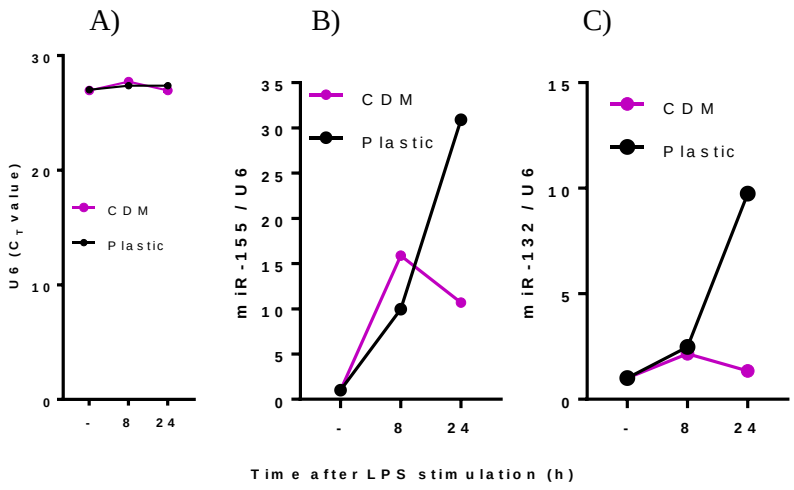


Figure 68. WT CDM and plastic substrates alter the expression of LPS-induced proinflammatory miRs by macrophages.
Monocytes were differentiated to macrophages on CDM or plastic, stimulated with 1 ng/ml of LPS at 0, 4, 8 and 24h. RNA was extracted at the indicated time points before RT-qPCR. Gene expression presented as fold-change normalised to U6 housekeeping gene and relative to unstimulated control (ΔΔct). Data from one donor conducted in triplicate.

Taking this one step further, the expression of the miR-155- regulated pro-inflammatory cytokine TNF-α [312] was assessed by ELISA using supernatant from LPS-stimulated macrophages on WT CDM or plastic. The production of TNF-α by macrophages seeded in 10 cm dishes showed striking differences between macrophage substrates which was in line with miR-155 expression (Figure 69 A). Both macrophages on plastic and macrophages on WT CDM exhibited an induction of TNF-α which peaked 8 hours after stimulation. The peak production was more pronounced in macrophages on plastic. In both conditions the production of TNF-α decreased only marginally by the 24 hour time point. The differences in TNF-α production were less evident in macrophages cultured on the different substrates in 96 well plates (Figure 69 C). The general pattern of TNF-α expression was similar with peak concentrations at 8 hour followed again by lower concentrations at 24 hours. However, these reductions were much larger than previously, 44.5 % compared to 9 %, and 55.2 % compared to 2. 2% in macrophages on CDM and macrophages on plastic, respectively. This reduction by 24 hours is expected due to the stringent regulation of pro-inflammatory molecules, to avoid a sustained inflammatory response that would result in tissue damage. Whilst the results show slight differences, together they indicate that macrophages cultured on CDMs retain their ability to mount an inflammatory response to LPS.

In both 10 cm dishes and 96 well plates, unstimulated macrophages on CDM have much lower basal production of TNF-α than macrophages on plastic (Figure 69 B, D). TNF-α production by macrophages on WT CDM at 0 hour was approximately 50% less than that of macrophages on plastic. Equally, in 96 well plates, TNF-α production at 0 hour in macrophages on WT CDM was 90% less than that of macrophages on plastic. This indicates artificial macrophage activation by plastic, which is absent in the CDM model, demonstrating the potential of CDMs as a substrate in producing more physiological responses, compared to tissue culture plastic.

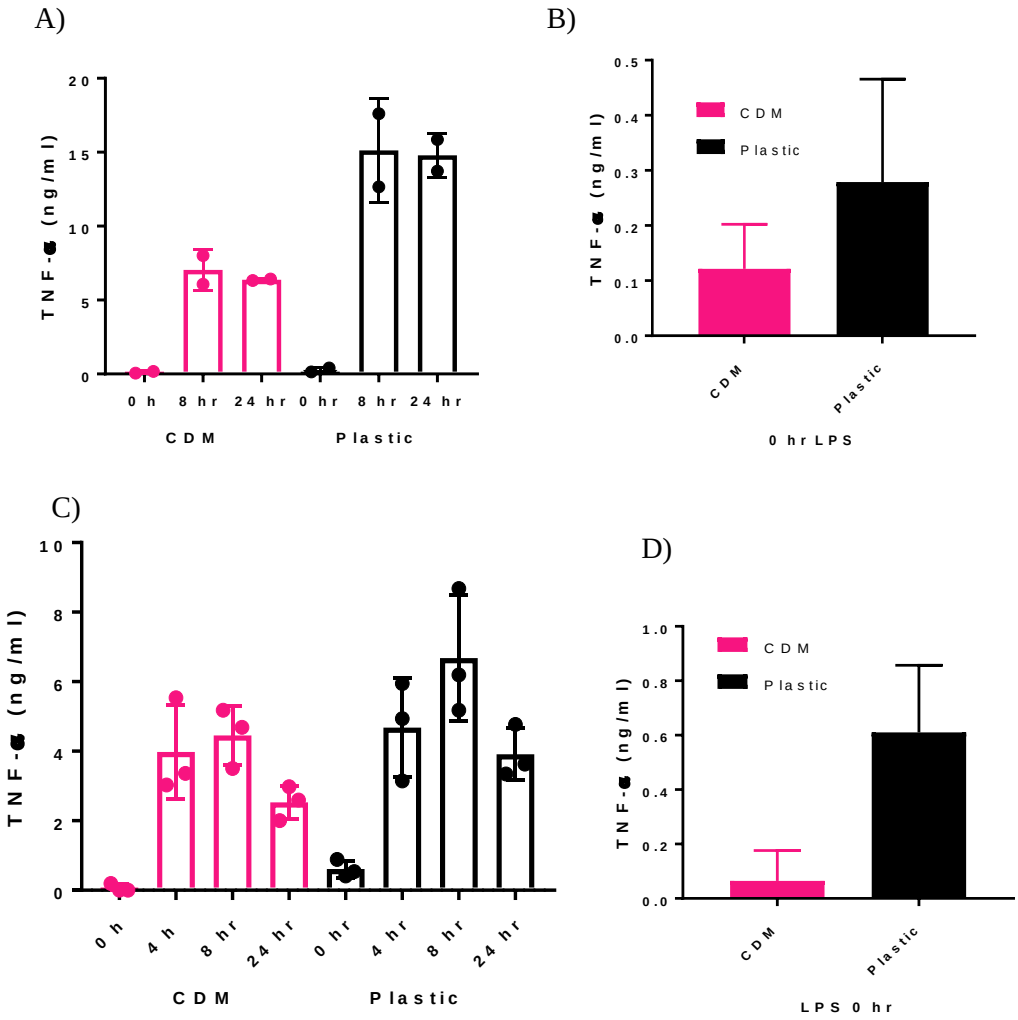


Figure 69. LPS- induced TNF-α production is differentially regulated from macrophages on plastic or macrophages on WT CDM.
A) TNF-α production by macrophages on plastic or macrophages on WT CDM from 10 cm dishes was assessed via an ELISA after 0, 4, 8 or 24 hrs LPS (1 ng/ml) stimulation. **B)** Production of TNF-α from unstimulated macrophages on plastic or WT CDM from 10 cm dishes. Results from 3 ELISA technical replicates from 1 human donor, mean ± SD. **C)** TNF-α production by macrophages on plastic or WT CDM from 96 wells was assessed via an ELISA after 0, 4, 8 or 24 hrs LPS stimulation. **D)** Production of TNF-α from unstimulated macrophages on plastic or WT CDM in 96 well plates. Results from 3 ELISA technical replicates from 1 humour donor, mean ± SD.

5.6 USING MODEL 1 TO ASSESS THE ROLE OF ECM- ASSOCIATED TN-C IN LPS-INDUCED MIR-155 EXPRESSION

After preliminary experiments to verify that CDMs were a suitable substrate for the culture of macrophages, WT CDM and KO CDM were utilised, to assess the role of ECM-assembled TN-C in gene regulation during the inflammatory response (Figure 70). Macrophages on WT CDM had a rapid miR-155 induction following LPS stimulation that increased linearly to a 24 hour peak. In contrast, the

peak induction of miR-155 expression in macrophages on KO CDM occurred at 8 hours after LPS stimulation that was followed by a decrease in expression by 24 hours. The induction of miR-155 was stronger throughout the time course from macrophages on WT CDM macrophages and hampered in macrophages on KO CDM. This demonstrates the role of ECM-associated TN-C in the regulation of miR-155 expression.

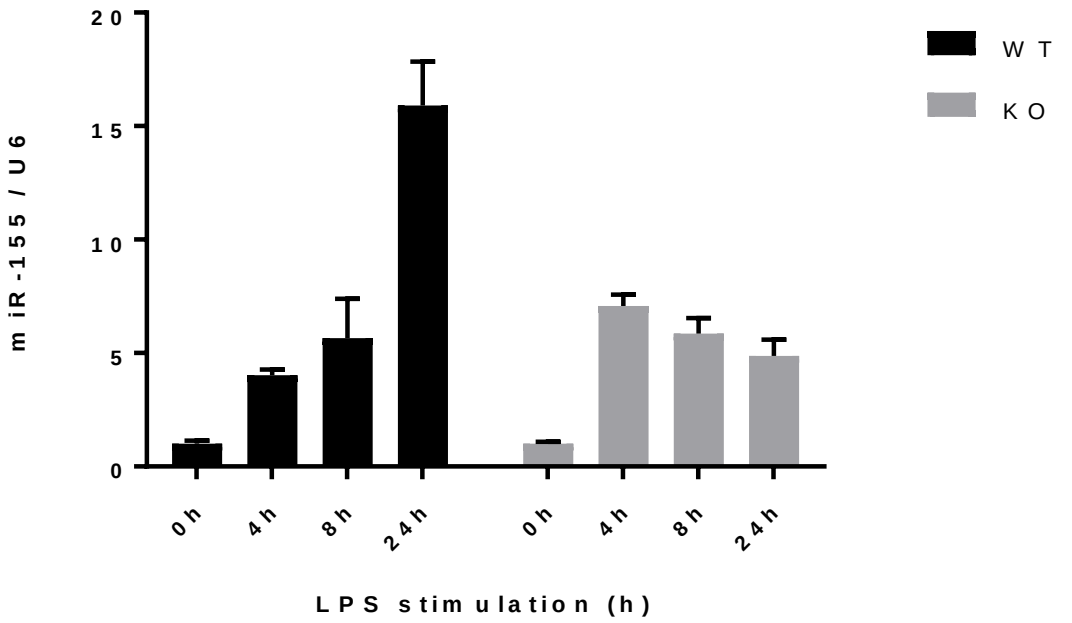


Figure 70. ECM assembled Tenascin- C regulates LPS- induced miR-155 expression in primary human macrophages.
miR-155-5p expression in macrophages on WT or KO CDM stimulated with LPS at 0, 4, 8 and 24h. RNA was extracted followed by reverse transcription and qPCR. Gene expression presented as fold-change normalised to U6 housekeeping gene and relative to unstimulated control (ΔΔct). Mean ± SD from 3 qPCR technical replicates from 1 human donor.

5.7 THE ROLE OF ECM- ASSOCIATED TN-C IN LPS-INDUCED TNF- α

Next, the production of the pro-inflammatory cytokine TNF-α was assessed by ELISA, using supernatant from the model 1 experiment. In line with the miR-155 data, an increase in TNF-α production in macrophages on WT-CDM occurs peaking at 8 hours, followed by a decrease at 24 hours (Figure 71). This decrease is not surprising given the destructive effects uncontrolled or sustained production of this cytokine has, hence the production of inflammatory cytokines is tightly regulated.

Overall, the data indicates that the ECM-assembled TN-C regulates miR-155 and TNF-α production in primary human macrophages.

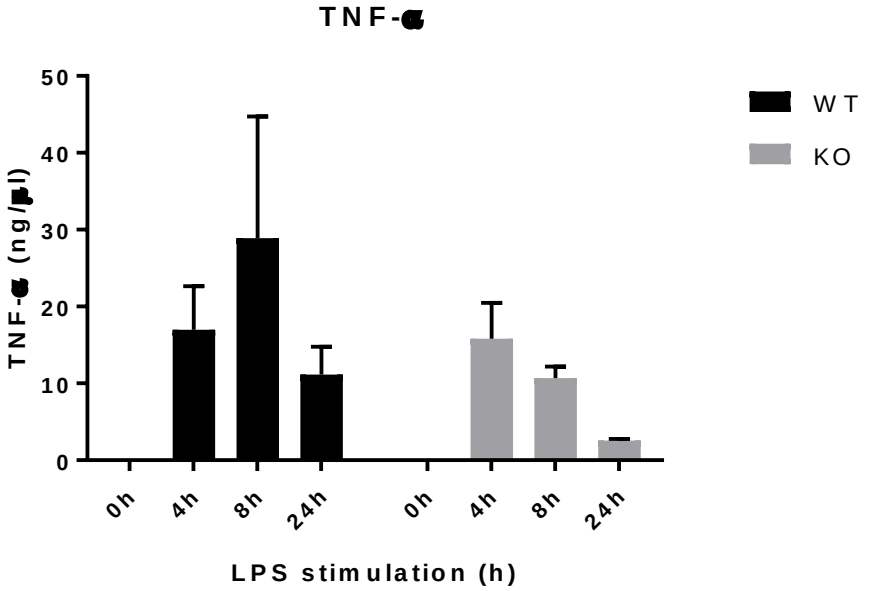


Figure 71. **LPS- induced TNF-α production from macrophages on WT CDM or macrophages on KO CDM is regulated by ECM incorporated TN-C**.
TNF-α production by macrophages on WT or KO CDM, using supernatant from 6 well plates was assessed via an ELISA 0, 4, 8 or 24 hrs after LPS (1 ng/ml) stimulation from 1 donor, 1 biological experiment conducted in triplicate showing mean ± SD.

5.8 GENERATION AND VALIDATION OF TN-C OVEREXPRESSION MODEL (TN-C$^+$ CDMS)

In order to fully assess the role of ECM associated TN-C, an overexpression model would be utilised that would be expected to show the opposite and complementary results to those generated using *TNC* KO models. Due to time constraints it was not possible to generate a CRISPR overexpression model. Therefore, to generate CDMs 'overexpressing' TN-C (TN-C$^+$ CDM), full-length, recombinant human TN-C was added to WT CDMs. In order to successfully investigate the role of ECM-associated (insoluble) TN-C, it was first necessary to confirm a) that the additional TN-C would successfully assemble into the ECM and b) the optimum concentration of TN-C to add to WT CDMs. Therefore three different concentrations (1.5 µg/ml, 3 µg/ml, 6 µg/ml) of TN-C in coating buffer were added to decellularised WT CDMs and incubated at 4°C overnight followed by washing, blocking and

immunostaining of TN-C (Figure 72). A dose-response of incorporation occurred in which TN-C incorporation was demonstrated to be more prevalent as the concentration increased. For clarity, edges of the coverslip are demonstrated with blue arrows, which demonstrate a good contrast between the glass and CDM with stained TN-C. White arrows indicate the CDM. WT CDM, with no TN-C added, was treated the same as TN-C addition conditions, and therefore any staining differences are genuine and not an artefact resulting from the TN-C coating protocol. Interestingly, and also in a dose-dependent manner, brighter clumps of staining are more ubiquitous upon increasing concentrations. This could likely be aggregated TN-C that has not been fully incorporated into the ECM, rather bound to the surface by one of their ECM protein binding domains. FN was also stained as a control, due to consistent and large deposition of this protein, to demonstrate that the addition of TN-C did not affect the generated CDM in some other way. FN was seen as a dense meshwork and did not appear altered upon visualisation by fluorescence microscopy. This demonstrated that additional TN-C could successfully be incorporated into WT CDM to generate TN-C^{+} CDMs.

Next, to determine the optimum concentration of TN-C to add to generate TN-C $^{+}$ CDMs, the expression of LPS-induced miR-155 from macrophages differentiated and cultured on WT CDM and TN-C^{+} CDMs at concentrations 1.5 µg/ml, 3 µg/ml, 6 µg/ml was assessed. The 24-hour time point was chosen as previous time courses demonstrated the largest induction of miR-155 expression at this time, compared to untreated controls (Figure 70) and the biggest differences when compared to the miR-155 induction in KO CDM-macrophages. Again, in line with the immunofluorescence images, a dose-response was observed with miR-155 expression increasing with increasing TN-C concentrations (Figure 73). However, a reduction in miR-155 induction was seen at the 6 µg/ml concentration. Collectively, these results demonstrate that full length TN-C can be assembled into an already deposited ECM and indicate a concentration up to 3 µg/ml is sufficient to induce regulatory changes in gene expression by macrophages.

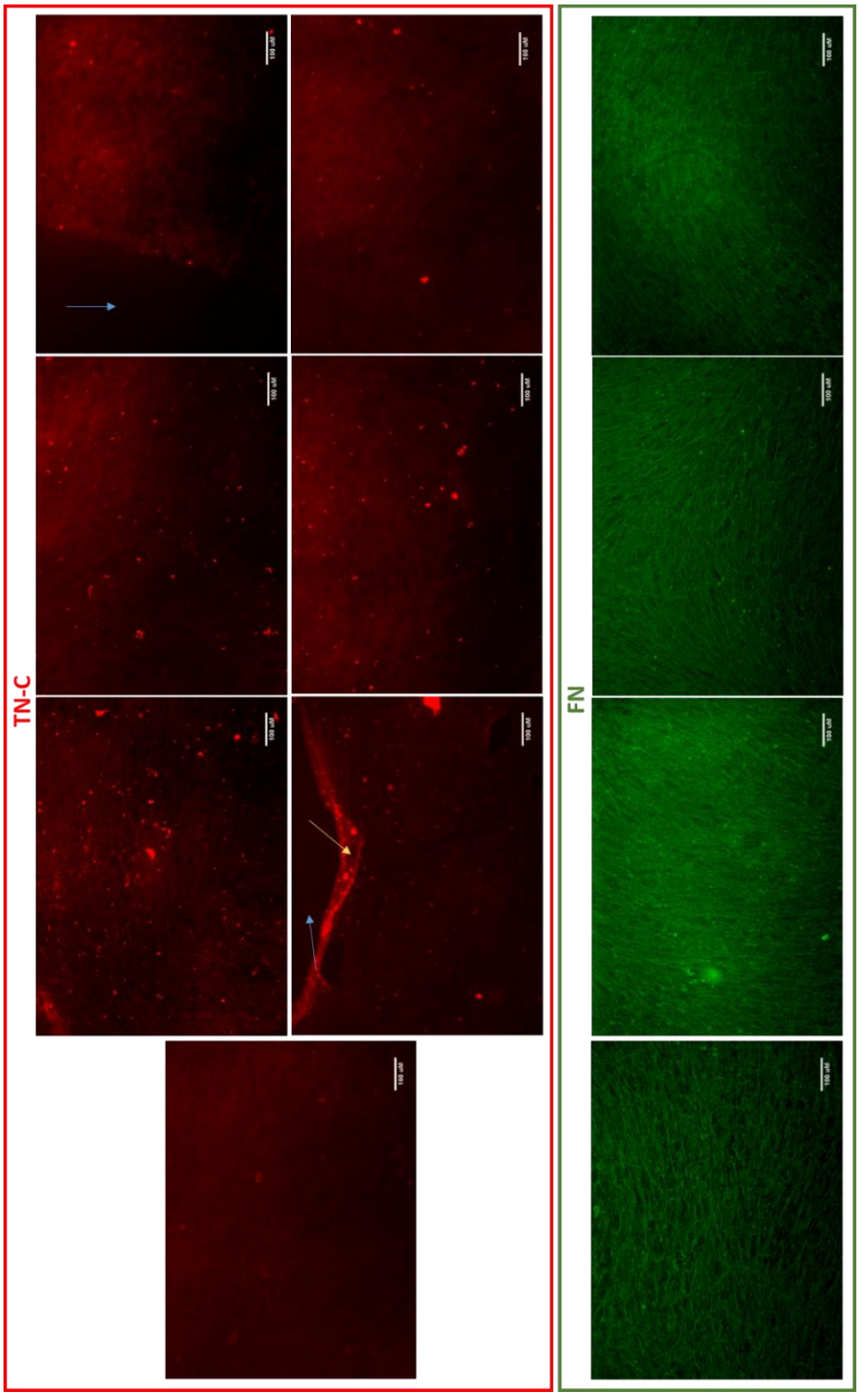


Figure 72. Additional full-length recombinant human- TN-C can successfully be incorporated into to already deposited WT CDMs. BJ-fibroblasts were allowed to deposit ECM for 14- days followed by decellularisation. Full-length recombinant human TN-C was added to WT CDMs at a concentration of 0 µg/ml (WT CDM), 1.5 µg/ml, 3 µg/ml or 6 µg/ml and allowed to incorporate, followed by immunostaining of TN-C (red) and FN(green). Scale bar 100 µm

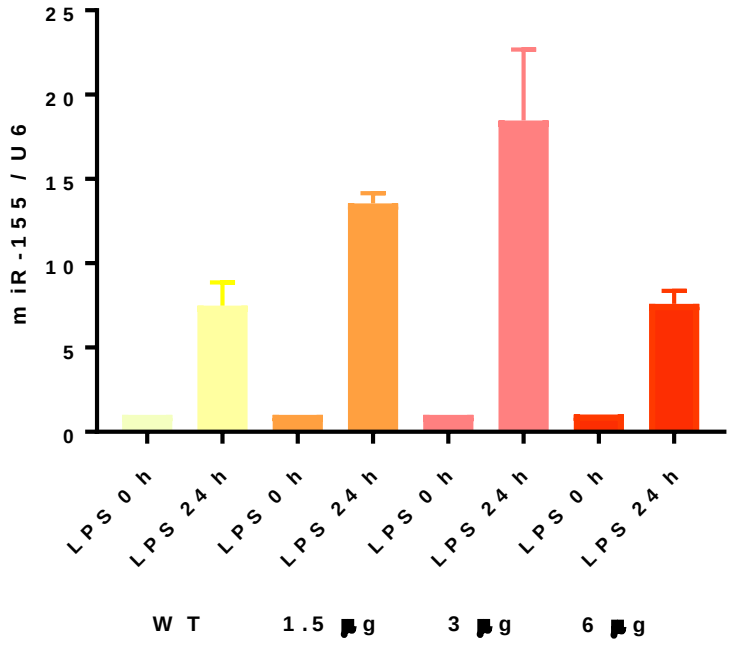


Figure 73. ECM-assembled Tenascin- C in TN-C $^{+}$-CDMs regulates LPS- induced miR-155 expression in primary human macrophages.
miR-155 expression from WT CDM-macrophages or TN-C^{+} 1.5 µg, 3 µg, 6µg CDM-macrophages, untreated (0) or stimulated with LPS for 24 hours. RNA was extracted followed by reverse transcription and qPCR. Mean ± SD from 3 qPCR technical replicates from 1 human donor. U6 was used as endogenous housekeeping gene and untreated samples as calibrator (ΔΔCt method).

5.9 SUMMARY

We have validated the *TNC* KO CDM, generated using *TNC* KO BJ fibroblasts from chapter 2, as a model to investigate how *TNC* in the ECM microenvironment affects gene regulation during the inflammatory response. qPCR and ELISA have shown, for the first time in humans, that ECM-assembled TN-C contributes to the regulation of LPS-induced miR-155 expression and subsequently TNF-α production. Furthermore, an overexpression model has been started to be developed, however this requires more validation.

5.10 DISCUSSION

The findings from this chapter support the previous findings in mice which demonstrated that TN-C posttranscriptionally regulated miR-155 production, and that *TNC*-KO mice had a significantly impaired LPS-induced inflammatory response in which miR-155 and TNF-α were significantly reduced

compared to WT mice [188]. Here, genetic ablation of TN-C from the ECM substrate drastically impaired LPS-induced miR-155 expression and TNF-α production in macrophages, as demonstrated by qPCR. For our experiments we have differentiated macrophages with M-CSF to M0 macrophages and stimulated with LPS obtaining inflammatory (M1-like) macrophages [559]. Raised levels of TN-C have been demonstrated in serum/plasma from patients with inflammatory conditions such as sepsis [560, 561] and rheumatoid arthritis [562] and equally, in arthritic mice models, genetic ablation of TN-C has shown to be protective against developing of the inflammatory disorder, whilst intra-articular injection of the FBG domain of TN-C resulted in joint inflammation [235]. Furthermore genetic ablation of TN-C in mice results in reduced production of inflammatory cytokines [324], indicating TN-Cs role in initiating and propagating a persistent inflammatory response.

It is established that LPS strongly induces miR-155 production in macrophages in response to LPS [188, 377, 563]. In line with this previous research, macrophages seeded on WT-*TNC* CDMs were indeed responsive to LPS stimulation, which elicited the production of miR-155. The rapid induction of TNF-α in response to LPS, followed by a reduction over time, particularly by 24 hours is documented in peripheral blood mononuclear cells (PBMCs), THP-1 macrophages and monocytes [564], and was also observed in our experiments.

There is the possibility that part of the LPS response in macrophages seeded on TN-C KO CDMs may be due to TN-C produced by macrophages [312, 565], however it would take some time for the gene to be expressed, translated, and assembled into the ECM. So this contribution here should be negligible.

There are limitations to the conclusions drawn from these experiments, mainly due to a lack of biological replicates. A second replicate utilising cells from a different donor was carried out to assess the effect of *TNC* KO CDM on miR-155 expression, but unfortunately the monocytes were unresponsive to LPS in this instance (data not shown). Due to Covid-19 it was not possible to perform the standard 4-5 biological replicates with different donors of human peripheral blood monocytes in order to determine a clear pattern; there is large intra-and inter-individual variation of the immune system at baseline and in response to stimuli such as vaccinations in human populations [566]. Even studies assessing the profiles of immune cell population frequencies, cytokine responses and serum

proteins in 210 identical twins have found large variations due to heritable and mainly non-heritable influences such as environmental factors. This variation increased over time [567]. The variation in immune responses in unrelated adult populations could therefore be far greater. Certainly, variation can be seen comparing miR-155 expression in the preliminary experiments from monocytes seeded on WT CDM in 10 cm dishes and the miR-155 expression from monocytes seeded on WT CDM in 6 well plates in 'model 1'. Although the seeding conditions may account for this, studies have demonstrated considerable heterogeneity within and between human macrophage donors. *In vitro* experiments (with LPS) have revealed that there is cell variability in response to infection. For instance, Salmonella-challenge of mouse bone marrow-derived macrophages results in a range of diverse phenotypes in which individual macrophages vary in their ability to phagocytose bacteria. Single-cell sequencing of cells sorted by FACS revealed cells with disparate phenotypic states reflecting macrophages responding to either extracellular bacterial exposure or intracellular signals of bacterial infection [568]. Whilst this study assesses the response to infection over the course of 24 hours, and not just the effects of LPS exposure, it gives insight into the multitude of responses from individual cells.

Successful TLR4 stimulation and induction of inflammatory cytokines by LPS requires the coordinated and timely action of multiple molecules. Soluble LPS binding protein (LBP) facilitates the association of LPS with CD14. [569] CD14 is essential for the transfer of LPS to the TLR4/MD-2 receptor complex, which in turn results in recognition [300, 570, 571]. Following recognition, TLR4 oligomerization occurs [572], followed by recruitment of downstream adaptor proteins that activate two main signalling pathways, the MyD88-dependent or independent pathway. In these pathways there are many steps where individual cellular response may vary. Recently, single molecule localisation microscopy has been used to investigate the spatial distribution of TLR4 receptors in primary human macrophages responding to LPS, and the effect of this on the immune response [295]. The presence of TLR4 clusters varied on unstimulated macrophages on individual cells and between cells. Of relevance to this study is that there was large donor-dependent variation in TLR4 clusters density; one unstimulated donor had a much lower density (~80 signals per μm^2) compared to 2 other donors (~250 and ~280 signals per μm^2). This basal cluster density was suggested to modulate the response to LPS; the donor with the

lowest basal clustering had the most pronounced signal increase following LPS stimulation (390%) that was not seen in the other 2 donors. This donor dependent differences highlights the need for large cohorts for studies with heterogeneous human samples.

Whilst the expression of CD68+ was assessed in a preliminary experiment this is not a comprehensive catalogue of markers to assess successful macrophage maturation. The original aim was to assess maturation of monocytes on plastic, WT-CDM or KO-CDM, unfortunately due to Covid-19 this was not possible. Detailed immunophenotyping would have assessed expression of the markers: CD68+, CD11b+, CD14+ and CD206+. The latter two markers can discriminate between monocytes and macrophages with CD206+ expression associated with M2 polarised macrophages [302]. Given that TN-C has been shown to bind to von Willebrand factor in co-precipitation experiments [573] and CD11b has a region encoded for the type A domain of von Willebrand factor [574], this may provide a mechanism for macrophage interaction with TN-C. Whilst polarisation was not an aspect of this study, expression of the markers CD86+ (or CD80+) /CD163- could be utilised to determine if macrophages were classically activated 'M1' macrophages after LPS stimulation.

On collection of data from at least 5 donors statistical analysis would have been carried out. This would have involved a 2-way ANOVA with Sidek's test for multiple comparisons to assess both the difference between the means for time points and the macrophage substrate (plastic, TN-C KO CDM or WT CDM matrix).

Staining of the TN-C+ overexpression model revealed that incorporation was only able to occur homogenously throughout the matrix up to a concentration of 3μg/ml. At 6μg/ml clusters of staining appeared, indicating that TN-C was associated in clusters at the surface of the matrix. Saturation may have occurred at 3μg/ml or, at higher concentrations an excess of TN-C resulted in self-association, due to the large number of compatible binding domains within each monomer of the hexabrachion. This is supported by the assessment of miR-155 from the overexpression models, in which miR-115 production was also decreased after 24 hours in the 6μg/ml, compared to the 3 μg/ml condition. This supports the

previous findings of this chapter that only ECM-assembled TN-C is able to propagate LPS-induced inflammatory signalling.

As a direct comparison CRISPRa would have ideally been utilised to overexpress TN-C. Briefly, CRISPRa works by the same mechanism as CRISPR/Cas9, with the exception that no DSBs are introduced. Instead, dCas9, a catalytically inactive Cas9, is guided by sgRNA to bind within 200bp upstream of the promoter region. This dCas9 mutant features the same mutation as Cas9n, in D10A, but also in H840A to inactivate the two nuclease domains [575]. Transcriptional activator domains can be fused to the inactive Cas9, to recruit a host of transcription proteins; the Synergistic Activation Mediator (SAM) complex is the most widely used and consists of a dCas9 fusion with VP64, and MS2-p65-HSF1 [576]. This method would require viral transduction in order to be stable and the length of time that overexpression via plasmid methods can be maintained for remains unclear, which may be an issue when CDM generation takes 14 days alone. However, SAM expression has been demonstrated by qPCR and PCR 20 days after transduction [577].

Activation of expression, using the native promoter also means expression is within range of endogenous levels. Other methods may result in expression of a protein at non-physiological levels, for example 6μg/ml of TN-C may not be within range of the amount found within tissues even during an inflammatory condition. Indeed, in serum from 95 patients with arterial hypertension and subsequent ventricular problems, TN-C containing domain B (FNIII), was found to circulate at levels ranging from ~100ng/ml to 2000ng/ml whilst TN-C containing domain C (FNIII) was found to circulate at levels ranging from 25ng/ml to 150ng/ml. This level was much reduced in samples from healthy patients; 100ng/ml to 1000ng/ml for TN-C containing B domain and 40ng/ml to 75ng/ml for TN-C containing C domain[578]. However, two other studies which assessed the large variant of TN-C containing the C domain of the FNIII repeats, found circulating levels to be much lower: 25ng/ml to 275ng/ml in samples from 63 patients with non-small cell lung cancer [579], and ~20ng/ml to 190ng/ml in 105 patients with myocardial infarction [580]. However, these studies do not assess all isoforms, and nor do the latter two studies evaluate samples resulting from infection and therefore an LPS response. In 167 patients with sepsis, TN-C serum levels ranged from 20pg/ml to 190pg/ml, although this was an uncontrolled

population in which sample collection/duration of sepsis was not identical across all samples within the study [560]. However, in cells collected from synovial membranes from patients with rheumatoid arthritis, TN-C levels ranged from 30ng/ml to 1500 ng/ml[224]. These studies highlight the variation in TN-C expression, between disease, during disease and in different tissues. It appears that concentrations up to 1.5 μg/ml may be best suited for further assessment of overexpression of TN-C. However, it is important to note that here, with the exogenous addition of recombinant TN-C, we have not assessed exactly how much has been incorporated and saturation will likely occur at 3 μg or above.

Whilst the current work demonstrates a role for ECM-assembled TN-C in the regulation of gene expression, further work is necessary for conclusions. Variation of immune responses is well documented and therefore replicates from additional donors are required for both models utilised in this study.

6 A ROLE FOR TENASCIN- C IN PREMATURE CELL AGING AND SENESCENCE?

6.1 INTRODUCTION

Observations of *TNC* knockout fibroblasts, generated by CRISPR/Cas9 genome editing, and partial *TNC* knockouts, generated by CRISPR/RNP, indicated that these cells exhibit slowed growth on defrosting for approximately 2-3 weeks or absent growth, yet do not undergo apoptosis, when compared to their wild type fibroblast counterparts. Observations of CRISPR–edited fibroblasts' phenotype by light microscopy were in line with hallmarks of senescence.

Cellular senescence is distinct from apoptosis as cells to do not undergo programmed cell death. They rather withdraw from the cell cycle and transition into an irreversible quiescent state, but still retain their metabolic activity. Senescence occurs when cells reach their replicative limit, an event known as 'replicative senescence' or the 'Hayflick limit' as determined by Hayflick *et al.* in the 1960s [581]. Hayflick *et al*. found that primary fibroblasts could replicate for a well-defined amount of population doublings (PDLs) before cellular division was blocked, in this way replicative senescence can be thought of as aging.

The PDL of cells *in vitro* can be calculated at each passage using the following formula:

PDL = 3.32 (log UCY - log l) + X,

where UCY = the cell number at passage, l = the cell number initially seeded at the previous passage, and X = the PDL of the previous passage.

This replicative limit is a consequence of progressive telomere shortening with each replication (Figure 74 A) leading to telomere uncapping and subsequent DNA damage and DNA damage response (DDR) signalling [582, 583], representing the most established and studied cause. Telomeric DNA naturally shortens with each S phase as proposed in the 'end replication problem' hypothesis in which the lagging

strand telomeres are unable to be extended by DNA polymerase during DNA replication [584, 585]. In support of this, fibroblasts exposed to mild oxidative stress have been shown to display accelerated telomere shortening and a replicative senescence phenotype, which can be counteracted by treatment with antioxidants [586, 587]. In further support of this, ectopic expression of the telomerase catalytic protein component (hTERT), which restores the telomere extension capability of telomerase, results in the immortalisation of primary foreskin fibroblasts [588]. Telomerase is expressed in embryonic stem cells and male germline cells, but is generally undetectable in somatic tissue with some exceptions [589]. Furthermore, a DNA damage response has been demonstrated in fibroblasts passaged until replicative senescence [590-593].

'Premature senescence' diverges from replicative senescence in that this occurs prematurely and can occur in a telomere-independent manner (Figure 74 B). In the former, this results from environmental stress prematurely shortening the telomeres or damaging the telomeres, irrespective of length. In both cases, senescence initiation is still reliant on a critical level of DNA damage in the form of DSBs, lesions and chromosomal instability which can be induced by a variety of stimuli, including oxidative stress, reactive oxygen species (ROS), inflammation and oncogenic stress [594]. Several studies now have reported that telomere dysfunction can occur regardless of length with persistent damage foci located at telomeres that are not critically short and it has been proposed that DSBs in telomeric sequences resist repair [595-597].

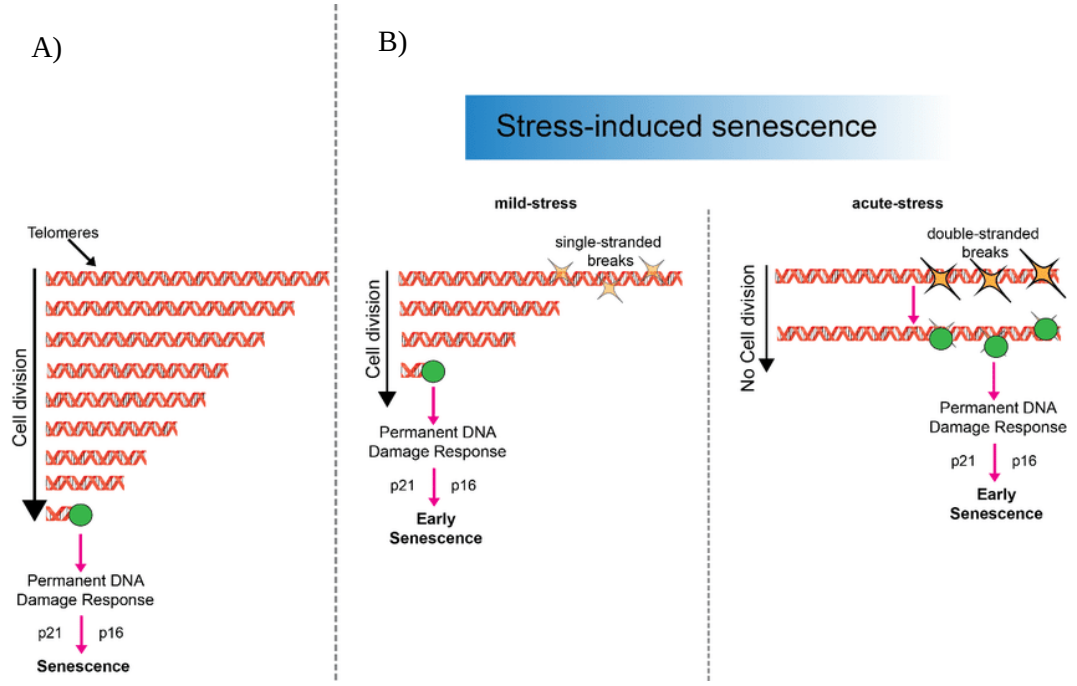


Figure 74. **Telomere length and senescence.**
A) Telomere shortening occurs naturally and gradually with each round of cell replication until the ends of chromosomes are exposed, triggering the DNA damage response (DDR). This is known as 'replicative senescence' or the 'Hayflick limit'. **B)** Premature senescence may occur resulting from stress-induced telomere shortening in which mild insults induce DSBs and DNA damage that shorten the telomere to a critical length. Acute stress may induce DSBs at telomeres which resist repair and result in a prolonged DDR leading to telomere-independent premature senescence. Adapted from Victorelli *et al.* [597].

Cell senescence is widely accepted as a tumour protective mechanism by arresting the cell cycle before malignant transformation can occur [594]. Both non-telomeric genomic damage and telomeric damage generate persistent DDR signalling that leads to the onset of senescence [598]. DNA DSBs usually induce senescence [599].

Once the critical level of DNA damage has occurred within the cell, cell cycle arrest occurs by one of three main tumour suppressive pathways controlled by p53, pRB and $p16^{INK4a}$ [600-603]. Senescent cells exhibit drastic changes in gene expression, metabolism, and a distinctive pro-inflammatory senescence associated secretory phenotype (SASP) that can promote tumorigenesis [603]. This is also known as the senescence-messaging secretome. This SASP is generally proinflammatory and includes overexpression of the proinflammatory cytokine interleukin 6 (IL-6) and the chemokine IL-8 [603], amongst many others. MMPs and serine proteases are also overexpressed. The SASP has also been linked to inflammation, given that many of these factors are also involved in inflammation. In addition

to their role as inflammatory mediators, IL-6 and IL-8 amongst others have been shown to propagate and maintain senescent growth arrest [604, 605]. Senescent fibroblasts and endothelial cells exhibit a >50-fold increase in plasminogen activator activity [606]. The expression of other molecules such as fibronectin are also increased in replicative and prematurely senescent fibroblasts [607]. Whilst the core expression of SASP molecules are well defined, SASPs compositions vary depending on the initiator of DNA damage and senescence with the exception of p16INK4-mediated senescence, which does not result in the development of a SASP. However, SASPs develop very rapidly after the onset of senescence [603].

During senescence cells undergo many morphological changes. Whilst the understanding of the impact of senescence on the F-actin cytoskeleton is poorly understood, the actin cytoskeleton undergoes profound rearrangements in senescent cells [608]. While senescent cells can have heterogeneous and highly variable phenotypes, common hallmarks of senescence include an irregular shaped body, enlarged cell size and generally flattened phenotype, and growth arrest (Hayflick limit) [581, 609].

6.1.1 Speculative role of *TNC* in cell senescence/proliferation pathways

TN-C can regulate many targets, indirectly via miR-155. These targets include both pro-inflammatory and anti-inflammatory proteins that act to regulate TLR4 signalling, as discussed earlier around the TN-C/miR-155/TNF-α axis (chapter 5) [188, 311]. Other targets are related to cell cycle progression and thus aging and senescence, although how specifically it is not clear. In macrophages, miR-155 downregulates Src homology 2 (SH2) domain-containing inositol 5'-phosphatase 1 (SHIP1) by direct interaction with its 3' UTR [610, 611]. SHIP1 is a PI3K phosphatase and therefore functions as tumour suppressor protein by modulating kinase signalling [612]. SHIP1 negatively regulates Akt, a target kinase downstream of the PI3K pathway, thereby inhibiting cell growth and migration [613]. Correspondingly, SHIP1 downregulation by miR-155 results in increased Akt kinase activity [611]. In support of this anti-proliferative role of SHIP1, and oncogenic function of miR-155, in B-cell lymphomas it has been established that miR-155-mediated downregulation of SHIP1 results in increased cell proliferation in a TNF-α-dependent manner [614, 615]. Furthermore, SHIP1 colocalises

with p53, indicating another layer of complexity in the role of this protein in cell cycle progression [616]. Increased Akt activation also leads to increased NF-kB activation, and consequently transcription of a host of inflammation related proteins. To add further complexity to this arm of regulation, Akt also negatively regulates miR-155 in murine macrophages in response to LPS [617]. This is indicative of a negative feedback loop, which, given the involvement of these proteins in the inflammatory response, likely serves a protective mechanism to halt sustained inflammation. Linked to this, miR-155 expression decreased collagen expression via targeting the HIF-1α/PI3K/Akt pathway in fibroblasts[618]. However, another study of cocultures of fibroblasts and macrophages found the opposite. Specifically, it found that miR-155 downregulation leads to reduced proliferation and collagen synthesis in fibroblasts via targeting of SHIP1 and subsequent modulation of the PI3K/Akt pathway [619]. This effect was specifically due to attenuation of IL-1β and TGF-β1 secreted by macrophages. Differences may have arisen from the different culture methods. Indeed, it is widely accepted that miR-155 drives fibrosis [620] and that fibrosis can be age-related disorganisation of the tissue maintenance process.

Specifically, Akt activity is tightly linked to cell proliferation and, by extension, senescence. The upstream inputs directly dictate what the kinase will do. For example, a growth factor will increase proliferation and reduce apoptosis [621], whereas a stress inducer will downregulate Akt signalling [622]. It has been shown that in human dermal fibroblasts (HDFs) and mouse embryonic fibroblasts (MEFs), Akt knockdown results in slower growing cells and is sufficient to delay senescence [623]. The PI3K/Akt pathway regulates cell proliferation and cell survival in part by phosphorylation of Akt's downstream target Forkhead box transcription factors (FOXO). There are many types of FOXO factors that are ubiquitously expressed in mammals in certain tissue niches [624]. FoxO phosphorylation by Akt inactivates FoxO, resulting in cytoplasmic accumulation and therefore inaction. If this phosphorylation does not occur, nuclear localisation of FoxO leads to cell cycle suspension at G1 and this is via transcriptional activation of the cyclin dependent kinase inhibitors (CDKi) $p27^{KIP1}$ and $p21^{WAF1/CIP1}$ [625, 626]. Quiescence, with the right signals such as low activation of p53, can then lead to

senescence [627]. This anti-proliferative nature of FOXO proteins is highlighted in triple knockout mice, which lack three principle FOXO members and develop lymphoproliferative diseases [628].

Very recently, FoxO3a, a negative regulator of Akt, has been shown to positively regulate TN-C expression in IL-1β treated rabbit chondrocytes [629]. FoxO3a has been extensively studied in regard to its role in inflammation and diseases such as cancer and fibrosis [630]. It has also been established that miR-155 regulates FoxO3a in gliomas to control cell proliferation by directly binding its 3' UTR [631].

Together, these studies indicate a complex signalling pathway that is regulated at each layer. The effects of the PI3K/Akt signalling pathway on cell cycle progression and aging, which are also linked to cancer development, are context dependent and highly variable; this level of detail is not covered here but has been extensively reviewed [632, 633].

Discerning the role of TN-C in the regulation of cell proliferation/senescence requires a thorough investigation and the expression of the genes and protein mentioned above are candidates that should be investigated in a genotype and age-related manner. In the present study, fibroblasts at different ages with a *TNC* gene knockout or fully expressing *TNC* were used. Several hallmarks of senescence were investigated in these cells; morphological changes, such as cell shape and size with regard to the actin cytoskeleton; nuclear morphological changes; the expression of senescence associated β-galactosidase (SA-BG) and the expression of SASP molecules. The previous studies indicate that *TNC* mediated premature senescence could occur through miR-155 expression acting directly through SHIP1 and FoxO3a to modulate the PI3K/Akt pathway.

6.2 RESULTS

6.2.1 Partial and complete *TNC* KO cells exhibit hallmarks of senescence by light microscopy

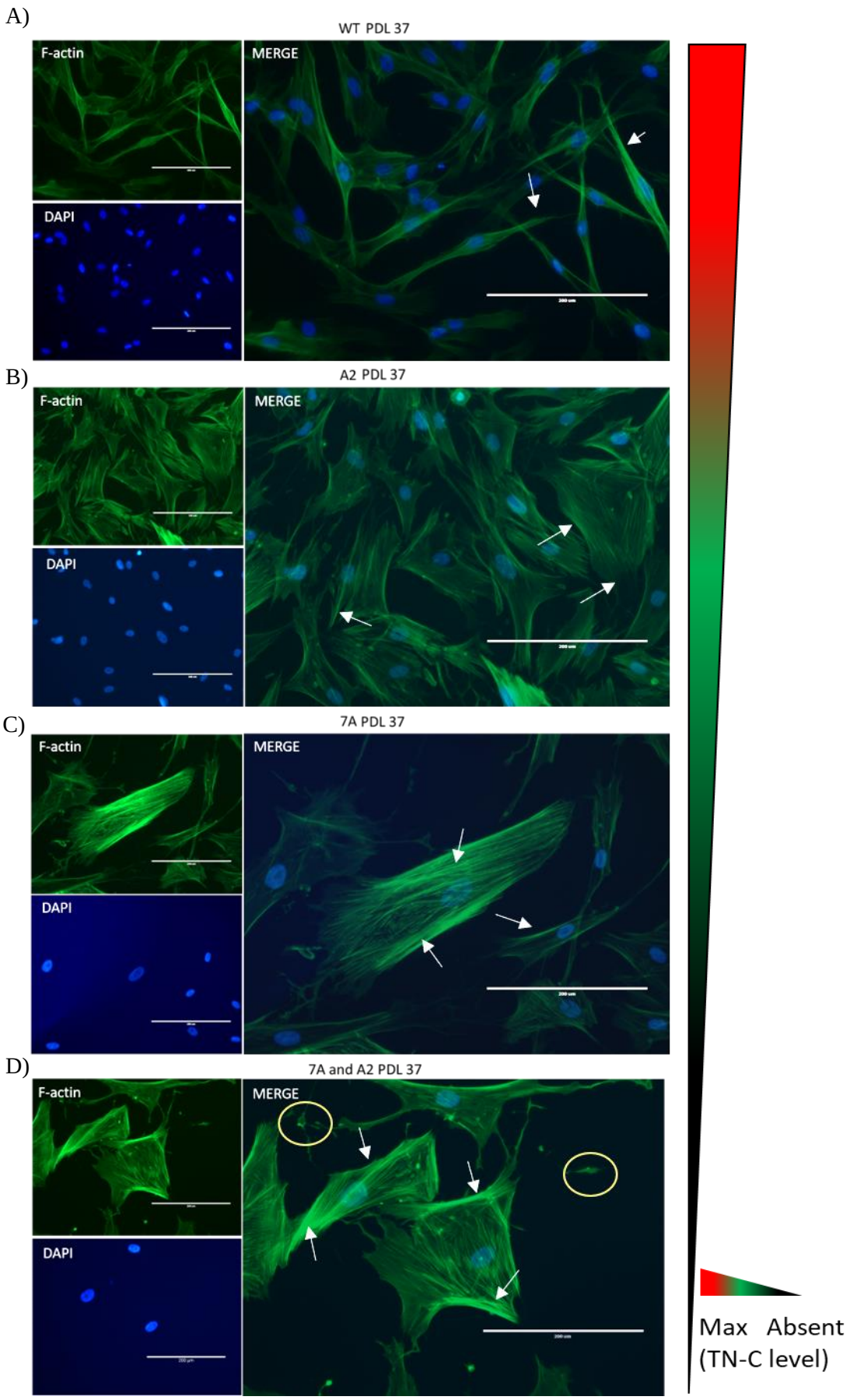
The pools of cells generated utilising CRISPR/Cas9 RNP, named after the sgRNAs they were transfected with (A2, 7A, 7A and A2), and the complete *TNC* KO cells had reduced or almost absent growth and were typically much larger in size, two hallmarks of senescence. Although the 'A2' (62%

TNC KO) fibroblasts had reduced growth compared to WT fibroblasts, as observed by light microscopy, their growth was increased compared to the '7A' (79% *TNC* KO) and '7A and A2' (79% *TNC* KO) fibroblasts. This is reflected in the cell density of the A2 cell population compared to the '7A' and '7A and A2' cell populations (Figure 75). Normal human fibroblasts acquired from the ATCC have a maximum of 72 PDLs before the onset of senescence. The CRISPR/Cas9 RNP partial *TNC* KO fibroblasts had a PDL of 37 before the cell growth slowed or became absent and this was not the case for WT fibroblasts at PDL 37. This slowed growth was also apparent in the *TNC* KO fibroblasts at PDL 55. WT fibroblasts at PDL 55 exhibited a slight decrease in growth rate, however, while this would be expected with aging, this did not occur to the same extent as the *TNC* KO cells in which the time taken for 1 PDL nearly doubled. The CRISPR/Cas9 *TNC* KO then exhibited a complete cessation of growth at PDL 63. The PDL was far higher in the latter cell line due to the fact that cells were single sorted and therefore this cell line originated from one cell requiring far more cell doublings for expansion of sufficient numbers of cells. Furthermore, the CRISPR/RNP cell populations and the *TNC* KO populations were difficult to image by light microscopy due to their flatness and enlarged size compared to WT cells. Due to the fact that the CRISPR/Cas9 RNP partial *TNC* KO populations exhibited the same phenotype with many characteristics of senescence, at a far earlier PDL than the CRISPR/Cas9 *TNC* KO cell line, premature senescence and not replicative senescence was predicted to be the likely culprit.

6.2.2 Cell age and TN-C affect actin cytoskeleton rearrangement

To start to test the hypothesis that *TNC* plays a role in the regulation of cellular senescence, the F-actin cytoskeleton of these cells was stained with phalloidin and imaged, based on the fact that the cytoskeleton undergoes profound rearrangements in senescent cells [608]. The F-actin cytoskeleton was analysed in both "young" and "old" CRISPR/RNP and *TNC* KO populations, and was compared to that of WT cells at the same age, which was tracked by their PDL. A detailed account of this analysis is reported in the next section. Briefly, clear differences between the actin cytoskeleton were observed between PDL 37 (young) WT cells and young CRISPR/Cas9 RNP cells (Figure 75). Differences between the complete *TNC* KO fibroblasts at PDL 55 (old) and old WT fibroblasts were also observed, although the differences were less striking, but nonetheless apparent (Figure 76). The cell images in

this chapter are presented in order from most to least TN-C expression: WT, A2/7A, 7A and A2, and *TNC* KO as semi-quantified by western blotting in chapter 3; in this way, a dose-response effect of *TNC* gene and age on cell phenotype dysregulation can be seen.


A)
WT PDL 37
F-actin
MERGE
DAPI
B)
A2 PDL 37
F-actin
MERGE
DAPI
C)
7A PDL 37
F-actin
MERGE
DAPI
D)
7A and A2 PDL 37
F-actin
MERGE
DAPI
Max Absent
(TN-C level)

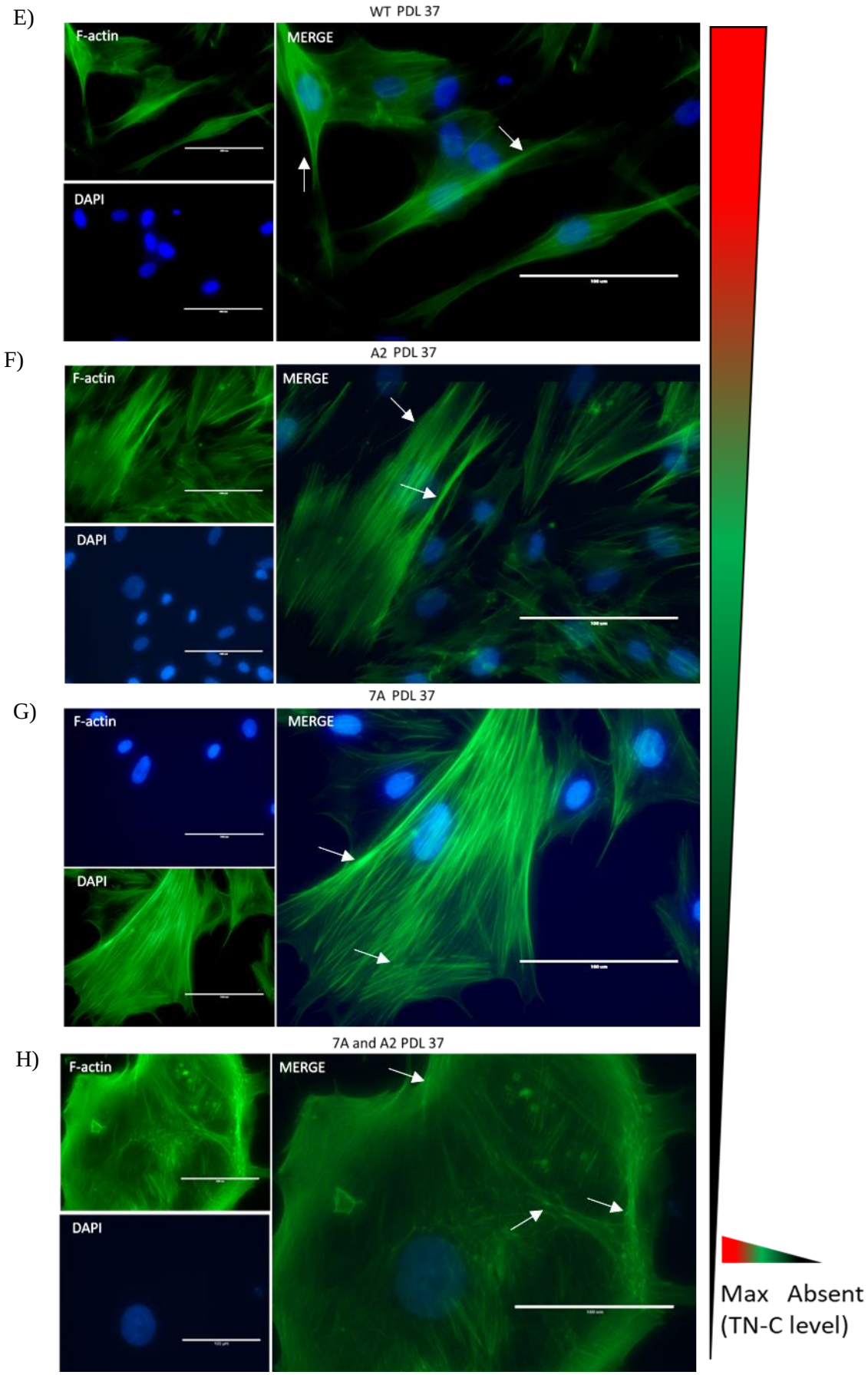

E)
WT PDL 37
F-actin
MERGE
DAPI
F)
A2 PDL 37
F-actin
MERGE
DAPI
G)
7A PDL 37
F-actin
MERGE
DAPI
H)
7A and A2 PDL 37
F-actin
MERGE
DAPI
Max
Absent
(TN-C level)

Figure 75. 'Young' CRISPR-Cas9/RNP partial *TNC* KO BJ fibroblasts at PDL 37 exhibit a dysregulated F-actin cytoskeleton compared to their 'young' WT counterparts.
Partial *TNC* KO fibroblasts were cultured and monitored by light microscopy for slowed/absent growth. Cells were fixed and fibrous actin (F-actin) and DNA stained with phalloidin and DAPI, respectively. CRISPR RNP cell populations are named after the CRISPR guides they were transfected with. **A)** and **E)** untransfected WT BJ fibroblasts at PDL 37, **B)** and **F)** 'A2' CRISPR/Cas9 RNP cells at PDL 37, **C)** and **G)** '7A' CRISPR/Cas9 RNP cells at PDL 37, **D)** and **H)** '7A and A2' CRISPR/Cas9 RNP at PDL 37. Scale bar: A-D, 200 µm, E-H, 100µm. White arrows indicate stress fibres. Representative images from each magnification and cell population are shown.

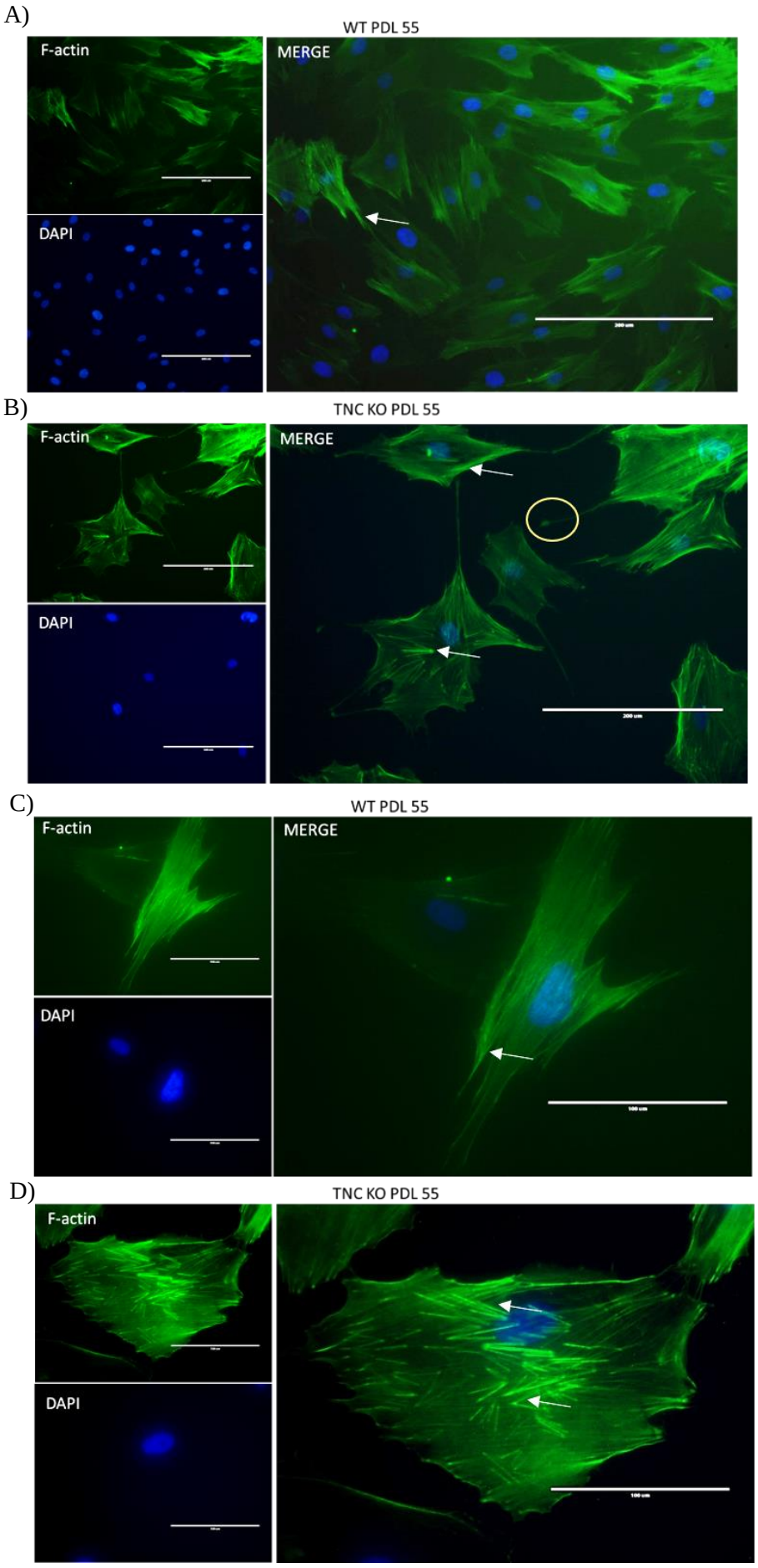

A)
WT PDL 55
F-actin
MERGE
DAPI
B)
TNC KO PDL 55
F-actin
MERGE
DAPI
C)
WT PDL 55
F-actin
MERGE
DAPI
D)
TNC KO PDL 55
F-actin
DAPI

Figure 76. 'Old' *TNC* KO BJ fibroblasts at PDL 55 exhibit a dysregulated F-actin cytoskeleton compared to their 'old' WT counterparts.
Partial and complete *TNC* KO fibroblasts were cultured and monitored by light microscopy for slowed/absent growth. Cells were fixed and fibrous actin (F-actin) and DNA stained with phalloidin and DAPI, respectively. **A)** and **C)** untransfected WT BJ fibroblasts at PDL 55, **B)** and **D)** *TNC* KO fibroblasts at PDL 55. Scale bar: A-B 200 µm, C-D 100µm. White arrows indicate stress fibres. Yellow circles highlight trailing actin cytoskeleton or cell. Representative images from each magnification and cell population are shown.

6.2.2.1 Cell age and TNC affect actin network formation

Visual analysis of the phalloidin-stained actin cytoskeleton shows that its arrangement within fibroblasts appears strikingly different depending on both age of the cells and *TNC* expression levels. Actin stress fibers appear more prevalent, pronounced (Figure 75, Figure 76 white arrows) and localised in different cellular regions, when comparing CRISPR-edited fibroblasts to WT cells. In young WT cells, stress fibers appear oriented in the elongated protruding cell appendages, however in the young partial *TNC* KOs these stress fibers appear throughout the cell and perpendicular to other stress fibers giving the appearance of a lattice. In old WT cells stress fibers appear similarly to those in young WT fibroblasts. In old *TNC* KO fibroblasts, stress fibers appear more pronounced and more frequently, mirroring the phenotype displayed by partial *TNC* KO fibroblasts. Increased stress fiber formation has been demonstrated in senescent human foreskin fibroblasts at PDL 60, in which these fibers were localised to the edges of the cells, rather than centralised [634]. Here, stress fibers appear thick and well aligned, something that has previously been demonstrated when plating fibroblasts on rigid surfaces [635]. An exception to this can be seen for the '7A and A2' fibroblasts (Figure 75 H) in which a particularly enlarged cell has showed loss of dense actin bundles. It may be that there is a similar actin content to other enlarged cells, just spread over a larger area; there is a limit to scaling gene expression in larger cells [636]. In the partial A2 *TNC* KO fibroblasts, the cytoskeleton on one cell appears to be linked to the cytoskeleton of a neighbouring cells (Figure 75 F), or it is very difficult to distinguish where one cell boundary ends and another begins. Actin cytoskeletons can become linked by adherens junctions to the neighbouring cells in endothelial, epithelial, mesenchymal and neural cells, whilst tricellular adherens junctions are common in all cell monolayers [637]. Parallel actin fiber bundles running perpendicular to the membrane, as seen in figure 2 F in particular, have been demonstrated in normal rat kidney fibroblasts. At the end of these fibers and at the tip of cellular processes, P-cadherin,

α-catenin, ZO-1 and vinculin were located indicating small adherens junctions [638]. It may be that actin cytoskeleton linkage in these cells is altered compared to WT cells, however no staining was carried out for specific proteins to assess this. Moreover, only in the *TNC* KO and '7A and 7B' partial *TNC* KO trailing cytoskeleton can sometimes be observed separate to a cell (Figure 75, yellow circles). Together, this indicates very prominent actin cytoskeleton dysregulation that results in an enlarged cell body and loss of morphological characteristics, which may be indicative of senescence, that occur in a TN-C and age dependent manner.

6.2.2.2 Age and TN-C genotype affect the cell phenotype of fibroblasts

Quantification of F-actin staining in the images presented in the previous section was performed using ImageJ to allow analysis of other cell features. Fibroblasts display a characteristic elongated spindle shape [639] and this was evident for both young and old WT fibroblasts (Figure 75 A and Figure 76 A). A score of 1 for 'circularity' in ImageJ indicates a perfect circle, and deviation away from 1 indicates an elongated spindle shape. The young partial *TNC* KO 'A2', '7A' and '7A and A2' fibroblasts exhibit a significantly rounder cell morphology ($P \leq 0.0001$) than young WT cells, as do the old WT and *TNC* KO fibroblasts ($P < 0.0001$: Figure 77). The old *TNC* KO fibroblasts were significantly rounder than the old WT fibroblasts ($P = 0.0003$), the 'A2' fibroblasts ($P = 0.062$) and '7A' fibroblasts ($P = 0.0343$) yet not the '7A and A2' fibroblasts ($P = 0.144$: Figure 77). Compared to the young partial *TNC* KO A2 fibroblasts ($P = 0.973$), 7A fibroblasts ($P = 0.788$) and '7A and A2 fibroblasts' ($P = 0.436$), the aged WT fibroblasts were similarly round (Figure 77). The similarities between old WT fibroblasts and young partial *TNC* KO fibroblasts and the similarity between young partial *TNC* KOs and old *TNC* KOs suggest both an age and *TNC* genotype effect; fibroblasts lose their characteristic spindle shape with decreasing TN-C expression and increasing age.

Accompanying the reduction in elongated morphology, the young partial *TNC* KO also displayed a larger cell area (μm^2) than young WT fibroblasts (Figure 77 A). Again this occurred in a TN-C expression level dependent manner with 'A2' cells presenting the smallest increase in area ($P = 0.068$), followed by '7A' ($P < 0.0001$) and '7A and A2' ($P= 0.0001$). This *TNC* genotype-dependent cell size

increase was also observed in old *TNC KO* cells, which were significantly larger than old WT cells ($P < 0.0001$) and young WT cells ($P < 0.0001$). The '*A2*' fibroblasts were significantly smaller than old *TNC* KO fibroblasts ($P < 0.0001$), yet the '7A' and '7A and A2' fibroblasts were not significantly different in size ($P = 0.729$ and 0. 773, respectively). This size increase appeared specifically dependent on *TNC* genotype, not age, given that old WT fibroblasts were not significantly larger than young WT fibroblasts ($P = 0.961$) and significantly smaller than the young '7A' and '7A and A2' partial TN-C fibroblasts ($P = 0.0001$) (Figure 77). Whilst there are obvious differences in size between the largest partial *TNC* KO and WT fibroblasts, some of the cell areas from the A2 partial *TNC* KOs were in range of 'normal' WT cell areas. This may be explained by the fact that these 'A2' partial KO cells expressed some residual TN-C due to these populations only being partial knockouts. In further support of a TN-C dependent effect, the old *TNC* KO fibroblasts did not display any cells with cell areas in the range of that of old WTs.

As expected, echoing the increased cell size, the cell perimeter (Figure 77 B) was also significantly larger in the A2 fibroblasts ($P = 0.0454$), even larger in the '7A' fibroblasts ($P= 0.0001$) and the largest in '7A and A2' fibroblasts ($P = 0.0001$). This perimeter difference was also apparent in the old *TNC* KO fibroblasts which were significantly larger than old WT fibroblasts ($P < 0.001$).

Together it appears that the fibroblast populations can be distinguished into two groups based on their phenotype: WT 'young' and 'old' cells (group 1), and the partial and complete knockouts (group 2). However, this clear division into groups is slightly lost when comparing the cell populations for spindle shape, inferred by roundness, in which WT 'young' fibroblasts would be grouped alone and partial and complete KO fibroblasts would be grouped with 'old' WT fibroblasts. Generally, it appears a dose-response of cell size increase occurs, with decreasing TN-C expression. These results also indicate that age related changes include loss of spindle shape and that a lack of TN-C expression is correlated with the loss of spindle shape at a much younger cell age, again in a *TNC* gene dose-dependent manner.

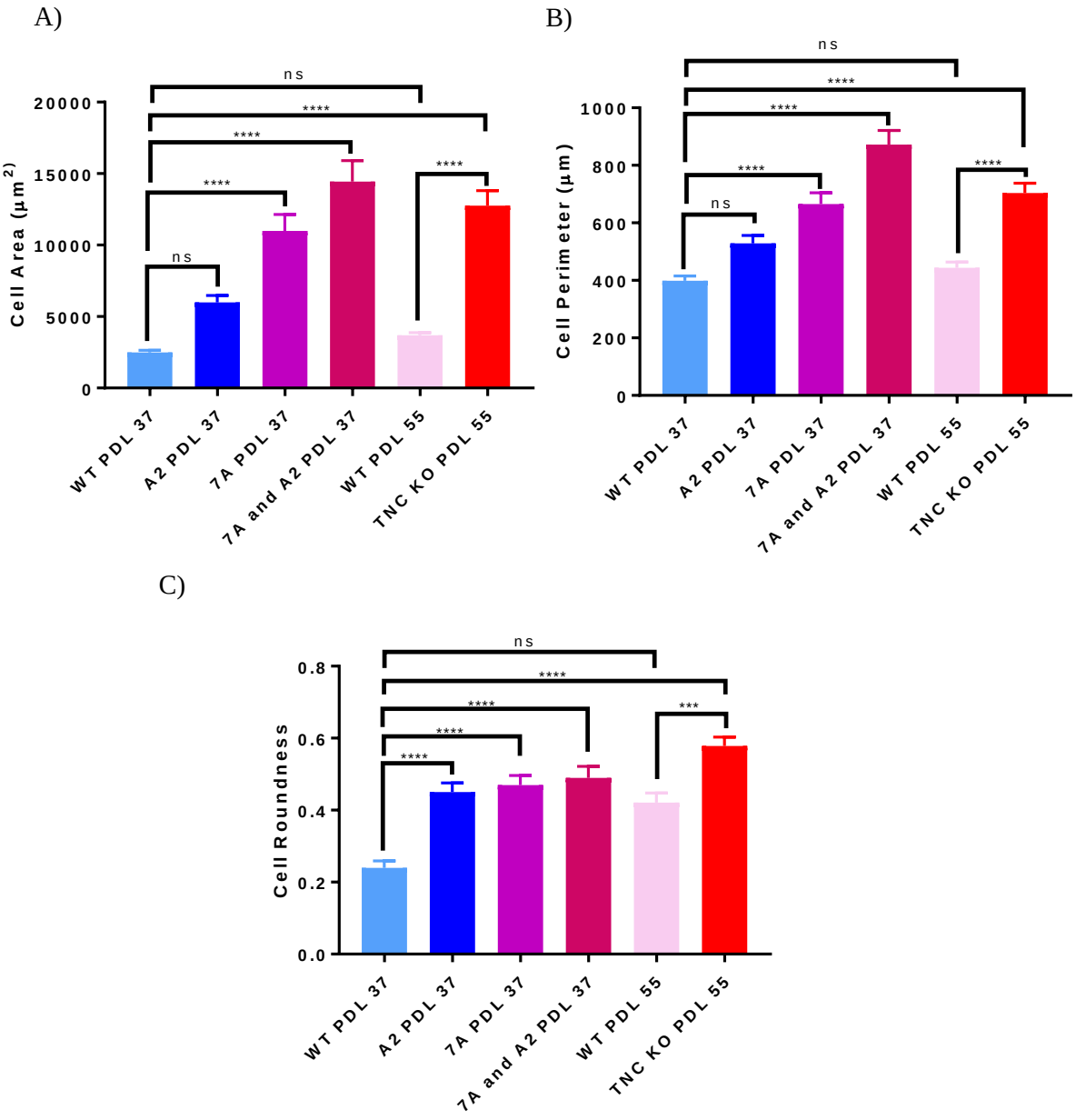


Figure 77. Quantification of phalloidin staining of the F-actin cytoskeleton revealed cell morphology is significantly altered in CRISPR/Cas9 RNP partial and complete *TNC* KO fibroblasts compared to their WT counterparts.
WT PDL 37 (light blue), CRISPR/Cas9 RNP cell populations: A2 (dark blue), 7A (purple) and 7A and A2 (dark pink), WT PDL 55 (light pink) and *TNC* KO (red). Cell area (µm2) (A), cell perimeter (µm) **(B)** and cell roundness **(C).** Images were analysed and data collected using ImageJ. Mean +/- SEM. N = 50, One-way ANOVA with Tukey's multiple comparison test to compare each group with every other group.

6.2.3 Nucleus morphology and location is altered in an age and TN-C genotype dependent fashion

Next, quantification of the nucleus using ImageJ was utilised to assess any nuclear morphological changes in the partial and complete *TNC* KO fibroblasts. Analysis of the DAPI stained nuclei demonstrated differences in the nucleus morphology of the different fibroblast cell populations (Figure 78). Young WT fibroblasts display the usual oval nucleus central to the spindle [639]. While in most cells of the CRISPR populations it is difficult to discern if the nucleus is in the centre, as the spindle shape is lost and cells exhibit a much rounder appearance, in some cells it is evident that the nucleus is localised off-centre (Figure 75 H, Figure 76 D). All partial *TNC* KO cells exhibited a significantly rounder nucleus than young WT fibroblasts (A2; $P = 0.0008$, 7A; $P < 0.0001$, 7A and A2; $P = 0.026$) (Figure 78 C). The old WT and old *TNC* KO nuclei were also significantly rounder than the young WT nuclei ($P < 0.0001$). However, the old WT fibroblasts were not significantly rounder than *TNC* KO fibroblasts nuclei ($P = 0.455$). This indicates that although nucleus rounding appears to be an age-related function, it is also confounded by the *TNC* genotype.

Regardless of location of the nucleus, the fibroblast populations could be grouped for nuclear phenotype as per the cell phenotypes; with WT and 'A2' fibroblasts in one group and '7A', 'A2 and 7A' and *TNC* KO fibroblasts together in a second group, in relation to nuclear size and perimeter. The 'A2' cells displayed a marginally larger nucleus area ($P = 0.188$) and perimeter ($P = 0.013$) than young WT fibroblasts. The '7A' and '7A and A2' nuclei were significantly enlarged in both nucleus area ($P < 0.0001$) and perimeter ($P < 0.0001$) compared to young WT fibroblasts. The old WT fibroblasts did not have nuclei significantly larger than young WT fibroblasts ($P = 0.984$) and this was true for the perimeter also ($P = 0.997$), suggesting that the *TNC* genotype, rather than the age span investigated, affects these features. In support of this, the old *TNC* KO fibroblasts displayed a significantly larger nucleus ($P < 0.0001$) and perimeter ($P < 0.0001$) compared to both old and young WT fibroblasts and compared to the '7A' ($P = 0.025$) and '7A and A2' ($P = 0.0037$) fibroblasts nuclei (Figure 78). Whilst the link between nucleus enlargement and roundness is not clear, it does appear that nuclear enlargement occurs with aging and that *TNC* genotype may override this age-related regulation.

Together, these populations may provide clues as to the mechanistic way in which TN-C contributes to regulation of these phenotypes. For example, in the young partial *TNC* KOs, and in the complete *TNC* KOs it appears that the fibroblasts displaying the most dysregulation of features are those in which exon 7 was targeted via CRISPR, as opposed to exon 2. Given that exon 7 contains the integrin-specific 'RGD' motif, it may be that TN-C regulates actin assembly and nuclear behaviour via interaction with integrins.

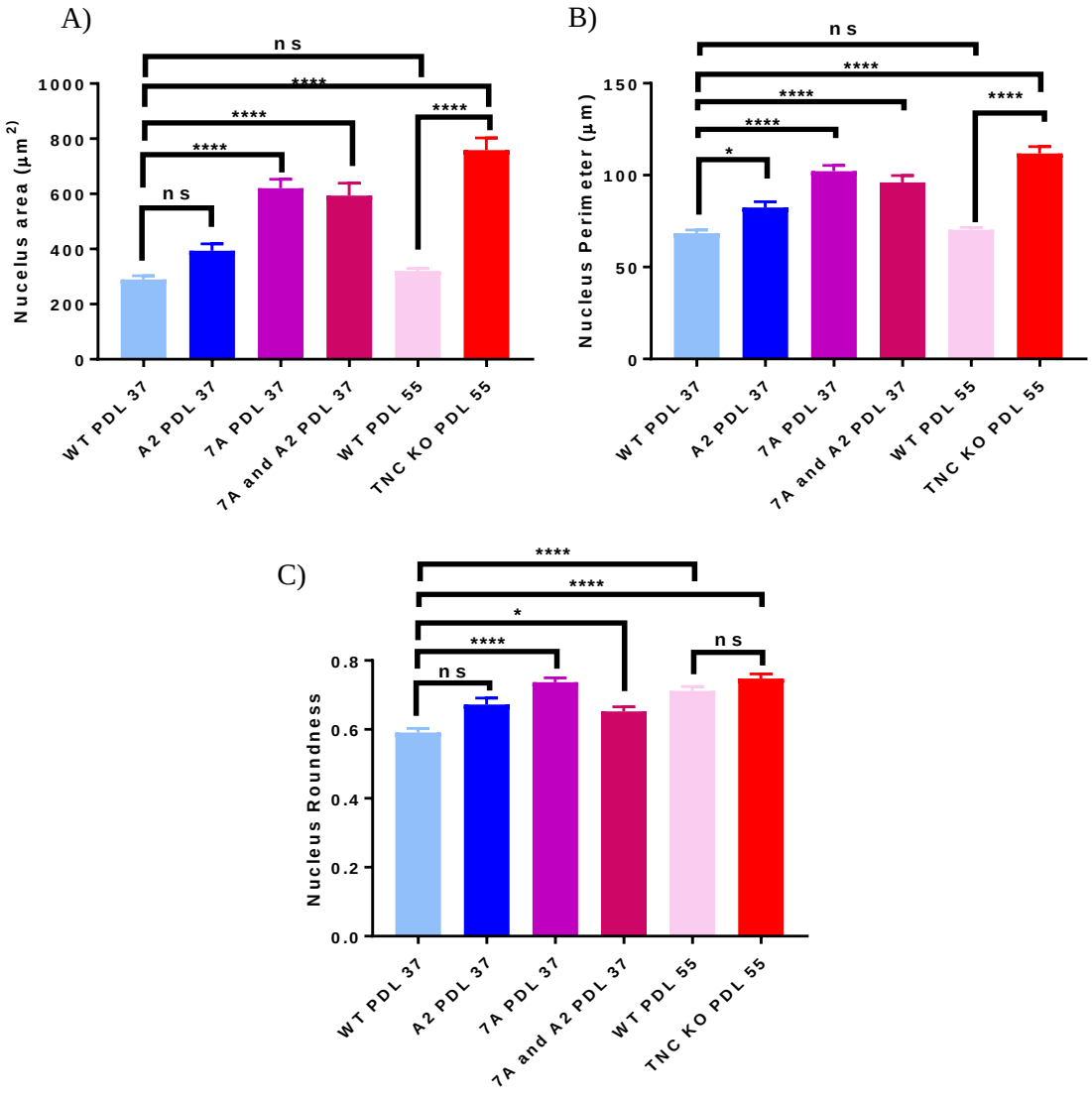


Figure 78. Quantification of DAPI staining of the nucleus reveals nucleus morphology is significantly altered in CRISPR/Cas9 RNP partial and complete *TNC* KO fibroblasts compared to their WT counterparts. WT PDL 37 (light blue), CRISPR/Cas9 RNP cell populations: A2 (dark blue), 7A (purple) and 7A and A2 (dark pink), WT PDL 55 (light pink) and *TNC* KO (red). Cell area (µm^2) **(A),** cell perimeter (µm) **(B)** and cell roundness **(C).** Images were analysed and data collected using ImageJ. Mean +/- SEM. N = 50, One-way ANOVA with Tukey's multiple comparison test to compare each group with every other group.

6.2.4 Senescent associated heterochromatin foci (SAHF) are not present in complete or partial *TNC* KO fibroblasts

As well as general assessment of nuclear morphology, DAPI staining also facilitates chromatin morphology analysis. Chromatin in the nucleus of senescent cells is remodelled to form senescence-associated heterochromatin foci (SAHF) [640]. These are regions of facultative heterochromatic DNA that lead to the silencing of pro-proliferation genes in senescent cells. Such regions are observed as darker punctate regions within the nuclei following DNA staining [640]. SAHF are specifically found in senescent nuclei and distinguishable from other constitutive heterochromatin- such as pericentromeric and telomeric heterochromatin [641]. As with many senescent characteristics, this is not a ubiquitous feature of all senescent cells and is cell-type specific as well as senescence-induction type specific [642]. Whilst some darker staining regions are observed (Figure 79) the punctate regions are not as defined as the SAHF regions shown in previous studies on IMR-90 fibroblasts [640]. This is in line with previous studies which demonstrate SAHF do not form in BJ fibroblasts [642].

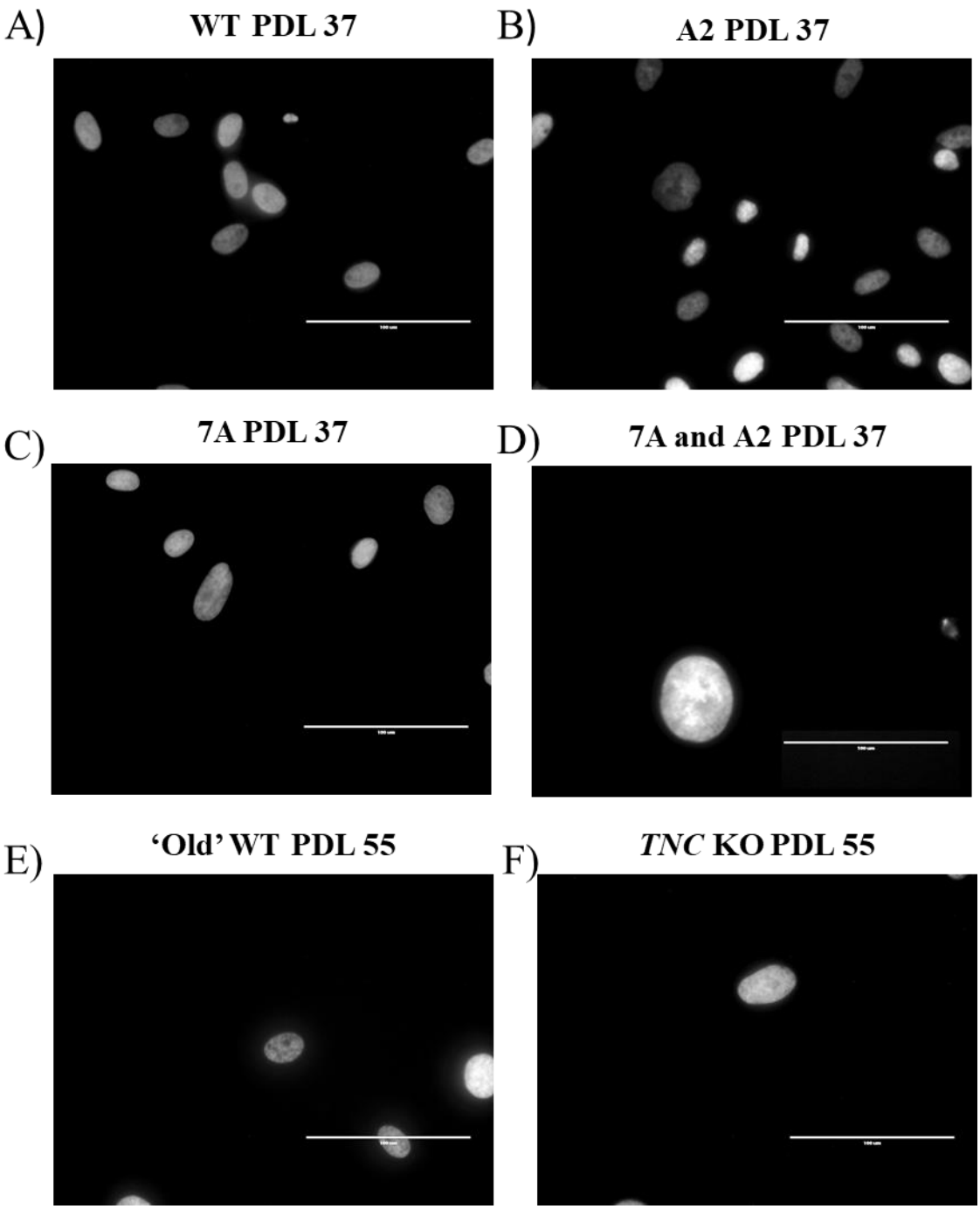


Figure 79. SAFH do not form in BJ fibroblasts regardless of cell age or *TNC* gene expression status. Images of cell nuclei stained with DAPI from: **A)** young WT fibroblasts, **B)** young partial knockout A2 fibroblasts, **C)** young partial knockout 7A fibroblasts, **D)** young partial knockout 7A and A2 fibroblasts, **E)** old WT fibroblasts, and **F)** old complete KO fibroblasts. Scale bar: 100 µm.

6.2.5 The relationship between nuclear and cell size is abrogated in a *TNC* genotype-specific manner

Correlation analysis of cell and nuclear area to probe the relationship between the two phenotypes further demonstrated that a nuclear size cap seems to occur in WT cells that is then progressively lost in a *TNC* gene-dose-dependent manner (Figure 80). In young fibroblasts there is no correlation ($r = 0.082$; $P = 0.574$) between cell and nucleus area. The direction of the relationship is limited by the apparent size limit on the nucleus size; only slight increases in the nucleus size occur regardless of the increase in cell size. A positive correlation ($r = 0.720$; $P < 0.0001$) is demonstrated in 'A2' fibroblasts, in which an increase in cell area is generally accompanied by an increase in nucleus area, although some exceptions can be seen. This indicates that the maximum nucleus and cell size may still be regulated in some fibroblasts within this population. Out of all the CRISPR-treated cells, this cell line expressed the most TN-C, indicating that the *TNC* genotype is likely to be involved in the regulation of these phenotypes. Similarly, a positive correlation occurs in the '7A' fibroblasts ($r = 0.635$; $P < 0.0001$) and '7A and A2' fibroblasts ($r = 0.812$; $P = 0.0001$) in which cell enlargement is generally accompanied by nuclear enlargement. These two populations exhibit much larger cell and nucleus sizes compared to both the 'A2' fibroblasts and young WT fibroblasts and express less TN-C than the A2 and WT fibroblasts. The relatively proportional increase in size of both nucleus and cell area indicates that size regulation for both cell aspects has been lost. Just as in the young WT fibroblasts, old WT fibroblasts do not display a relationship ($r = 0.121$, $P = 0.402$) between cell and nuclear size; a limit still seems to occur on nuclear size in aged cells. This size cap appears to be slightly larger in older fibroblasts compared to younger fibroblasts, yet far smaller than in any CRISPR-treated fibroblasts. Correspondingly, the old *TNC* knockout fibroblasts show a strong positive relationship ($r = 0.705$, $P = <0.0001$) between nuclear and cell area supporting the theory that CRISPR treatment and the resulting *TNC* gene knockout or reduction, in the case of CRISPR RNP-treated cells, influences both cell and nuclear size regulation, although to what extent these observation are confounded by age is unknown. In terms of cell and nuclear size it appears that area and perimeter are influenced by TN-C in a gene-dose-response way, whereas cell and nuclear roundness are also influenced by age.

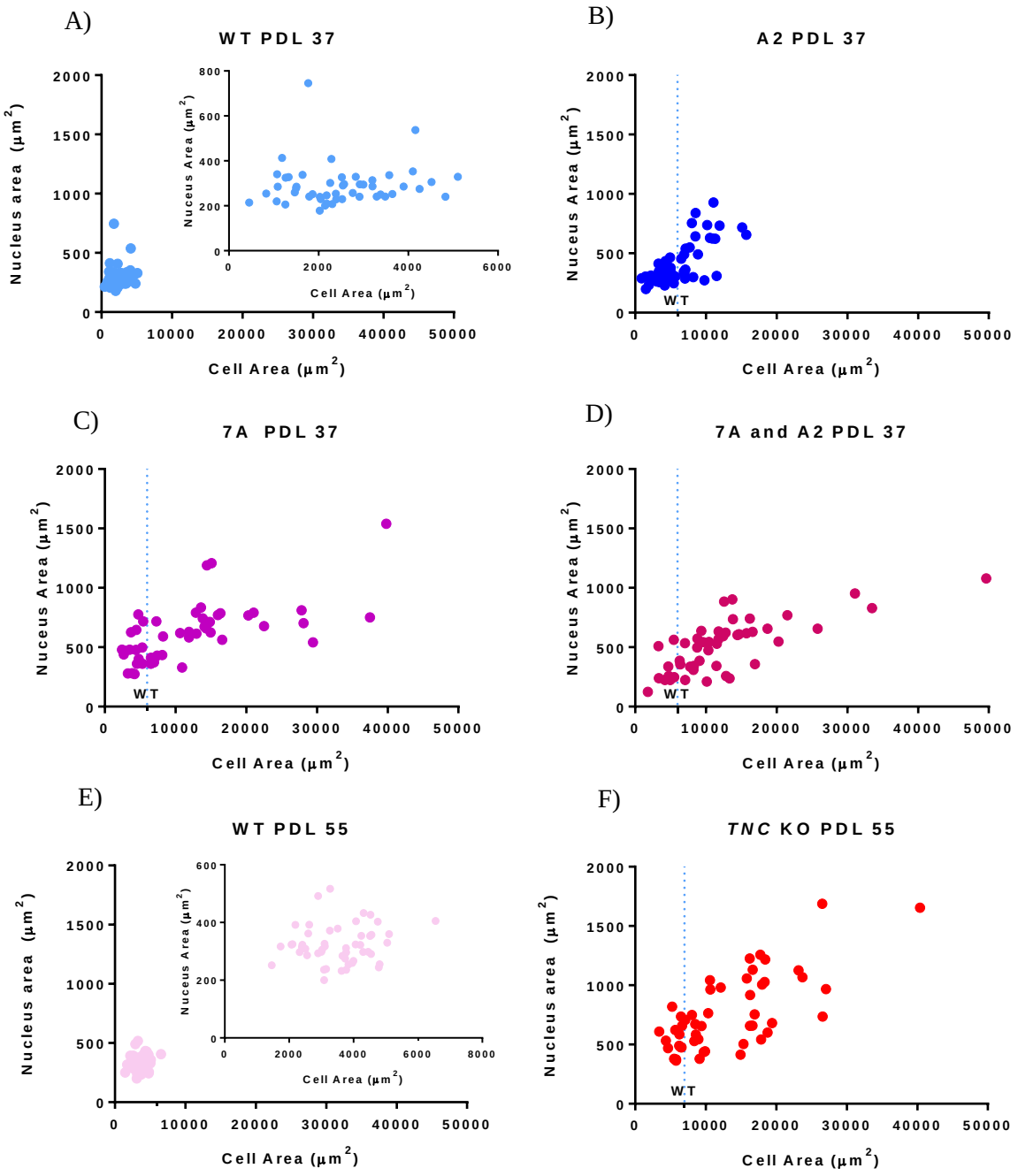


Figure 80. Partial and complete *TNC* KO fibroblasts lose the nucleus size limit and have an altered relationship between cell area and nucleus area.

The relationship between cell area and nucleus area was investigated by measuring the area (µm^2) of phalloidin stained fibroblasts and their corresponding DAPI stained nucleus in ImageJ. WT cells **(A)** or partial *TNC* KO CRISPR/Cas9 RNP transfected fibroblasts 'A2' **(B)**, '7A' **(C),** '7A and A2' **(D)** at PDL 37, or WT fibroblasts and complete *TNC* KO fibroblasts at PDL 55 were assessed. The vertical dashed blue line on graph B, C and D indicates where the WT PDL 37 cell area size limit occurs. The vertical blue dashed line on graph **F)** indicates were the size area limit occurred for WT PDL 55. As figure A) and E) cell and nuclei sizes are smaller than in other panels, clarity of the sample point measurements are lost when utilising scale bars matching the other panels, therefore inset panels display the same data with a smaller more representative scale. Spearman's correlation, N =50.

6.2.6 β-galactosidase staining for senescence indicates different levels of senescent cells in old WT and old *TNC* KO fibroblasts

Following the morphological assessment which supports the hypothesis that partial and complete *TNC* KO cells undergo senescence earlier than WT cells, senescence associated (SA) β-galactosidase (βG) assay was next used to detect senescence biochemically. Only the complete *TNC* KO cells were analysed from here on, as this would allow the most clear analysis when compared with WT fibroblasts, given that no TN-C was expressed. The other cell populations would of course be useful, specifically in providing evidence of receptor involvement given the different CRISPR/Cas9 target loci. β-galactosidase is a family of enzymes located in the lysosome that is reported to be specifically expressed by senescent cells at pH 6.0 in contrast to acidic βG present in all proliferating cells at pH 4.0 [643]. The expression of this specific SA-βG is correlated with aging in many cell types, tissues and organisms, including skin fibroblasts [643], mouse embryonic fibroblasts [644], Hela cells, hepatocytes[645], human and rabbit blood vessels [646] and *C. elegans* [644]. This method is now cited in thousands of peer-reviewed scientific articles as the most commonly used method to verify cellular senescence [647, 648]. Old WT fibroblasts, old *TNC* KO fibroblasts and young WT fibroblasts were seeded at 1 x 10^4 cells in 12 well plates and cultured for 24 hours before staining for SA-βG activity (Figure 81). Young WT fibroblasts were also treated with H_2O_2 and cultured for a further 72 hours to induce senescence before SA-βG staining. H_2O_2 treatment is known to cause oxidative stress and induce stress-induced premature senescence (SIPS) in fibroblasts 3 days after treatment [649, 650] and was therefore the positive control. Fibroblasts treated with H_2O_2 displayed very frequent and intense staining in the majority of cells indicating a successful assay. Young WT fibroblasts displayed infrequent staining, which is to be expected, however this was no near as prevalent or intense as in the positive control, or in the old WT fibroblasts. As young WT fibroblasts continue to double during *in vitro* culture, this staining is expected to increase in intensity and prevalence. In the old WT fibroblasts SA-βG staining was very common and intense, and slightly reduced compared to SA-βG staining in the positive control. A stark difference was apparent in the level of staining between the WT and *TNC* KO fibroblasts, in which KO fibroblasts displayed more

prevalent and larger stained areas – possibly linking to slightly enlarged nucleus size. Critically, cell confluence can affect the results of this assay. Specifically, if cells are too confluent then false positive staining can occur hence cells were seeded at lower densities than would usually be utilised for cell experiments. The slightly lower confluence of fibroblasts in the H_2O_2 condition resulted from some cells dying during the treatment. Overall, SA-BG staining indicated that whilst both 'old' WT and *TNC* KO cells exhibited staining, there was far more staining in the *TNC* KO cell line, indicative of higher levels of senescence within this cell line.

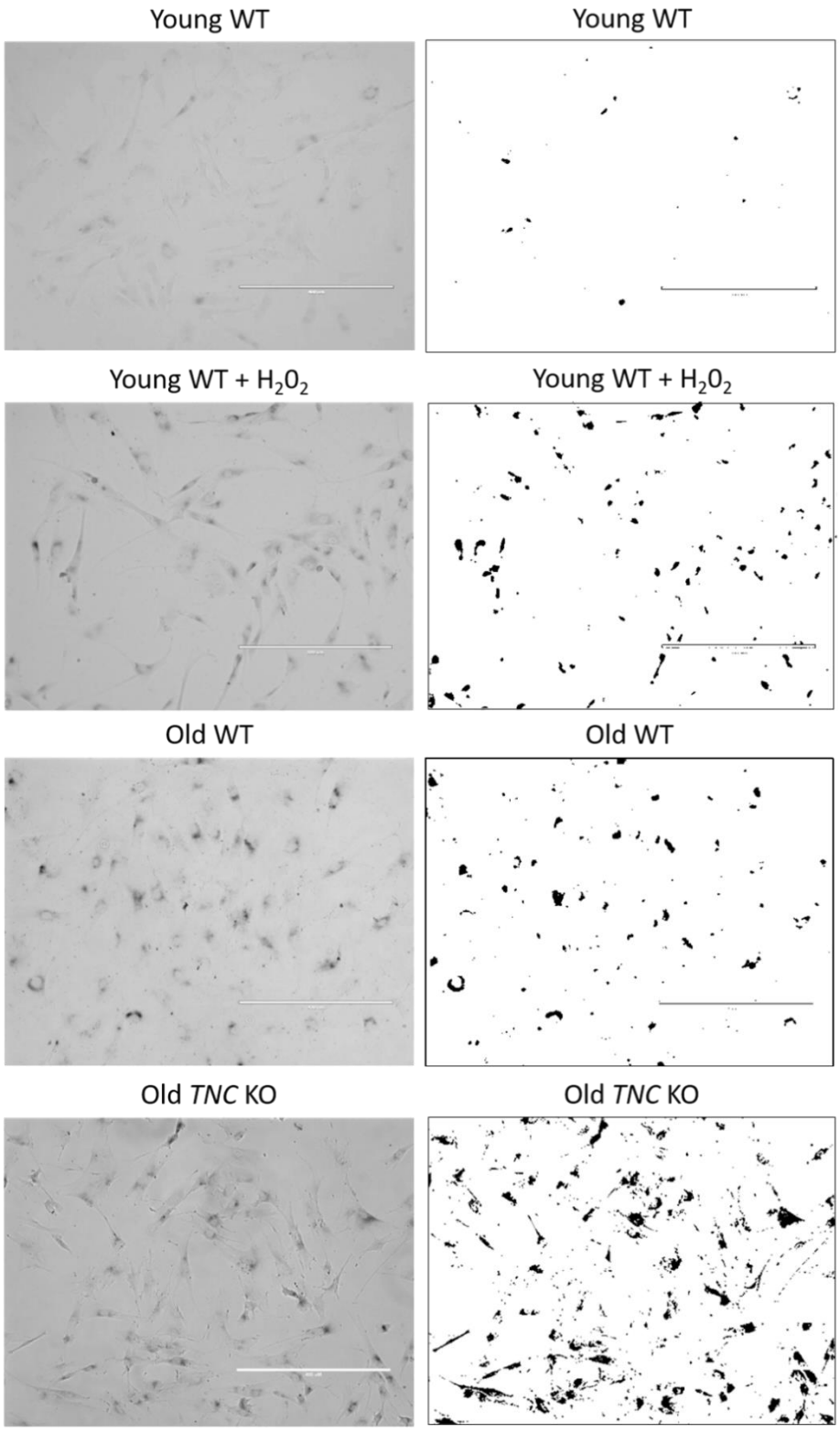


Figure 81. Old WT fibroblasts and old *TNC* KO fibroblasts display different levels of SA-BG staining. Cells were seeded at 1 x 10^4 in triplicate wells and cultured for 24 hours before staining for SA-βG utilising the Cell Signalling Technology™ SA-βG kit. Young WT cells, utilised as a negative control, were also treated with H_2O_2 to induce senescence, and therefore act as a positive control. The left panel of figures are the original microscope images, whilst images on the right panels are the processed images (black and white) to allow higher contrast between cells and nucleus, for clarity. Representative images of each sample are shown. Scale bar 400 µm.

6.2.7 The inflammatory-associated SASP components IL-6 and IL-8 are expressed differentially in old WT and old *TNC* knockout fibroblasts

Next, the presence of two prominent components of the SASP in cell supernatant was investigated. Interleukin-6 (IL-6) and interleukin-8 (IL-8) are two pro-inflammatory soluble factors of the SASP. Whilst the exact composition of the SASP varies depending on tissue origin and triggering factor, these two components are consistently present in SASPs and responsible for propagation and maintenance of the SASP [594, 603, 651]. Supernatant was harvested from old WT or old *TNC* KO fibroblasts after 24 hours of culture and analysed for IL-6 and IL-8 content by ELISA (Figure 82). *TNC* KO cells produced significantly less of both components: 30 % of WT IL-6 (P = 0.0082) and 43% of WT IL–8 (P = 0.0003).

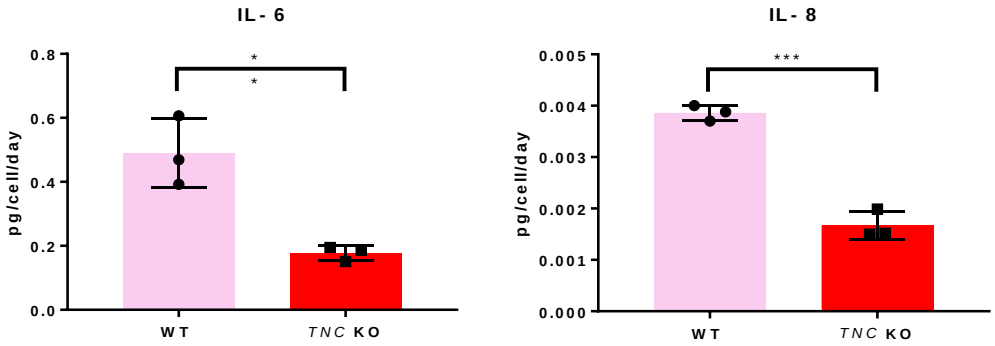


Figure 82. 'Old' *TNC* KO fibroblasts produce less IL-6 and IL-8 than 'old' WT fibroblasts.
'Old' WT and *TNC* KO fibroblasts at PDL 55 were seeded at ~ 1 x 10^6 cells per 10 cm dish and cultured for 24 hours before supernatant was collected and analysed for the presence of the SASP components IL-6 and IL-8 (calculated as pg/cell/day) by ELISA. Mean +/- SEM. N = 3 biological replicates in triplicate. Data were analysed by unpaired T- test.

6.2.8 Pro-inflammatory miR-155 expression is altered in old WT and old *TNC* KO fibroblasts

Following the assessment of two pro-inflammatory cytokines, basal expression of the pro-inflammatory miR-155 in unstimulated fibroblasts was next investigated (Figure 83). miR-155 was chosen as it is specific to lymphoid organs and thus the immune response and it has previously been

demonstrated to be downregulated in macrophages from *TNC* $^{-/-}$ KO mice [188]. In the previous chapter it was also found that miR-155 expression is downregulated when macrophages are seeded on *TNC* KO matrices, and that ECM-associated TN-C is responsible for this. However, it is not known whether TN-C regulates the expression of miR-155 in other cell types. Moreover, miR-155 has also been implicated in senescence, with studies showing its downregulation during senescence in aged WI-38 fibroblasts [652], senescent BJ fibroblasts [653] and in aged humans [654]. miR-155 expression in *TNC* KO fibroblasts was significantly lower (P = 0.0325) than in WT fibroblasts. These results were not confirmed by small RNA-Seq which revealed no significant differences in the expression of miR-155-5p (P > 0.5).

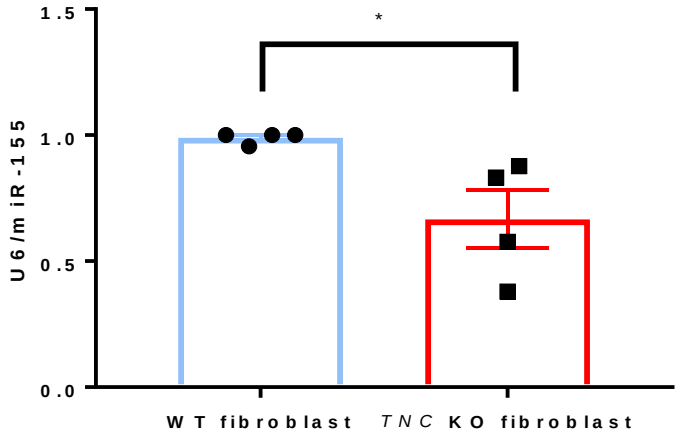


Figure 83. miR-155 expression is reduced in 'old' *TNC* KO fibroblasts compared to 'old' WT fibroblasts. Total RNA was extracted from unstimulated 'old' WT and 'old' *TNC* KO fibroblasts at PDL 55, cultured at 1 x 10 ^{6} cells per 10 cm dish. Expression levels of miR-155 were measured by qPCR. sn*U6* was used as endogenous housekeeping gene and untreated samples were used as calibrator (ΔΔCt method). Mean +/- SEM. N = 4 biological replicates conducted in triplicate. Data were analysed by unpaired T-test.

6.2.9 RNA-Seq of WT and *TNC* KO fibroblasts revealed no significant differences in miR-155 expression

RNA-Seq of WT and *TNC* KO fibroblasts revealed no significant differences in the expression of miR-155-5p or miR-155-3p (adjusted P value > 0.5, as determined by Novogene). Some very small differences occurred: miR-155-3p was downregulated by 0.29 log2Fc whilst miR-155-5p was upregulated by 0.08 log2Fc in *TNC* KO fibroblasts compared to WT fibroblasts, although these contradict the qPCR results. The number preceding the 'p' indicates whether the functional mature

miRNAs are processed from the 5' or 3' arm of the pre-miR-155. Both mature miRNAs have roles in the regulation of inflammation, however miR-155-5p is the predominant form [655], assessed by qPCR previously in this work.

6.2.10 RNA sequencing of 'old' *TNC* KO and WT fibroblasts reveals differences in gene expression between the two cell populations

The whole transcriptome of 'old' *TNC* KO and WT fibroblasts at PDL 55 was subsequently analysed. Three biological replicates for each cell genotype were analysed. The RNA samples corresponded those used for the phalloidin, SA -βG staining, ELISA and qPCR analysis of inflammatory factors. Sequencing and bioinformatics analysis was carried out by Novogene, Cambridge.

6.2.10.1 *The expression of SASP associated genes is altered between TNC* KO and WT

On the topic of the 'SASP', it is widely accepted that this has a highly variable phenotype that is cell-type and senescence specific [437, 438, 609]. This topic has been the focus of many studies that utilise different methods to induce senescence with the aim of defining the general SASP. This has spawned many catalogues of SASP signatures, some specifically for fibroblasts, which generally do not match [436-438]. Based on this, three SASP signatures were assessed from three independent sources; the SASP from the Reactome [436], the fibroblast-specific SASP atlas signature [438] and a fibroblast-specific SASP as determined by a meta- analysis study [437]. Surprisingly, there was no overlap between the signatures from these 3 sources (Figure 86), apart from for one gene present both on the Reactome and SASP atlas list, the chemokine (C-X-C motif) ligand -8 (*CXCL8)* although this gene was not detected in either samples.

A set of 81 SASP annotated genes was retrieved from Reactome [436] as previously done by Martyanov *et al.* [656]. Only 54 out of these 88 were present and exhibited altered gene expression between WT and *TNC* KO fibroblasts (Figure 84 A), however euchromatic histone-lysine N-methyltransferase 1 (*EHMT1)* was removed from the list due to exhibiting a + infinity 1 log2Fc change (Table S1 appendix) which resulted from a complete lack of detection in the WT fibroblasts. This could indicate that *TNC*

inhibits the expression of this gene. In the 53 remaining hits, the changes were small (generally < 1 log2Fc) with the exception of cyclin A2 (*CCNA2*; -3.9), anaphase promoting complex subunit 11 (*ANAPC11;* -2.3), cell division cycle (*CDC16*; -1.7) Ubiquitin-conjugating enzyme E2 E1 (*UBE2E1*; -1.6), Nuclear Factor Kappa B Subunit 1 (*NFKB1*; -1.1), ubiquitin B (*UBB*; +1.16) and cyclin–dependent kinase 4 inhibitor C (*CDKN2C*; + 1.8). The differences were not significant (all adjusted P values were > 0.1 with the majority above 0.4, as determined by Novogene), except for *H2AX* which was downregulated in the KO fibroblasts (P = 0.041). Notably absent from the detected genes in both cell lines were, *IL8*, *IL6* and *CDKN2A* ($p16i^{nk4a}$), long accepted SASP gene signatures.

Since the Reactome provides a general SASP signature, additional genes were investigated. These had been identified as fibroblasts SASP specific from a meta-analysis of studies that included BJ fibroblasts, lung fibroblasts (IMR90), human foreskin fibroblasts, MRC5, WI38 and HCA-2 from 5 independent labs and 3 different senescence-inducing stimuli: replicative senescence, oncogene-induced senescence, and ionizing radiation-induced senescence [437]. Of the 55 from this list only 24 were detected. Of these 24, 14 were downregulated and 10 were upregulated in senescent fibroblasts (Figure 84 B). Here, 3 out of the expected 4 were downregulated, and only 9 out of the expected 20 were upregulated, with none of these differences being statistically significant (adjusted P value > 0.05). Nucleolar protein 3 (*NOL3*) was removed from the analysis as it displayed an infinity log2Fc change resulting from complete lack of detection in the WT cell line (table S1 appendix), possibly indicating that *TNC* inhibits it's expression. Generally, the fold changes were small (> 0.5 log2Fc). The largest difference was seen in solute carrier family 10 member 3 (*SLC10A3*) which was downregulated -1.25 log2Fc and zinc finger HIT-type containing 1 (*ZNHIT1*) which was upregulated + 1.5 log2Fc in the KO compared to WT. Whilst there are some matches with differentially expressed genes, in the expected direction, generally the gene signature of the *TNC* KO fibroblasts does not match the expected SASP from the meta-analysis (Figure 86).

A very recent study (SASP Atlas) utilised the expression of the classical SASP factors: IL-6, p16 and SA-BG (*GLB1)* to confirm the induction of senescence in primary lung fibroblasts, although stated the caveat that there is not always a change in expression of these factors [438]. This study also compiled

another different senescence signature, specific for primary lung fibroblasts, which assessed: chemokine (C-X-C motif) ligand 1 (*CXCL1)* and -8 (*CXCL8*) which were not detected, High mobility group box 1 protein (*HMGB1*), IGFBP2/3/4/5/7 (which was identified in the other SASP lists), *MMP1* and *2*, *LAMB1*, *TIMP1* and *2*, growth differentiation factor 15 (*GDF15*), stanniocalcin 1 (*STC1*), and serine protease inhibitors (*SERPIN1*). These genes were also investigated (Figure 84 C) with the final addition of p53 (*TP53)* which has long been indicated in the literature as SASP associated [657]. In line with the literature, TP53 was upregulated by 0.24 log2Fc in *TNC* KO fibroblasts although this was not significant (adjusted P value > 0.05). Of the 19 genes investigated from this study (SASP Atlas), 14 were detected in the WT and *TNC* KO fibroblasts with statistically significant differences in *IGFBP2 (*adjusted P = 0.02*), SERPINE1* (adjusted P = 0.01) and *LAMC (*adjusted P = 0.02) . All the genes were expected to be upregulated with senescence [438], however only 6 out of 14 genes were upregulated in *TNC* KO fibroblasts compared to WT fibroblasts: *LAMB1*, *HMGB1*, *TIMP1*, *GLB1*, *STC1* and *GDF15*. The biggest differences in gene expression were in *LAMB1* and *HBMG1* which were upregulated 5.2, 3.5 and log2Fc respectively. *IGFBP2*, *TIMP2* and *IGFBP3* exhibited a downregulation of -3.6, -2, -1.4 log2Fc, respectively. *MMP1* expression was reduced by – 1 log2Fc and all other changes were small (>1 log2Fc). As mentioned previously, notably absent from the detected genes in both populations were *IL6* and *CDKN2A*. Again, the *TNC* KO gene signature is very different from the SASP Atlas gene signature (Figure 86).

Together, these studies highlight that the *TNC* KO fibroblast population does exhibit some of the SASP gene signatures identified across multiple studies in the literature. However, these findings also highlight the caveat that there is still no consensus on the SASP (Figure 86). This is expressed in a temporal and cell specific manner and therefore SASP signatures need to be validated over time after the initial onset of senescence. The difference between gene signatures also likely results from the different senescence inducing methods, including drug treatments and radiation. It is difficult to find comprehensive studies which assess the SASP following senescence induced by serial passaging, which is more applicable for comparison with the WT and *TNC* KO cell lines.

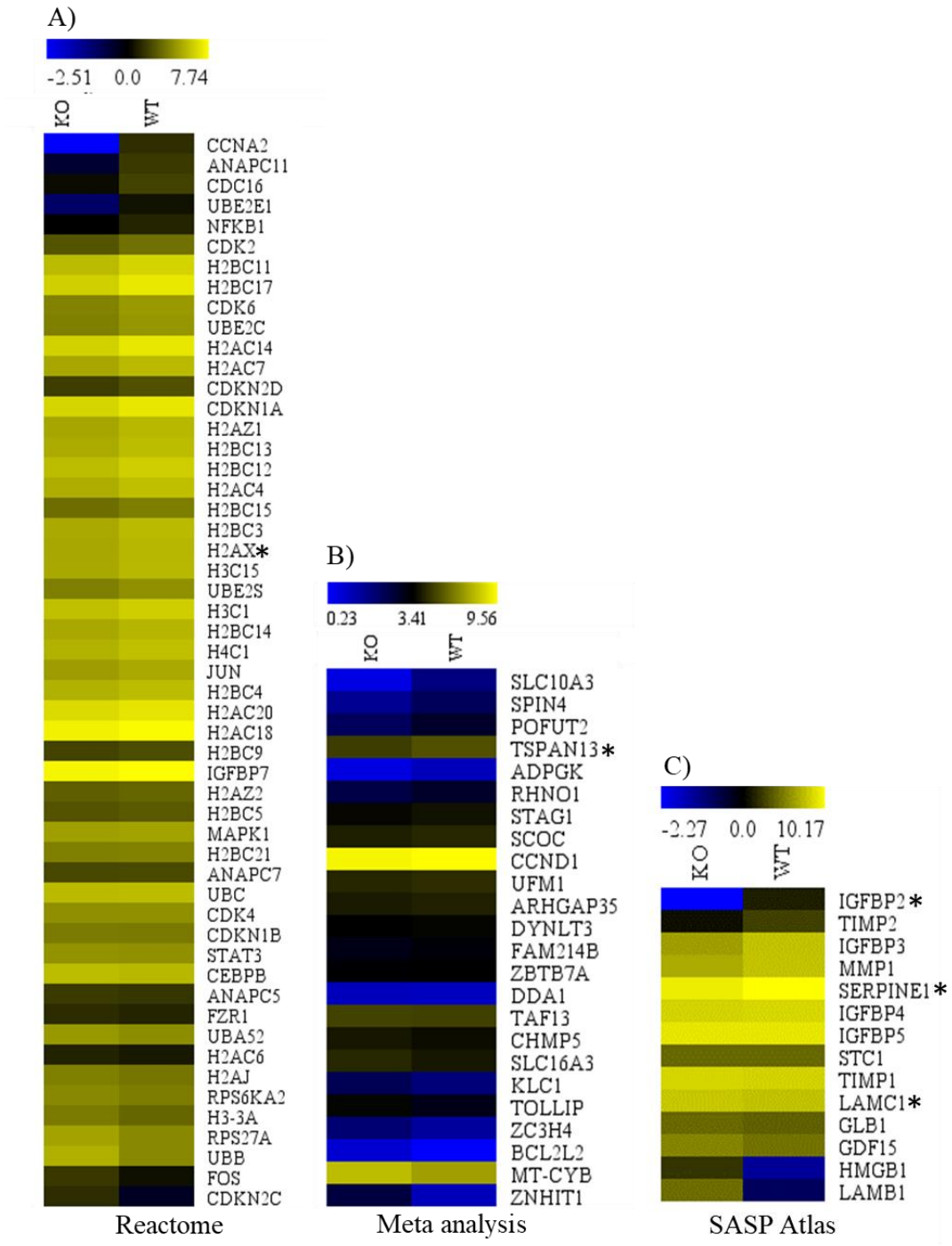


Figure 84. Heatmaps showing differentially expressed SASP genes, utilising SASP signatures from three different sources, in 'old' PDL 55 *TNC* KO fibroblasts relative to 'old' PDL 55 WT fibroblasts.
Upregulated genes are denoted in yellow, whilst downregulated genes are denoted in blue (mRNA abundance on a $\log_2$ scale). Gene signatures were analysed from three sources with the numbers found in both WT and *TNC* fibroblasts indicated: **A)** 53 of the 88 SASP annotated genes from Reactome [436] **B)** 24 of 55 genes indicated in fibroblast specific SASP genes from a meta-analysis [437], **C)** 14 of 19 genes indicated in fibroblast specific SASP from SASP Atlas [438]. Genes were not significantly differentially expressed (adjusted P value > 0.05 as determine by Novogene) except for *H2AX*, *IGFBP2*, *SERPINE1*, *LAMC1* and *TSPAN13* (indicated by an *). Data from three independent experiments. Heatmaps were generated using MeV. Transcripts were assessed for their ability to produce a protein using Ensemble [658]. Abbreviations: signal transducer and activator of transcription 3 (*STAT3*), anaphase promoting complex subunit 5 (*ANAPC5*), cyclin-dependent kinase 2 (*CDK2*), H2B clustered histone 17 (*H2BC17*), cyclin dependent kinase 6 (*CDK6*), ubiquitin conjugating enzyme E2 C (*UBE2C*), H2A clustered histone 14 (*H2AC14*), H2A clustered histone 7 (*H2AC7*), cyclin dependent kinase inhibitor 2D

(*CDKN2D*), ubiquitin A-52 residue ribosomal protein fusion product 1 (*UBA52*), $p21^{CIP1}$ cyclin dependent kinase inhibitor 1A (*CDKN1A*), H2B clustered histone 13 (*H2BC13*), H2B clustered histone 12 (*H2BC12*), H2A clustered histone 4 (*H2A4*), H2B clustered histone 15 (*H2AC15*), H2B clustered histone 3 (*H2B3*), H3 clustered histone 15 (*H3C15*), ubiquitin conjugating enzyme E2 S (*UBE2S*), H3 clustered histone 1 (*H3C1*), H2A. X variant histone (*H2AX*), H2B clustered histone 4 (*H2B4*), H2A clustered histone 20 (*H2AC20*), H2A clustered histone 18 (H2AC18), H2B clustered histone 9 (*H2BC9*), H2A.Z variant histone 1 (*H2AZ1*), cell division cycle 27 (*CDC27*), H2A.Z variant histone 2 (*H2AZ2*), insulin like growth factor binding protein 7 *(IGFBP7*), H2B clustered histone 5 (*H2BC5*), mitogen-activated protein kinase 1 (*MAPK1*), H2B clustered histone 21 (*H2BC21*), ubiquitin C (*UBC*), H2B clustered histone 11 (*H2BC11*), H2A clustered histone 6 (*H2AC6*), H2A.J histone (*H2AJ*), ribosomal protein 56 kinase A2 (*RSP6KA2*), H3.3 histone A (*H3-3A*), CCAAT enhancer binding protein beta (*CEBPB*), anaphase promoting complex subunit 7 (*ANAPC7*), Fos proto-oncogene, AP-1 transcription factor subunit (*FOS*), cell division cycle 16 (*CDC16*), ubiquitin B (*UBB*), cyclin dependent kinase inhibitor 2C (*CDKN2C*), Meis homeobox 3 (*MEIS1*), Rho GTPase activating protein 35 (ARHGAP35), zinc finger CCCH-type containing 4 (*ZC3H4*), spindling family member 4 (*SPIN4*), stromal antigen 1 (STAG1), RAD9-HUS1-RAD1 interaction nuclear orphan 1 (*RHNO1*), solute carrier family 16 member 3 (*SLC16A3*), tertraspanin 13 (*TSPAN13*), cyclin D1 (*CCD1*), zinc finger and BTB domain containing 7A (*ZBTB7A*), short coiled-coil protein (*SSOC*), ubiquitin fold modifier 1 (*UFM1*), protein- O- fructosyltransferase 2 (*POFUT2*), TATA-box binding protein associated factor 13 (*TAF13*), ADP dependent glucokinase (*ADPGK*), DET1 and DDB1 associated 1 (*DDA1*), zinc finger HIT-type containing 1 (*ZNHIT1*), charged multivesicular body protein 5 (*CHMP5*), toll interaction protein (*TOLLIP*), kinesin light chain 1 (*KLC1*), BCL2 like 2 (*BCL2L2*), dynein light chain T type-3 (*DYNLT3*), family with sequence similarity 214 member B (*FAM214B*), mitochondrially encoded cytochrome B (MT-CYB).

Due to the lack of consensus on the SASP between the literature and the signatures assessed here, and the very tight link between aging and the onset of senescence, one last database was utilised to investigate whether the cells displayed an 'aging' replicative senescence gene signature. Entering of all the differentially expressed genes into the 'Cell Age' database [439] resulted in 37 hits for genes which have previously been shown to be up or downregulated during replicative senescence of human cells (Figure 85). Of this list only one gene, *TSPAN13* was present on the previous SASP meta-analysis list. SH3 and PX domains 2A *(SH3PXD2A)* was removed due to an infinity log2Fc (table S1 appendix) change resulted from lack of detection in the KO cell line, indicating that *TNC* may regulate the expression of this gene and leaving 36 hits. The majority of the genes were significantly differentially expressed (adjusted P value < 0.05) with the exception of cytochrome b5 domain containing 2 (*CYB5D2), STAT1,* chromatin licensing and DNA replication factor 1 *(CDT1),* ADP ribosylation factor (*AFR4),* ANTXR cell adhesion molecule 2 (*ANTXR1),* stearoyl-CoA desaturase 5 (*SCD5)* and *LAMB2 (*indicated by a – on the heatmap). Furthermore, the direction of the regulation, i.e. up or down in KO cell line, matched the expected direction documented in Cell Age for replicative senescent cells for 25/36 genes. The largest down regulation in the KO cell line in the expected direction for replicative senescence was displayed by stathmin 1 (*STMN1;* -6.5 log2Fc*)* whilst the largest up regulation in the

expected direction, in the KO cell line, was demonstrated by MX dynamin like GTPase 1 (*MX1;* 7.3 log2Fc*).* Only 11 out of the 36 exhibited differential expression in the unexpected direction: serpin family B member 2 (*SERPINB2), CYB5D2,* L1 cell adhesion molecule *(L1CAM), STAT1, ARF4, TSPAN13,* transglutaminase 2 (*TGM2),* TIMP3, SPARC, *ANTXR1* and HAUS augmin like complex subunit 7 (*HAUS7).* Interestingly, there was overlap between the non-significantly differentially expressed genes and those that were differentially expressed in the wrong direction. This could indicate that these genes may not have as much of an effect as those that were differentially expressed in the expected direction. Whilst there were many hits from this list, there was negligible overlap with the SASP signatures assessed earlier (Figure 86). However, overall, assessment of this catalogue of genes indicates that the KO fibroblasts are exhibiting a gene signature in line with replicative senescence.

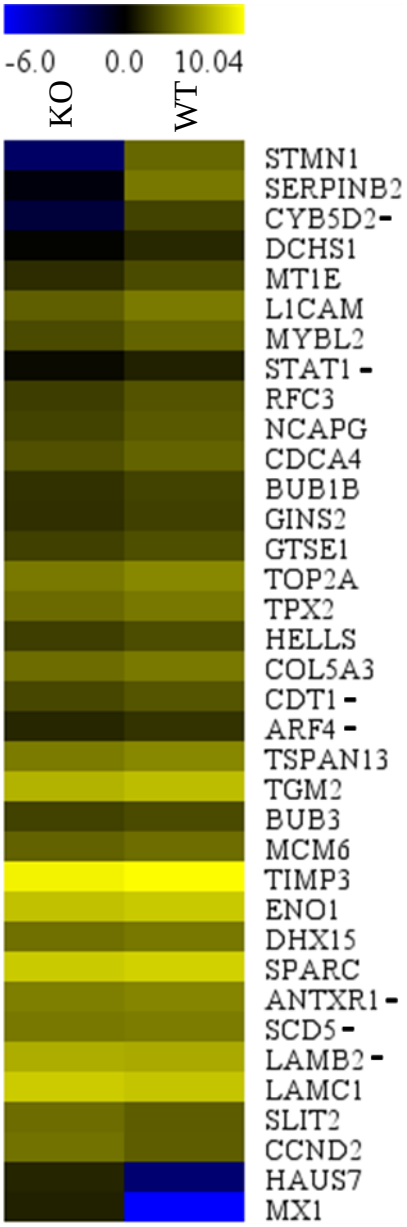


Figure 85. Heat map of the 36 differentially expressed genes in 'old' PDL 55 *TNC* KO fibroblasts relative to 'old' PDL 55 WT fibroblasts unstimulated cells that matched with 'Cell Age' database [439] for the gene signature for replicative senescence.

Upregulated genes are shown in yellow, whilst downregulated genes are shown in blue (mRNA abundance on a $\log_2$ scale). All changes were statistically significant (adjusted P value < 0.05 as determined by Novogene) except in those genes marked with (-). Data are from three independent experiments. SH3 and PX domains 2A (SH3PXD2A) was removed due to an infinity log2Fc Abbreviations: dachsous cadherin-related 1 *(DCHS1*), metallothionein 1E (*MT1E*), MYB proto-oncogene like 2 (*MYBL2*), replication factor C subunit 3 (*RFC3*), non-SMC condensing I complex subunit G (*NCAPG*), BUB1 mitotic checkpoint serine/threonine kinase B (*BUB1B*), GINS complex subunit 2 (*GINS2*), DNA topoisomerase II alpha (*TOP2A*), TPX microtubule nucleation (*TPX2*), helicase, lymphoid specific (*HELLS*), BUB3 mitotic checkpoint protein (*BUB3*), minichromosomal maintenance complex component 6 (*MCM6*), DEAH-box helicase 15 (*DHX15*), slit guidance ligand 2 (*SLIT2*), MX dynamin like GTPase1 (*MX1*).

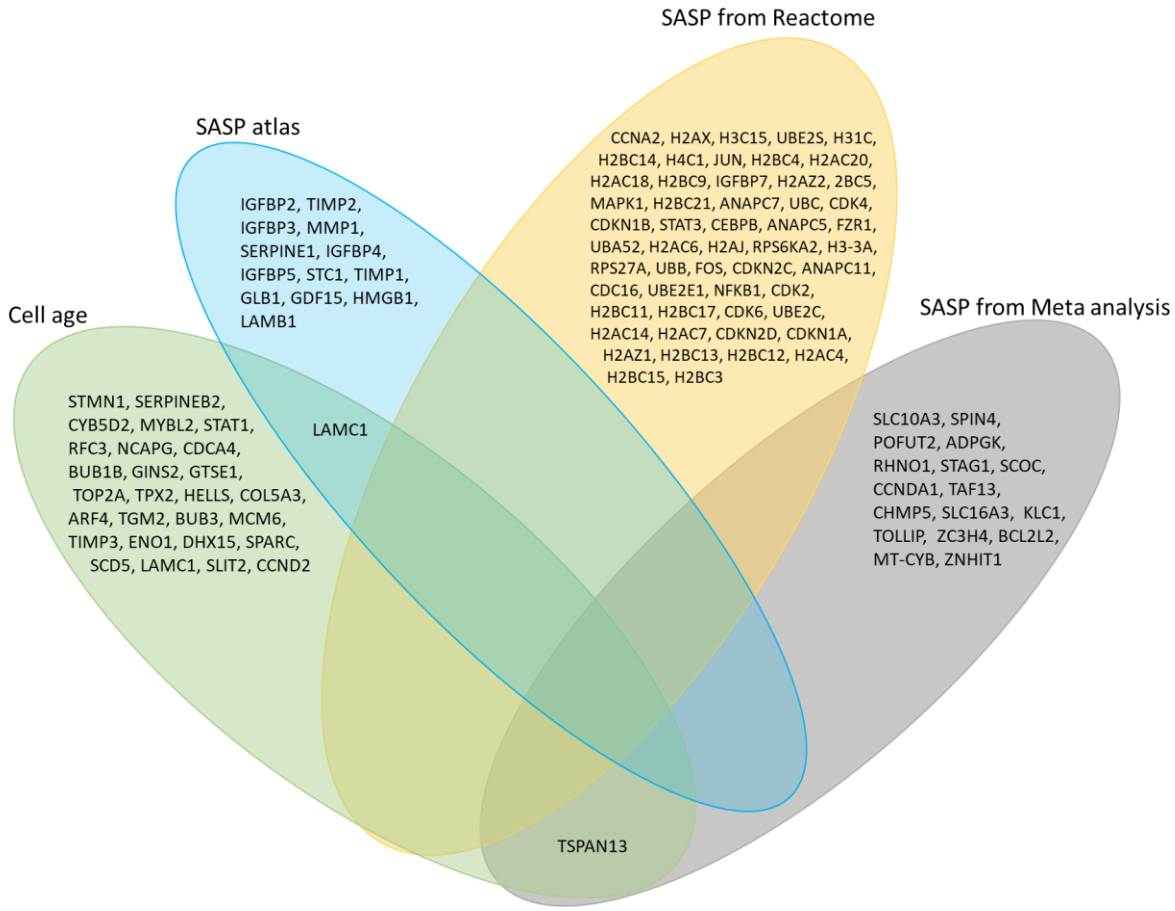


Figure 86. Little overlap between the gene signatures from 3 SASP and 1 replicative senescence database occurs when assessing the differentially expressed genes from 'old' PDL 55 *TNC* KO and 'old' PDL 55 WT fibroblasts.
As previously, gene signatures for the SASP were compiled from 3 online databases: 'The Reactome'[436], 'SASP Atlas'[438] and a SASP compiled from a meta-analysis in fibroblasts cells[437]. A further database 'Cell age'[439] was utilised which catalogued the gene signature for replicative senescence**.** The differentially expressed genes list was screened for these signatures (as detailed in Figure 83 and Figure 84) and displayed in the venn diagram.

6.2.10.2 Significantly differentially expressed genes confer important biological functions relating to actin cytoskeleton regulation

Focussing on significantly differentially expressed genes (adjusted P value < 0.05) between 'old' *TNC* knockout fibroblasts and WT fibroblasts revealed 94 downregulated genes and 70 upregulated genes (Figure 87). This list was filtered to remove genes with: a log2Fc change < 1 or > -1 which removes genes with less than 2- fold changes, and also genes with a + or – infinity log change (supplementary table S1, appendix) which resulted from detection of the gene in only one cell line. This left a list of 31 downregulated genes and 21 upregulated genes (Figure 88). In this list, the largest changes in

expression were exhibited by amyloid beta precursor like protein 2 (*APLP2*), SS nuclear autoantigen 1 (*SSNA1*), and GH 3 domain containing (*GHDC*), which were downregulated by 11, 11 and 9.7 log2Fc, respectively, in the KO compared to the WT. The largest upregulation was seen in reticulon 4 (*RTN4*), cofilin 2 (*CFL2*) and MX dynamin like GTPase1 (*MX1*), of 10.6, 8.9 and 7.3 log2Fc, respectively.

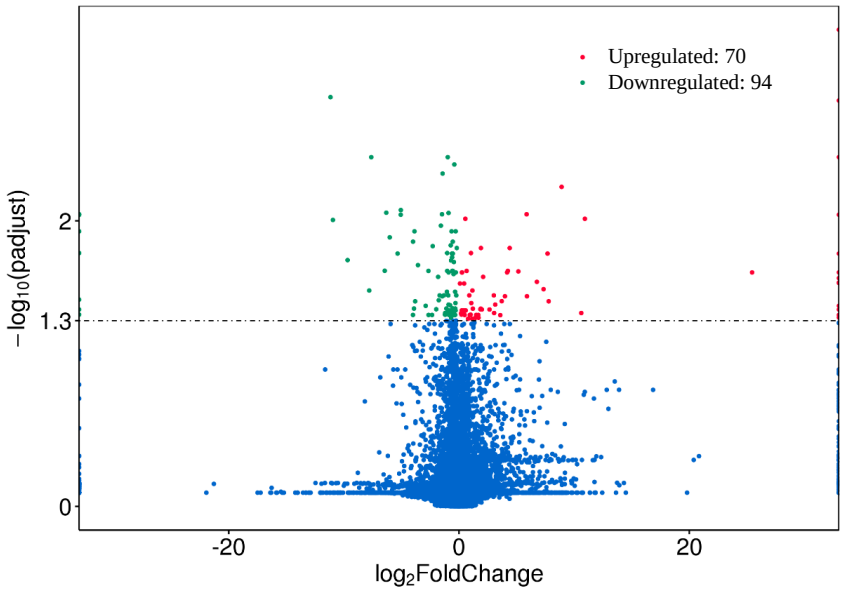


Figure 87: Volcano plot indicated the number of significantly differentially expressed genes in 'old' PDL 55 TNC KO fibroblasts compared to 'old' PDL 55 WT fibroblasts.
Upregulated genes are shown in red whilst downregulated genes are shown in green. Genes with no differences are shown in blue. This list was filtered to remove genes with: a log2Fc change < 1 or > -1 which removes genes with less than 2- fold changes, and also genes with a + or – infinity log change.

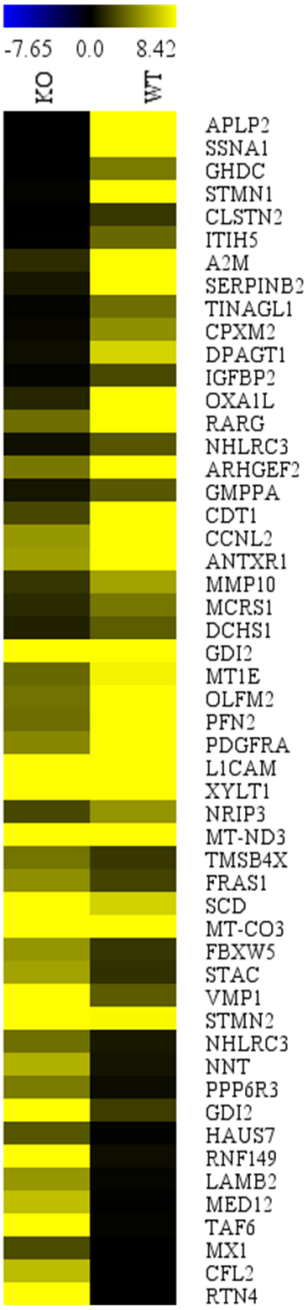


Figure 88. Heat map of the 52 significantly differentially expressed genes with upregulated (yellow) or downregulated (blue) expression (mRNA abundance on a $\log_2$ scale) in 'old' PDL 55 *TNC* KO fibroblasts relative to 'old' PDL 55 WT fibroblasts.

All changes were statistically significant (adjusted P value < 0.05, as determined by Novogene). Data are from three independent experiments. Abbreviations: calsyntenin 2 (*CLSTN2*), inter-alpha-trypsin inhibitor heavy chain 5 (*ITIH5*), alpha-2-macroglobulin (*A2M*), tubulointerstitial nephritis antigen like 1 (*TINAGL1*), carboxypeptidase X M14 family member 2 (*CPXM2*), dolichyl-phosphate N-acetylglucosamine phosphotransferase 1 (*DPAGT1*), OXA1L mitochondrial inner membrane protein (*OXA1L*), retinoic acid receptor gamma (*RARG*), NHL repeat containing 3 (*NHLRC3*), Rho/Rac guanine nucleotide exchange factor 2 (*ARHGEF2*), GDP-mannose pyrophosphorylase A (*GMPPA*), matrix metallopeptidase 10 (*MMP10*), microspherule protein 1 (*MCRS1*), GDP dissociation inhibitor 2 (*GDI2*), olfactomedin 2 (*OLFM2*), profilin 2 (*PFN2*), platelet derived growth factor receptor alpha (PDGFRA), xylosyltransferase 1 (XYLT1), nuclear receptor interacting protein 3 (*NRIP3*), mitochondrially encoded NADH:ubiquinone oxidoreductase core subunit 3 (*MT-ND3*), thymosin beta 4 X-linked (*TMSB4X*), fraser extracellular matrix complex subunit 1 (*FRAS1*), mitochondrially encoded cytochrome c oxidase III (*MT-CO3*), F-box and WD repeat domain containing 5 (*FBXW5)*, SH3 and cysteine rich domain (*STAC*), vacuole membrane protein 1 (*VMP1*), stathmin 2 (*STMN2*), NHL repeat containing 3 (*NHLRC3*),

nicotinamide nucleotide transhydrogenase (*NNT*), protein phosphatase 6 regulatory subunit 3 (*PPP6R3*), GDP dissociation inhibitor 2 (*GDI2*), HAUS augmin like complex subunit 7 (*HAUS7*), ring finger protein 149 (*RNF149*), mediator complex subunit 12 (*MED12*), TATA-box binding protein associated factor 6 (*TAF6*).

Over representation analysis (ORA) or gene set enrichment analysis (GSEA) is a method that determines whether a functional group of differentially expressed genes (belonging to a specific KEGG pathway) are enriched in the sample population more or less than would be expected. It is important to assess not only the expression of genes, but also how they work together since multiple genes contribute to each biological function; slight changes in one gene in the pathway may cause significant alterations overall. To further investigate the transcriptome differences between cell lines, GSEA was carried out on up- and down-regulated genes to identify enriched pathways in gene ontology biological processes using the Kyoto encyclopaedia of genes and genomes (KEGG) on Webgestalt [440]. This revealed that significantly differentially expressed genes were enriched in many pathways, including regulation of actin cytoskeleton, oxidative phosphorylation and HIPPO pathway signalling, all of which are of particular interest in progression of the cell cycle (table 24).

Within the actin cytoskeleton KEGG pathway, 4 genes were differentially expressed, *PDGFRA*, a receptor tyrosine kinase, *PFN2*, *TMSB4* and *CFL2*. The direct roles of *PFN2* (downregulated in KO) and *CFL2* (upregulated in KO) in F-actin network construction and destruction [659] support the dysregulation status seen in partial and complete *TNC* KO fibroblasts.

The HIPPO pathway is an evolutionary conserved, tumour suppressor pathway that regulates organ size and development through cell proliferation and growth [660]. This also has implication for cell cycle regulation in the *TNC* KO fibroblasts. However, KEGG pathway analysis results should be interpreted with caution given that only one gene hit is required to indicate a pathway, as is the case for the HIPPO pathway in which *DCHS1* was altered in gene expression.

Table 24. Over representation analysis of differentially expressed genes enriched in KEGG pathways. The FDR was automatically altered to > 0.05. Generated using Webgestalt.org [440].

Description	Ratio	P Value	FDR
Regulation of actin cytoskeleton	5.1949	0.006741	1
Complement and coagulation cascades	7.0033	0.032686	1
Glycosaminoglycan biosynthesis	13.831	0.069955	1
Protein export	12.027	0.080033	1
Oxidative phosphorylation	4.1598	0.082738	1
Glycosaminoglycan biosynthesis	11.526	0.083369	1
Parkinson disease	3.8962	0.09253	1
Biosynthesis of unsaturated fatty acids	10.246	0.093307	1
Hippo signalling pathway	9.539	0.099874	1

Further analysis using Webgestalt and gene ontology (GO) enrichment analysis, which analysed enriched genes for their biological functions based on annotations within the repository, indicated key biological functions that may be altered (Figure 89). Of particular relevance is cell proliferation, cell growth, biological regulation and cellular compartment organisation. The latter two of this list had some of the highest proportions of input genes annotated to them, 37/50 and 29/50.

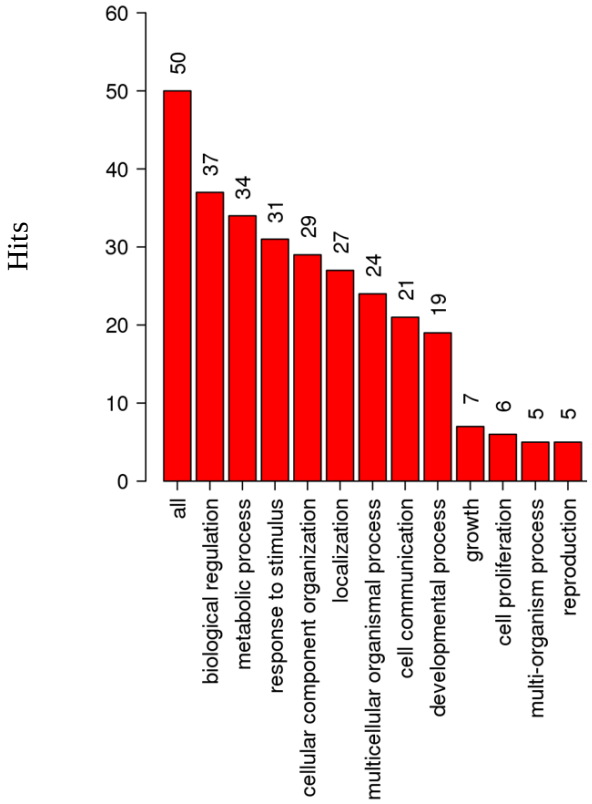


Figure 89. Gene ontology ORA revealed differentially expressed genes are enriched in pathways relating to important biological functions.
The first bar 'all' indicates the total number of differentially expressed input into the analysis. The other bars represent the number of input genes in each pathway.

6.2.10.3 Significantly differentially expressed miRNAs target genes conferring important biological functions relating to cell cycle and other signalling pathways

Next, miRNA expression from 'old' *TNC* KO and 'old' WT fibroblasts was assessed. Again, significantly differentially expressed miRNAs (adjusted P value < 0.05, as determined by Novogene) were filtered to remove any with log2Fc changes between 1 and -1 as this would leave miRNAs that were at least 2 fold up or downregulated. This left a list of 30 differentially expressed miRNAs from the original list of 89 (table 25) of which 23 were downregulated and 7 upregulated. The largest log2Fc changes were exhibited by miR-615-3p, miR-618 and miR-483-3p, which were downregulated by a log2Fc change of 3. In comparison, the largest upregulation was 1.5 log2Fc exhibited by miR-548p, miR-548ah-3p and miR-548am-3p. Notably absent from the significantly differentially expressed list was miR-155-5p; expression remained very similar (log2Fc change 0.08466, adjusted P value = 0.57583) between *TNC* KO and WT fibroblasts.

Table 25. Significantly differentially expressed miRNAs in *TNC* KO fibroblasts compared to WT fibroblasts presenting from most significant to least. Adjusted P values were determined by bioinformatic analysis by Novogene.

miRNA	Up or down regulated in the *TNC* KO relative to WT fibroblasts	Log2Fc change relative to WT	Adjusted P value
miR-615-3p	↓	-3.2649	1.06E-61
miR-618	↓	-3.2475	2.60E-42
miR-483-3p	↓	-3.1279	7.74E-14
miR-514a-3p	↓	-2.9084	6.85E-14
miR-509-3-5p	↓	-2.8674	4.76E-09
miR-196a-5p	↓	-2.5727	2.36E-35
miR-490-3p	↓	-2.4403	3.40E-54
miR-490-5p	↓	-2.3223	1.22E-31
miR-335-3p	↓	-2.2461	2.29E-19
miR-137-3p	↓	-2.1916	2.26E-12
miR-335-5p	↓	-2.1294	1.46E-17
miR-10b-3p	↓	-2.0356	2.82E-09
miR-10401-3p	↓	-1.9579	7.10E-06
miR-483-5p	↓	-1.8181	0.001902
miR-2682-5p	↓	-1.6682	1.84E-05
miR-582-3p	↓	-1.379	0.014406
miR-10b-5p	↓	-1.2936	1.78E-14
miR-129-2-3p	↓	-1.2283	0.000823
miR-219a-2-3p	↓	-1.1906	0.0264
miR-219b-5p	↓	-1.1906	0.0264
miR-224-3p	↓	-1.1553	3.08E-05
miR-556-5p	↓	-1.1312	0.034381

miR-744-5p	↓	-1.0172	3.67E-15
miR-576-3p	↑	1.0174	0.001339
miR-708-3p	↑	1.1273	0.001006
miR-548j-3p	↑	1.1843	0.006909
miR-95-3p	↑	1.4402	9.50E-11
miR-548p	↑	1.5194	7.58E-06
miR-548ah-3p	↑	1.5357	6.49E-06
miR-548am-3p	↑	1.5614	3.45E-05

Inputting of the list of the 30 miRNAs into MirWalk [661], a bioinformatic software to mine target genes, filtered for validated interactions with targets, returned over 1200 gene targets (including different transcripts of the same genes). GSEA of the these miRNA targets using KEGG revealed enrichment for targets in 139 pathways, including several of the pathways that were mentioned in the previous section, including regulation of the actin cytoskeleton and HIPPO signalling. However, these pathways were above the significance cut off level (miRWalk P value = 0.12 and 0.14, respectively). The most relevant pathways are summarised in table 26; this is not an exhaustive list of the enriched pathways. Interestingly, cell cycle/cell senescence and the FOXO3 pathway arose in this analysis, supporting the original hypothesis in this chapter of TN-C effects on cell fate. However, a direct link has not been elucidated in this data set yet. The ErbB signalling pathways are tightly linked to MAPK and PI3K pathways and therefore has key roles in cell fate [662]. This linkage may also indicate the mechanism by which the previously Akt signalling may be indirectly affected by TN-C. As aging is linked to senescence in the sense that aging will eventually lead to senescence, the longevity regulating pathway is also of interest. Many targets are enriched in multiple pathways from the 8 listed in table 26, including; RAC-beta serine/threonine- protein kinase *(AKT2)* in 6 pathways, cyclin D2 *(CCND2*) in 5 pathways, cyclin D3 *(CCND3)*, myc proto oncogene (*MYC*) and mouse double minite 2 (*MDM2)* in 4 pathways and cyclin division protein kinase 6 (*CDK6)* in 3 pathways. These genes have key roles in

cell cycle proliferation, and progression or regulation of the cell cycle. *CCND2* was significantly differentially expressed and should therefore be prioritised for further investigation, along with the other targets common to multiple pathways. Overall, RNA-Seq data indicates a transcriptional effect of *TNC* knockout that requires validation of key targets, summarised in the table, by qPCR and western blotting.

Table 26. KEGG pathway enrichment analysis of the target genes of differentially expressed miRNAs. Hits indicates the number of genes targeted in the pathway that are indicated in the next column. One pathway with a P value just over the significance cut off was included due to the relevance of this pathway. Targets that were differentially expressed ($P < 0.05$, determined by KEGG analysis on Webgestalt) are underlined, and highlighted to denote upregulation (yellow) or downregulation (blue) in *TNC* KO fibroblasts compared to WT fibroblasts. Plain black font indicates a gene not detected in either WT or *TNC* KO fibroblasts whilst TCONS transcripts are denoted in bold. Red highlight indicates an infinity log change due to lack of detection in the KO cell line.

Name	Hits	Genes	P value
Cell cycle	15	ORC4; **GSK3B**; E2F3; CCND3; CDK6; YWHAZ; MYC; CDKN2A; CDC14B; CCND2; MDM2; SMAD4; CDC25B; RBL1; SMC1A	0.0028
p53 signalling pathway	11	CASP3; CCND3; CDK6; CDKN2A; SESN3; EI24; CCND2; MDM2; IGF1; THBS1; BBC3	0.0028
MAPK signalling pathway	24	STMN1; ELK4; MAPKAPK2; MAP3K2; FGF2; PDGFC; **RAPGEF2**; CASP3; SRF; MAP3K7; BRAF; MYC; IGF1; FGF9; IGF1R; **MAPK3**; CRK; ARRB2; MAPK7; AKT2; CDC25B; RPS6KA6; DUSP9; IRAK1	0.0046
Focal adhesion	18	CTNNB1; **GSK3B**; MYLK; PDGFC; CCND3; BRAF; VCL; PARVA; CCND2; IGF1; ARHGAP5; THBS1; IGF1R; **MAPK3**; CRK; PAK4; AKT2; XIAP	0.0069
ErbB signalling pathway	10	ABL2; **GSK3B**; BRAF ; MYC; CAMK2G; **MAPK3**; CRK; PAK4; AKT2; PLCG1	0.012
FoxO signalling pathway	12	S1PR1; SETD7; SOD2; BRAF; CCND2; MDM2; IGF1; PRKAB1; IGF1R; **MAPK3**; SMAD4; AKT2	0.0192
Cellular senescence	12	MAPKAPK2; E2F3; CCND3; CDK6; HIPK2; MYC; CDKN2A; CCND2; MDM2; **MAPK3**; AKT2; RBL1	0.0483
Longevity regulating pathway	8	ATG5; SOD2; SESN3; ADCY6; IGF1; PRKAB1; IGF1R; AKT2	0.0512*

Abbreviations: origin recognition complex subunit 4 (*ORC4*), glycogen synthase kinase 3 beta (*GSK3B)*, transcription factor E2F3 (*E2F3)*, Tyrosine 3-Monooxygenase/Tryptophan 5-Monooxygenase Activation Protein Zeta *(YWHAZ)*, cell division cycle 14B (*CDC14B)*, SMAD family member 4 *(SMAD4)*, cell division cycle 25B *(CDC25B)*, RB transcriptional corepressor like 1 (*RBL1)*, structural maintenance of chromosome 1A *(SMC1A)*, caspase 3 (*CASP3)*, sestrin 3 (*SESN3)*,EI24 autophagy associated transmembrane protein *(EI24)*, insulin- like growth factor 1 (*IGF1)*, thrombospondin 1 *(THBS1)*, BCL2 binding component 3 *(BBC3)*, ETS transcription factor 4 (*ELK4)*, mitogen- activated protein kinase – activated protein kinase 2 (*MAPKAPK2)*, mitogen-activated protein kinase kinase kinase 2 *(MAP3K2)*, platelet derived growth factor C (*PDGFC)*, rap guanine nucleotide

exchange factor 2 (*RAPGEF2).* Serum response factor *(SRF),* mitogen-activated protein kinase kinase kinase 7 *(MAP3K7),* B-Raf- proto- oncogene *(BRAF),* insulin-like growth factor 1 receptor (*IGF1R),* mitogen-activated protein kinase 3 (*MAPK3);* adaptor molecule crk (*CRK),* arrestin beta 2 *(ARRB2),* mitogen-activated kinase 7 (*MAPK7),* ribosomal protein S6 kinase *(RPS6KA6),*dual specificity phosphatase 9 *(DUSP9),* interleukin 1 receptor associated kinase 1 *(IRAK1),* catenin beta 1 (*CTNNB1),* myosin light chain kinase *(MYLK),* platelet derived growth factor C (*PDGFC),* vinculin (*VCL),* parvin alpha *(PARVA),* Rho GTPase activating protein 5 *(ARHGAP5),* P21 (RAC1) activated kinase 4 (*PAK4),* X- linked inhibitor of apoptosis protein (*XIAP),* ABL proto-oncogene 2 (*ABL2),* calcium/calmodulin-dependent protein kinase type II gamma (*CAMK2G),* phospholipase C (*PLCG1),* sphingosine-1 phosphate receptor 1 (*S1PR1),* histone lysine methyltranferase SETD7 *(SETD7),* superoxide dismutase 2 *(SOD2),* 5'-AMP, activated protein kinase subunit beta-1 (*PRKAB1),* mothers againt decapentaplegic homolog 4 *(SMAD4),* homeodomain interacting protein kinase 2 (*HIPK2),* autophagy related gene 5 (*ATG5) SESN3,* adenylyl cyclase type 6 (*ADCY6).*

6.3 SUMMARY

Key phenotypic changes arise from either partial or complete KO of *TNC*. These changes include aberrant organisation of the actin cytoskeleton that ultimately results in cell enlargement and a loss of spindle morphology. This is supported by RNA-Seq data which indicates changes in expression of specific actin network related genes. These changes appear to be compounded by age in some instances, however it is clear that the genotype of these cells ultimately dictates cell phenotype. Clear differences occur in the expression of inflammatory molecules – the expression of which is hampered in *TNC* KO fibroblasts. RNA-Seq data indicates that *TNC* knockout causes transcriptional effects, including downregulation of many miRNAs that have targets in pathways relating to cell cycle, FoxO signalling and cell senescence. These targets will need to be validated by qPCR and other methods such as western blotting.

6.4 DISCUSSION

Many features of cells are assessed in order to ascertain whether cells are actively proliferating or senescent. These commonly include assessment of morphological characteristics, including flattening of the cell body, enlargement and loss of spindle morphology. Here, we have shown that the *TNC* genotype correlates with drastic changes to cell morphology. Fibroblasts lacking *TNC* are enlarged and lose their characteristic spindle shape, with a concurrent enlarged nucleus, compared to WT *TNC*-expressing fibroblasts. Cellular enlargement and flattening, and an increase in cell volume in aging and senescence have been documented in many cell types, although none more so than fibroblasts,

particularly WI-38 fibroblasts which are popular due to their low PDL before the onset of senescence [581, 609, 650, 663, 664]. Consistent with our studies, fibroblast volume enlargement during senescence is roughly four-fold and senescence of normal human diploid fibroblasts appeared to occur at PDL 55, the age of both the old WT and old *TNC* KO fibroblasts [665]. This study attributed senescence as the result of oxidative stress resulting from mitochondrial dysfunction.

WI-38 fibroblasts and HT-p21 cells enlarge during senescence and have impaired gene expression and cell proliferation [666]. Maximum mitochondrial activity and cell fitness is achieved at intermediate cell sizes suggesting the existence of an optimal cell size [667]. Similarly, cell organelles, RNA, mRNA, and protein cell content have been shown to scale linearly, to a certain extent, to cell size [636]. Studies on IMR-90 fibroblasts show that the DNA: cytoplasm ratio characterizes the optimum cell size and the biosynthetic potential of protein and RNA expression necessary to maintain gene expression and cell signalling to sustain survival or result in senescence [668]. The drastic increase in size in fibroblasts exhibiting features of senescent cells reinforces this; cell enlargement and resulting dilution of cell cytoplasm outside the optimum range becomes detrimental to cell viability.

Consistent with our findings, studies have demonstrated heterogeneity of cell size even in 'old' PDL 68 senescent populations that have been passaged to senescence, with both very large and smaller cells being present in the population. This heterogeneity is not seen in cells treated with compounds to induce senescence which suggests a link to growth-arrest pathways [669]. This is in agreement with the general acceptance that cells enter senescence at different time points in culture, and the increase in senescent cells over time contributes to a decreasing PDL. The reason for the increase in cell size and heterogeneity is not well understood. However, it is thought that uncoupling between signalling pathways and cell cycle progression occurs, and that mammalian target of rapamycin (mTOR) activity remains high which results in continued protein synthesis without cell cycle progression [670]. Indeed, upregulation of PI3K, an upstream activator of mTOR, has been reported in fibroblasts induced to senesce with H_2O_2. Inhibition of PI3K was shown to abrogate H_2O_2-induced senescence-associated morphological changes, but this had no effect on H_2O_2-induced cell cycle arrest [671]. However, protein and RNA synthesis are generally reported to decrease with age, which does not support the

increased volume theories. Although it could be that decreased degradation occurs. It is clear that further study is required to determine the pathways specifically modulated in senescence.

The idea that protein synthesis continues, yet cell cycle progression is halted, resulting in increased cell volume that is not controlled by mitosis, is supported by the observations that increased nuclear expansion occurs in Hutchinson-Gilford progeria syndrome (HGPS) cells; HGPS is a fatal premature aging disorder [672]. This study also demonstrated that nuclear expansion occurs in aged primary human fibroblasts supporting our results. There was a concurrent increase in protein synthesis and cell cycling in HGPS cells, possibly in an attempt to cope with the increase in protein synthesis and maintain the nucleus and cell size within normal ranges. This suggests an optimum nucleus size and that large sizes are detrimental to cell viability. In support of this, a small nucleus size predicts longevity in *C. elegans* [673]. An enlarged nucleus has been demonstrated in populations of slowly or non-replicating WI-38 fibroblasts, compared to rapidly replicating populations of WI-38 fibroblasts [674]. TIG-7 fibroblasts have also been shown to display nuclear swelling following replication inhibition with thymidine to induce premature senescence [675]. This is also in line with research that scaling of cell components occurs with cell enlargement, to a certain extent [636]. In line with our findings, a link between the actin cytoskeleton, cell swelling and nuclear swelling has been demonstrated [676]. Other cells undergo nuclear enlargement upon senescence, however this is not a ubiquitous feature of all cells [677]. In this study, in young partial and old complete *TN-C* KO fibroblasts, which were morphologically senescent, the nucleus was increased in size compared to young WT. Here, only a modest increase in nuclear size between young and old WT fibroblasts occurred. It may be that further culturing would result in increased enlargement.

Nuclear and cell enlargement was seen specifically in a *TNC* gene-dose-dependent manner. Thus, a lack of *TNC* may result in acceleration of aging. Comparing the young partial KOs with the old complete KO indicated that age in the *TNC* KO appeared to enhance these morphological changes, or it could be that residual TN-C expression abrogated any effects. However this is indicative of a link between *TNC* and aging. Supporting this, very recently a study proposed a link between *TNC* and aging in which TN-C contributes to the maintenance of tissue integrity via TGF-β signalling [522]. In our study, RNA-Seq

showed a small increase in TGFβ1 and a larger increase in TGFβR2 in the *TNC* KO cell line, however this was not significant.

Actin cytoskeletal rearrangements have been demonstrated in senescent cells, although the effect is not well-defined. In lung fibroblasts thick actin stress fibres have been demonstrated, supporting the phenotype observed here [634, 678]. Other studies have demonstrated reduced F-actin protein deposition and reduced stress fiber formation in senescent/aged fibroblasts [664, 679]. Linked to this, mouse fibroblasts plated on TN-C-containing matrix display reduced stress fibers, compared to fibroblasts plated on a fibronectin matrix, that inhibition experiments determined was due to TN-C-mediated suppression of Rho [680]. A further study on senescent fibroblasts has demonstrated trailing F-actin [664] as seen in the partial KOs with the most reduced TN-C expression and in the complete *TNC* KOs. Another of the three filament systems comprising the cytoskeleton, vimentin intermediate filaments, has been linked to changes during aging and senescence. Reduced motility, increased adhesion and cell spreading in aged fibroblasts has been attributed to increased vimentin intermediate filament assembly [664, 681]. Supporting this, here, vimentin expression was upregulated in *TNC* KO fibroblasts by 4.1 log2Fc, although this was not significant.

RNA-Seq data and GSEA analysis also demonstrated that actin regulation was altered. Two genes directly involved in the polymerisation (profilin-2) and depolymerisation (cofilin-2) of the actin cytoskeleton [659] were significantly differentially expressed. The changes were slightly surprising – cofilin severs actin filaments and this was upregulated, whilst profilin, responsible for polymerisation, was downregulated. Actin construction and deconstruction (actin treadmilling) is a normal process that facilitates lamellipodium formation and cell cycle progression and this is a complicated process with many levels of regulation [682]. The increase in cofilin expression may explain the lattice effect seen – severed F-actin filaments could accumulate in the cell. These severed filament may also explain the unjointed cytoskeletal ‘blebs’ seen in the complete knockout fibroblast and the partial knockout with the largest knockdown. Furthermore, when cofilin proteins are at high concentration in the cell, they can promote spontaneous nucleation: polymerisation of monomeric G-actin to F-actin filaments *and* stabilise existing filaments [683, 684]. This up-regulation of cofilin-1 results in decreased actin

depolymerisation and morphological enlargement of senescent cells; over expression of cofilin promotes cell senescence in WI-38 fibroblasts, and this was not dependent on p53 or p16 [685], which is in line with the finding of this study in which these two latter genes were not differentially expressed. Knockdown of profilin-2 has also been shown to reduce lung cancer cell proliferation [686] suggesting that a reduction of this protein in the *TNC* knockout cell line may further contribute to reduced proliferation. In addition, in this study *APLP2* exhibited the largest downregulation, and in pancreatic cancer cells, downregulation of *APLP2* resulted in significantly reduced cortical actin, and increased intracellular actin filament accumulation, mirroring the phenotype observed in the present study [687]. In MCF10A breast cancer cells, utilising BrdU incorporation to determine cell cycle progression, and Arp2/3 inhibitors to inhibit cortical actin network formation, it was demonstrated that cortical branched actin actually controls cell cycle progression [688]. This work built on the decades old study, which utilised cytochalasin D and latrunculin A to inhibit actin polymerisation and halt the cell cycle in mouse fibroblasts [689]. Furthermore, in yeast fission studies, a final check point in the cell cycle monitors the actin cytoskeleton, ensuring proper organisation; if the cytoskeleton is not organised correctly, the yeast cells do not divide [690]. Cell cycle arrest then typically leads to senescence [691]. *ENAH* was one of the largest upregulated genes – this protein is known to promote actin filament extension [692], explaining the much longer actin filaments seen in the *TNC* KO fibroblasts. Together, these studies explain the increased and dysregulated actin network seen in the complete *TNC* KO fibroblasts (and likely the partial KOs too) actin networks. It appears that *TNC* controls morphological characteristics via regulation of actin related genes. These studies also demonstrate why these *TNC* mediated changes in gene expression likely result in cell cycle arrest and this arrest can typically lead to senescence.

Actin stress fiber formation has been linked to yes-associated-protein (*YAP*) and transcriptional coactivator with PDZ-binding motif *(TAZ*), which are transcriptional cofactors that can associate with several transcription factors, although mainly TEA domain family member (*TEAD*)[693]. These proteins are the main components of the Hippo pathway, which phosphorylates YAP/TAZ thus leading to their cytoplasmic retention and degradation. Actin stress fibers inhibit the Hippo signalling pathway and subsequent phosphorylation of YAP leading to its accumulation in the nucleus and subsequent YAP

activity [694]. Similarly, and of particular relevance here, TN-C inhibits actin stress fiber formation and abolishes YAP activity via binding of α9β1 integrin, in human osteosarcoma cells. The same authors also demonstrated that YAP participates in a feedback loop in which inactive YAP downregulates TN-C [695]. Supporting this, RNA-Seq data demonstrated increased YAP expression in *TNC* KO fibroblasts. Dupont *et al*. established that YAP/TAZ regulation occurred independently to the Hippo pathway, however it was still dependent on actin stress fiber formation and cytoskeletal tension [696]. TN-C-mediated abrogation of actin stress fiber formation is supported by our results in which partial and complete *TN-C* KOs display dramatic increases in stress fiber formation. In addition, YAP/TAZ has been shown to activate FOXD1 and regulate senescence in human stem cells [697]. Together, this suggests a potential role for TN-C-mediated YAP/TAZ/TEAD activity in the generation of actin stress fibers, aging, enlarged phenotypes and senescence. However, unlike *YAP*, *TAZ* and *FOXD1* were downregulated by 0.4 and 0.3 log2Fc in *TNC* KO fibroblasts compared to WT fibroblasts. KEGG GSEA revealed that differentially expressed genes were enriched in the hippo signalling pathway; other proteins upstream of YAP/TAZ in the pathway may be modulated to result in an overall pathway change. As senescence is a process that develops over time, possibly this signalling pathway was in the process of upregulation, but had not yet reached that point yet. Alternatively, WT fibroblasts were also starting to age and thus the aging signalling pathways were upregulated in WT fibroblasts already, thus hiding a difference. Ideally, a young WT control would be utilised to assess this.

Other differences in the cytoskeleton may be attributable to integrin linked kinase (*ILK*). This is a factor that links the cytoskeleton to cell signalling and adhesion. In aged rat fibroblasts ILK levels are increased compared to young rat fibroblasts, and ILK expression was correlated with a senescent phenotype and increased vimentin filament formation [698]. However, RNA-Seq indicated that, conversely, ILK expression was downregulated 0.12 log2Fc. Overexpression of *ILK* in young rat fibroblast resulted in a senescent cell phenotype as measured by cell enlargement, increased vimentin filaments, SA-BG staining and the induction of the cell cycle repressors p53 and p21. P53 was upregulated although the difference was not significant, whilst p21 (*CDK1)* was not detected. P53 will

be discussed in more detail later. The fact that differences are not seen in line with the literature could be that these cells were not fully senescent yet.

SA-BG staining of old WT and old *TNC* KO fibroblasts at PDL 55 to assess senescence revealed that there was a marked difference in staining prevalence and intensity between the two fibroblast populations. This result supported the cell-phenotype observations in which old *TNC* KO fibroblasts were significantly more enlarged and rounder than old WT fibroblasts, whilst old WT fibroblasts were similar in size and roundness to young WT fibroblasts. Furthermore, RNA-Seq data indicates an upregulation of beta-galactosidase in *TNC* KO fibroblasts supporting the phenotype seen in the staining assay.

Whilst SA-βG detection is considered a gold standard and utilised by thousands of studies of senescence, this method is not without limitations. High SA-βG activity has been demonstrated in conditions unrelated to senescence at pH 6.0, for example in confluent cultures and in serum starved fibroblasts [699-701]. There is also a strong evidence supporting the argument that SA-βG activity is actually just lysosomal β-Gal activity that is increased enough, due to increased lysosome number, to allow detection at suboptimal pH 6 [702, 703], [643, 703]. Supporting this, lysosome number increases with age in fibroblasts [704] and young fibroblasts displayed strong staining after being left in SA-βG staining for an extended period of time [700]. In lysosomes in aged cells, increased amounts of indigestible and fluorescent lipofuscin accumulate. This is seen as a ‘wear and tear’ product of lysosomal digestion. These lysosomes are suspected to be less efficient at turning over damaged organelles, such as mitochondria, which result in increased ROS and contribute to accelerated aging [705]. Whilst SA-βG is widely used to infer senescence, this method should not be the sole experimental approach to determine senescence, however the SA-βG staining do support the other experiments indicating a senescence phenotype.

Assessment of the proinflammatory miR-155 demonstrated a slight discrepancy in that the RNA-Seq data did not demonstrate the same reduction in miR-155 that was demonstrated by qPCR. However, RNA-Seq data always requires validation by individual gene expression assays and studies have shown that differences in miRNA detection can occur even between small RNA-Seq platforms [706, 707].

Assessment of the proinflammatory SASP factors IL-6 and IL-8 by ELISA indicated that *TN-C* fibroblasts produced less inflammatory components. This was also supported by the RNA-Seq data. The reduction in inflammatory components does not correlate with the morphological changes seen, in which *TNC* KO fibroblasts appeared to have started to undergo senescence and therefore should produce more IL-6 and IL-8. The decreased cytokine production may seem surprising, given that the *TNC* genotype drastically affects cell characteristics such as cell size and shape. However, genetic ablation of *TNC* in animal models has been shown to result in impaired cytokine synthesis in whole mouse joints (where synovial fibroblasts reside) and primary cells isolated from these joints [235, 324]. Conversely, but supporting this, treatment of cells with the FBG domain of *TNC* induces cytokine synthesis [235, 522]. It is well known that TN-C regulates miR-155 production via TNF-α [188] and TNF-α regulates the expression of IL-6 and IL-8 [708]. As such, *TNC* KO fibroblasts, senescent or otherwise, are likely to display very different SASP characteristics compared to typical senescent fibroblasts, due to a reduced inflammatory signalling capability.

Supporting this, the fibroblasts did not display a classical SASP gene signature seen in one general SASP catalogue [436], one fibroblasts specific SASP catalogue or a fibroblasts specific meta-analysis study of senescence [437]. Whilst some factors were differentially regulated these often were not in the expected direction. However, the studies utilised as references indicate that the SASP changes throughout time, and whilst classical factors like IL-6, IL-8 are usually upregulated in expression, the lack of TN-C production is likely to result in impaired cytokine production, as explained above. This means that senescence cannot be ruled out based on the SASP profile, and the *TNC* KO fibroblasts are likely exhibiting a *TNC* KO specific SASP, with other morphological changes in line with previous studies. It seems highly likely that the *TNC* KO fibroblasts are exhibiting a premature aging and replicative senescent phenotype, based on the results from comparing the gene signature of the *TNC* fibroblasts, to that documented for replicative senescence in the Cell Age database.

Indeed, this conclusion is supported by the RNA-Seq data which indicates that at the transcriptional (mRNAs) and posttranscriptional (miRNAs with their target genes) levels the genes that are differentially expressed are enriched in specific senescence or cell cycle pathways, but also pathways

directly linked to these pathways, including p53, and FoxO signalling. The latter is of particular interest, given the recent paper by Wang *et al.* that indicated TN-C regulation by FoxO3a [629]. FoxO3a and p53 also interact directly [629]. The MAPK pathways are also of particular interest, given that increased MAPK signalling (demonstrated by RNA-Seq data for both differentially regulated mRNA and enriched targets of differentially expressed miRNAs) can lead to senescence when the cell cycle is arrested, as confirmed with assessment of the actin cytoskeletal dysregulation [670]. The fact such pathways are indicated supports the theory of senescence in the *TNC* KO fibroblasts.

A literature search of the top three downregulated and upregulated miRNAs did not result in any direct links to senescence, however each of these miRNAs has hundreds of gene targets. Notably, miR-615-3p has been linked to replicative aging due to its role in supressing hTERT expression, however this is not directly relevant here given that hTERT expression is barely detectable in adult somatic cells [589]. A study on human mesenchymal stem cells demonstrated miR-618-3p was downregulated, however many more miRNAs, none of which match those seen here, were downregulated [709]. Again this is not surprising since senescence is cell-type specific. miR-196a and miR-490 are downregulated in senescent IMR90 fibroblasts[710], whilst miR483 is downregulated in senescent WI-38 fibroblasts[711], matching the expression profile seen here in the *TNC* KO fibroblasts. Another interesting miRNA that was downregulated in the *TNC* KO fibroblasts is miR-335, which is known to target *TNC* in cancer [226]. This may be a good candidate miRNA to specifically assess first, given that cancers result from dysregulated and uncontrolled cell cycle progression. The largest differentially expressed miRNAs are also not documented to have any direct targets directly affecting the cell cycle, such as cyclins or CDKIs [712]. Although, key targets generated *in silico* using MirWalk [661] repeatedly identified genes enriched in various pathways relating to the cell cycle and cell proliferation, including *AKT2*, cyclin D2 and D3. Cyclin D2 should be prioritised for investigation since this was significantly differentially expressed. Progression of cells through the cycle requires the formation and activation of cyclin-dependent kinases (CDKs); the type – D cyclins bind to CDK4 and CDK6 to form these complexes[713]. Interestingly, cyclin D2 and D3 have been shown to have opposing roles in the development of tumours in mouse skin [714]. Overexpression of one resulted in downregulation of the

other (confirmed here in this study), with cyclin D3 inhibiting malignant progression, whilst cyclin D2 drove malignant progression. This supports another study which demonstrated that in human induced pluripotent stem cell–derived cardiomyocytes (hPSC-CMs) overexpression of *CCND2* activates cell cycle progression[715]. These studies do not support the observations here in *TNC* KO fibroblasts. Given that a cumulative effect of multiple genes is necessary for many biological processes, the role of the other genes that were repeatedly identified by MirWalk should not be ignored either. For example, cyclic D3 was downregulated in *TNC* KO fibroblasts and in human colon cancer cells *HTC116, CCND3* knockdown using miRNA mimics has been shown to result in cell cycle arrest [716]. Nevertheless, it seems that actin cytoskeletal changes are likely the driving factors towards arrest in the *TNC* knockout fibroblasts. Many targets have been identified that require further probing and validation.

The lack of specific links of the cell cycle genes or senescence pathways directly to *TNC* could indicate another mechanism caused these phenotypic changes and the onset of cell cycle arrest/senescence. A new area of research, resulting from the surge in CRISPR/Cas9 experiments, is focussing specifically on DBS breaks and the effect this has on cell cycle progression. It has been demonstrated in immortalised (hTERT) retinal epithelial cells (RTE) that delayed processing of double strand breaks induced by irradiation can result in cell cycle exit and the onset of senescence [717]. This occurred in a p21 dependent manner – which we did not observe significantly increased expression of. Furthermore, this study also established that it is not solely the number of breaks that determines the fate as cells with < 10 breaks did not recover whilst cells with >20 breaks did. However, using a doxycycline-inducible CRISPR/Cas9 system in RTE hTERT cells, it has been demonstrated that a single DSB is enough to induce cell cycle arrest [718]. This may be due to CRISPR/Cas9 induced breaks taking longer to repair than irradiation breaks [718] possibly because Cas9 remains bound to broken DNA [719]. However, it was also shown that this cell cycle arrest did not lead to cell cycle exit and was simply a pause to facilitate DNA repair before progression through the cell cycle [46]. These two studies also show this cell cycle arrest occurs very short terms after DSBs are induced (less than a week). Therefore this seems unlikely to be the case in *TNC* knockout fibroblasts which can be cultured for several weeks before growth arrest occurs.

Generally, the pathways indicated all have links to aging. A very recent study by Choi *et al.* investigated TN-C expression in fibroblasts with regard to ECM maintenance and aging [522]. Comparing tissues from young and aged mice and humans revealed, for the first time, that *TNC* expression is downregulated during intrinsic skin aging. This study also found that treatment of fibroblasts with TN-C resulted in an upregulation of *COL1A1 and COL1A2* and a downregulation of *MMP1*. This is not surprising given TN-Cs role in the wound healing response [177]. Conversely, in *TNC* KO fibroblasts, *COL1A1* and *COL1A2* were upregulated, whilst *MMP1* was downregulated, but none of these differences were significant. However, this study utilised the addition of recombinant full length TN-C to fibroblasts, rather than genetically ablating TN-C as done here, and this may explain the differences. This was also not carried out in replicative aged fibroblasts and so is not directly comparable with the study here. This study identifies that TN-C regulates age related changes, the exact mechanism is yet to be defined. It would be interesting to see if the fibroblasts from the aged tissues in this study displayed altered actin cytoskeleton networks.

Taken together, it appears that age-related changes occur in fibroblasts but that *TNC* ablation results in accelerated aging-related morphological changes. It is still difficult to determine whether these cells have undergone premature senescence or just accelerated aging – the latter would eventually lead to senescence. Whilst many of the classical characteristics of the SASP are missing, this is likely due to the fact that *TNC* directly and indirectly regulates the production of inflammatory 'SASP' cytokine synthesis, and as such, *TNC* KO cells are incapable of producing the classical SASP signature. The stark dysregulation of the actin cytoskeleton points to an arrested cell cycle at the very least, which could result in senescence if it is prolonged. It may be that *TNC* ablation results in an accelerated aging-phenotype compared to WT cells of the same age, resulting from a dysregulated actin cytoskeleton that then leads to cell cycle arrest. This prolonged cell cycle arrest is then indistinguishable from senescence, since the *TNC* KO fibroblasts are incapable of producing a classical SASP. However, it may be that the fibroblasts have not yet entered senescence and have been studied at the time point of prolonged cell cycle arrest.

7 CONCLUDING THOUGHTS

The ECM plays key roles in the regulation of genes that direct and influence all aspects of cell behaviour. More recently, this regulatory role has been extended to the inflammatory response, in which TN-C, a matricellular ECM protein, was shown to post-transcriptionally regulate the expression of LPS-induced miR-155 and TNF-α production in mice, both of which were required for an appropriate inflammatory response [188]. Many questions arose from this recent finding, including on the specific mechanism by which this occurs as it was not clear whether this effect was due to ECM-assembled, insoluble TN-C or secreted, soluble TN-C, which becomes elevated during infection or inflammation in serum of patients [187, 720]. It was also unknown whether this TN-C/ miR-155/TNF-α regulatory axis occurs in humans. Countless options of ECM models are currently available [389], however these are subject to many limitations which exclude these from use in our study. Such limitations include being derived from non-human sources, and thus immunogenic, or from tumours, being synthetic, containing microbial contaminations, or simply failing to recapitulate the natural complexity of the ECM [389].

To overcome these limitations and allow the elucidation of the role of the ECM protein TN-C in the LPS-induced inflammatory response, I addressed three key objectives. The first was to develop a robust and reproducible, physiologically relevant, BJ-derived ECM model, that lacked any of the key limitations mentioned above. The second objective was to create a BJ-derived ECM model lacking TN-C. To do this, I utilised CRISPR/Cas9 genome editing to produce a *TNC* KO BJ fibroblast cell line. The third objective was to use these generated ECM models, WT and KO, as a substrate for the culture of macrophages to assess LPS-induced inflammatory gene expression. Together, these would allow the elucidation of how the TN-C regulates inflammatory gene expression in macrophages, and address my project ‘engineering extracellular matrix factories to study how the extracellular microenvironment influences gene expression’.

With regard to the first objective, I developed a robust and reproducible BJ-derived ECM model. *In vitro*, cells deposit ECM components with a similar construction and organisation to that of their *in vivo* ECM [389, 479]. BJ cells were utilised due to their inherent ability to secrete ECM physiologically similar to interstitial ECM, which is found throughout the body and is comprised of the structural ECM proteins including collagens and fibronectin, as well as proteoglycans, growth factors and remodelling enzymes such as MMPs [479]. I built on previous work from our group which characterised this model and confirmed, by immunofluorescence, that these ECMs were consistently comprised of collagen 1, collagen 3 and fibronectin, core structural ECM proteins, as well as rich in GAGs, and that this was a 3D ECM. One of the key steps to the generation of this ECM model was the removal of cells and DNA that would otherwise act as antigens and generate an immune response from macrophages seeded on this substrate [14]. To do this, I employed a decellularisation buffer containing Triton X-100 followed by an additional step of DNase I treatment which successfully removed cells and DNA without stripping GAGs, which can occur when using Triton-based decellularisation buffers [466]. I also characterised the thickness of the ECM by immunofluorescence, a widely used technique, and additionally by optical profiling. The latter has the advantage of a smaller nm range, thus resulting in more accurate measurements than confocal, which measures in μm slices. Moreover, optical profiling can assess every protein within the CDM, as opposed to the former, which assessed one highly prevalent protein, such as fibronectin or collagen [447, 473-476]. This is likely why the thickness measurement I attained was far thicker than those already described in the literature from fibroblasts that relied on the immunofluorescence method [497]. Optical profiling also overcame issues of CDMs shrinkage caused by fixation that has been documented before [497].

Linked to this first objective, the generation of BJ-derived CDMs lacking TN-C was essential to determine whether ECM-assembled TN-C regulates LPS-induced miR-155 expression in human macrophages; a comparison of macrophage gene expression on both BJ-derived CDMs would allow clarification of differences. This approach is similar to a reverse genetics approach in which the gene of interest is knocked out in order to ascertain the gene function. I designed guides targeting specific regions of the *TNC* gene in exon 2 and 7, and successfully cloned these into the CRISPR/Cas9 vectors,

using both Cas9 WT and nickase plasmids methods and finally the RNP method. I optimised every step of the CRISPR/Cas9 method, from the number of cells and plasmid quantity for transfection/nucleofection of cells to the optimum conditions for maximum single cell recovery, following sorting of GFP positive, Cas9 expressing cells. Using CRISPR/Cas9, I successfully generated a *TNC* KO fibroblast cell line by transfecting with two plasmids targeting two exons (2 and 7) of the *TNC* gene simultaneously. This resulted in a 12kb inversion spanning the DSB from both target sites and flipping of top and bottom DNA strands within this inversion, thus knocking out protein production as confirmed by western blotting. Large inversions of several hundred Kb to Mb in a range of cell types and whole organisms have been documented with the transfection of pairs of guides targeting two regions in the gene concurrently [408, 524, 526, 527]. This seems like a highly successful approach to gene silencing, even if an inversion does not occur, two DSBs that have the potential for two sites of INDELs are likely to knock out the gene.

I also carried out a comprehensive screening of predicted off-targets utilising the Benchling software, which utilised the algorithms from the MIT software [435], which was recently closed down, plus some additional algorithms to generate scores for on- and off- target scores; the higher the score the more likely the cleavage with the highest score possible 100 for the targets within the *TNC* gene. Whilst the 'NGG' PAM sequence is a prerequisite for Cas9 cleavage, it has been demonstrated that 'NAG' PAM sequences are tolerated albeit with less cleavage efficiency [406, 523]. Therefore, all guides 'NGG' predicted off-targets and the highest scoring 'NAG' target was assessed by PCR amplification and sequencing. The scores for 'NGG' targets were very low and 'NAG' targets almost negligible. No off-targets were detected in any predicted off-targets. Criticisms may arise in the form of suggesting the amplicons were too small, however the primers were designed so that the predicted off-target site was in the middle of the amplicon, and both this work and others demonstrate specific cleavage consistently 3bp upstream of the PAM site [526]. Thus, I am confident these screening regions were appropriate. Furthermore, RNA-Seq of WT and *TNC* KO fibroblasts was utilised to probe other predicted off-targets. Unfortunately, not all off-target transcripts were detected in either cell line, however those that were, were not significantly differentially expressed between WT and KO fibroblasts, indicating no large

mutations had occurred, and did not contradict the PCR results. Together, with these two aforementioned objectives I have generated a biological ECM tool to allow the study of healthy human ECM, as well as ECM lacking a specific ECM molecule, and its impact on cell behaviour.

The third objective was to utilise the biological ECMs as a substrate for other cells. Primary human macrophages were used here due to their pivotal role in the inflammatory response, constantly sampling the ECM environment and recognizing both exogenous PAMPs and endogenous DAMPs [235, 282, 283, 285, 721, 722]. Preliminary work indicated the BJ-derived CDM as a suitable substrate that allowed normal macrophage maturation. Furthermore, macrophages exhibit more physiological behaviour on the CDM substrate compared to plastic, with adhesion and inflammatory cytokine expression profiles more similar to that observed *in vivo,* with stringent regulation of the latter. Initial work of LPS stimulation of macrophages on both *TNC* KO CDM and TN-C+ CDM demonstrated, for the first time in humans, that *TNC* is involved in the regulation of LPS-induced inflammatory gene expression – specifically miR-155 and subsequent production of TNF-α expression. This confirms the previous work that demonstrated a *TNC*/miR-155 axis in the mouse [312] and previous work demonstrating miR-155 as a pivotal, proinflammatory cytokine [376, 377, 563, 723].

Together these three objectives allowed the development and validation of a biological ECM tool utilised as a model of inflammation in humans. The development of this tool also overcomes many, if not all of the key limitations to the ECM models currently on the market to study healthy ECM, including inconsistency between batches, synthetic/animal or tumorous origins and structural disparity between native ECM [389]. This study also highlights the scope of this ECM tool: this ECM model could be genetically modified in a number of ways that, in conjunction with the culture of diverse cell types, could allow the elucidation of the ECMs gene modulatory role in disease pathogenesis with the potential to highlight novel targets for therapies. Potential modifications could include knockout of other ECM proteins, or addition of other proteins, as was carried out with the ‘overexpression’ model.

In addition to successfully completing the original objectives, I demonstrated a novel finding that genetic ablation of *TNC* resulted in stark morphological changes, specifically concerning the actin cytoskeleton network arrangement that was highly dysregulated. KO of *TNC* appeared to accelerate the

aging phenotype of WT fibroblasts and result in cell cycle arrest, as confirmed by a number of approaches, including RNA-Seq analysis. Unfortunately, confirmation of whether this was a direct effect of *TNC* ablation was not possible, however I identified a transcriptional effect with the downregulation of numerous miRNAs and altered regulation of many mRNAs that requires subsequent validation. This effect was novel, while *TNC* has been previously implicated in actin cytoskeleton reorganisation; suppressing actin stress fiber formation in mouse 3T3 fibroblasts [680] and T98G:TM1 cancer cells, as well as inhibiting spreading [724] similar to that seen here, there is very little is known about TN-C and premature aging (more on this later). However, a recent paper did identify that TN-C contributes to the maintenance of normal ECM, via the upregulation of collagen expression, and that in the aged skin tissue of humans and mice TN-C and collagen expression is decreased [522]. This confirmed the findings of a previous study which also demonstrated *TNC* expression was downregulated in aged mouse and human skin tissue and fibroblasts from aged tissue, proposing that *TNC* downregulation is a characteristic of aged skin [725]. This therefore begs the questions: does TN-C promote the maintenance of cells and tissue, and does TN-C ablation result in premature aging? The results of this study support the premature aging theory, both in the complete and partial *TNC* KOs. However, RNA-Seq data did not replicate the findings that type I collagen expression was downregulated with *TNC* knockout; an upregulation was demonstrated here. But the experimental systems by Choi *et al.* were different to those here – the complete knockout of TN-C *in vitro* as opposed to a natural reduction *in vivo* in aged human and mouse tissue likely results in differences in gene regulation. RNA-Seq of WT or *TNC* KO fibroblasts at a younger age, for comparison with the 'old' fibroblasts may well have revealed starker differences. In this instance, the partial knockouts could provide the basis for further investigation, and could be utilised to assess whether differences in collagen expression occurs between the cell populations in a TN-C dose-response effect. Furthermore, whilst there were no significant differences detected between WT and *TNC* KO fibroblasts, this assumes that transcription and protein production correlate well. Whilst I have confirmed the successful deposition of collagen protein into the *TNC* KO–derived ECM by immunofluorescence, this was at a younger age (1 PDL younger) than the *TNC* KO fibroblasts assessed by RNA-Seq, Thus, very slight differences may have occurred, however I would expect these to be minor.

The final objective was to determine whether TN-C plays a role in cellular senescence and assessment of key senescent characteristics, including cell size, actin cytoskeleton arrangement and SA-βG staining supported the theory that *TNC* KO cells had entered a senescent state [581, 608, 609, 643, 647]. Assessment of the SASP blurred the conclusion of senescence. The SASP is characterised by an inflammatory profile; common components are IL-6 and IL- 8, however a consensus is lacking on the specific profile between and within cell types as this can be a highly variable phenotype [437, 438, 603]. It was not possible to confirm the acquirement of a SASP in the *TN*-C KO fibroblasts given that *TNC* drives inflammatory cytokine synthesis [235, 309, 311]. Correspondingly, mice lacking TN-C have impaired inflammatory cytokine production [324] and supporting this the KO fibroblasts produced less inflammatory cytokines than WT fibroblasts. A classical SASP is therefore unlikely to be produced by these fibroblasts, even with additional stressors applied known to induce the SASP.

On the topic of the aging phenotype demonstrated in the *TNC* KO fibroblasts, aging is the largest risk factor for the development of a myriad of diseases in humans and mice. Aging is the accumulation of detrimental alterations at the molecular level, leading to cellular degeneration, functional decline and eventual morbidity [726]. One such aging related disease (along with other factors) is osteoarthritis (OA), which *TNC* KO mice are shown to spontaneously develop [727]. Using surgically induced OA in both KO and WT mice led the authors of this study to identify that TN-C is required for cartilage regeneration. This echoes the study by Choi *et al*. [522], which demonstrated that collagen expression and subsequent ECM integrity is compromised with age and a natural reduction in *TNC* expression. An additional study revealed that chondrogenesis is impaired in *TNC* KO mice, further highlighting the notion of TN-Cs role in ECM integrity and instability [728], and the lack of which results in aging and degenerative changes to the ECM that are propagated to the resident cells.

Following this *TNC*-regulated premature aging theory further, as *TNC* is crucial in the generation and propagation of an inflammatory response, is premature aging caused because of a reduced inflammatory signalling capability, or is this an effect of aging, or both? The degenerative process of aging is interwoven with inflammation and the former question has already warranted consideration [729]. However, it has generally been shown that chronic inflammation is associated with age related

degenerative neurological diseases such as Alzheimer's and Parkinson's [730, 731] and levels of inflammation components typically increase with age, in the absence of infection or other stress [732]. An accumulation of oxidative-induced molecular and cellular damage are thought to result in immunosenescence, which is the steady decline of the immune system with age. This is of particular relevance with the current pandemic. In aged individuals, improper immune function combined with exposure to antigens and pathogenic insults then results in chronic low grade inflammation, which is exacerbated by the production of inflammatory components from senescent cells and their SASP [732, 733]. In this way, aging and inflammation both cause and contribute to the other. There are many other theories on aging and inflammation, however I have focussed on the most relevant. Clearly, a combination of many factors and pathways contribute to the complex process of aging that requires further investigation.

Overall, I have developed a biological ECM tool that started to elucidate the molecular mechanisms by which the ECM protein TN-C regulates the expression of LPS-induced inflammatory cytokine synthesis. I have also made a potential novel finding that TN-C is involved in the regulation of cell cycle progression and subsequently senescence, and contributes to maintenance of the cell architecture. Further work is required to fully understand by what precise mechanism this occurs.

8 FUTURE WORK

The overall aim of this thesis was to investigate the role of the ECM glycoprotein TN-C in the regulation of gene expression. To do this, I developed a reproducible, acellular, human ECM model, utilising BJ-derived CDM. To assess the specific role of TN-C in the BJ-derived CDM, I genetically ablated *TNC* using CRISPR/Cas9 and utilised these CDMs as substrates for the culture of primary human macrophages. LPS stimulation was utilised to simulate an inflammatory response, at which point the expression of miR-155 was assessed via qPCR. Together, this indicated, for the first time in humans, that TN-C drives the expression of miR-155 and subsequent TNF-α production. Furthermore, an unexpected finding, resulting from *TNC* KO in fibroblasts, was that cells underwent morphological and biochemical changes that were highly indicative of senescence. However, future work is required on each of these topics.

The current KO cell line should be utilised to generate more replicates of the experiments utilising the ECM models to assess inflammatory gene expression from macrophages by assessing miR-155 expression and TNF-α production. The current N of one donor is not enough to draw any conclusions; a minimum of 5 biological replicates are needed. Furthermore, the assessment of other inflammatory miRNAs should be carried out. To build on the preliminary work, miR-132 should be one such candidate as this miRNA is indicated as anti-inflammatory [734]. In this way, a more thorough evaluation of the inflammatory response, including both the initiation and the resolution, would be possible. This approach could include any miRNA with key roles during the inflammatory response. A cytokine array could be utilised to determine any other specific TN-C targets. As an additional layer, the macrophages could also be seeded on the partial *TNC* KO fibroblast-derived-CDMs to confirm if a dose-response of inflammatory genes expression was observed with increased reduction in *TNC* expression. However, this would first require assessment of the partial KO CDMs, specifically TN-C incorporation. The use of the partial KOs to generate CDMs could also provide specific clues about receptor involvement given the targets in different exons. As the 7A guide targeted exon 7 in the RGD

domain, whilst the A2 guide targeted the start codon, these two could be compared to determine if integrin-mediated binding results in dysregulation of the actin cytoskeleton.

An additional layer, once these experiments have been completed, would be to develop the model even more physiological by utilising titrated bacteria to induce the inflammatory response, instead of LPS. This type of experiment would also provide the stage to study ECM degradation resulting from infection, for example by bacterial proteases. Candidate bacteria should be clinically relevant and normally colonise the skin niche, for example *Pseudomonas aeruginosa*.

Whilst a significant amount of work investigating the composition of the CDM has been carried out, mass spectrometry analysis could provide a fully comprehensive analysis of the constituent components. Given that the *TNC* KO fibroblasts appear to have an accelerated aging phenotype and arrested cell cycle and that aging is usually associated with a functional decline, it would be interesting to specifically assess whether ECM deposition is affected. Whilst the immunostaining and RNA-Seq did not indicate there would be significant changes here, this requires further confirmation at the protein level. Unfortunately, due to Covid-19 there was not enough time to assess the ECM deposited by the partial knockouts which may have provided a deeper understanding of both genotype and age related changes, since these partial knockout cell populations are far younger than the *TNC* KO fibroblasts. Initial work should start with these cells to confirm the successful production and deposition of collagen 1and 3 and FN. Furthermore, given the wide ranging and far-reaching effects of stiffness on gene regulation and cell fate [505], the stiffness of the deposited CDM should be assessed. I had started to assess the stiffness of the CDM by atomic force microscopy (AFM), however due to the complexity of the substrate (hydrophilic, adhesive/sticky) and a lack of time due to Covid-19 this experiment did not develop far past initial training.

The overexpression model, in which full-length TN-C was incorporated into WT CDMs, supported the data from the *TNC* KO fibroblasts and previous work in mice [188, 735], indicating ECM assembled TN-C regulates the expression of macrophage produced miR-155 in response to LPS. Immunofluorescence also confirmed successful incorporation into the matrix at all tested concentrations. An optimum concentration should be decided and then this experiment repeated to

determine the expression of LPS-induced miR-155 from macrophages seeded on these matrices. Furthermore, the production of TNF-α should also be assessed. Whilst this looks to be a simple, yet promising 'over expression' model, in order to be directly comparable with the *TNC* KO fibroblasts results an overexpression model could be generated using CRISPR/Cas9 activation (CRISPRa). This method uses a nuclease-deficient *S. pyogenes* Cas9 (dCas9), complexed with a variety of transcriptional activator domains (dCas9-VP16 or -VPR), or a scaffold which recruits activator proteins (SunTag), or an activator and a tagged gRNA which recruits other activators (SAM), with guides that are complementary to the gene promoter region, thus facilitating the initiation of transcription [736]. The addition of specific RNA components further enhances the initiation ability [737]. However, whichever activator domains are used, plasmid transfection and single cell sorting methods should be avoided due to the previous difficulties in generating a clonal cell in this way. Lentiviral methods would be the most efficient, however rounds of RNP nucleofection would likely result in high percentage 'overexpression' populations. Additionally, inducible variants of these activators are also possible; with light inducible systems this allows temporal and possibly dose control of overexpression [738].

Given the apparent cell cycle arrest/onset of senescence in partial and complete *TNC* KO fibroblasts, future work could utilise immortalised BJ fibroblasts (hTERT) to generate *TNC* knockout fibroblasts with exactly the same guides and utilising the same methods as done in this thesis. The creation of WT hTERT and *TNC* KO hTERT BJ fibroblasts would allow the assessment of several things. First, the current work did not determine if telomere or DNA damage resulted in cell cycle arrest. Utilising BJ fibroblasts that are able to replicate telomeres and assessing if the same morphological changes occur, or not, would allow elucidation as to whether telomere shortening results in morphological changes, in addition to TN-C ablation. Secondly, given that BJ fibroblasts have a finite lifespan, the generation of an immortalised ECM-depositing cell line would allow for further investigation into the effects of the ECM microenvironment on inflammatory gene expression, when the CDMs from these cells are utilised as a substrate for macrophages.

On the topic of cell senescence, whilst the experiments indicate cell cycle arrest and most probably senescence, it is difficult to confirm in fibroblasts that are incapable of producing a SASP. Many

candidate genes were identified in the senescence/cell cycle pathway/ FoxO pathway by RNA-Seq and indicated as targeted by differentially expressed miRNAs using KEGG pathway analysis. The genes in these pathways should be assessed by qPCR and western blotting in order to determine the specific effect of *TNC* ablation on these genes and their encoded proteins. To properly confirm cell cycle arrest, a BrdU incorporation assay could be utilised to compare cell proliferation between the partial and complete *TNC* KOs fibroblasts and WT fibroblasts. Cell cycle analysis with flow cytometry and PI or Hoechst could also be utilised. Furthermore, analysis of p53, p16 and p21 and phosphorylation of the histone H2AX should be assessed by blotting/qPCR and immunofluorescence.

Thinking about BJ-derived-CDMs in a clinical application, CDMs could potentially be utilised as an ECM bioscaffold for chronic wounds. A huge range of such types of 'patches' are currently available on the market, however, as with most ECM models, these are derived from animals, specifically of porcine origin. These are therefore not suitable for people with allergies. However, patches like Alloderm are derived from donated human skin, decellularised and currently utilised as a sheet in reconstruction surgeries [739] or for the management of full thickness burns [740], providing structural support as well as encouraging fibroblast proliferation and angiogenesis. Taken one step further, fibroblasts could be isolated from the patient and utilised to create a fully personalised, compatible ECM scaffold. The potential for this biological tool is clear, although the above work needs to be completed in order to progress this to a validated model.

9 REFERENCES

1. Frantz, C., K.M. Stewart, and V.M. Weaver, *The extracellular matrix at a glance.* J Cell Sci, 2010. **123**(Pt 24): p. 4195-200.
2. Naba, A., et al., *The extracellular matrix: Tools and insights for the "omics" era.* Matrix Biol, 2016. **49**: p. 10-24.
3. Hynes, R.O. and A. Naba, *Overview of the Matrisome-An Inventory of Extracellular Matrix Constituents and Functions.* Cold Spring Harbor Perspectives in Biology, 2012. **4**(1).
4. Dunsmore, S.E., *Treatment of COPD: A matrix perspective.* International Journal of Chronic Obstructive Pulmonary Disease 2008. **3**(1): p. 9.
5. Bianco, P., et al., *Expression of bone sialoprotein (BSP) in developing human tissues.* Calcif Tissue Int, 1991. **49**(6): p. 421-6.
6. Termine, J.D., et al., *Osteonectin, a bone-specific protein linking mineral to collagen.* Cell, 1981. **26**(1 Pt 1): p. 99-105.
7. Holliday, L.S., et al., *Initiation of osteoclast bone resorption by interstitial collagenase.* J Biol Chem, 1997. **272**(35): p. 22053-8.
8. Deák, F., et al., *The matrilins: a novel family of oligomeric extracellular matrix proteins.* Matrix Biol, 1999. **18**(1): p. 55-64.
9. Mithieux, S.M. and A.S. Weiss, *Elastin.* Adv Protein Chem, 2005. **70**: p. 437-61.
10. Roberts, D., *Emerging functions of matricellular proteins.* Cellular and Molecular Life Sciences, 2011. **68**(19): p. 3133-3136.
11. Hall, Z.W. and J.R. Sanes, *Synaptic structure and development: the neuromuscular junction.* Cell, 1993. **72 Suppl**: p. 99-121.
12. Sanes, J.R., *The basement membrane/basal lamina of skeletal muscle.* J Biol Chem, 2003. **278**(15): p. 12601-4.
13. Yurchenco, P.D. and B.L. Patton, *Developmental and Pathogenic Mechanisms of Basement Membrane Assembly.* Curr Pharm Des, 2009. **15**(12): p. 1277-94.
14. Mouw, J., G. Ou, and V. Weaver, *Extracellular matrix assembly: a multiscale deconstruction.* Nature Reviews Molecular Cell Biology, 2014. **15**(12): p. 771-785.
15. Iozzo, R. and L. Schaefer, *Proteoglycan form and function: A comprehensive nomenclature of proteoglycans.* Matrix Biology, 2015. **42**: p. 11-55.
16. Jayadev, R. and D.R. Sherwood, *Basement membranes.* Curr Biol, 2017. **27**(6): p. R207-r211.
17. Walko, G., M.J. Castañón, and G. Wiche, *Molecular architecture and function of the hemidesmosome*, in *Cell Tissue Res*. 2015. p. 529-44.
18. Soltanoff, C.S., et al., *Signaling Networks that Control the Lineage Commitment and Differentiation of Bone Cells.* Crit Rev Eukaryot Gene Expr, 2009. **19**(1): p. 1-46.
19. Buckwalter, J.A. and H.J. Mankin, *Articular cartilage: tissue design and chondrocyte-matrix interactions.* Instr Course Lect, 1998. **47**: p. 477-86.
20. Ruiz-Meana, M., et al., *Propagation of cardiomyocyte hypercontracture by passage of Na(+) through gap junctions.* Circ Res, 1999. **85**(3): p. 280-7.
21. Harburger, D.S. and D.A. Calderwood, *Integrin signalling at a glance*, in *J Cell Sci*. 2009. p. 159-63.
22. Vogel, W.F., R. Abdulhussein, and C.E. Ford, *Sensing extracellular matrix: an update on discoidin domain receptor function.* Cell Signal, 2006. **18**(8): p. 1108-16.
23. Stepp, M.A., et al., *Syndecan-1 and Its Expanding List of Contacts.* Adv Wound Care (New Rochelle), 2015. **4**(4): p. 235-249.
24. Ponta, H., L. Sherman, and P.A. Herrlich, *CD44: from adhesion molecules to signalling regulators.* Nat Rev Mol Cell Biol, 2003. **4**(1): p. 33-45.

25. Jokinen, J., et al., *Integrin-mediated cell adhesion to type I collagen fibrils.* Journal of Biological Chemistry, 2004. **279**(30): p. 31956-31963.
26. Belkin, A. and M. Stepp, *Integrins as receptors for laminins.* Microscopy Research and Technique, 2000. **51**(3): p. 280-301.
27. Li, S., et al., *The role of laminin in embryonic cell polarization and tissue organization.* Developmental Cell, 2003. **4**(5): p. 613-624.
28. Hynes, R., *Cell adhesion: old and new questions (Reprinted from Trends in Biochemical Science, vol 12, Dec., 1999).* Trends in Genetics, 1999. **15**(12): p. M33-M37.
29. Danen, E., et al., *The fibronectin-binding integrins alpha 5 beta 1 and alpha v beta 3 differentially modulate RhoA-GTP loading, organization of cell matrix adhesions, and fibronectin fibrillogenesis.* Journal of Cell Biology, 2002. **159**(6): p. 1071-1086.
30. Kim, A., N. Lakshman, and W. Petroll, *Quantitative Assessment of Local Collagen Matrix Remodeling in 3-D Culture: The Role of Rho Kinase.* Exp Cell Res, 2006. **312**(18): p. 3683-92.
31. Provenzano, P.P., et al., *Contact guidance mediated three-dimensional cell migration is regulated by Rho/ROCK-dependent matrix reorganization.* Biophys J, 2008. **95**(11): p. 5374-84.
32. Sander, E.A., V.H. Barocas, and R.T. Tranquillo, *Initial fiber alignment pattern alters extracellular matrix synthesis in fibroblast-populated fibrin gel cruciforms and correlates with predicted tension.* Ann Biomed Eng, 2011. **39**(2): p. 714-29.
33. Wang, J.H., et al., *Cell orientation determines the alignment of cell-produced collagenous matrix.* J Biomech, 2003. **36**(1): p. 97-102.
34. Yang, N., et al., *Syndecan-1 in breast cancer stroma fibroblasts regulates extracellular matrix fiber organization and carcinoma cell motility.* Am J Pathol, 2011. **178**(1): p. 325-35.
35. Edmondson, R., et al., *Three-Dimensional Cell Culture Systems and Their Applications in Drug Discovery and Cell-Based Biosensors.* Assay and Drug Development Technologies, 2014. **12**(4): p. 207-218.
36. Knapp, D.M., E.F. Helou, and R.T. Tranquillo, *A fibrin or collagen gel assay for tissue cell chemotaxis: assessment of fibroblast chemotaxis to GRGDSP.* Exp Cell Res, 1999. **247**(2): p. 543-53.
37. Melvin, A., et al., *In Chemotaxing Fibroblasts, Both High-Fidelity and Weakly Biased Cell Movements Track the Localization of PI3K Signaling*, in *Biophys J*. 2011. p. 1893-901.
38. Rozario, T. and D. DeSimone, *The extracellular matrix in development and morphogenesis: A dynamic view.* Developmental Biology, 2010. **341**(1): p. 126-140.
39. Rifkin, D., *Latent transforming growth factor-beta (TGF-beta) binding proteins: Orchestrators of TGF-beta availability.* Journal of Biological Chemistry, 2005. **280**(9): p. 7409-7412.
40. ten Dijke, P. and C.S. Hill, *New insights into TGF-β–Smad signalling.* Trends in biochemical sciences, 2004. **29**(5): p. 265-273.
41. Zhu, H., et al., *A SMAD ubiquitin ligase targets the BMP pathway and affects embryonic pattern formation.* Nature, 1999. **400**(6745): p. 687-93.
42. Verrecchia, F., M.-L. Chu, and A. Mauviel, *Identification of novel TGF-β/Smad gene targets in dermal fibroblasts using a combined cDNA microarray/promoter transactivation approach.* Journal of Biological Chemistry, 2001. **276**(20): p. 17058-17062.
43. Yuan, W. and J. Varga, *Transforming growth factor-beta repression of matrix metalloproteinase-1 in dermal fibroblasts involves Smad3.* J Biol Chem, 2001. **276**(42): p. 38502-10.
44. Chen, G., C. Deng, and Y.P. Li, *TGF-β and BMP signaling in osteoblast differentiation and bone formation.* Int J Biol Sci, 2012. **8**(2): p. 272-88.
45. Akhurst, R.J. and A. Hata, *Targeting the TGFβ signalling pathway in disease.* Nature reviews Drug discovery, 2012. **11**(10): p. 790-811.
46. van den Bosch, M.H., et al., *Canonical Wnt signaling skews TGF-β signaling in chondrocytes towards signaling via ALK1 and Smad 1/5/8.* Cell Signal, 2014. **26**(5): p. 951-8.

47. Muñoz-Félix, J.M., et al., *Identification of bone morphogenetic protein 9 (BMP9) as a novel profibrotic factor in vitro.* Cell Signal, 2016. **28**(9): p. 1252-61.
48. Coricor, G. and R. Serra, *TGF-β regulates phosphorylation and stabilization of Sox9 protein in chondrocytes through p38 and Smad dependent mechanisms.* Sci Rep, 2016. **6**: p. 38616.
49. Mott, J. and Z. Werb, *Regulation of matrix biology by matrix metalloproteinases.* Current Opinion in Cell Biology, 2004. **16**(5): p. 558-564.
50. Nagase, H., R. Visse, and G. Murphy, *Structure and function of matrix metalloproteinases and TIMPs.* Cardiovasc Res, 2006. **69**(3): p. 562-73.
51. Lu, P., et al., *Extracellular Matrix Degradation and Remodeling in Development and Disease*, in *Cold Spring Harb Perspect Biol*. 2011.
52. Kuno, K., et al., *Molecular cloning of a gene encoding a new type of metalloproteinase-disintegrin family protein with thrombospondin motifs as an inflammation associated gene.* J Biol Chem, 1997. **272**(1): p. 556-62.
53. Tetsunaga, T., et al., *Regulation of mechanical stress-induced MMP-13 and ADAMTS-5 expression by RUNX-2 transcriptional factor in SW1353 chondrocyte-like cells.* Osteoarthritis Cartilage, 2011. **19**(2): p. 222-32.
54. Udeze, C.P., et al., *An in vitro investigation into the effects of 10 Hz cyclic loading on tenocyte metabolism.* Scand J Med Sci Sports, 2019. **29**(10): p. 1511-1520.
55. Nelson, C.M. and J.P. Gleghorn, *Sculpting organs: mechanical regulation of tissue development.* Annu Rev Biomed Eng, 2012. **14**: p. 129-54.
56. Docking, S., et al., *Relationship between compressive loading and ECM changes in tendons.* Muscles Ligaments Tendons J, 2013. **3**(1): p. 7-11.
57. Cawston, T.E. and D.A. Young, *Proteinases involved in matrix turnover during cartilage and bone breakdown.* Cell Tissue Res, 2010. **339**(1): p. 221-35.
58. Goldfinger, L.E., M.S. Stack, and J.C. Jones, *Processing of laminin-5 and its functional consequences: role of plasmin and tissue-type plasminogen activator.* J Cell Biol, 1998. **141**(1): p. 255-65.
59. Goldfinger, L.E., et al., *Spatial regulation and activity modulation of plasmin by high affinity binding to the G domain of the alpha 3 subunit of laminin-5.* J Biol Chem, 2000. **275**(45): p. 34887-93.
60. Horowitz, J.C., et al., *Plasminogen Activation–Induced Pericellular Fibronectin Proteolysis Promotes Fibroblast Apoptosis*, in *Am J Respir Cell Mol Biol*. 2008. p. 78-87.
61. Lyons, R.M., et al., *Mechanism of activation of latent recombinant transforming growth factor beta 1 by plasmin.* J Cell Biol, 1990. **110**(4): p. 1361-7.
62. Berg, R.A., et al., *Lysosomal function in the degradation of defective collagen in cultured lung fibroblasts.* Biochemistry, 1984. **23**(10): p. 2134-8.
63. Panetti, T.S. and P.J. McKeown-Longo, *The alpha v beta 5 integrin receptor regulates receptor-mediated endocytosis of vitronectin.* J Biol Chem, 1993. **268**(16): p. 11492-5.
64. Sottile, J. and J. Chandler, *Fibronectin Matrix Turnover Occurs through a Caveolin-1–dependent Process.* Mol Biol Cell, 2005. **16**(2): p. 757-68.
65. Egeberg, M., et al., *Internalization and stepwise degradation of heparan sulfate proteoglycans in rat hepatocytes.* Biochim Biophys Acta, 2001. **1541**(3): p. 135-49.
66. Yamazaki, Y., et al., *Development of collagen fibres and lysyl oxidase expression in the presumptive dermis of chick limb bud.* Anat Histol Embryol, 2012. **41**(1): p. 68-74.
67. Lucero, H. and H. Kagan, *Lysyl oxidase: an oxidative enzyme and effector of cell function.* Cellular and Molecular Life Sciences, 2006. **63**(19-20): p. 2304-2316.
68. Hynes, R., *The Extracellular Matrix: Not Just Pretty Fibrils.* Science, 2009. **326**(5957): p. 1216-1219.
69. Theocharis, A.D., et al., *Proteoglycans in health and disease: novel roles for proteoglycans in malignancy and their pharmacological targeting.* FEBS J, 2010. **277**(19): p. 3904-23.

70. Syrokou, A., et al., *Effects of glycosaminoglycans on proliferation of epithelial and fibroblast human malignant mesothelioma cells: a structure-function relationship.* Cell Prolif, 1999. **32**(2-3): p. 85-99.
71. Roughley, P.J. and J.S. Mort, *The role of aggrecan in normal and osteoarthritic cartilage.* J Exp Orthop, 2014. **1**(1): p. 8.
72. Yamagata, M., et al., *Chondroitin sulfate proteoglycan (PG-M-like proteoglycan) is involved in the binding of hyaluronic acid to cellular fibronectin.* J Biol Chem, 1986. **261**(29): p. 13526-35.
73. Kim, S., et al., *Carcinoma Produced Factors Activate Myeloid Cells via TLR2 to Stimulate Metastasis.* Nature, 2009. **457**(7225): p. 102-6.
74. Wight, T.N., et al., *Versican-A Critical Extracellular Matrix Regulator of Immunity and Inflammation.* Front Immunol, 2020. **11**: p. 512.
75. Carol, A., et al., *Poly I: C induces mononuclear leukocyte-adhesive hyaluronan structures on colon smooth muscle cells: IαI and versican facilitate adhesion*, in *Hyaluronan*. 2002, Elsevier. p. 381-388.
76. Chang, M.Y., et al., *Versican is produced by Trif-and type I interferon-dependent signaling in macrophages and contributes to fine control of innate immunity in lungs.* American Journal of Physiology-Lung Cellular and Molecular Physiology, 2017. **313**(6): p. L1069-L1086.
77. Day, J.M., et al., *Alternative splicing in the aggrecan G3 domain influences binding interactions with tenascin-C and other extracellular matrix proteins.* J Biol Chem, 2004. **279**(13): p. 12511-8.
78. Aspberg, A., et al., *The C-type lectin domains of lecticans, a family of aggregating chondroitin sulfate proteoglycans, bind tenascin-R by protein-protein interactions independent of carbohydrate moiety.* Proc Natl Acad Sci U S A, 1997. **94**(19): p. 10116-21.
79. Hu, B., et al., *The proteoglycan brevican binds to fibronectin after proteolytic cleavage and promotes glioma cell motility.* J Biol Chem, 2008. **283**(36): p. 24848-59.
80. Schaefer, L. and R. Schaefer, *Proteoglycans: from structural compounds to signaling molecules.* Cell and Tissue Research, 2010. **339**(1): p. 237-246.
81. Kalamajski, S. and A. Oldberg, *The role of small leucine-rich proteoglycans in collagen fibrillogenesis.* Matrix Biol, 2010. **29**(4): p. 248-53.
82. Krusius, T. and E. Ruoslahti, *Primary structure of an extracellular matrix proteoglycan core protein deduced from cloned cDNA.* Proc Natl Acad Sci U S A, 1986. **83**(20): p. 7683-7.
83. Danielson, K.G., et al., *Targeted Disruption of Decorin Leads to Abnormal Collagen Fibril Morphology and Skin Fragility.* J Cell Biol, 1997. **136**(3): p. 729-43.
84. Zhang, G., et al., *Decorin regulates assembly of collagen fibrils and acquisition of biomechanical properties during tendon development.* J Cell Biochem, 2006. **98**(6): p. 1436-49.
85. Yamaguchi, Y., D.M. Mann, and E. Ruoslahti, *Negative regulation of transforming growth factor-beta by the proteoglycan decorin.* Nature, 1990. **346**(6281): p. 281-4.
86. Nili, N., et al., *Decorin inhibition of PDGF-stimulated vascular smooth muscle cell function: potential mechanism for inhibition of intimal hyperplasia after balloon angioplasty.* Am J Pathol, 2003. **163**(3): p. 869-78.
87. Csordás, G., et al., *Sustained Down-regulation of the Epidermal Growth Factor Receptor by Decorin.* 2000.
88. Khan, G.A., et al., *Decorin is a novel VEGFR-2-binding antagonist for the human extravillous trophoblast.* Mol Endocrinol, 2011. **25**(8): p. 1431-43.
89. Neill, T., et al., *Decorin antagonizes the angiogenic network: concurrent inhibition of Met, hypoxia inducible factor 1alpha, vascular endothelial growth factor A, and induction of thrombospondin-1 and TIMP3.* J Biol Chem, 2012. **287**(8): p. 5492-506.
90. Schaefer, L., et al., *The matrix component biglycan is proinflammatory and signals through Toll-like receptors 4 and 2 in macrophages.* J Clin Invest, 2005. **115**(8): p. 2223-33.

91. Kolb, M., et al., *Proteoglycans decorin and biglycan differentially modulate TGF-beta-mediated fibrotic responses in the lung.* Am J Physiol Lung Cell Mol Physiol, 2001. **280**(6): p. L1327-34.
92. Bianco, P., et al., *Expression and localization of the two small proteoglycans biglycan and decorin in developing human skeletal and non-skeletal tissues.* J Histochem Cytochem, 1990. **38**(11): p. 1549-63.
93. Hynes, R.O., *The extracellular matrix: not just pretty fibrils.* Science, 2009. **326**(5957): p. 1216-9.
94. Aviezer, D., et al., *Perlecan, basal lamina proteoglycan, promotes basic fibroblast growth factor-receptor binding, mitogenesis, and angiogenesis.* Cell, 1994. **79**(6): p. 1005-13.
95. Li, S., et al., *Integrin and dystroglycan compensate each other to mediate laminin-dependent basement membrane assembly and epiblast polarization.* Matrix Biol, 2017. **57-58**: p. 272-84.
96. Aumailley, M., et al., *Binding of nidogen and the laminin-nidogen complex to basement membrane collagen type IV.* Eur J Biochem, 1989. **184**(1): p. 241-8.
97. Kwon, M.J., et al., *Syndecans play dual roles as cell adhesion receptors and docking receptors.* FEBS Lett, 2012. **586**(16): p. 2207-11.
98. Carey, D.J., *Syndecans: multifunctional cell-surface co-receptors.* Biochem J, 1997. **327 (Pt 1)**: p. 1-16.
99. Bass, M.D., et al., *Syndecan-4–dependent Rac1 regulation determines directional migration in response to the extracellular matrix.* Journal of Cell Biology, 2007. **177**(3): p. 527-538.
100. Midwood, K.S., et al., *Coregulation of Fibronectin Signaling and Matrix Contraction by Tenascin-C and Syndecan-4.* Mol Biol Cell, 2004. **15**(12): p. 5670-7.
101. Mayor, S. and H. Riezman, *Sorting GPI-anchored proteins.* Nat Rev Mol Cell Biol, 2004. **5**(2): p. 110-20.
102. Grisaru, S., et al., *Glypican-3 modulates BMP-and FGF-mediated effects during renal branching morphogenesis.* Developmental biology, 2001. **231**(1): p. 31-46.
103. Capurro, M.I., et al., *Glypican-3 promotes the growth of hepatocellular carcinoma by stimulating canonical Wnt signaling.* Cancer Res, 2005. **65**(14): p. 6245-54.
104. Ricard-Blum, S., *The Collagen Family*, in *Cold Spring Harb Perspect Biol*. 2011.
105. Kadler, K., A. Hill, and E. Canty-Laird, *Collagen fibrillogenesis: fibronectin, integrins, and minor collagens as organizers and nucleators.* Current Opinion in Cell Biology, 2008. **20**(5): p. 495-501.
106. Brazel, D., et al., *Completion of the amino acid sequence of the alpha 1 chain of human basement membrane collagen (type IV) reveals 21 non-triplet interruptions located within the collagenous domain.* Eur J Biochem, 1987. **168**(3): p. 529-36.
107. Khoshnoodi, J., et al., *Molecular recognition in the assembly of collagens: terminal noncollagenous domains are key recognition modules in the formation of triple helical protomers.* J Biol Chem, 2006. **281**(50): p. 38117-21.
108. Wu, J.J., et al., *Differences in Chain Usage and Cross-linking Specificities of Cartilage Type V/XI Collagen Isoforms with Age and Tissue*.* J Biol Chem, 2009. **284**(9): p. 5539-45.
109. Makareeva, E., et al., *Carcinomas contain an MMP-resistant isoform of type I collagen exerting selective support to invasion.* Cancer Res, 2010. **70**(11): p. 4366-74.
110. Jimenez, S.A., et al., *Identification of collagen alpha1(I) trimer in embryonic chick tendons and calvaria.* Biochem Biophys Res Commun, 1977. **78**(4): p. 1354-61.
111. Pace, J.M., et al., *Defective C-propeptides of the proalpha2(I) chain of type I procollagen impede molecular assembly and result in osteogenesis imperfecta.* J Biol Chem, 2008. **283**(23): p. 16061-7.
112. Malfait, F., et al., *Total absence of the alpha2(I) chain of collagen type I causes a rare form of Ehlers-Danlos syndrome with hypermobility and propensity to cardiac valvular problems*, in *J Med Genet*. 2006: England. p. e36.

113. Han, S., et al., *Molecular Mechanism of Type I Collagen Homotrimer Resistance to Mammalian Collagenases*.* J Biol Chem, 2010. **285**(29): p. 22276-81.
114. Bateman, J., R. Boot-Handford, and S. Lamande, *Genetic diseases of connective tissues: cellular and extracellular effects of ECM mutations.* Nature Reviews Genetics, 2009. **10**(3): p. 173-183.
115. Marini, J.C., et al., *Consortium for osteogenesis imperfecta mutations in the helical domain of type I collagen: regions rich in lethal mutations align with collagen binding sites for integrins and proteoglycans.* Hum Mutat, 2007. **28**(3): p. 209-21.
116. Shoulders, M.D. and R.T. Raines, *Collagen structure and stability.* Annu Rev Biochem, 2009. **78**: p. 929-58.
117. Epstein, E.H., Jr. and N.H. Munderloh, *Human skin collagen. Presence of type I and type III at all levels of the dermis.* J Biol Chem, 1978. **253**(5): p. 1336-7.
118. Holmes, D.F. and K.E. Kadler, *The 10+4 microfibril structure of thin cartilage fibrils.* Proc Natl Acad Sci U S A, 2006. **103**(46): p. 17249-54.
119. Orgel, J.P., et al., *Microfibrillar structure of type I collagen in situ.* Proc Natl Acad Sci U S A, 2006. **103**(24): p. 9001-5.
120. Kadler, K.E., Y. Hojima, and D.J. Prockop, *Assembly of collagen fibrils de novo by cleavage of the type I pC-collagen with procollagen C-proteinase. Assay of critical concentration demonstrates that collagen self-assembly is a classical example of an entropy-driven process.* J Biol Chem, 1987. **262**(32): p. 15696-701.
121. D'Aniello, C., et al., *Vitamin C in Stem Cell Biology: Impact on Extracellular Matrix Homeostasis and Epigenetics.* Stem Cells Int, 2017. **2017**: p. 8936156.
122. Eyre, D.R., M.A. Weis, and J.J. Wu, *Advances in collagen cross-link analysis.* Methods, 2008. **45**(1): p. 65-74.
123. Halász, K., et al., *COMP acts as a catalyst in collagen fibrillogenesis.* J Biol Chem, 2007. **282**(43): p. 31166-73.
124. Reed, C.C. and R.V. Iozzo, *The role of decorin in collagen fibrillogenesis and skin homeostasis.* Glycoconj J, 2002. **19**(4-5): p. 249-55.
125. Bella, J., et al., *Conformational effects of Gly-X-Gly interruptions in the collagen triple helix.* J Mol Biol, 2006. **362**(2): p. 298-311.
126. Young, B.B., M.K. Gordon, and D.E. Birk, *Expression of type XIV collagen in developing chicken tendons: association with assembly and growth of collagen fibrils.* Dev Dyn, 2000. **217**(4): p. 430-9.
127. Zhang, G., B.B. Young, and D.E. Birk, *Differential expression of type XII collagen in developing chicken metatarsal tendons.* J Anat, 2003. **202**(5): p. 411-20.
128. Sakai, L.Y., et al., *Type VII collagen is a major structural component of anchoring fibrils.* J Cell Biol, 1986. **103**(4): p. 1577-86.
129. Pöschl, E., et al., *Collagen IV is essential for basement membrane stability but dispensable for initiation of its assembly during early development.* Development, 2004. **131**(7): p. 1619-28.
130. Sutmuller, M., J.A. Bruijn, and E. de Heer, *Collagen types VIII and X, two non-fibrillar, short-chain collagens. Structure homologies, functions and involvement in pathology.* Histol Histopathol, 1997. **12**(2): p. 557-66.
131. Kwan, K.M., et al., *Abnormal compartmentalization of cartilage matrix components in mice lacking collagen X: implications for function.* J Cell Biol, 1997. **136**(2): p. 459-71.
132. Dennis, J., et al., *Collagen XIII Induced in Vascular Endothelium Mediates α1β1 Integrin-Dependent Transmigration of Monocytes in Renal Fibrosis.* Am J Pathol, 2010. **177**(5): p. 2527-40.
133. Franzke, C.W., et al., *Shedding of collagen XVII/BP180: structural motifs influence cleavage from cell surface.* J Biol Chem, 2004. **279**(23): p. 24521-9.
134. Nishie, W., et al., *Dynamic Interactions of Epidermal Collagen XVII with the Extracellular Matrix: Laminin 332 as a Major Binding Partner.* Am J Pathol, 2011. **179**(2): p. 829-37.

135. Baldock, C., et al., *The supramolecular organization of collagen VI microfibrils.* J Mol Biol, 2003. **330**(2): p. 297-307.
136. Kuo, H.J., et al., *Type VI collagen anchors endothelial basement membranes by interacting with type IV collagen.* J Biol Chem, 1997. **272**(42): p. 26522-9.
137. Bidanset, D.J., et al., *Binding of the proteoglycan decorin to collagen type VI.* J Biol Chem, 1992. **267**(8): p. 5250-6.
138. Wiberg, C., et al., *Biglycan and decorin bind close to the n-terminal region of the collagen VI triple helix.* J Biol Chem, 2001. **276**(22): p. 18947-52.
139. Tillet, E., et al., *Recombinant expression and structural and binding properties of alpha 1(VI) and alpha 2(VI) chains of human collagen type VI.* Eur J Biochem, 1994. **221**(1): p. 177-85.
140. Aumailley, M., et al., *Cell attachment properties of collagen type VI and Arg-Gly-Asp dependent binding to its alpha 2(VI) and alpha 3(VI) chains.* Exp Cell Res, 1989. **181**(2): p. 463-74.
141. Schnoor, M., et al., *Production of type VI collagen by human macrophages: a new dimension in macrophage functional heterogeneity.* J Immunol, 2008. **180**(8): p. 5707-19.
142. Wagenseil, J.E. and R.P. Mecham, *New insights into elastic fiber assembly.* Birth Defects Res C Embryo Today, 2007. **81**(4): p. 229-40.
143. Hayward, C. and D.J. Brock, *Fibrillin-1 mutations in Marfan syndrome and other type-1 fibrillinopathies.* Hum Mutat, 1997. **10**(6): p. 415-23.
144. Putnam, E.A., et al., *Fibrillin-2 (FBN2) mutations result in the Marfan-like disorder, congenital contractural arachnodactyly.* Nat Genet, 1995. **11**(4): p. 456-8.
145. Yanagisawa, H. and E.C. Davis, *Unraveling the mechanism of elastic fiber assembly: The roles of short fibulins.* Int J Biochem Cell Biol, 2010. **42**(7): p. 1084-93.
146. Craft, C.S., T.J. Broekelmann, and R.P. Mecham, *Microfibril-associated glycoproteins MAGP-1 and MAGP-2 in Disease.* Matrix Biol, 2018. **71-72**: p. 100-11.
147. Timpl, R., et al., *Fibulins: a versatile family of extracellular matrix proteins.* Nat Rev Mol Cell Biol, 2003. **4**(6): p. 479-89.
148. Djokic, J., et al., *Fibulin-3, -4, and -5 are highly susceptible to proteolysis, interact with cells and heparin, and form multimers.* J Biol Chem, 2013. **288**(31): p. 22821-35.
149. Sabatier, L., et al., *Fibrillin Assembly Requires Fibronectin.* Molecular Biology of the Cell, 2009. **20**(3): p. 846-858.
150. George, E.L., et al., *Defects in mesoderm, neural tube and vascular development in mouse embryos lacking fibronectin.* Development, 1993. **119**(4): p. 1079-91.
151. Hynes, R.O. and K.M. Yamada, *Fibronectins: multifunctional modular glycoproteins.* J Cell Biol, 1982. **95**(2 Pt 1): p. 369-77.
152. Kornblihtt, A.R., K. Vibe-Pedersen, and F.E. Baralle, *Human fibronectin: cell specific alternative mRNA splicing generates polypeptide chains differing in the number of internal repeats.* Nucleic Acids Res, 1984. **12**(14): p. 5853-68.
153. Main, A.L., et al., *The three-dimensional structure of the tenth type III module of fibronectin: an insight into RGD-mediated interactions.* Cell, 1992. **71**(4): p. 671-8.
154. Krammer, A., et al., *Forced unfolding of the fibronectin type III module reveals a tensile molecular recognition switch.* Proc Natl Acad Sci U S A, 1999. **96**(4): p. 1351-6.
155. Carnemolla, B., et al., *The inclusion of the type III repeat ED-B in the fibronectin molecule generates conformational modifications that unmask a cryptic sequence.* J Biol Chem, 1992. **267**(34): p. 24689-92.
156. Moretti, F.A., et al., *A Major Fraction of Fibronectin Present in the Extracellular Matrix of Tissues Is Plasma-derived.* 2007.
157. Aguirre, K.M., R.J. McCormick, and J.E. Schwarzbauer, *Fibronectin self-association is mediated by complementary sites within the amino-terminal one-third of the molecule.* J Biol Chem, 1994. **269**(45): p. 27863-8.

158. Cho, C., et al., *Cryptic activity within the Type III1 domain of fibronectin regulates tissue inflammation and angiogenesis.* Current topics in peptide & protein research, 2015. **16**: p. 37.
159. Dallas, S., et al., *Fibronectin regulates latent transforming growth factor-beta (TGF beta) by controlling matrix assembly of latent TGF beta-binding protein-1.* Journal of Biological Chemistry, 2005. **280**(19): p. 18871-18880.
160. Chung, C. and H. Erickson, *Glycosaminoglycans modulate fibronectin matrix assembly and are essential for matrix incorporation of tenascin-C.* Journal of Cell Science, 1997. **110**: p. 1413-1419.
161. Singh, P., et al., *Assembly of Fibronectin Extracellular Matrix.* Annual Review of Cell and Developmental Biology, Vol 26, 2010. **26**: p. 397-419.
162. Fogelgren, B., et al., *Cellular fibronectin binds to lysyl oxidase with high affinity and is critical for its proteolytic activation.* J Biol Chem, 2005. **280**(26): p. 24690-7.
163. Huang, G., et al., *Fibronectin Binds and Enhances the Activity of Bone Morphogenetic Protein 1**, in *J Biol Chem*. 2009. p. 25879-88.
164. Zollinger, A.J. and M.L. Smith, *Fibronectin, the extracellular glue.* Matrix Biol, 2017. **60-61**: p. 27-37.
165. Lin, F., et al., *Fibronectin growth factor-binding domains are required for fibroblast survival.* J Invest Dermatol, 2011. **131**(1): p. 84-98.
166. Schwartz, M., *Integrins and Extracellular Matrix in Mechanotransduction.* Cold Spring Harbor Perspectives in Biology, 2010. **2**(12).
167. Champliaud, M.F., et al., *Posttranslational modifications and beta/gamma chain associations of human laminin alpha1 and laminin alpha5 chains: purification of laminin-3 from placenta.* Exp Cell Res, 2000. **259**(2): p. 326-35.
168. Harrison, D., et al., *Crystal structure and cell surface anchorage sites of laminin alpha1LG4-5.* J Biol Chem, 2007. **282**(15): p. 11573-81.
169. Yurchenco, P.D., et al., *Laminin polymerization in vitro. Evidence for a two-step assembly with domain specificity.* J Biol Chem, 1985. **260**(12): p. 7636-44.
170. Stetefeld, J., et al., *Crystal structure of three consecutive laminin-type epidermal growth factor-like (LE) modules of laminin gamma1 chain harboring the nidogen binding site.* J Mol Biol, 1996. **257**(3): p. 644-57.
171. Mascarenhas, J.B., et al., *Mapping of the laminin-binding site of the N-terminal agrin domain (NtA).* EMBO J, 2003. **22**(3): p. 529-36.
172. Sung, U., J.J. O'Rear, and P.D. Yurchenco, *Cell and heparin binding in the distal long arm of laminin: identification of active and cryptic sites with recombinant and hybrid glycoprotein.* J Cell Biol, 1993. **123**(5): p. 1255-68.
173. Aumailley, M., et al., *A simplified laminin nomenclature.* Matrix Biol, 2005. **24**(5): p. 326-32.
174. Sage, E.H. and P. Bornstein, *Extracellular Proteins That Modulate Cell-Matrix Interactions-Sparc, Tenascin, And Thrombospondin.* Journal of Biological Chemistry, 1991. **266**(23): p. 14831-14834.
175. Bornstein, P. and E. Sage, *Matricellular proteins: extracellular modulators of cell function.* Current Opinion in Cell Biology, 2002. **14**(5): p. 608-616.
176. Bornstein, P., *Matricellular proteins: an overview.* Journal of Cell Communication and Signaling, 2009. **3**(3-4): p. 163-165.
177. Midwood, K., L. Williams, and J. Schwarzbauer, *Tissue repair and the dynamics of the extracellular matrix.* International Journal of Biochemistry & Cell Biology, 2004. **36**(6): p. 1031-1037.
178. Bourdon, M., et al., *Human Glioma-Mesenchymal Extracellular-Matrix Antigen Defined By Monoclonal-Antibody.* Cancer Research, 1983. **43**(6): p. 2796-2805.
179. Erickson, H. and J. Inglesias, *A 6-Armed Oligomer Isolated*

From Cell-Surface Fibronectin Preparations. Nature. Nature, 1984. **311**(5983): p. 267-269.

180. Grumet, M., et al., *Cytotactin, An*

Extracellular-Matrix Protein Of Neural And Non-Neural Tissues That Mediates

Glia Neuron Interaction. Proceedings of the National Academy of Sciences of the United States of America, 1985. **82**(23): p. 8075-8079.

181. Rettig, W., T. Triche, and P. Garinchesa, *Stimulation Of Human*

Neuronectin Secretion By Brain-Derived Growth-Factors. Brain Research, 1989. **487**(1): p. 171-177.

182. Chiquet-Ehrismann , R., et al., *Tenascin - An Extracellular-Matrix Protein Involved In Tissue Interactions During Fetal Development And Oncogenesis.* Cell, 1986. **47**(1): p. 131-139.

183. Midwood, K. and G. Orend, *The role of tenascin-C in tissue injury and tumorigenesis.* Journal of Cell Communication and Signaling, 2009. **3**(3-4): p. 287-310.

184. Midwood, K., et al., *Advances in tenascin-C biology.* Cellular and Molecular Life Sciences, 2011. **68**(19): p. 3175-3199.

185. Orend, G. and R. Chiquet-Ehrismann, *Tenascin-C induced signaling in cancer.* Cancer Letters, 2006. **244**(2): p. 143-163.

186. Martin, J.A., et al., *The role of tenascin-C in adaptation of tendons to compressive loading.* Biorheology, 2003. **40**(1-3): p. 321-9.

187. Midwood, K.S., et al., *Tenascin-C at a glance.* J Cell Sci, 2016. **129**(23): p. 4321-4327.

188. Piccinini, A. and K. Midwood, *Endogenous Control of Immunity against Infection: Tenascin-C Regulates TLR4-Mediated Inflammation via MicroRNA-155.* Cell Reports, 2012. **2**(4): p. 914-926.

189. Hsia, H.C. and J.E. Schwarzbauer, *Meet the tenascins: multifunctional and mysterious.* J Biol Chem, 2005. **280**(29): p. 26641-4.

190. Rathjen, F.G., J. Wolff, and R. Chiquet-Ehrismann, *Restrictin: a chick neural extracellular matrix protein involved in cell attachment co-purifies with the cell recognition molecule F11.* Development, 1991. **113**(1): p. 151-164.

191. Valcourt, U., et al., *Tenascin-X: beyond the architectural function.* Cell adhesion & migration, 2015. **9**(1-2): p. 154-165.

192. Scherberich, A., et al., *Murine tenascin-W: a novel mammalian tenascin expressed in kidney and at sites of bone and smooth muscle development.* J Cell Sci, 2004. **117**(Pt 4): p. 571-81.

193. Chiovaro, F., R. Chiquet-Ehrismann, and M. Chiquet, *Transcriptional regulation of tenascin genes.* Cell Adhesion & Migration, 2015. **9**(1-2): p. 34-47.

194. Moritz, S., et al., *An induction gene trap screen in neural stem cells reveals an instructive function of the niche and identifies the splicing regulator sam68 as a tenascin-C-regulated target gene.* Stem Cells, 2008. **26**(9): p. 2321-31.

195. Latijnhouwers, M.A., et al., *Expression of tenascin-C splice variants by human skin cells.* Arch Dermatol Res, 2000. **292**(9): p. 446-54.

196. Jones, F., et al., *A Detailed Structural Model Of*

Cytotactin - Protein Homologies, Alternative Rna Splicing, And Binding Regions. Proceedings of the National Academy of Sciences of the United States of America, 1989. **86**(6): p. 1905-1909.

197. Conway, J. and D. Parry, *3-Stranded Alpha-Fibrous Proteins-The Heptad Repeat And Its Implications For Structure.* International Journal of Biological Macromolecules, 1991. **13**(1): p. 14-16.

198. Kammerer, R., et al., *Tenascin-C hexabrachion assembly is a sequential two-step process initiated by coiled-coil alpha-helices.* Journal of Biological Chemistry, 1998. **273**(17): p. 10602-10608.

199. Redick, S. and J. Schwarzbauer, *Rapid Intracellular Assembly Of*

Tenascin Hexabrachions Suggests A Novel Cotranslational Process. Journal of Cell Science, 1995. **108**: p. 1761-1769.

200. Giblin, S.P. and K.S. Midwood, *Tenascin-C: Form versus function.* Cell Adhesion & Migration, 2015. **9**(1-2): p. 48-82.
201. Jones, P. and F. Jones, *Tenascin-C in development and disease: gene regulation and cell function.* Matrix Biology, 2000. **19**(7): p. 581-596.
202. Erickson, H. and J. Inglesias, *A 6-Armed Oligomer Isolated From Cell-Surface Fibronectin Preparations.* Nature, 1984. **311**(5983): p. 267-269.
203. Van Obberghen-Schilling, E., et al., *Fibronectin and tenascin-C: accomplices in vascular morphogenesis during development and tumor growth.* Int J Dev Biol, 2011. **55**(4-5): p. 511-25.
204. Jones, F.S., et al., *Activation of the cytotactin promoter by the homeobox-containing gene Evx-1.* Proc Natl Acad Sci U S A, 1992. **89**(6): p. 2091-5.
205. Copertino, D.W., G.M. Edelman, and F.S. Jones, *Multiple promoter elements differentially regulate the expression of the mouse tenascin gene.* Proc Natl Acad Sci U S A, 1997. **94**(5): p. 1846-51.
206. Ihida-Stansbury, K., et al., *Paired-related homeobox gene Prx1 is required for pulmonary vascular development.* Circ Res, 2004. **94**(11): p. 1507-14.
207. Sarasa-Renedo, A. and M. Chiquet, *Mechanical signals regulating extracellular matrix gene expression in fibroblasts.* Scandinavian Journal of Medicine & Science in Sports, 2005. **15**(4): p. 223-230.
208. Fluck, M., V. Tunc-Civelek, and M. Chiquet, *Rapid and reciprocal regulation of tenascin-C and tenascin-Y expression by loading of skeletal muscle.* Journal of Cell Science, 2000. **113**(20): p. 3583-3591.
209. Chiquet, M., A. Sarasa-Renedo, and V. Tunc-Civelek, *Induction of tenascin-C by cyclic tensile strain versus growth factors: distinct contributions by Rho/ROCK and MAPK signaling pathways.* Biochimica Et Biophysica Acta-Molecular Cell Research, 2004. **1693**(3): p. 193-204.
210. Sarasa-Renedo, A., V. Tunc-Civelek, and M. Chiquet, *Role of RhoA/ROCK-dependent actin contractility in the induction of tenascin-C by cyclic tensile strain.* Experimental Cell Research, 2006. **312**(8): p. 1361-1370.
211. Chiquet, M., V. Tunc-Civelek, and A. Sarasa-Renedo, *Gene regulation by mechanotransduction in fibroblasts.* Applied Physiology Nutrition and Metabolism-Physiologie Appliquee Nutrition Et Metabolisme, 2007. **32**(5): p. 967-973.
212. Lutz, R., T. Sakai, and M. Chiquet, *Pericellular fibronectin is required for RhoA-dependent responses to cyclic strain in fibroblasts.* J Cell Sci, 2010. **123**(Pt 9): p. 1511-21.
213. Asparuhova, M.B., et al., *The transcriptional regulator megakaryoblastic leukemia-1 mediates serum response factor-independent activation of tenascin-C transcription by mechanical stress.* Faseb j, 2011. **25**(10): p. 3477-88.
214. Mikic, B., et al., *Mechanical modulation of tenascin-C and collagen-XII expression during avian synovial joint formation.* Journal of Orthopaedic Research, 2000. **18**(3): p. 406-415.
215. Maqbool, A., et al., *Interleukin-1 has opposing effects on connective tissue growth factor and tenascin-C expression in human cardiac fibroblasts.* Matrix Biology, 2013. **32**(3-4): p. 208-214.
216. Chevillard, G., et al., *Identification of interleukin-1beta regulated genes in uterine smooth muscle cells.* Reproduction, 2007. **134**(6): p. 811-22.
217. Jinnin, M., et al., *Upregulation of tenascin-C expression by IL-13 in human dermal fibroblasts via the phosphoinositide 3-kinase/Akt and the protein kinase C signaling pathways.* Journal of Investigative Dermatology, 2006. **126**(3): p. 551-560.
218. Ogawa, K., et al., *Tenascin-C is upregulated in the skin lesions of patients with atopic dermatitis.* J Dermatol Sci, 2005. **40**(1): p. 35-41.
219. Pearson, C., et al., *Tenascin - cDNA Cloning And Induction By TGF-Beta.* Embo Journal, 1988. **7**(10): p. 2977-2981.

220. Jinnin, M., et al., *Platelet derived growth factor induced tenascin-C transcription is phosphoinositide 3-kinase/Akt-dependent and mediated by Ets family transcription factors.* Journal of Cellular Physiology, 2006. **206**(3): p. 718-727.
221. Jinnin, M., et al., *Tenascin-C upregulation by transforming growth factor-beta in human dermal fibroblasts involves Smad3, Sp1, and Ets1.* Oncogene, 2004. **23**(9): p. 1656-67.
222. Ekblom, M., et al., *Downregulation of tenascin expression by glucocorticoids in bone marrow stromal cells and in fibroblasts.* J Cell Biol, 1993. **123**(4): p. 1037-45.
223. Ghatnekar, A. and M. Trojanowska, *GATA-6 is a novel transcriptional repressor of the human Tenascin-C gene expression in fibroblasts.* Biochim Biophys Acta, 2008. **1779**(3): p. 145-51.
224. Goh, F.G., et al., *Transcriptional Regulation of the Endogenous Danger Signal Tenascin-C: A Novel Autocrine Loop in Inflammation.* Journal of Immunology, 2010. **184**(5): p. 2655-2662.
225. Giblin, S.P., A. Schwenzer, and K.S. Midwood, *Alternative splicing controls cell lineage-specific responses to endogenous innate immune triggers within the extracellular matrix.* Matrix Biol, 2020. **93**: p. 95-114.
226. Tavazoie, S.F., et al., *Endogenous human microRNAs that suppress breast cancer metastasis.* Nature, 2008. **451**(7175): p. 147-52.
227. Siri, A., et al., *Different susceptibility of small and large human tenascin-C isoforms to degradation by matrix metalloproteinases.* J Biol Chem, 1995. **270**(15): p. 8650-4.
228. Gueders, M.M., et al., *Matrix metalloproteinase-19 deficiency promotes tenascin-C accumulation and allergen-induced airway inflammation.* Am J Respir Cell Mol Biol, 2010. **43**(3): p. 286-95.
229. Nakahara, H., et al., *Deficiency of tenascin C attenuates allergen-induced bronchial asthma in the mouse.* Eur J Immunol, 2006. **36**(12): p. 3334-45.
230. Kasbaoui, L., et al., *Differences in glycosylation state of fibronectin from two rat colon carcinoma cell lines in relation to tumoral progressiveness.* Cancer Res, 1989. **49**(19): p. 5317-22.
231. Saga, Y., et al., *Mice develop normally without tenascin.* Genes Dev, 1992. **6**(10): p. 1821-31.
232. Forsberg, E., et al., *Skin wounds and severed nerves heal normally in mice lacking tenascin-C.* Proc Natl Acad Sci U S A, 1996. **93**(13): p. 6594-9.
233. Tucker, R.P. and M. Degen, *The Expression and Possible Functions of Tenascin-W During Development and Disease.* Front Cell Dev Biol, 2019. **7**.
234. Nakao, N., et al., *Tenascin-C promotes healing of Habu-snake venom-induced glomerulonephritis: studies in knockout congenic mice and in culture.* Am J Pathol, 1998. **152**(5): p. 1237-45.
235. Midwood, K., et al., *Tenascin-C is an endogenous activator of Toll-like receptor 4 that is essential for maintaining inflammation in arthritic joint disease.* Nat Med, 2009. **15**(7): p. 774-80.
236. Tanaka, K., et al., *Tenascin-C regulates angiogenesis in tumor through the regulation of vascular endothelial growth factor expression.* Int J Cancer, 2004. **108**(1): p. 31-40.
237. Jallo, G.I., et al., *Tenascin-C expression in the cyst wall and fluid of human brain tumors correlates with angiogenesis.* Neurosurgery, 1997. **41**(5): p. 1052-9.
238. Castellon, R., et al., *Effects of tenascin-C on normal and diabetic retinal endothelial cells in culture.* Invest Ophthalmol Vis Sci, 2002. **43**(8): p. 2758-66.
239. Orend, G. and R. Chiquet-Ehrismann, *Adhesion modulation by antiadhesive molecules of the extracellular matrix.* Experimental Cell Research, 2000. **261**(1): p. 104-110.
240. Wehrle, B. and M. Chiquet, *Tenascin Is Accumulated Along Developing Peripheral-Nerves And Allows Neurite Outgrowth Invitro.* Development, 1990. **110**(2): p. 401-415.
241. Ambort, D., et al., *Specific processing of tenascin-C by the metalloprotease meprinbeta neutralizes its inhibition of cell spreading.* Matrix Biol, 2010. **29**(1): p. 31-42.

242. Talts, J.F., et al., *Tenascin-C modulates tumor stroma and monocyte/macrophage recruitment but not tumor growth or metastasis in a mouse strain with spontaneous mammary cancer.* J Cell Sci, 1999. **112 (Pt 12)**: p. 1855-64.
243. Tucker, R.P. and R. Chiquet-Ehrismann, *Tenascin-C: its functions as an integrin ligand.* The international journal of biochemistry & cell biology, 2015. **65**: p. 165-168.
244. Hauzenberger, D., et al., *Tenascin-C inhibits beta1 integrin-dependent T lymphocyte adhesion to fibronectin through the binding of its fnIII 1-5 repeats to fibronectin.* Eur J Immunol, 1999. **29**(5): p. 1435-47.
245. El-Karef, A., et al., *Deficiency of tenascin-C attenuates liver fibrosis in immune-mediated chronic hepatitis in mice.* J Pathol, 2007. **211**(1): p. 86-94.
246. Huang, W., et al., *Interference of tenascin-C with syndecan-4 binding to fibronectin blocks cell adhesion and stimulates tumor cell proliferation.* Cancer Res, 2001. **61**(23): p. 8586-94.
247. Schalkwijk, J., et al., *Tenascin expression in human dermis is related to epidermal proliferation.* Am J Pathol, 1991. **139**(5): p. 1143-50.
248. Gerritsen, M.J., et al., *Recruitment of cycling epidermal cells and expression of filaggrin, involucrin and tenascin in the margin of the active psoriatic plaque, in the uninvolved skin of psoriatic patients and in the normal healthy skin.* J Dermatol Sci, 1997. **14**(3): p. 179-88.
249. Schenk, S., L. Bruckner-Tuderman, and R. Chiquet-Ehrismann, *Dermo-epidermal separation is associated with induced tenascin expression in human skin.* Br J Dermatol, 1995. **133**(1): p. 13-22.
250. Swindle, C.S., et al., *Epidermal growth factor (EGF)-like repeats of human tenascin-C as ligands for EGF receptor.* J Cell Biol, 2001. **154**(2): p. 459-68.
251. Iyer, A.K.V., et al., *Cell surface restriction of EGFR by a tenascin cytotactin-encoded EGF-like repeat is preferential for motility-related signaling.* Journal of cellular physiology, 2008. **214**(2): p. 504-512.
252. Wallner, K., et al., *EGF-Like domain of tenascin-C is proapoptotic for cultured smooth muscle cells.* Arterioscler Thromb Vasc Biol, 2004. **24**(8): p. 1416-21.
253. To, W.S. and K.S. Midwood, *Cryptic domains of tenascin-C differentially control fibronectin fibrillogenesis.* Matrix Biol, 2010. **29**(7): p. 573-85.
254. Saito, Y., et al., *A peptide derived from tenascin-C induces beta1 integrin activation through syndecan-4.* J Biol Chem, 2007. **282**(48): p. 34929-37.
255. Kim, S.H., J. Turnbull, and S. Guimond, *Extracellular matrix and cell signalling: the dynamic cooperation of integrin, proteoglycan and growth factor receptor.* J Endocrinol, 2011. **209**(2): p. 139-51.
256. O'Toole, T.E., et al., *Integrin cytoplasmic domains mediate inside-out signal transduction.* J Cell Biol, 1994. **124**(6): p. 1047-59.
257. Woods, A. and J.R. Couchman, *Syndecan-4 and focal adhesion function.* Current opinion in cell biology, 2001. **13**(5): p. 578-583.
258. Aruffo, A., et al., *CD44 is the principal cell surface receptor for hyaluronate.* Cell, 1990. **61**(7): p. 1303-13.
259. Vogel, W., et al., *The discoidin domain receptor tyrosine kinases are activated by collagen.* Mol Cell, 1997. **1**(1): p. 13-23.
260. Morrissey, M.A., N. Kern, and R.D. Vale, *CD47 Ligation Repositions the Inhibitory Receptor SIRPA to Suppress Integrin Activation and Phagocytosis.* Immunity, 2020. **53**(2): p. 290-302.e6.
261. Dai, S., et al., *Fibroblast Growth Factor Receptors (FGFRs): Structures and Small Molecule Inhibitors.* Cells, 2019. **8**(6).
262. Herbst, R.S., *Review of epidermal growth factor receptor biology.* Int J Radiat Oncol Biol Phys, 2004. **59**(2 Suppl): p. 21-6.
263. Pasquale, E.B., *Eph receptors and ephrins in cancer: bidirectional signalling and beyond.* Nat Rev Cancer, 2010. **10**(3): p. 165-80.

264. Nakada, M., et al., *Receptor Tyrosine Kinases: Principles and Functions in Glioma Invasion.* Adv Exp Med Biol, 2020. **1202**: p. 151-178.
265. Clause, K.C. and T.H. Barker, *Extracellular matrix signaling in morphogenesis and repair.* Curr Opin Biotechnol, 2013. **24**(5): p. 830-3.
266. Hynes, R.O., *Integrins: bidirectional, allosteric signaling machines.* Cell, 2002. **110**(6): p. 673-87.
267. de Melker, A.A. and A. Sonnenberg, *Integrins: alternative splicing as a mechanism to regulate ligand binding and integrin signaling events.* Bioessays, 1999. **21**(6): p. 499-509.
268. Ruoslahti, E., *RGD and other recognition sequences for integrins.* Annual review of cell and developmental biology, 1996. **12**(1): p. 697-715.
269. Wipff, P.-J. and B. Hinz, *Integrins and the activation of latent transforming growth factor β1–an intimate relationship.* European journal of cell biology, 2008. **87**(8-9): p. 601-615.
270. Lee, J.O., et al., *Crystal structure of the A domain from the alpha subunit of integrin CR3 (CD11b/CD18).* Cell, 1995. **80**(4): p. 631-8.
271. Gullberg, D.E. and E. Lundgren-Åkerlund, *Collagen-binding I domain integrins—what do they do?* Progress in histochemistry and cytochemistry, 2002. **37**(1): p. 3-54.
272. Larson, R.S., et al., *Primary structure of the leukocyte function-associated molecule-1 alpha subunit: an integrin with an embedded domain defining a protein superfamily.* J Cell Biol, 1989. **108**(2): p. 703-12.
273. Goldfinger, L.E., et al., *The alpha3 laminin subunit, alpha6beta4 and alpha3beta1 integrin coordinately regulate wound healing in cultured epithelial cells and in the skin.* Journal of Cell Science, 1999. **112**(16): p. 2615-2629.
274. Takagi, J., et al., *Global conformational rearrangements in integrin extracellular domains in outside-in and inside-out signaling.* Cell, 2002. **110**(5): p. 599-11.
275. Watanabe, N., et al., *Mechanisms and consequences of agonist-induced talin recruitment to platelet integrin alphaIIbbeta3.* J Cell Biol, 2008. **181**(7): p. 1211-22.
276. Critchley, D.R. and A.R. Gingras, *Talin at a glance.* J Cell Sci, 2008. **121**(Pt 9): p. 1345-7.
277. Brown, M.C., J.A. Perrotta, and C.E. Turner, *Identification of LIM3 as the principal determinant of paxillin focal adhesion localization and characterization of a novel motif on paxillin directing vinculin and focal adhesion kinase binding.* J Cell Biol, 1996. **135**(4): p. 1109-23.
278. Lawson, C., et al., *FAK promotes recruitment of talin to nascent adhesions to control cell motility.* J Cell Biol, 2012. **196**(2): p. 223-32.
279. Renshaw, M.W., X.D. Ren, and M.A. Schwartz, *Growth factor activation of MAP kinase requires cell adhesion.* EMBO J, 1997. **16**(18): p. 5592-9.
280. Gerber, E.E., et al., *Integrin-modulating therapy prevents fibrosis and autoimmunity in mouse models of scleroderma.* Nature, 2013. **503**(7474): p. 126-130.
281. Raetz, C.R. and C. Whitfield, *Lipopolysaccharide endotoxins.* Annu Rev Biochem, 2002. **71**: p. 635-700.
282. Matzinger, P., *Tolerance, danger, and the extended family.* Annu Rev Immunol, 1994. **12**: p. 991-1045.
283. Janeway, C.A., Jr., *Approaching the asymptote? Evolution and revolution in immunology.* Cold Spring Harb Symp Quant Biol, 1989. **54 Pt 1**: p. 1-13.
284. Sokolove, J., et al., *Immune complexes containing citrullinated fibrinogen costimulate macrophages via Toll-like receptor 4 and Fcgamma receptor.* Arthritis Rheum, 2011. **63**(1): p. 53-62.
285. Okamura, Y., et al., *The extra domain A of fibronectin activates Toll-like receptor 4.* J Biol Chem, 2001. **276**(13): p. 10229-33.
286. Julier, Z., et al., *The TLR4 Agonist Fibronectin Extra Domain A is Cryptic, Exposed by Elastase-2; use in a fibrin matrix cancer vaccine.* Sci Rep, 2015. **5**.

287. Johnson, G.B., et al., *Receptor-mediated monitoring of tissue well-being via detection of soluble heparan sulfate by Toll-like receptor 4.* J Immunol, 2002. **168**(10): p. 5233-9.
288. Schaefer, L., et al., *The matrix component biglycan is proinflammatory and signals through Toll-like receptors 4 and 2 in macrophages*, in *J Clin Invest*. 2005. p. 2223-33.
289. Merline, R., et al., *Signaling by the Matrix Proteoglycan Decorin Controls Inflammation and Cancer Through PDCD4 and MicroRNA-21.* Science Signaling, 2011. **4**(199).
290. Valesini, G., et al., *Citrullination and autoimmunity.* Autoimmun Rev, 2015. **14**(6): p. 490-7.
291. Akira, S., S. Uematsu, and O. Takeuchi, *Pathogen recognition and innate immunity.* Cell, 2006. **124**(4): p. 783-801.
292. O'Neill, L.A., et al., *Mal and MyD88: adapter proteins involved in signal transduction by Toll-like receptors.* J Endotoxin Res, 2003. **9**(1): p. 55-9.
293. Kawai, T., et al., *Unresponsiveness of MyD88-deficient mice to endotoxin.* Immunity, 1999. **11**(1): p. 115-22.
294. Sakai, J., et al., *Lipopolysaccharide-induced NF-κB nuclear translocation is primarily dependent on MyD88, but TNFα expression requires TRIF and MyD88.* Scientific reports, 2017. **7**(1): p. 1428-1428.
295. Neumann, J., et al., *Nanoscale distribution of TLR4 on primary human macrophages stimulated with LPS and ATI.* Nanoscale, 2019. **11**(19): p. 9769-9779.
296. Wynn, T.A., A. Chawla, and J.W. Pollard, *Macrophage biology in development, homeostasis and disease.* Nature, 2013. **496**(7446): p. 445-55.
297. Bartocci, A., et al., *Macrophages specifically regulate the concentration of their own growth factor in the circulation.* Proc Natl Acad Sci U S A, 1987. **84**(17): p. 6179-83.
298. Murray, P.J. and T.A. Wynn, *Protective and pathogenic functions of macrophage subsets.* Nat Rev Immunol. **11**(11): p. 723-37.
299. Holness, C.L. and D.L. Simmons, *Molecular cloning of CD68, a human macrophage marker related to lysosomal glycoproteins.* Blood, 1993. **81**(6): p. 1607-13.
300. Wright, S.D., et al., *CD14, a receptor for complexes of lipopolysaccharide (LPS) and LPS binding protein.* Science, 1990. **249**(4975): p. 1431-3.
301. Martinez-Pomares, L., *The mannose receptor.* J Leukoc Biol, 2012. **92**(6): p. 1177-86.
302. Stein, M., et al., *Interleukin 4 potently enhances murine macrophage mannose receptor activity: a marker of alternative immunologic macrophage activation.* J Exp Med, 1992. **176**(1): p. 287-92.
303. Stout, R.D., et al., *Macrophages sequentially change their functional phenotype in response to changes in microenvironmental influences.* J Immunol, 2005. **175**(1): p. 342-9.
304. Malyshev, I. and Y. Malyshev, *Current Concept and Update of the Macrophage Plasticity Concept: Intracellular Mechanisms of Reprogramming and M3 Macrophage "Switch" Phenotype.* Biomed Res Int, 2015. **2015**.
305. Hamilton, J.A., *Colony-stimulating factors in inflammation and autoimmunity.* Nat Rev Immunol, 2008. **8**(7): p. 533-44.
306. Chiquet-Ehrismann, R. and M. Chiquet, *Tenascins: regulation and putative functions during pathological stress.* Journal of Pathology, 2003. **200**(4): p. 488-499.
307. Willems, I., J. Arends, and M. Daemen, *Tenascin and fibronectin expression in healing human myocardial scars.* Journal of Pathology, 1996. **179**(3): p. 321-325.
308. Duerr, G., et al., *Comparison of Myocardial Remodeling between Cryoinfarction and Reperfused Infarction in Mice.* Journal of Biomedicine and Biotechnology, 2011.
309. Page, T.H., et al., *Raised circulating tenascin-C in rheumatoid arthritis.* Arthritis Research & Therapy, 2012. **14**(6).
310. Zuliani-Alvarez, L., et al., *Mapping tenascin-C interaction with toll-like receptor 4 reveals a new subset of endogenous inflammatory triggers.* Nat Commun, 2017. **8**(1): p. 1595.
311. Piccinini, A.M., et al., *Distinct microenvironmental cues stimulate divergent TLR4-mediated signaling pathways in macrophages.* Sci Signal, 2016. **9**(443): p. ra86.

312. Piccinini, A.M. and K.S. Midwood, *Endogenous control of immunity against infection: tenascin-C regulates TLR4-mediated inflammation via microRNA-155.* Cell Rep, 2012. **2**(4): p. 914-26.
313. Patel, L., et al., *Tenascin-C induces inflammatory mediators and matrix degradation in osteoarthritic cartilage.* BMC Musculoskelet Disord, 2011. **12**: p. 164.
314. Benbow, J.H., et al., *Diet-Induced Obesity Enhances Progression of Hepatocellular Carcinoma through Tenascin-C/Toll-Like Receptor 4 Signaling.* Am J Pathol, 2016. **186**(1): p. 145-58.
315. Kanayama, M., et al., *α9 integrin and its ligands constitute critical joint microenvironments for development of autoimmune arthritis.* The Journal of Immunology, 2009. **182**(12): p. 8015-8025.
316. Kanayama, M., et al., *α9β1 integrin-mediated signaling serves as an intrinsic regulator of pathogenic Th17 cell generation.* The Journal of Immunology, 2011. **187**(11): p. 5851-5864.
317. Shimojo, N., et al., *Tenascin-C may accelerate cardiac fibrosis by activating macrophages via the integrin αVβ3/nuclear factor–κB/interleukin-6 axis.* Hypertension, 2015. **66**(4): p. 757-766.
318. Prieto, A.L., G.M. Edelman, and K.L. Crossin, *Multiple integrins mediate cell attachment to cytotactin/tenascin.* Proc Natl Acad Sci U S A, 1993. **90**(21): p. 10154-8.
319. Yokoyama, K., et al., *Identification of amino acid sequences in fibrinogen gamma -chain and tenascin C C-terminal domains critical for binding to integrin alpha vbeta 3.* J Biol Chem, 2000. **275**(22): p. 16891-8.
320. Hesse, J., et al., *CD73-derived adenosine and tenascin-C control cytokine production by epicardium-derived cells formed after myocardial infarction.* FASEB J, 2017. **31**(7): p. 3040-3053.
321. Martinon, F., K. Burns, and J. Tschopp, *The inflammasome: a molecular platform triggering activation of inflammatory caspases and processing of proIL-β.* Molecular cell, 2002. **10**(2): p. 417-426.
322. Palenski, T.L., C.M. Sorenson, and N. Sheibani, *Inflammatory cytokine-specific alterations in retinal endothelial cell function.* Microvasc Res, 2013. **89**: p. 57-69.
323. Machino-Ohtsuka, T., et al., *Tenascin-C aggravates autoimmune myocarditis via dendritic cell activation and Th17 cell differentiation.* J Am Heart Assoc, 2014. **3**(6): p. e001052.
324. Ruhmann, M., et al., *Endogenous activation of adaptive immunity: tenascin-C drives interleukin-17 synthesis in murine arthritic joint disease.* Arthritis Rheum, 2012. **64**(7): p. 2179-90.
325. Bhattacharyya, S., et al., *Tenascin-C drives persistence of organ fibrosis.* Nature communications, 2016. **7**(1): p. 1-14.
326. Kuriyama, N., et al., *Tenascin-C: a novel mediator of hepatic ischemia and reperfusion injury.* Hepatology, 2011. **54**(6): p. 2125-36.
327. Maqbool, A., et al., *Tenascin C upregulates interleukin-6 expression in human cardiac myofibroblasts via toll-like receptor 4.* World J Cardiol, 2016. **8**(5): p. 340-50.
328. De Laporte, L., et al., *Tenascin C promiscuously binds growth factors via its fifth fibronectin type III-like domain.* PLoS One, 2013. **8**(4): p. e62076.
329. Aluwihare, P., et al., *Mice that lack activity of αvβ6-and αvβ8-integrins reproduce the abnormalities of Tgfb1-and Tgfb3-null mice.* Journal of cell science, 2009. **122**(2): p. 227-232.
330. Kelly, A., et al., *Regulation of innate and adaptive immunity by TGFβ*, in *Advances in immunology*. 2017, Elsevier. p. 137-233.
331. Rüegg, C.R., R. Chiquet-Ehrismann, and S.S. Alkan, *Tenascin, an extracellular matrix protein, exerts immunomodulatory activities.* Proc Natl Acad Sci U S A, 1989. **86**(19): p. 7437-41.
332. Jachetti, E., et al., *Tenascin-C protects cancer stem–like cells from immune surveillance by arresting T-cell activation.* Cancer research, 2015. **75**(10): p. 2095-2108.

333. Puente Navazo, M.D., D. Valmori, and C. Rüegg, *The alternatively spliced domain TnFnIII A1A2 of the extracellular matrix protein tenascin-C suppresses activation-induced T lymphocyte proliferation and cytokine production.* J Immunol, 2001. **167**(11): p. 6431-40.
334. Humphrey, J.D., E.R. Dufresne, and M.A. Schwartz, *Mechanotransduction and extracellular matrix homeostasis.* Nat Rev Mol Cell Biol, 2014. **15**(12): p. 802-12.
335. Meredith Jr, J., B. Fazeli, and M. Schwartz, *The extracellular matrix as a cell survival factor.* Molecular biology of the cell, 1993. **4**(9): p. 953-961.
336. Peyton, S.R. and A.J. Putnam, *Extracellular matrix rigidity governs smooth muscle cell motility in a biphasic fashion.* J Cell Physiol, 2005. **204**(1): p. 198-209.
337. Wang, H.-B., M. Dembo, and Y.-L. Wang, *Substrate flexibility regulates growth and apoptosis of normal but not transformed cells.* American Journal of Physiology-Cell Physiology, 2000. **279**(5): p. C1345-C1350.
338. Pelham, R.J. and Y.-l. Wang, *Cell locomotion and focal adhesions are regulated by substrate flexibility.* Proceedings of the National Academy of Sciences, 1997. **94**(25): p. 13661-13665.
339. Klein, E.A., et al., *Cell-cycle control by physiological matrix elasticity and in vivo tissue stiffening.* Curr Biol, 2009. **19**(18): p. 1511-8.
340. Wang, T., et al., *Extracellular matrix stiffness and cell contractility control RNA localization to promote cell migration.* Nat Commun, 2017. **8**(1): p. 896.
341. Yui, S., et al., *YAP/TAZ-Dependent Reprogramming of Colonic Epithelium Links ECM Remodeling to Tissue Regeneration.* Cell Stem Cell, 2018. **22**(1): p. 35-49.e7.
342. Karnik, S.K., et al., *A critical role for elastin signaling in vascular morphogenesis and disease.* Development, 2003. **130**(2): p. 411-23.
343. Karamichos, D., R.A. Brown, and V. Mudera, *Collagen stiffness regulates cellular contraction and matrix remodeling gene expression.* J Biomed Mater Res A, 2007. **83**(3): p. 887-94.
344. Edgar, L.T., et al., *Extracellular matrix density regulates the rate of neovessel growth and branching in sprouting angiogenesis.* PLoS One, 2014. **9**(1): p. e85178.
345. Wen, J.H., et al., *Interplay of matrix stiffness and protein tethering in stem cell differentiation.* Nat Mater, 2014. **13**(10): p. 979-87.
346. Storm, C., et al., *Nonlinear elasticity in biological gels.* Nature, 2005. **435**(7039): p. 191-4.
347. Spencer, V.A., R. Xu, and M.J. Bissell, *Extracellular matrix, nuclear and chromatin structure, and gene expression in normal tissues and malignant tumors: a work in progress.* Adv Cancer Res, 2007. **97**: p. 275-94.
348. Handwerger, K.E. and J.G. Gall, *Subnuclear organelles: new insights into form and function.* Trends Cell Biol, 2006. **16**(1): p. 19-26.
349. Starr, D.A. and H.N. Fridolfsson, *Interactions between nuclei and the cytoskeleton are mediated by SUN-KASH nuclear-envelope bridges.* Annu Rev Cell Dev Biol, 2010. **26**: p. 421-44.
350. Petersen, O.W., et al., *Interaction with basement membrane serves to rapidly distinguish growth and differentiation pattern of normal and malignant human breast epithelial cells.* Proceedings of the National Academy of Sciences, 1992. **89**(19): p. 9064-9068.
351. Lelièvre, S.A., et al., *Tissue phenotype depends on reciprocal interactions between the extracellular matrix and the structural organization of the nucleus.* Proc Natl Acad Sci U S A, 1998. **95**(25): p. 14711-6.
352. Schmidhauser, C., et al., *Extracellular matrix and hormones transcriptionally regulate bovine beta-casein 5'sequences in stably transfected mouse mammary cells.* Proceedings of the National Academy of Sciences, 1990. **87**(23): p. 9118-9122.
353. Pujuguet, P., et al., *Trichostatin a inhibits β-casein expression in mammary epithelial cells.* Journal of cellular biochemistry, 2001. **83**(4): p. 660-670.
354. Plachot, C. and S.A. Lelièvre, *DNA methylation control of tissue polarity and cellular differentiation in the mammary epithelium.* Exp Cell Res, 2004. **298**(1): p. 122-32.

355. Parker, M.W., et al., *Fibrotic extracellular matrix activates a profibrotic positive feedback loop.* J Clin Invest, 2014. **124**(4): p. 1622-35.
356. Bartel, D., *MicroRNAs: Target Recognition and Regulatory Functions.* Cell, 2009. **136**(2): p. 215-233.
357. Morlando, M., et al., *Primary microRNA transcripts are processed co-transcriptionally.* Nat Struct Mol Biol, 2008. **15**(9): p. 902-9.
358. Finnegan, E.F. and A.E. Pasquinelli, *MicroRNA biogenesis: regulating the regulators.* Critical reviews in biochemistry and molecular biology, 2013. **48**(1): p. 51-68.
359. Rutnam, Z., T. Wight, and B. Yang, *miRNAs regulate expression and function of extracellular matrix molecules.* Matrix Biology, 2013. **32**(2): p. 74-85.
360. Piccinini, A. and K. Midwood, *Illustrating the interplay between the extracellular matrix and microRNAs.* International Journal of Experimental Pathology, 2014. **95**(3): p. 158-180.
361. Shan, S., et al., *MicroRNA MiR-17 retards tissue growth and represses fibronectin expression.* Nature Cell Biology, 2009. **11**(8): p. 1031-U278.
362. Morton, S.U., et al., *microRNA-138 modulates cardiac patterning during embryonic development.* Proc Natl Acad Sci U S A, 2008. **105**(46): p. 17830-5.
363. Mouw, J.K., et al., *Tissue mechanics modulate microRNA-dependent PTEN expression to regulate malignant progression.* Nature Medicine, 2014. **20**(4): p. 360-+.
364. Shi, W., et al., *MicroRNA-301 mediates proliferation and invasion in human breast cancer.* Cancer Res, 2011. **71**(8): p. 2926-37.
365. Li, C., et al., *Post-transcriptional up-regulation of miR-21 by type I collagen.* Mol Carcinog, 2011. **50**(7): p. 563-70.
366. Sengupta, S., et al., *MicroRNA 29c is down-regulated in nasopharyngeal carcinomas, up-regulating mRNAs encoding extracellular matrix proteins.* Proc Natl Acad Sci U S A, 2008. **105**(15): p. 5874-8.
367. Maurer, B., et al., *MicroRNA-29, a key regulator of collagen expression in systemic sclerosis.* Arthritis Rheum, 2010. **62**(6): p. 1733-43.
368. Fang, J.H., et al., *MicroRNA-29b suppresses tumor angiogenesis, invasion, and metastasis by regulating matrix metalloproteinase 2 expression.* Hepatology, 2011. **54**(5): p. 1729-40.
369. Shi, J., et al., *miR-328-3p mediates the anti-tumor effect in osteosarcoma via directly targeting MMP-16*, in *Cancer Cell Int*. 2019.
370. Yuan, C., *miR-616 promotes breast cancer migration and invasion by targeting TIMP2 and regulating MMP signaling.* Oncol Lett, 2019. **18**(3): p. 2348-55.
371. Ucar, A., et al., *miR-212 and miR-132 are required for epithelial stromal interactions necessary for mouse mammary gland development.* Nat Genet, 2010. **42**(12): p. 1101-8.
372. Koshizuka, K., et al., *Inhibition of integrin β1-mediated oncogenic signalling by the antitumor microRNA-29 family in head and neck squamous cell carcinoma.* Oncotarget, 2018. **9**(3): p. 3663-76.
373. Augoff, K., et al., *miR-31 Is a Broad Regulator of beta 1-Integrin Expression and Function in Cancer Cells.* Molecular Cancer Research, 2011. **9**(11): p. 1500-1508.
374. Fang, L., et al., *MicroRNA miR-93 promotes tumor growth and angiogenesis by targeting integrin-beta8.* Oncogene, 2011. **30**(7): p. 806-21.
375. Zhang, X., et al., *Up-regulated microRNA-143 transcribed by nuclear factor kappa B enhances hepatocarcinoma metastasis by repressing fibronectin expression.* Hepatology, 2009. **50**(2): p. 490-9.
376. Bala, S., et al., *Up-regulation of microRNA-155 in macrophages contributes to increased tumor necrosis factor {alpha} (TNF{alpha}) production via increased mRNA half-life in alcoholic liver disease.* J Biol Chem, 2011. **286**(2): p. 1436-44.
377. Tili, E., et al., *Modulation of miR-155 and miR-125b levels following lipopolysaccharide/TNF-alpha stimulation and their possible roles in regulating the response to endotoxin shock.* J Immunol, 2007. **179**(8): p. 5082-9.

378. Kontoyiannis, D., et al., *Impaired on/off regulation of TNF biosynthesis in mice lacking TNF AU-rich elements: implications for joint and gut-associated immunopathologies.* Immunity, 1999. **10**(3): p. 387-398.
379. Birgersdotter, A., R. Sandberg, and I. Ernberg, *Gene expression perturbation in vitro - A growing case for three-dimensional (3D) culture systems.* Seminars in Cancer Biology, 2005. **15**(5): p. 405-412.
380. Badylak, S.F., D.O. Freytes, and T.W. Gilbert, *Extracellular matrix as a biological scaffold material: Structure and function.* Acta Biomater, 2009. **5**(1): p. 1-13.
381. Kleinman, H. and G. Martin, *Matrigel: Basement membrane matrix with biological activity.* Seminars in Cancer Biology, 2005. **15**(5): p. 378-386.
382. Xu, C., et al., *Feeder-free growth of undifferentiated human embryonic stem cells.* Nat Biotechnol, 2001. **19**(10): p. 971-4.
383. Schwartz, R.E., et al., *Multipotent adult progenitor cells from bone marrow differentiate into functional hepatocyte-like cells.* J Clin Invest, 2002. **109**(10): p. 1291-302.
384. Elisseeff, J., et al., *Biological response of chondrocytes to hydrogels.* Reparative Medicine: Growing Tissues and Organs, 2002. **961**: p. 118-122.
385. Hughes, C., L. Postovit, and G. Lajoie, *Matrigel: A complex protein mixture required for optimal growth of cell culture.* Proteomics, 2010. **10**(9): p. 1886-1890.
386. Taub, M., et al., *Epidermal Growth-Factor Or*

Transforming Growth Factor-Alpha Is Required For Kidney Tubulogenesis In

Matrigel Cultures In Serum-Free Medium. Proceedings of the National Academy of Sciences of the United States of America, 1990. **87**(10): p. 4002-4006.
387. Doyle, A. and K. Yamada, *Mechanosensing via cell-matrix adhesions in 3D microenvironments.* Experimental Cell Research, 2016. **343**(1): p. 60-66.
388. Baker, B. and C. Chen, *Deconstructing the third dimension - how 3D culture microenvironments alter cellular cues.* Journal of Cell Science, 2012. **125**(13): p. 3015-3024.
389. Fitzpatrick, L. and T. McDevitt, *Cell-derived matrices for tissue engineering and regenerative medicine applications.* Biomaterials Science, 2015. **3**(1): p. 12-24.
390. Badylak, S., D. Freytes, and T. Gilbert, *Extracellular matrix as a biological scaffold material: Structure and function.* Acta Biomaterialia, 2009. **5**(1): p. 1-13.
391. Miller, J., et al., *Bioactive hydrogels made from step-growth derived PEG-peptide macromers.* Biomaterials, 2010. **31**(13): p. 3736-3743.
392. Ott, H., et al., *Perfusion-decellularized matrix: using nature's platform to engineer a bioartificial heart.* Nature Medicine, 2008. **14**(2): p. 213-221.
393. Uygun, B., et al., *Organ reengineering through development of a transplantable recellularized liver graft using decellularized liver matrix.* Nature Medicine, 2010. **16**(7): p. 814-U120.
394. Petersen, T., et al., *Tissue-Engineered Lungs for in Vivo Implantation.* Science, 2010. **329**(5991): p. 538-541.
395. Hoshiba, T., et al., *Maintenance of cartilaginous gene expression on extracellular matrix derived from serially passaged chondrocytes during in vitro chondrocyte expansion.* Journal of Biomedical Materials Research Part a, 2012. **100A**(3): p. 694-702.
396. Lu, H., et al., *Autologous extracellular matrix scaffolds for tissue engineering.* Biomaterials, 2011. **32**(10): p. 2489-2499.
397. Gilbert, T., T. Sellaro, and S. Badylak, *Decellularization of tissues and organs.* Biomaterials, 2006. **27**(19): p. 3675-3683.
398. Klein, T. and R. Bischoff, *Physiology and pathophysiology of matrix metalloproteases.* Amino acids, 2011. **41**(2): p. 271-290.

399. Lai, Y., et al., *Reconstitution of Marrow-Derived Extracellular Matrix Ex Vivo: A Robust Culture System for Expanding Large-Scale Highly Functional Human Mesenchymal Stem Cells.* Stem Cells and Development, 2010. **19**(7): p. 1095-1107.
400. Cha, M., et al., *Induction of Re-Differentiation of Passaged Rat Chondrocytes Using a Naturally Obtained Extracellular Matrix Microenvironment.* Tissue Engineering Part a, 2013. **19**(7-8): p. 978-988.
401. Sun, Y., et al., *Rescuing replication and osteogenesis of aged mesenchymal stem cells by exposure to a young extracellular matrix.* Faseb Journal, 2011. **25**(5): p. 1474-1485.
402. Ahringer, J., *Reverse genetics*, in *WormBook: The Online Review of C. elegans Biology [Internet]*. 2006, WormBook.
403. Hsu, P., et al., *DNA targeting specificity of RNA-guided Cas9 nucleases.* Nature Biotechnology, 2013. **31**(9): p. 827-+.
404. Mali, P., et al., *RNA-Guided Human Genome Engineering via Cas9.* Science, 2013. **339**(6121): p. 823-826.
405. Westra, E., A. Buckling, and P. Fineran, *CRISPR-Cas systems: beyond adaptive immunity.* Nature Reviews Microbiology, 2014. **12**(5): p. 317-326.
406. Hsu, P., E. Lander, and F. Zhang, *Development and Applications of CRISPR-Cas9 for Genome Engineering.* Cell, 2014. **157**(6): p. 1262-1278.
407. Sander, J. and J. Joung, *CRISPR-Cas systems for editing, regulating and targeting genomes.* Nature Biotechnology, 2014. **32**(4): p. 347-355.
408. Choi, P.S. and M. Meyerson, *Targeted genomic rearrangements using CRISPR/Cas technology.* Nat Commun, 2014. **5**: p. 3728.
409. Kim, S., et al., *Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins.* Genome Res, 2014. **24**(6): p. 1012-9.
410. Liang, X., et al., *Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection.* J Biotechnol, 2015. **208**: p. 44-53.
411. Dorsett, Y. and T. Tuschl, *siRNAs: applications in functional genomics and potential as therapeutics.* Nat Rev Drug Discov, 2004. **3**(4): p. 318-29.
412. Dykxhoorn, D.M., C.D. Novina, and P.A. Sharp, *Killing the messenger: short RNAs that silence gene expression.* Nat Rev Mol Cell Biol, 2003. **4**(6): p. 457-67.
413. Jinek, M., et al., *A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity.* Science, 2012. **337**(6096): p. 816-821.
414. Wiedenheft, B., S. Sternberg, and J. Doudna, *RNA-guided genetic silencing systems in bacteria and archaea.* Nature, 2012. **482**(7385): p. 331-338.
415. Haurwitz, R., et al., *Sequence- and Structure-Specific RNA Processing by a CRISPR Endonuclease.* Science, 2010. **329**(5997): p. 1355-1358.
416. Brouns, S., et al., *Small CRISPR RNAs guide antiviral defense in prokaryotes.* Science, 2008. **321**(5891): p. 960-964.
417. Deltcheva, E., et al., *CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III.* Nature, 2011. **471**(7340): p. 602-+.
418. Staals, R., et al., *RNA Targeting by the Type III-A CRISPR-Cas Csm Complex of Thermus thermophilus.* Molecular Cell, 2014. **56**(4): p. 518-530.
419. Deveau, H., et al., *Phage response to CRISPR-Encoded resistance in Streptococcus thermophilus.* Journal of Bacteriology, 2008. **190**(4): p. 1390-1400.
420. Wells, D.J., *Gene therapy progress and prospects: electroporation and other physical methods.* Gene Ther, 2004. **11**(18): p. 1363-9.
421. Baum, C., et al., *Mutagenesis and oncogenesis by chromosomal insertion of gene transfer vectors.* Hum Gene Ther, 2006. **17**(3): p. 253-63.
422. Bessis, N., F.J. GarciaCozar, and M.C. Boissier, *Immune responses to gene therapy vectors: influence on vector function and effector mechanisms.* Gene Ther, 2004. **11 Suppl 1**: p. S10-7.

423. Gaj, T., et al., *Targeted gene knockout by direct delivery of zinc-finger nuclease proteins.* Nature methods, 2012. **9**(8): p. 805-807.
424. Ran, F., et al., *Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity.* Cell, 2013. **154**(6): p. 1380-1389.
425. Fu, Y., et al., *High frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells.* Nat Biotechnol, 2013. **31**(9): p. 822-6.
426. Moreno-Mateos, M.A., et al., *CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo.* Nature methods, 2015. **12**(10): p. 982-988.
427. Wang, T., et al., *Genetic screens in human cells using the CRISPR-Cas9 system.* Science, 2014. **343**(6166): p. 80-84.
428. Modzelewski, A.J., et al., *Efficient mouse genome engineering by CRISPR-EZ technology.* Nat Protoc, 2018. **13**(6): p. 1253-1274.
429. Lin, S., et al., *Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery.* Elife, 2014. **3**.
430. Filippo, J., P. Sung, and H. Klein, *Mechanism of eukaryotic homologous recombination.* Annual Review of Biochemistry, 2008. **77**: p. 229-257.
431. Kaulich, M. and S. Dowdy, *Combining CRISPR/Cas9 and rAAV Templates for Efficient Gene Editing.* Nucleic Acid Therapeutics, 2015. **25**(6): p. 287-296.
432. Bottcher, R., et al., *Efficient chromosomal gene modification with CRISPR/cas9 and PCR-based homologous recombination donors in cultured Drosophila cells.* Nucleic Acids Research, 2014. **42**(11).
433. Rong, Z., et al., *Homologous recombination in human embryonic stem cells using CRISPR/Cas9 nickase and a long DNA donor template.* Protein & Cell, 2014. **5**(4): p. 258-260.
434. Lieber, M., *The mechanism of human nonhomologous DNA end joining.* Journal of Biological Chemistry, 2008. **283**(1): p. 1-5.
435. Doench, J.G., et al., *Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9.* Nat Biotechnol, 2016. **34**(2): p. 184-191.
436. Fabregat, A., et al., *The Reactome Pathway Knowledgebase.* Nucleic Acids Res, 2018. **46**(D1): p. D649-D655.
437. Hernandez-Segura, A., et al., *Unmasking Transcriptional Heterogeneity in Senescent Cells.* Curr Biol, 2017. **27**(17): p. 2652-2660.e4.
438. Basisty, N., et al., *A proteomic atlas of senescence-associated secretomes for aging biomarker development.* PLoS Biol, 2020. **18**(1): p. e3000599.
439. Avelar, R.A., et al., *A multidimensional systems biology analysis of cellular senescence in aging and disease.* Genome Biol, 2020. **21**(1): p. 91.
440. Liao, Y., et al., *WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs.* Nucleic Acids Res, 2019. **47**(W1): p. W199-W205.
441. Maniotis, A., C. Chen, and D. Ingber, *Demonstration of mechanical connections between integrins, cytoskeletal filaments, and nucleoplasm that stabilize nuclear structure*, in *Proc Natl Acad Sci U S A*. 1997. p. 849-54.
442. Tomlin, H. and A.M. Piccinini, *A complex interplay between the extracellular matrix and the innate immune response to microbial pathogens.* Immunology, 2018. **155**(2): p. 186-201.
443. Hughes, C.S., L.M. Postovit, and G.A. Lajoie, *Matrigel: a complex protein mixture required for optimal growth of cell culture.* Proteomics, 2010. **10**(9): p. 1886-90.
444. Gjorevski, N., et al., *Designer matrices for intestinal stem cell and organoid culture.* Nature, 2016. **539**(7630): p. 560-564.
445. Vukicevic, S., et al., *Identification of multiple active growth factors in basement membrane Matrigel suggests caution in interpretation of cellular activity related to extracellular matrix components.* Exp Cell Res, 1992. **202**(1): p. 1-8.
446. Elsdale, T.R., *Parallel orientation of fibroblasts in vitro.* Exp Cell Res, 1968. **51**(2-3): p. 439-50.

447. Kaukonen, R., et al., *Cell-derived matrices for studying cell proliferation and directional migration in a complex 3D microenvironment.* Nature Protocols, 2017. **12**(11): p. 2376-2390.
448. Wang, Y.M., R.C. Gallant, and H.Y. Ni, *Extracellular matrix proteins in the regulation of thrombus formation.* Current Opinion in Hematology, 2016. **23**(3): p. 280-287.
449. Chaubaroux, C., et al., *Cell Alignment Driven by Mechanically Induced Collagen Fiber Alignment in Collagen/Alginate Coatings.* Tissue Eng Part C Methods, 2015. **21**(9): p. 881-8.
450. Sechler, J.L., S.A. Corbett, and J.E. Schwarzbauer, *Modulatory roles for integrin activation and the synergy site of fibronectin during matrix assembly.* Mol Biol Cell, 1997. **8**(12): p. 2563-73.
451. Kii, I., et al., *Incorporation of Tenascin-C into the Extracellular Matrix by Periostin Underlies an Extracellular Meshwork Architecture**, in *J Biol Chem*. 2010. p. 2028-39.
452. Van Robertson, W.B. and B. Schwartz, *Ascorbic acid and the formation of collagen.* J Biol Chem, 1953. **201**(2): p. 689-96.
453. Murad, S., et al., *Regulation of collagen synthesis by ascorbic acid.* Proc Natl Acad Sci U S A, 1981. **78**(5): p. 2879-82.
454. Chojkier, M., et al., *Stimulation of collagen gene expression by ascorbic acid in cultured human fibroblasts. A role for lipid peroxidation?* J Biol Chem, 1989. **264**(28): p. 16957-62.
455. Asgari, M., et al., *In vitro fibrillogenesis of tropocollagen type III in collagen type I affects its relative fibrillar topology and mechanics*, in *Sci Rep*. 2017.
456. Scherzer, M.T., et al., *Fibroblast-Derived Extracellular Matrices: An Alternative Cell Culture System That Increases Metastatic Cellular Properties*, in *PLoS One*. 2015.
457. Gelain, F., et al., *Designer self-assembling peptide nanofiber scaffolds for adult mouse neural stem cell 3-dimensional cultures.* PLoS One, 2006. **1**: p. e119.
458. Zheng, M.H., et al., *Porcine small intestine submucosa (SIS) is not an acellular collagenous matrix and contains porcine DNA: possible implications in human implantation.* J Biomed Mater Res B Appl Biomater, 2005. **73**(1): p. 61-7.
459. Chen, G.Y. and G. Nuñez, *Sterile inflammation: sensing and reacting to damage.* Nat Rev Immunol, 2010. **10**(12): p. 826-37.
460. Hwang, H.S., et al., *NOD2 signaling pathway is involved in fibronectin fragment-induced pro-catabolic factor expressions in human articular chondrocytes*, in *BMB Rep*. 2019. p. 373-8.
461. Scheibner, K.A., et al., *Hyaluronan fragments act as an endogenous danger signal by engaging TLR2.* J Immunol, 2006. **177**(2): p. 1272-81.
462. Jiang, D., et al., *Regulation of lung injury and repair by Toll-like receptors and hyaluronan.* Nat Med, 2005. **11**(11): p. 1173-9.
463. Thomas, A.H., E.R. Edelman, and C.M. Stultz, *Collagen fragments modulate innate immunity.* Exp Biol Med (Maywood), 2007. **232**(3): p. 406-11.
464. Nichols, J.E., et al., *Production and Assessment of Decellularized Pig and Human Lung Scaffolds*, in *Tissue Eng Part A*. 2013. p. 2045-62.
465. Hudson, T.W., S.Y. Liu, and C.E. Schmidt, *Engineering an improved acellular nerve graft via optimized chemical processing.* Tissue Eng, 2004. **10**(9-10): p. 1346-58.
466. Xu, H., et al., *Comparison of Decellularization Protocols for Preparing a Decellularized Porcine Annulus Fibrosus Scaffold*, in *PLoS One*. 2014.
467. Scott, J.E. and J. Dorling, *Differential staining of acid glycosaminoglycans (mucopolysaccharides) by alcian blue in salt solutions.* Histochemie, 1965. **5**(3): p. 221-33.
468. Weaver, V.M., et al., *Reversion of the malignant phenotype of human breast cells in three-dimensional culture and in vivo by integrin blocking antibodies.* J Cell Biol, 1997. **137**(1): p. 231-45.
469. Pickl, M. and C.H. Ries, *Comparison of 3D and 2D tumor models reveals enhanced HER2 activation in 3D associated with an increased response to trastuzumab.* Oncogene, 2009. **28**(3): p. 461-8.

470. Hait, W.N., *Anticancer drug development: the grand challenges.* Nat Rev Drug Discov, 2010. **9**(4): p. 253-4.
471. Brown, J.M. and W.R. Wilson, *Exploiting tumour hypoxia in cancer treatment.* Nat Rev Cancer, 2004. **4**(6): p. 437-47.
472. Tung, Y.C., et al., *High-throughput 3D spheroid culture and drug testing using a 384 hanging drop array.* Analyst, 2011. **136**(3): p. 473-8.
473. Soucy, P.A., et al., *Microelastic properties of lung cell-derived extracellular matrix.* Acta Biomater, 2011. **7**(1): p. 96-105.
474. Cukierman, E., et al., *Taking cell-matrix adhesions to the third dimension.* Science, 2001. **294**(5547): p. 1708-1712.
475. Kusuma, G.D., et al., *Decellularized extracellular matrices produced from immortal cell lines derived from different parts of the placenta support primary mesenchymal stem cell expansion.* PLoS One, 2017. **12**(2): p. e0171488.
476. Franco-Barraza, J., et al., *Preparation of extracellular matrices produced by cultured and primary fibroblasts.* Curr Protoc Cell Biol, 2016. **71**: p. 10 9 1-10 9 34.
477. Kaur, A., et al., *Remodeling of the Collagen Matrix in Aging Skin Promotes Melanoma Metastasis and Affects Immune Cell Motility.* Cancer Discov, 2019. **9**(1): p. 64-81.
478. Woodfin, A., et al., *Endothelial cell activation leads to neutrophil transmigration as supported by the sequential roles of ICAM-2, JAM-A, and PECAM-1.* Blood, 2009. **113**(24): p. 6246-57.
479. Ahlfors, J.E. and K.L. Billiar, *Biomechanical and biochemical characteristics of a human fibroblast-produced and remodeled matrix.* Biomaterials, 2007. **28**(13): p. 2183-91.
480. Rezakhaniha, R., et al., *Experimental investigation of collagen waviness and orientation in the arterial adventitia using confocal laser scanning microscopy.* Biomech Model Mechanobiol, 2012. **11**(3-4): p. 461-73.
481. Püspöki, Z., et al., *Transforms and Operators for Directional Bioimage Analysis: A Survey.* Adv Anat Embryol Cell Biol, 2016. **219**: p. 69-93.
482. Adebayo, O., et al., *Self-assembled smooth muscle cell tissue rings exhibit greater tensile strength than cell-seeded fibrin or collagen gel rings.* Journal of biomedical materials research Part A, 2013. **101**(2): p. 428-437.
483. Lynch, H.A., et al., *Effect of fiber orientation and strain rate on the nonlinear uniaxial tensile material properties of tendon.* J Biomech Eng, 2003. **125**(5): p. 726-31.
484. Duclos, G., et al., *Perfect nematic order in confined monolayers of spindle-shaped cells.* Soft Matter, 2014. **10**(14): p. 2346-53.
485. Isenberg, B.C., et al., *Micropatterned cell sheets with defined cell and extracellular matrix orientation exhibit anisotropic mechanical properties.* J Biomech, 2012. **45**(5): p. 756-61.
486. Neidlinger-Wilke, C., et al., *Fibroblast orientation to stretch begins within three hours.* J Orthop Res, 2002. **20**(5): p. 953-6.
487. Rens, E.G. and R.M.H. Merks, *Cell Contractility Facilitates Alignment of Cells and Tissues to Static Uniaxial Stretch.* Biophys J, 2017. **112**(4): p. 755-766.
488. Collinsworth, A.M., et al., *Orientation and length of mammalian skeletal myocytes in response to a unidirectional stretch.* Cell Tissue Res, 2000. **302**(2): p. 243-51.
489. Liu, C., et al., *Effect of Static Pre-stretch Induced Surface Anisotropy on Orientation of Mesenchymal Stem Cells.* Cell Mol Bioeng, 2014. **7**(1): p. 106-121.
490. van der Schaft, D.W., et al., *Mechanoregulation of vascularization in aligned tissue-engineered muscle: a role for vascular endothelial growth factor.* Tissue Eng Part A, 2011. **17**(21-22): p. 2857-65.
491. Siedlik, M.J. and C.M. Nelson, *Regulation of tissue morphodynamics: an important role for actomyosin contractility.* Curr Opin Genet Dev, 2015. **32**: p. 80-5.
492. Vader, D., et al., *Strain-induced alignment in collagen gels.* PLoS One, 2009. **4**(6): p. e5902.

493. Klebe, R.J., H. Caldwell, and S. Milam, *Cells transmit spatial information by orienting collagen fibers.* Matrix, 1989. **9**(6): p. 451-8.
494. Tondon, A. and R. Kaunas, *The direction of stretch-induced cell and stress fiber orientation depends on collagen matrix stress.* PLoS One, 2014. **9**(2): p. e89592.
495. Lo, C.M., et al., *Cell movement is guided by the rigidity of the substrate.* Biophys J, 2000. **79**(1): p. 144-52.
496. Singh, P., C. Carraher, and J.E. Schwarzbauer, *Assembly of fibronectin extracellular matrix.* Annu Rev Cell Dev Biol, 2010. **26**: p. 397-419.
497. Mao, Y. and J.E. Schwarzbauer, *Stimulatory effects of a three-dimensional microenvironment on cell-mediated fibronectin fibrillogenesis.* J Cell Sci, 2005. **118**(Pt 19): p. 4427-36.
498. Hata, R., et al., *Regulation of collagen metabolism and cell growth by epidermal growth factor and ascorbate in cultured human skin fibroblasts.* Eur J Biochem, 1988. **173**(2): p. 261-7.
499. Ornitz, D.M. and N. Itoh, *Fibroblast growth factors.* Genome biology, 2001. **2**(3): p. 1-12.
500. Bettger, W.J., et al., *Rapid clonal growth and serial passage of human diploid fibroblasts in a lipid-enriched synthetic medium supplemented with epidermal growth factor, insulin, and dexamethasone.* Proc Natl Acad Sci U S A, 1981. **78**(9): p. 5588-92.
501. Warshamana, G.S., et al., *Dexamethasone activates expression of the PDGF-alpha receptor and induces lung fibroblast proliferation.* Am J Physiol, 1998. **274**(4): p. L499-507.
502. Zoppi, N., et al., *Effect of dexamethasone on the assembly of the matrix of fibronectin and on its receptors organization in Ehlers-Danlos syndrome skin fibroblasts.* Cell Biol Int, 1998. **22**(7-8): p. 499-508.
503. Brenner, R.E., et al., *Effects of dexamethasone on proliferation, chemotaxis, collagen I, and fibronectin-metabolism of human fetal lung fibroblasts.* Pediatr Pulmonol, 2001. **32**(1): p. 1-7.
504. Daian, T., et al., *Insulin-like growth factor-I enhances transforming growth factor-beta-induced extracellular matrix protein production through the P38/activating transcription factor-2 signaling pathway in keloid fibroblasts.* J Invest Dermatol, 2003. **120**(6): p. 956-62.
505. Handorf, A.M., et al., *Tissue stiffness dictates development, homeostasis, and disease progression.* Organogenesis, 2015. **11**(1): p. 1-15.
506. Discher, D.E., P. Janmey, and Y.L. Wang, *Tissue cells feel and respond to the stiffness of their substrate.* Science, 2005. **310**(5751): p. 1139-43.
507. Engler, A.J., et al., *Matrix elasticity directs stem cell lineage specification.* Cell, 2006. **126**(4): p. 677-89.
508. Georges, P.C. and P.A. Janmey, *Cell type-specific response to growth on soft materials.* J Appl Physiol (1985), 2005. **98**(4): p. 1547-53.
509. Pathak, A. and S. Kumar, *Independent regulation of tumor cell migration by matrix stiffness and confinement.* Proc Natl Acad Sci U S A, 2012. **109**(26): p. 10334-9.
510. Seewaldt, V., *ECM stiffness paves the way for tumor cells.* Nat Med, 2014. **20**(4): p. 332-3.
511. Soofi, S.S., et al., *The elastic modulus of Matrigel as determined by atomic force microscopy.* J Struct Biol, 2009. **167**(3): p. 216-9.
512. Wells, R.G., *The role of matrix stiffness in regulating cell behavior.* Hepatology, 2008. **47**(4): p. 1394-400.
513. Kaukonen, R., et al., *Normal stroma suppresses cancer cell proliferation via mechanosensitive regulation of JMJD1a-mediated transcription.* Nat Commun, 2016. **7**: p. 12237.
514. Swift, J., et al., *Nuclear lamin-A scales with tissue stiffness and enhances matrix-directed differentiation.* Science, 2013. **341**(6149): p. 1240104.
515. Akira, S. and K. Takeda, *Toll-like receptor signalling.* Nat Rev Immunol, 2004. **4**(7): p. 499-511.
516. Yu, C., et al., *Small molecules enhance CRISPR genome editing in pluripotent stem cells.* Cell Stem Cell, 2015. **16**(2): p. 142-7.

517. Orend, G., *Potential oncogenic action of tenascin-C in tumorigenesis.* Int J Biochem Cell Biol, 2005. **37**(5): p. 1066-83.
518. Yang, H.A., et al., *Reduced expression of Toll-like receptor 4 inhibits human breast cancer cells proliferation and inflammatory cytokines secretion.* Journal of Experimental & Clinical Cancer Research, 2010. **29**.
519. Grosjean, F., et al., *S-phase synchronized CHO cells show elevated transfection efficiency and expression using CaPi.* Cytotechnology, 2002. **38**(1-2): p. 57-62.
520. Suzuki, T., et al., *Comprehensive gene expression profile of LPS-stimulated human monocytes by SAGE.* Blood, 2000. **96**(7): p. 2584-2591.
521. Naba, A., et al., *Characterization of the Extracellular Matrix of Normal and Diseased Tissues Using Proteomics.* J Proteome Res, 2017. **16**(8): p. 3083-3091.
522. Choi, Y.E., et al., *Effects of Tenascin C on the Integrity of Extracellular Matrix and Skin Aging.* Int J Mol Sci, 2020. **21**(22).
523. Bikard, D., et al., *Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system.* Nucleic Acids Research, 2013. **41**(15): p. 7429-7437.
524. BeltCappellino, A., et al., *CRISPR/Cas9-Mediated Knockout and In Situ Inversion of the ORF57 Gene from All Copies of the Kaposi's Sarcoma-Associated Herpesvirus Genome in BCBL-1 Cells.* J Virol, 2019. **93**(21).
525. Soda, M., et al., *Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer.* Nature, 2007. **448**(7153): p. 561-6.
526. Li, J., et al., *Efficient inversions and duplications of mammalian regulatory DNA elements and gene clusters by CRISPR/Cas9.* J Mol Cell Biol, 2015. **7**(4): p. 284-98.
527. Canver, M.C., et al., *Characterization of Genomic Deletion Efficiency Mediated by Clustered Regularly Interspaced Palindromic Repeats (CRISPR)/Cas9 Nuclease System in Mammalian Cells*.* J Biol Chem, 2014. **289**(31): p. 21312-24.
528. Kraft, K., et al., *Deletions, Inversions, Duplications: Engineering of Structural Variants using CRISPR/Cas in Mice.* Cell Rep, 2015. **10**(5): p. 833-839.
529. Mettananda, S., et al., *Editing an alpha-globin enhancer in primary human hematopoietic stem cells as a treatment for beta-thalassemia.* Nat Commun, 2017. **8**(1): p. 424.
530. Korablev, A.N., I.A. Serova, and O.L. Serov, *Generation of megabase-scale deletions, inversions and duplications involving the Contactin-6 gene in mice by CRISPR/Cas9 technology.* BMC Genet, 2017. **18**(Suppl 1).
531. Schaefer, K.A., et al., *Unexpected mutations after CRISPR-Cas9 editing in vivo.* Nat Methods, 2017. **14**(6): p. 547-548.
532. Alanis-Lobato, G., et al., *Frequent loss-of-heterozygosity in CRISPR-Cas9-edited early human embryos.* bioRxiv, 2020.
533. Zuccaro, M.V., et al., *Reading frame restoration at the EYS locus, and allele-specific chromosome removal after Cas9 cleavage in human embryos.* bioRxiv, 2020.
534. Mitalipov, S., *FREQUENT GENE CONVERSION IN HUMAN EMBRYOS INDUCED BY DOUBLE STRAND BREAKS.* bioRxiv, 2020.
535. Cong, L., et al., *Multiplex genome engineering using CRISPR/Cas systems.* Science, 2013. **339**(6121): p. 819-23.
536. Cho, S.W., et al., *Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases.* Genome Res, 2014. **24**(1): p. 132-41.
537. Pattanayak, V., et al., *High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity.* Nat Biotechnol, 2013. **31**(9): p. 839-43.
538. Seidl, C.I., T.A. Fulga, and C.L. Murphy, *CRISPR-Cas9 targeting of MMP13 in human chondrocytes leads to significantly reduced levels of the metalloproteinase and enhanced type II collagen accumulation.* Osteoarthritis Cartilage, 2019. **27**(1): p. 140-147.
539. Chen, X., et al., *Probing the impact of chromatin conformation on genome editing tools.* Nucleic Acids Res, 2016. **44**(13): p. 6482-92.

540. Horlbeck, M.A., et al., *Nucleosomes impede Cas9 access to DNA in vivo and in vitro.* eLife, 2016. **5**.
541. Kuscu, C., et al., *Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease.* Nat Biotechnol, 2014. **32**(7): p. 677-83.
542. Pausch, P., et al., *CRISPR-CasΦ from huge phages is a hypercompact genome editor.* Science, 2020. **369**(6501): p. 333-337.
543. Richardson, C.D., et al., *Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA.* Nat Biotechnol, 2016. **34**(3): p. 339-44.
544. Brinkman, E.K., et al., *Kinetics and Fidelity of the Repair of Cas9-Induced Double-Strand DNA Breaks.* Mol Cell, 2018. **70**(5): p. 801-813 e6.
545. Gopalappa, R., et al., *Paired D10A Cas9 nickases are sometimes more efficient than individual nucleases for gene disruption.* Nucleic Acids Res, 2018. **46**(12): p. e71.
546. Certo, M.T., et al., *Tracking genome engineering outcome at individual DNA breakpoints.* Nat Methods, 2011. **8**(8): p. 671-6.
547. Ihry, R.J., et al., *p53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells.* Nat Med, 2018. **24**(7): p. 939-946.
548. Haapaniemi, E., et al., *CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response.* Nat Med, 2018. **24**(7): p. 927-930.
549. Carroll, D., *p53 Throws CRISPR a Curve.* Trends Pharmacol Sci, 2018. **39**(9): p. 783-784.
550. Poltorak, A., et al., *Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene.* Science, 1998. **282**(5396): p. 2085-8.
551. Poltorak, A., et al., *Genetic and physical mapping of the Lps locus: identification of the toll-4 receptor as a candidate gene in the critical region.* Blood Cells Mol Dis, 1998. **24**(3): p. 340-55.
552. Cook, D.N., D.S. Pisetsky, and D.A. Schwartz, *Toll-like receptors in the pathogenesis of human disease.* Nat Immunol, 2004. **5**(10): p. 975-9.
553. Bala, S., et al., *Up-regulation of MicroRNA-155 in Macrophages Contributes to Increased Tumor Necrosis Factor alpha (TNF alpha) Production via Increased mRNA Half-life in Alcoholic Liver Disease.* Journal of Biological Chemistry, 2011. **286**(2): p. 1436-1444.
554. Koyama, Y., et al., *Type I collagen is a non-adhesive extracellular matrix for macrophages.* Arch Histol Cytol, 2000. **63**(1): p. 71-9.
555. Cui, K., et al., *Distinct Migratory Properties of M1, M2, and Resident Macrophages Are Regulated by αDβ2 and αMβ2 Integrin-Mediated Adhesion.* Front Immunol, 2018. **9**.
556. Tai, S.S. and X.M. Tang, *Manipulating biological samples for environmental scanning electron microscopy observation.* Scanning, 2001. **23**(4): p. 267-72.
557. Donald, A.M., *The use of environmental scanning electron microscopy for imaging wet and insulating materials.* Nat Mater, 2003. **2**(8): p. 511-6.
558. Quinn, S.R. and L.A. O'Neill, *A trio of microRNAs that control Toll-like receptor signalling.* International Immunology, 2011. **23**(7): p. 421-425.
559. Murray, P.J., et al., *Macrophage activation and polarization: nomenclature and experimental guidelines.* Immunity, 2014. **41**(1): p. 14-20.
560. Yuan, W., et al., *Clinical significance and prognosis of serum tenascin-C in patients with sepsis.* BMC Anesthesiol, 2018. **18**(1): p. 170.
561. Schenk, S., et al., *Tenascin-C in serum: a questionable tumor marker.* Int J Cancer, 1995. **61**(4): p. 443-9.
562. Page, T.H., et al., *Raised circulating tenascin-C in rheumatoid arthritis.* Arthritis Res Ther, 2012. **14**(6): p. R260.
563. O'Connell, R.M., et al., *MicroRNA-155 is induced during the macrophage inflammatory response.* Proc Natl Acad Sci U S A, 2007. **104**(5): p. 1604-9.

564. Schildberger, A., et al., *Monocytes, peripheral blood mononuclear cells, and THP-1 cells exhibit different cytokine expression patterns following stimulation with lipopolysaccharide.* Mediators Inflamm, 2013. **2013**: p. 697972.
565. Wallner, K., et al., *Tenascin-C is expressed in macrophage-rich human coronary atherosclerotic plaque.* Circulation, 1999. **99**(10): p. 1284-9.
566. Tsang, J.S., et al., *Global analyses of human immune variation reveal baseline predictors of postvaccination responses.* Cell, 2014. **157**(2): p. 499-513.
567. Brodin, P., et al., *Variation in the human immune system is largely driven by non-heritable influences.* Cell, 2015. **160**(1-2): p. 37-47.
568. Avraham, R., et al., *Pathogen Cell-to-Cell Variability Drives Heterogeneity in Host Immune Responses.* Cell, 2015. **162**(6): p. 1309-21.
569. Wright, S.D., et al., *Lipopolysaccharide (LPS) binding protein opsonizes LPS-bearing particles for recognition by a novel receptor on macrophages.* J Exp Med, 1989. **170**(4): p. 1231-41.
570. Shimazu, R., et al., *MD-2, a molecule that confers lipopolysaccharide responsiveness on Toll-like receptor 4.* J Exp Med, 1999. **189**(11): p. 1777-82.
571. Nagai, Y., et al., *Essential role of MD-2 in LPS responsiveness and TLR4 distribution.* Nat Immunol, 2002. **3**(7): p. 667-72.
572. Saitoh, S., et al., *Ligand-dependent Toll-like receptor 4 (TLR4)-oligomerization is directly linked with TLR4-signaling.* J Endotoxin Res, 2004. **10**(4): p. 257-60.
573. Schaff, M., et al., *Novel function of tenascin-C, a matrix protein relevant to atherosclerosis, in platelet recruitment and activation under flow.* Arterioscler Thromb Vasc Biol, 2011. **31**(1): p. 117-24.
574. Corbi, A.L., et al., *The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.* J Biol Chem, 1988. **263**(25): p. 12403-11.
575. Jinek, M., et al., *A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.* Science, 2012. **337**(6096): p. 816-21.
576. Konermann, S., et al., *Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex.* Nature, 2015. **517**(7536): p. 583-8.
577. Xiong, K., et al., *RNA-Guided Activation of Pluripotency Genes in Human Fibroblasts.* Cell Reprogram, 2017. **19**(3): p. 189-198.
578. Franz, M., et al., *Serum levels of large tenascin-C variants, matrix metalloproteinase-9, and tissue inhibitors of matrix metalloproteinases in concentric versus eccentric left ventricular hypertrophy.* European journal of heart failure, 2009. **11**(11): p. 1057-1062.
579. Ishiwata, T., et al., *Serum tenascin-C as a potential predictive marker of angiogenesis in non-small cell lung cancer.* Anticancer Res, 2005. **25**(1B): p. 489-95.
580. Sato, A., et al., *Serum tenascin-C might be a novel predictor of left ventricular remodeling and prognosis after acute myocardial infarction.* J Am Coll Cardiol, 2006. **47**(11): p. 2319-25.
581. HAYFLICK, L. and P.S. MOORHEAD, *The serial cultivation of human diploid cell strains.* Exp Cell Res, 1961. **25**: p. 585-621.
582. Yu, G.L., et al., *In vivo alteration of telomere sequences and senescence caused by mutated Tetrahymena telomerase RNAs.* Nature, 1990. **344**(6262): p. 126-32.
583. Harley, C.B., A.B. Futcher, and C.W. Greider, *Telomeres shorten during ageing of human fibroblasts.* Nature, 1990. **345**(6274): p. 458-60.
584. Olovnikov, A., *Principle of marginotomy in the synthesis of polynucleotides at a template.* Doklady Biochem. and Biophys., 1971. **201**: p. 394-397.
585. Watson, J.D., *Origin of concatemeric T7 DNA.* Nat New Biol, 1972. **239**(94): p. 197-201.
586. Serra, V., et al., *Extracellular superoxide dismutase is a major antioxidant in human fibroblasts and slows telomere shortening.* J Biol Chem, 2003. **278**(9): p. 6824-30.
587. Saretzki, G., M.P. Murphy, and T. von Zglinicki, *MitoQ counteracts telomere shortening and elongates lifespan of fibroblasts under mild oxidative stress.* Aging Cell, 2003. **2**(2): p. 141-3.

588. Bodnar, A.G., et al., *Extension of life-span by introduction of telomerase into normal human cells.* Science, 1998. **279**(5349): p. 349-52.
589. Wright, W.E., et al., *Telomerase activity in human germline and embryonic tissues and cells.* Dev Genet, 1996. **18**(2): p. 173-9.
590. d'Adda di Fagagna, F., et al., *A DNA damage checkpoint response in telomere-initiated senescence.* Nature, 2003. **426**(6963): p. 194-8.
591. Gire, V., et al., *DNA damage checkpoint kinase Chk2 triggers replicative senescence.* EMBO J, 2004. **23**(13): p. 2554-63.
592. Herbig, U., et al., *Telomere shortening triggers senescence of human cells through a pathway involving ATM, p53, and p21(CIP1), but not p16(INK4a).* Mol Cell, 2004. **14**(4): p. 501-13.
593. Sedelnikova, O.A., et al., *Senescing human cells and ageing mice accumulate DNA lesions with unrepairable double-strand breaks.* Nat Cell Biol, 2004. **6**(2): p. 168-70.
594. Coppé, J.P., et al., *The Senescence-Associated Secretory Phenotype: The Dark Side of Tumor Suppression.* Annu Rev Pathol, 2010. **5**: p. 99-118.
595. Fumagalli, M., et al., *Telomeric DNA damage is irreparable and causes persistent DNA-damage-response activation.* Nat Cell Biol, 2012. **14**(4): p. 355-65.
596. Hewitt, G., et al., *Telomeres are favoured targets of a persistent DNA damage response in ageing and stress-induced senescence.* Nat Commun, 2012. **3**: p. 708.
597. Victorelli, S. and J.F. Passos, *Telomeres and Cell Senescence - Size Matters Not.* EBioMedicine, 2017. **21**: p. 14-20.
598. Nakamura, A.J., et al., *Both telomeric and non-telomeric DNA damage are determinants of mammalian cellular senescence.* Epigenetics Chromatin, 2008. **1**(1): p. 6.
599. Di Leonardo, A., et al., *DNA damage triggers a prolonged p53-dependent G1 arrest and long-term induction of Cip1 in normal human fibroblasts.* Genes Dev, 1994. **8**(21): p. 2540-51.
600. Chen, Z., et al., *Crucial role of p53-dependent cellular senescence in suppression of Pten-deficient tumorigenesis.* Nature, 2005. **436**(7051): p. 725-30.
601. Collado, M., et al., *Tumour biology: senescence in premalignant tumours.* Nature, 2005. **436**(7051): p. 642.
602. Braig, M., et al., *Oncogene-induced senescence as an initial barrier in lymphoma development.* Nature, 2005. **436**(7051): p. 660-5.
603. Coppé, J.P., et al., *Senescence-associated secretory phenotypes reveal cell-nonautonomous functions of oncogenic RAS and the p53 tumor suppressor.* PLoS Biol, 2008. **6**(12): p. 2853-68.
604. Kuilman, T., et al., *Oncogene-induced senescence relayed by an interleukin-dependent inflammatory network.* Cell, 2008. **133**(6): p. 1019-31.
605. Acosta, J.C., et al., *Chemokine signaling via the CXCR2 receptor reinforces senescence.* Cell, 2008. **133**(6): p. 1006-18.
606. West, M.D., et al., *Altered expression of plasminogen activator and plasminogen activator inhibitor during cellular senescence.* Exp Gerontol, 1996. **31**(1-2): p. 175-93.
607. Rasoamanantena, P., et al., *Altered regulation of fibronectin gene expression in Werner syndrome fibroblasts.* Exp Cell Res, 1994. **213**(1): p. 121-7.
608. Moujaber, O., et al., *Cellular senescence is associated with reorganization of the microtubule cytoskeleton.* Cell Mol Life Sci, 2019. **76**(6): p. 1169-1183.
609. Hernandez-Segura, A., J. Nehme, and M. Demaria, *Hallmarks of Cellular Senescence.* Trends Cell Biol, 2018. **28**(6): p. 436-453.
610. Cremer, T.J., et al., *MiR-155 induction by F. novicida but not the virulent F. tularensis results in SHIP down-regulation and enhanced pro-inflammatory cytokine response.* PloS one, 2009. **4**(12): p. e8508.
611. O'Connell, R.M., et al., *Inositol phosphatase SHIP1 is a primary target of miR-155.* Proc Natl Acad Sci U S A, 2009. **106**(17): p. 7113-8.
612. Ware, M.D., et al., *Cloning and characterization of human SHIP, the 145-kD inositol 5-phosphatase that associates with SHC after cytokine stimulation.* 1996.

613. Fu, Q., et al., *SHIP1 inhibits cell growth, migration, and invasion in non-small cell lung cancer through the PI3K/AKT pathway.* Oncology reports, 2019. **41**(4): p. 2337-2350.
614. Pedersen, I.M., et al., *Onco-miR-155 targets SHIP1 to promote TNFα-dependent growth of B cell lymphomas.* EMBO molecular medicine, 2009. **1**(5): p. 288-295.
615. Costinean, S., et al., *Src homology 2 domain-containing inositol-5-phosphatase and CCAAT enhancer-binding protein beta are targeted by miR-155 in B cells of Emicro-MiR-155 transgenic mice.* Blood, 2009. **114**(7): p. 1374-82.
616. Ehm, P., et al., *The tumor suppressor SHIP1 colocalizes in nucleolar cavities with p53 and components of PML nuclear bodies.* Nucleus, 2015. **6**(2): p. 154-164.
617. Androulidaki, A., et al., *The kinase Akt1 controls macrophage response to lipopolysaccharide by regulating microRNAs.* Immunity, 2009. **31**(2): p. 220-231.
618. Wu, X., et al., *miR-155 inhibits the formation of hypertrophic scar fibroblasts by targeting HIF-1α via PI3K/AKT pathway.* Journal of molecular histology, 2018. **49**(4): p. 377-387.
619. Yang, L., et al., *Acute downregulation of miR-155 leads to a reduced collagen synthesis through attenuating macrophages inflammatory factor secretion by targeting SHIP1.* Journal of molecular histology, 2018. **49**(2): p. 165-174.
620. Eissa, M.G. and C.M. Artlett, *The microRNA miR-155 is essential in fibrosis.* Non-coding RNA, 2019. **5**(1): p. 23.
621. Dudek, H., et al., *Regulation of neuronal survival by the serine-threonine protein kinase Akt.* Science, 1997. **275**(5300): p. 661-5.
622. Qin, L., et al., *ER stress negatively regulates AKT/TSC/mTOR pathway to enhance autophagy.* Autophagy, 2010. **6**(2): p. 239-47.
623. Nogueira, V., et al., *Akt determines replicative senescence and oxidative or oncogenic premature senescence and sensitizes cells to oxidative apoptosis.* Cancer cell, 2008. **14**(6): p. 458-470.
624. Zhang, X., et al., *Akt, FoxO and regulation of apoptosis.* Biochimica et Biophysica Acta (BBA)-Molecular Cell Research, 2011. **1813**(11): p. 1978-1986.
625. Medema, R.H., et al., *AFX-like Forkhead transcription factors mediate cell-cycle regulation by Ras and PKB through p27 kip1.* Nature, 2000. **404**(6779): p. 782-787.
626. Chapuis, N., et al., *IκB kinase overcomes PI3K/Akt and ERK/MAPK to control FOXO3a activity in acute myeloid leukemia.* Blood, The Journal of the American Society of Hematology, 2010. **116**(20): p. 4240-4250.
627. Serrano, M., *Shifting senescence into quiescence by turning up p53.* Cell Cycle, 2010. **9**(21): p. 4256-4257.
628. So, C.W. and M.L. Cleary, *MLL-AFX requires the transcriptional effector domains of AFX to transform myeloid progenitors and transdominantly interfere with forkhead protein function.* Molecular and cellular biology, 2002. **22**(18): p. 6542-6552.
629. Wang, F., et al., *Foxo3a aggravates inflammation and induces apoptosis in IL-1-treated rabbit chondrocytes via positively regulating tenascin-c.* Folia Histochem Cytobiol, 2020. **58**(1): p. 1-8.
630. Nho, R.S. and P. Hergert, *FoxO3a and disease progression.* World J Biol Chem, 2014. **5**(3): p. 346-54.
631. Ling, N., et al., *microRNA-155 regulates cell proliferation and invasion by targeting FOXO3a in glioma.* Oncol Rep, 2013. **30**(5): p. 2111-8.
632. Hemmings, B.A. and D.F. Restuccia, *The PI3K-PKB/Akt pathway.* Cold Spring Harb Perspect Biol, 2015. **7**(4).
633. Steelman, L.S., et al., *Roles of the Raf/MEK/ERK and PI3K/PTEN/Akt/mTOR pathways in controlling growth and sensitivity to therapy-implications for cancer and aging.* Aging (Albany NY), 2011. **3**(3): p. 192-222.
634. Cho, K.A., et al., *Morphological adjustment of senescent cells by modulating caveolin-1 status.* J Biol Chem, 2004. **279**(40): p. 42270-8.

635. Prager-Khoutorsky, M., et al., *Fibroblast polarization is a matrix-rigidity-dependent process controlled by focal adhesion mechanosensing.* Nat Cell Biol, 2011. **13**(12): p. 1457-65.
636. Schmoller, K.M. and J.M. Skotheim, *The Biosynthetic Basis of Cell Size Control.* Trends Cell Biol, 2015. **25**(12): p. 793-802.
637. Franke, W.W., *Discovering the molecular components of intercellular junctions--a historical view.* Cold Spring Harbor perspectives in biology, 2009. **1**(3): p. a003061-a003061.
638. Yonemura, S., et al., *Cell-to-cell adherens junction formation and actin filament organization: similarities and differences between non-polarized fibroblasts and polarized epithelial cells.* J Cell Sci, 1995. **108 (Pt 1)**: p. 127-42.
639. Ravikanth, M., et al., *Heterogenecity of fibroblasts.* J Oral Maxillofac Pathol, 2011. **15**(2): p. 247-50.
640. Narita, M., et al., *Rb-mediated heterochromatin formation and silencing of E2F target genes during cellular senescence.* Cell, 2003. **113**(6): p. 703-16.
641. Zhang, R., W. Chen, and P.D. Adams, *Molecular dissection of formation of senescence-associated heterochromatin foci.* Mol Cell Biol, 2007. **27**(6): p. 2343-58.
642. Kosar, M., et al., *Senescence-associated heterochromatin foci are dispensable for cellular senescence, occur in a cell type- and insult-dependent manner and follow expression of p16(ink4a).* Cell Cycle, 2011. **10**(3): p. 457-68.
643. Dimri, G.P., et al., *A biomarker that identifies senescent human cells in culture and in aging skin in vivo.* Proc Natl Acad Sci U S A, 1995. **92**(20): p. 9363-7.
644. Dmitrieva, N.I. and M.B. Burg, *High NaCl promotes cellular senescence.* Cell Cycle, 2007. **6**(24): p. 3108-13.
645. Paradis, V., et al., *Replicative senescence in normal liver, chronic hepatitis C, and hepatocellular carcinomas.* Hum Pathol, 2001. **32**(3): p. 327-32.
646. van der Loo, B., M.J. Fenton, and J.D. Erusalimsky, *Cytochemical detection of a senescence-associated beta-galactosidase in endothelial and smooth muscle cells from human and rabbit blood vessels.* Exp Cell Res, 1998. **241**(2): p. 309-15.
647. Hooten, N.N. and M.K. Evans, *Techniques to induce and quantify cellular senescence.* JoVE (Journal of Visualized Experiments), 2017(123): p. e55533.
648. Debacq-Chainiaux, F., et al., *Protocols to detect senescence-associated beta-galactosidase (SA-βgal) activity, a biomarker of senescent cells in culture and in vivo.* Nature protocols, 2009. **4**(12): p. 1798.
649. Frippiat, C., et al., *Subcytotoxic H2O2 stress triggers a release of transforming growth factor-β1, which induces biomarkers of cellular senescence of human diploid fibroblasts.* Journal of Biological Chemistry, 2001. **276**(4): p. 2531-2537.
650. Chen, Q. and B.N. Ames, *Senescence-like growth arrest induced by hydrogen peroxide in human diploid fibroblast F65 cells.* Proc Natl Acad Sci U S A, 1994. **91**(10): p. 4130-4.
651. Ortiz-Montero, P., A. Londoño-Vallejo, and J.P. Vernot, *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, 2017. **15**(1): p. 17.
652. Brosh, R., et al., *p53-repressed miRNAs are involved with E2F in a feed-forward loop promoting proliferation.* Mol Syst Biol, 2008. **4**: p. 229.
653. Bonifacio, L.N. and M.B. Jarstfer, *MiRNA profile associated with replicative senescence, extended cell culture, and ectopic telomerase expression in human foreskin fibroblasts.* PLoS One, 2010. **5**(9).
654. Noren Hooten, N., et al., *microRNA expression patterns reveal differential expression of target genes with age.* PLoS One, 2010. **5**(5): p. e10724.
655. Mahesh, G. and R. Biswas, *MicroRNA-155: A Master Regulator of Inflammation.* J Interferon Cytokine Res, 2019. **39**(6): p. 321-330.

656. Martyanov, V., M.L. Whitfield, and J. Varga, *Senescence Signature in Skin Biopsies From Systemic Sclerosis Patients Treated With Senolytic Therapy: Potential Predictor of Clinical Response?* Arthritis Rheumatol, 2019. **71**(10): p. 1766-1767.
657. Serrano, M., et al., *Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a.* Cell, 1997. **88**(5): p. 593-602.
658. Hunt, S.E., et al., *Ensembl variation resources.* Database (Oxford), 2018. **2018**.
659. Didry, D., M.F. Carlier, and D. Pantaloni, *Synergy between actin depolymerizing factor/cofilin and profilin in increasing actin filament turnover.* J Biol Chem, 1998. **273**(40): p. 25602-11.
660. Johnson, R. and G. Halder, *The two faces of Hippo: targeting the Hippo pathway for regenerative medicine and cancer treatment.* Nat Rev Drug Discov, 2014. **13**(1): p. 63-79.
661. Sticht, C., et al., *miRWalk: An online resource for prediction of microRNA binding sites.* PLoS One, 2018. **13**(10): p. e0206239.
662. Yarden, Y. and M.X. Sliwkowski, *Untangling the ErbB signalling network.* Nat Rev Mol Cell Biol, 2001. **2**(2): p. 127-37.
663. Cristofalo, V.J. and D. Kritchevsky, *Cell size and nucleic acid content in the diploid human cell line WI-38 during aging.* Med Exp Int J Exp Med, 1969. **19**(6): p. 313-20.
664. Sliogeryte, K. and N. Gavara, *Vimentin Plays a Crucial Role in Fibroblast Ageing by Regulating Biophysical Properties and Cell Migration.* Cells, 2019. **8**(10).
665. Hutter, E., et al., *Senescence-associated changes in respiration and oxidative phosphorylation in primary human fibroblasts.* Biochem J, 2004. **380**(Pt 3): p. 919-28.
666. Demidenko, Z.N. and M.V. Blagosklonny, *Growth stimulation leads to cellular senescence when the cell cycle is blocked.* Cell Cycle, 2008. **7**(21): p. 3355-61.
667. Miettinen, T.P. and M. Björklund, *Cellular Allometry of Mitochondrial Functionality Establishes the Optimal Cell Size.* Dev Cell, 2016. **39**(3): p. 370-382.
668. Neurohr, G.E., et al., *Excessive Cell Growth Causes Cytoplasm Dilution And Contributes to Senescence.* Cell, 2019. **176**(5): p. 1083-1097.e18.
669. Song, Y.S., B.Y. Lee, and E.S. Hwang, *Dinstinct ROS and biochemical profiles in cells undergoing DNA damage-induced senescence and apoptosis.* Mech Ageing Dev, 2005. **126**(5): p. 580-90.
670. Blagosklonny, M.V., *Cell senescence: hypertrophic arrest beyond the restriction point.* J Cell Physiol, 2006. **209**(3): p. 592-7.
671. Wang, Y., A. Meng, and D. Zhou, *Inhibition of phosphatidylinostol 3-kinase uncouples H2O2-induced senescent phenotype and cell cycle arrest in normal human diploid fibroblasts.* Exp Cell Res, 2004. **298**(1): p. 188-96.
672. Buchwalter, A. and M.W. Hetzer, *Nucleolar expansion and elevated protein translation in premature aging.* Nat Commun, 2017. **8**(1): p. 328.
673. Tiku, V., et al., *Small nucleoli are a cellular hallmark of longevity.* Nat Commun, 2017. **8**: p. 16083.
674. Mitsui, Y. and E.L. Schneider, *Increased nuclear sizes in senescent human diploid fibroblast cultures.* Exp Cell Res, 1976. **100**(1): p. 147-52.
675. Kobayashi, Y., et al., *Nuclear swelling occurs during premature senescence mediated by MAP kinases in normal human fibroblasts.* Biosci Biotechnol Biochem, 2008. **72**(4): p. 1122-5.
676. Gruenbaum, Y., et al., *The nuclear lamina comes of age.* Nat Rev Mol Cell Biol, 2005. **6**(1): p. 21-31.
677. Yoon, K.B., et al., *Induction of Nuclear Enlargement and Senescence by Sirtuin Inhibitors in Glioblastoma Cells.* Immune Netw, 2016. **16**(3): p. 183-8.
678. Chen, Q.M., et al., *Involvement of Rb family proteins, focal adhesion proteins and protein synthesis in senescent morphogenesis induced by hydrogen peroxide.* J Cell Sci, 2000. **113 (Pt 22)**: p. 4087-97.

679. Oender, K., et al., *Cytokeratin-related loss of cellular integrity is not a major driving force of human intrinsic skin aging.* Mechanisms of ageing and development, 2008. **129**(10): p. 563-571.
680. Wenk, M.B., K.S. Midwood, and J.E. Schwarzbauer, *Tenascin-C suppresses Rho activation.* J Cell Biol, 2000. **150**(4): p. 913-20.
681. Chen, X., et al., *Integrin-linked kinase induces both senescence-associated alterations and extracellular fibronectin assembly in aging cardiac fibroblasts.* J Gerontol A Biol Sci Med Sci, 2006. **61**(12): p. 1232-45.
682. Heng, Y.W. and C.G. Koh, *Actin cytoskeleton dynamics and the cell division cycle.* Int J Biochem Cell Biol, 2010. **42**(10): p. 1622-33.
683. Kudryashov, D.S., et al., *Cofilin cross-bridges adjacent actin protomers and replaces part of the longitudinal F-actin interface.* J Mol Biol, 2006. **358**(3): p. 785-97.
684. Andrianantoandro, E. and T.D. Pollard, *Mechanism of actin filament turnover by severing and nucleation at different concentrations of ADF/cofilin.* Mol Cell, 2006. **24**(1): p. 13-23.
685. Tsai, C.H., et al., *Up-regulation of cofilin-1 in cell senescence associates with morphological change and p27.* Aging Cell, 2021. **20**(1): p. e13288.
686. Zhou, K., et al., *Profilin 2 Promotes Proliferation and Metastasis of Head and Neck Cancer Cells by Regulating PI3K/AKT/β-Catenin Signaling Pathway.* Oncol Res, 2019. **27**(9): p. 1079-1088.
687. Pandey, P., et al., *Amyloid precursor-like protein 2 (APLP2) affects the actin cytoskeleton and increases pancreatic cancer growth and metastasis.* Oncotarget, 2015. **6**(4): p. 2064-75.
688. Molinie, N., et al., *Cortical branched actin determines cell cycle progression.* Cell Res, 2019. **29**(6): p. 432-445.
689. Maness, P.F. and R.C. Walsh, *Dihydrocytochalasin B disorganizes actin cytoarchitecture and inhibits initiation of DNA synthesis in 3T3 cells.* Cell, 1982. **30**(1): p. 253-62.
690. Gachet, Y., et al., *A MAP kinase-dependent actin checkpoint ensures proper spindle orientation in fission yeast.* Nature, 2001. **412**(6844): p. 352-5.
691. Blagosklonny, M.V., *Cell cycle arrest is not senescence.* Aging, 2011. **3**(2): p. 94-101.
692. Reinhard, M., T. Jarchau, and U. Walter, *Actin-based motility: stop and go with Ena/VASP proteins.* Trends Biochem Sci, 2001. **26**(4): p. 243-9.
693. Pan, D., *The hippo signaling pathway in development and cancer.* Dev Cell, 2010. **19**(4): p. 491-505.
694. Wada, K., et al., *Hippo pathway regulation by cell morphology and stress fibers.* Development, 2011. **138**(18): p. 3907-14.
695. Sun, Z., et al., *Tenascin-C promotes tumor cell migration and metastasis through integrin α9β1–mediated yap inhibition.* Cancer research, 2018. **78**(4): p. 950-961.
696. Dupont, S., et al., *Role of YAP/TAZ in mechanotransduction.* Nature, 2011. **474**(7350): p. 179-83.
697. Fu, L., et al., *Up-regulation of FOXD1 by YAP alleviates senescence and osteoarthritis.* PLoS Biol, 2019. **17**(4): p. e3000201.
698. Wu, C. and S. Dedhar, *Integrin-linked kinase (ILK) and its interactors: a new paradigm for the coupling of extracellular matrix to actin cytoskeleton and signaling complexes.* J Cell Biol, 2001. **155**(4): p. 505-10.
699. Severino, J., et al., *Is beta-galactosidase staining a marker of senescence in vitro and in vivo?* Exp Cell Res, 2000. **257**(1): p. 162-71.
700. Georgakopoulou, E.A., et al., *Specific lipofuscin staining as a novel biomarker to detect replicative and stress-induced senescence. A method applicable in cryo-preserved and archival tissues.* Aging (Albany NY), 2013. **5**(1): p. 37-50.
701. Yang, N.C. and M.L. Hu, *The limitations and validities of senescence associated-beta-galactosidase activity as an aging marker for human foreskin fibroblast Hs68 cells.* Exp Gerontol, 2005. **40**(10): p. 813-9.

702. Lee, B.Y., et al., *Senescence-associated beta-galactosidase is lysosomal beta-galactosidase.* Aging Cell, 2006. **5**(2): p. 187-95.
703. Kurz, D.J., et al., *Senescence-associated (beta)-galactosidase reflects an increase in lysosomal mass during replicative ageing of human endothelial cells.* J Cell Sci, 2000. **113 (Pt 20)**: p. 3613-22.
704. Lipetz, J. and V.J. Cristofalo, *Ultrastructural changes accompanying the aging of human diploid cells in culture.* J Ultrastruct Res, 1972. **39**(1): p. 43-56.
705. Kurz, T., et al., *Lysosomes and oxidative stress in aging and apoptosis.* Biochim Biophys Acta, 2008. **1780**(11): p. 1291-303.
706. Leshkowitz, D., et al., *Differences in microRNA detection levels are technology and sequence dependent.* RNA, 2013. **19**(4): p. 527-38.
707. Godoy, P.M., et al., *Comparison of Reproducibility, Accuracy, Sensitivity, and Specificity of miRNA Quantification Platforms.* Cell Rep, 2019. **29**(12): p. 4212-4222.e5.
708. Zalewska, A., et al., *Interleukin 6 and 8 levels in plasma and fibroblast cultures in psoriasis.* Mediators Inflamm, 2006. **2006**(1): p. 81767.
709. Yoo, J.K., et al., *Discovery and characterization of miRNA during cellular senescence in bone marrow-derived human mesenchymal stem cells.* Exp Gerontol, 2014. **58**: p. 139-45.
710. Dhahbi, J.M., et al., *Deep sequencing reveals novel microRNAs and regulation of microRNA expression during cell senescence.* PLoS One, 2011. **6**(5): p. e20509.
711. Wang, Y., et al., *MicroRNA regulation of ionizing radiation-induced premature senescence.* Int J Radiat Oncol Biol Phys, 2011. **81**(3): p. 839-48.
712. Bueno, M.J. and M. Malumbres, *MicroRNAs and the cell cycle.* Biochim Biophys Acta, 2011. **1812**(5): p. 592-601.
713. Sherr, C.J., *D-type cyclins.* Trends Biochem Sci, 1995. **20**(5): p. 187-90.
714. Rojas, P., et al., *Cyclin D2 and cyclin D3 play opposite roles in mouse skin carcinogenesis.* Oncogene, 2007. **26**(12): p. 1723-30.
715. Zhu, W., et al., *CCND2 Overexpression Enhances the Regenerative Potency of Human Induced Pluripotent Stem Cell-Derived Cardiomyocytes: Remuscularization of Injured Ventricle.* Circulation research, 2018. **122**(1): p. 88-96.
716. Koo, K.H. and H. Kwon, *MicroRNA miR-4779 suppresses tumor growth by inducing apoptosis and cell cycle arrest through direct targeting of PAK2 and CCND3.* Cell Death Dis, 2018. **9**(2): p. 77.
717. Feringa, F.M., et al., *Persistent repair intermediates induce senescence.* Nat Commun, 2018. **9**(1): p. 3923.
718. van den Berg, J., et al., *A limited number of double-strand DNA breaks is sufficient to delay cell cycle progression.* Nucleic Acids Res, 2018. **46**(19): p. 10132-10144.
719. Sternberg, S., et al., *DNA interrogation by the CRISPR RNA-guided endonuclease Cas9.* Nature, 2014. **507**(7490): p. 62-+.
720. Kurowska-Stolarska, M., et al., *MicroRNA-155 as a proinflammatory regulator in clinical and experimental arthritis.* Proc Natl Acad Sci U S A, 2011. **108**(27): p. 11193-8.
721. Mogensen, T., *Pathogen Recognition and Inflammatory Signaling in Innate Immune Defenses.* Clinical Microbiology Reviews, 2009. **22**(2): p. 240-+.
722. Janeway, C.A., Jr. and R. Medzhitov, *Innate immune recognition.* Annu Rev Immunol, 2002. **20**: p. 197-216.
723. Yee, D., et al., *MicroRNA-155 induction via TNF- and IFN- suppresses expression of programmed death ligand-1 (PD-L1) in human primary cells.* Journal of Biological Chemistry, 2017. **292**(50): p. 20683-20693.
724. Ruiz, C., et al., *Differential gene expression analysis reveals activation of growth promoting signaling pathways by tenascin-C.* Cancer research, 2004. **64**(20): p. 7377-7385.
725. Park, J.S., et al., *Expression of Tenascin-C is down-regulated during intrinsic skin aging.* Journal of Dermatological Science, 2017. **86**(2): p. e92-e93.

726. Mitchell, S.J., et al., *Animal models of aging research: implications for human aging and age-related diseases.* Annu Rev Anim Biosci, 2015. **3**: p. 283-303.
727. Okamura, N., et al., *Deficiency of tenascin-C delays articular cartilage repair in mice.* Osteoarthritis Cartilage, 2010. **18**(6): p. 839-48.
728. Gruber, B.L., et al., *Tenascin-C expression controls the maturation of articular cartilage in mice.* BMC Res Notes, 2020. **13**(1): p. 78.
729. Jenny, N.S., *Inflammation in aging: cause, effect, or both?* Discovery medicine, 2012. **13**(73): p. 451-460.
730. McGeer, E.G. and P.L. McGeer, *Inflammatory processes in Alzheimer's disease.* Prog Neuropsychopharmacol Biol Psychiatry, 2003. **27**(5): p. 741-9.
731. McGeer, P.L. and E.G. McGeer, *Inflammation and neurodegeneration in Parkinson's disease.* Parkinsonism & related disorders, 2004. **10**: p. S3-S7.
732. Singh, T. and A.B. Newman, *Inflammatory markers in population studies of aging.* Ageing Res Rev, 2011. **10**(3): p. 319-29.
733. Freund, A., et al., *Inflammatory networks during cellular senescence: causes and consequences.* Trends Mol Med, 2010. **16**(5): p. 238-46.
734. Liu, F., et al., *miR-132 inhibits lipopolysaccharide-induced inflammation in alveolar macrophages by the cholinergic anti-inflammatory pathway.* Exp Lung Res, 2015. **41**(5): p. 261-9.
735. Piccinini, A.M. and K.S. Midwood, *Cross-talk between endogenous danger signals and microRNAs: tenascin-C fine-tunes the immune response to infection via induction of miR-155.* International Journal of Molecular Medicine, 2013. **32**: p. S9-S9.
736. Tanenbaum, M.E., et al., *A protein-tagging system for signal amplification in gene expression and fluorescence imaging.* Cell, 2014. **159**(3): p. 635-46.
737. Mali, P., et al., *CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering.* Nat Biotechnol, 2013. **31**(9): p. 833-8.
738. Polstein, L.R. and C.A. Gersbach, *A light-inducible CRISPR-Cas9 system for control of endogenous gene activation.* Nat Chem Biol, 2015. **11**(3): p. 198-200.
739. Misra, S., et al., *Results of AlloDerm use in abdominal hernia repair.* Hernia, 2008. **12**(3): p. 247-50.
740. Wainwright, D.J., *Use of an acellular allograft dermal matrix (AlloDerm) in the management of full-thickness burns.* Burns, 1995. **21**(4): p. 243-8.

10 APPENDIX

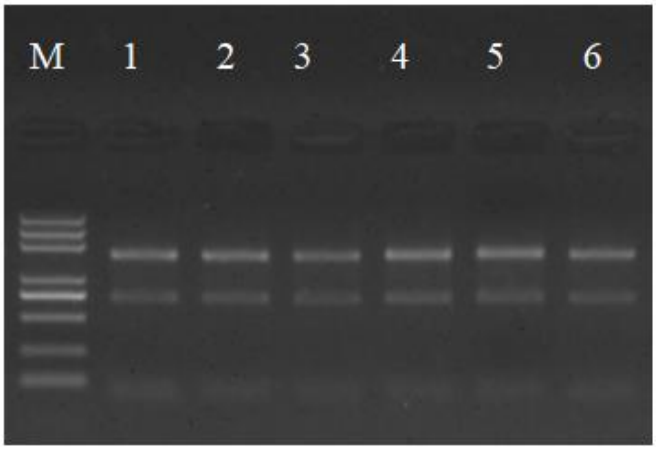


Figure S1: RNA samples passed the quality control check at Novogene as indicated by presence of 28S and 18S ribosomal RNA bands in agarose gel. Fibroblasts (WT and *TNC* KO) were seeded at 1 x 10^6 and cultured for 24 hourss before RNA extraction using ReliaPrep™ miRNA Cell and Tissue Kit (Promega). Quality and purity was assessed via bioanalyzer and agarose gel. M: 2K plus DNA ladder; 1-3, *TNC* KO RNA; 4-6 WT fibroblast RNA.

Table S1. Differentially expressed genes with a log2Fc change of + or - infinity that were removed from analysis. Of note is that the TCONS *TNC* transcript present in the KO cell line but not the wildtype is not translated into protein (as indicated on Ensemble). The 'ENST' *TNC* transcript is also not translated into protein based on a database search via ensemble. Significantly differentially expressed genes are highlighted in blue whilst differentially expressed genes that were not significant are highlighted in green.

Gene ID	Transcript ID	Cell line unable to detect in	log2Fcchange	P Value
TNC	TCONS_00026897	WT	+ Infinity	0.000458
ARNTL2	TCONS_00005878	WT	+ Infinity	0.001435
NDUFB8	ENST00000370322	WT	+ Infinity	0.00358
TMEM200A	ENST00000296978	KO	- Infinity	0.008995
TNC	ENST00000473855	KO	- Infinity	0.00901
STAT1	ENST00000673858	WT	+ Infinity	0.009057
R3HCC1	TCONS_00025160	KO	- Infinity	0.009078
IVD	ENST00000479013	KO	- Infinity	0.011842
SH3PXD2A	ENST00000355946	KO	- Infinity	0.016813
ATN1	TCONS_00005785	WT	+ Infinity	0.016954
MCL1	ENST00000307940	WT	+ Infinity	0.022977
HEATR3	TCONS_00010033	WT	+ Infinity	0.025176
RNF44	TCONS_00022451	WT	+ Infinity	0.027058

ATG2A	ENST00000418259	WT	+ Infinity	0.027248
MORF4L1	ENST00000559158	KO	- Infinity	0.03563
POGLUT2	ENST00000460338	WT	+ Infinity	0.039374
TNRC6B	TCONS_00018265	WT	+ Infinity	0.041786
RPL7A	ENST00000426651	KO	- Infinity	0.041786
PPP6R3	ENST00000527403	KO	- Infinity	0.045399
ATAD3A	TCONS_00000039	WT	+ Infinity	0.045622
TMF1	TCONS_00019869	WT	+ Infinity	0.047532
EHMT	ENST00000637748	WT	+ infinity	0.676716

Abbreviations: NADH: ubiquinone oxoreductase subunit B8 (*NDFUB8*), transmembrane protein 200A (*TMEM200A*), *TNC*, signal transducer and activator of transcription 1 (*STAT1*), Isovaleryl-CoA dehydrogenase (*IVD*), SH3 and PX domain- containing protein 2A (*SH3PXD2A*), induced myeloid leukemia cell differentiation protein 1 (*MCL1*), autophagy related 2A (*ATG2A*), mortality factor 4- like protein 1 (*MORF4L1)*, protein glucosyltranferase (*POGLUTU2*), 60S ribosomal protein L7a (*RPL7A*), protein phosphatase 6 regulatory subunit 3 *(PPP6R3*), euchromatic histone lysine methyltransferase 1 (*EHMT)*

Table S2. The full list of differentially expressed genes analysed by RNA-Seq from 'old' PDL 55 *TNC* KO fibroblasts and 'old' PDL 55 WT fibroblasts. Significantly differentially expressed genes are highlighted in blue whilst differentially expressed genes that were not significant are highlighted in green.

Gene Name	Transcript	KO expression value	WT expression value	log2Fc change	padjust
APLP2	ENST00000278756	0.035360667	81.21884133	-11.1655	0.00136
MED12	TCONS_00028277	0.03047	6.008548667	-7.62348	0.00358
SF3B1	ENST00000424674	0.899719333	1.772337333	-0.97811	0.00358
TIMP3	ENST00000266085	790.7693683	1049.554688	-0.40845	0.004028
TIMP3	ENST00000266085	790.7693683	1049.554688	-0.40845	0.004028
DCHS1	ENST00000299441	1.144794333	3.075705	-1.42583	0.004662
CFL2	ENST00000341223	6.214281	0.012908667	8.911104	0.005781
SERPINB2	ENST00000299502	0.820780333	27.27901333	-5.05465	0.008399
ANTXR1	TCONS_00005806	0.026289667	2.103032333	-6.32183	0.008768
MOXD1	ENST00000367963	8.681228667	16.31956667	-0.91063	0.008806
U52112.1	TCONS_00028464	2.032287	5.655420333	-1.47653	0.008995
MED12	ENST00000374080	6.371052	0.108896667	5.8705	0.008995
A2M	ENST00000318602	1.488917667	49.72663867	-5.06168	0.009057
BX322234.1	TCONS_00023096	12.79617667	8.750404	0.548291	0.009674
FP236383.5	TCONS_00017889	8.241338	0.004217333	10.93233	0.009674
SSNA1	ENST00000322310	0.007773333	15.443919	-10.9562	0.009859
MMP10	ENST00000279441	1.79701	5.360904333	-1.57688	0.010811
SERPINE1	ENST00000223095	736.508667	1154.421834	-0.6484	0.011842
TCP11L1	TCONS_00004176	0.172353333	2.519402667	-3.86964	0.011842

SERPINE1	ENST00000223095	736.508667	1154.421834	-0.6484	0.011842
ARSB	ENST00000264914	14.43130467	17.68106767	-0.293	0.011842
CLSTN2	ENST00000458420	0.028914667	1.877084667	-6.02055	0.01305
CPXM2	ENST00000241305	0.299386	4.774225	-3.99519	0.013987
CCN2	ENST00000367976	380.125061	553.6201173	-0.54242	0.013987
BUB1B	ENST00000287598	3.92424	6.473833333	-0.72221	0.01485
ARHGEF2	ENST00000313667	4.008845	19.54632467	-2.28564	0.015028
STMN2	ENST00000220876	30.91396467	8.269124667	1.902452	0.015504
VDAC3	ENST00000022615	25.68986867	29.45714933	-0.19742	0.015504
STMN2	ENST00000220876	30.91396467	8.269124667	1.902452	0.015504
LAMB2	ENST00000488638	5.004496667	0.236895	4.400905	0.015504
XYLT1	ENST00000261381	17.674349	36.17425533	-1.03331	0.016813
TMSB4X	ENST00000380636	3.876360333	1.890644667	1.035824	0.016813
TOP2A	ENST00000423485	27.96245467	42.78199767	-0.61351	0.016954
ITIH5	ENST00000397146	0.086137333	3.485636667	-5.33864	0.016954
TOP2A	ENST00000423485	27.96245467	42.78199767	-0.61351	0.016954
FCL	TCONS_00027973	7.638325333	11.12051733	-0.5419	0.016954
TCP11L1	TCONS_00004174	2.209423667	0.010709667	7.688613	0.016954
IQGAP3	ENST00000361170	8.901335	13.48351667	-0.5991	0.018014
GHDC	ENST00000587427	0.004967	4.076355667	-9.68069	0.01884
AL133346.1	TCONS_00022983	36.041167	58.30300533	-0.69392	0.018956
TGM2	ENST00000361475	133.8162077	182.705694	-0.44927	0.019442
IGFBP2	ENST00000233809	0.206825667	2.450803	-3.56677	0.020424
STMN1	ENST00000399728	0.195192667	17.12027933	-6.45466	0.022379
COL5A3	ENST00000264828	19.64730567	28.56118767	-0.53972	0.022426
TPX2	ENST00000300403	18.54521233	27.89378333	-0.5889	0.022426
RARG	ENST00000338561	3.680840667	23.20778267	-2.6565	0.022426
NRIP3	ENST00000309166	2.426647333	4.942637	-1.02632	0.022426
TPX2	ENST00000300403	18.54521233	27.89378333	-0.5889	0.022426
COL5A3	ENST00000264828	19.64730567	28.56118767	-0.53972	0.022426
MXRA5	ENST00000217939	29.88956333	18.911256	0.660397	0.022426
RNF149	ENST00000295317	9.760182667	0.517007667	4.238651	0.022426
ATP6V1F	TCONS_00024256	13.93942267	0.391745	5.153112	0.02261
LAMC1	ENST00000258341	262.9870047	220.1878357	0.256257	0.022964
HAUS7	ENST00000370210	2.889733667	0.160250333	4.172537	0.022977
MCM6	ENST00000264156	15.20827	20.33178267	-0.41888	0.022977
SPARC	ENST00000231061	257.818573	314.0782573	-0.28477	0.022977
HEBP1	ENST00000647702	1.449914667	2.714488	-0.90471	0.022977
MCM6	ENST00000264156	15.20827	20.33178267	-0.41888	0.022977
RELL1	ENST00000454158	10.57137933	13.270648	-0.32808	0.022977
SPARC	ENST00000231061	257.818573	314.0782573	-0.28477	0.022977
HAUS7	ENST00000370210	2.889733667	0.160250333	4.172537	0.022977
SEPTIN7	TCONS_00023850	15.34184267	3.33E-07	25.45593	0.022977
ARID5B	ENST00000279873	18.806921	15.72537233	0.258169	0.023012
TSPAN13	ENST00000262067	30.10658967	41.822862	-0.47421	0.023595
CDT1	ENST00000562747	2.438615333	8.593848	-1.81724	0.024687
NHLRC3	ENST00000379600	3.718002333	0.870755	2.094189	0.024687

FZD7	TCONS_00015708	4.906171333	0.045208	6.761876	0.026751
GHITM	ENST00000372134	50.53476833	47.71887867	0.082716	0.027484
GAL	ENST00000265643	32.48367633	23.92466867	0.441216	0.027484
MX1	ENST00000398598	2.539043333	0.015646667	7.342286	0.030119
PHLDB1	TCONS_00004760	0.017393	3.878927333	-7.80101	0.030774
KRT86	TCONS_00006011	15.49385433	6.921130333	1.162617	0.030774
L1CAM	ENST00000370055	13.868534	28.714385	-1.04996	0.031404
SPARC	ENST00000520687	855.9222617	1087.021484	-0.34483	0.031404
ANTXR1	ENST00000409829	5.21736	16.78714833	-1.68597	0.033279
GDI2	ENST00000380191	29.48161	73.48571267	-1.31765	0.033279
SRGN	ENST00000242465	25.48369867	48.65514233	-0.93302	0.033279
DCAF12	ENST00000361264	10.862107	12.79670567	-0.23647	0.033279
SLIT3	ENST00000519560	15.40511167	8.322653333	0.888294	0.033279
NNT	ENST00000344920	5.866643333	0.722347333	3.021771	0.033279
MCRS1	ENST00000552596	1.439380333	3.963542667	-1.46134	0.03376
KANK2	TCONS_00014307	4.522997	0.285451333	3.985962	0.03376
TAF6	ENST00000344095	11.84943733	0.198555667	5.899131	0.03376
CCNL2	ENST00000418865	5.033604667	17.071235	-1.7619	0.03599
AC087623.1	TCONS_00025231	1.577419667	0.120350333	3.712254	0.036552
DPAGT1	ENST00000472016	0.501716333	7.035550333	-3.80972	0.036624
AF196969.1	TCONS_00028207	6.733628667	0.030459667	7.78834	0.036624
DHX15	ENST00000336812	22.090804	27.37042133	-0.30917	0.03768
FRAS1	ENST00000512123	4.755198333	2.272974333	1.064924	0.03768
BOLA3	ENST00000327428	5.624423333	10.54544833	-0.90684	0.038623
PPP6R3	ENST00000393800	4.069075333	0.460722	3.142733	0.038623
CDCA4	ENST00000336219	9.155248	15.74128267	-0.78188	0.038983
OXA1L	ENST00000481218	1.260828	9.389360333	-2.89666	0.039374
ENO1	ENST00000464920	202.8523713	254.274297	-0.32596	0.041154
HELLS	TCONS_00003328	5.733174	8.610898667	-0.58683	0.041154
AC005746.2	TCONS_00012419	0.16012	2.412920667	-3.91356	0.041154
PDGFRA	ENST00000257290	4.503236	9.634088667	-1.09719	0.041154
HELLS	TCONS_00003328	5.733174	8.610898667	-0.58683	0.041154
1B14	TCONS_00027831	17.91165567	26.196179	-0.54846	0.041154
ENO1	ENST00000464920	202.8523713	254.274297	-0.32596	0.041154
H2AX	ENST00000530167	34.69038367	48.780224	-0.49176	0.041226
MT-CO3	ENST00000362079	342.9992777	149.1812897	1.201139	0.041226
GMPPA	ENST00000313597	0.753229	2.970793667	-1.97969	0.041613
VMP1	ENST00000592790	11.07388433	3.062174333	1.854533	0.041613
ENAH	TCONS_00002868	7.879034333	1.995371667	1.981361	0.041786
R3HCC1	TCONS_00025161	39.119559	6.248232	2.64637	0.041786
GINS2	ENST00000253462	3.756146	5.936635333	-0.66039	0.042434
MYBL2	ENST00000217026	7.831147333	15.43780667	-0.97917	0.042434
SLIT2	ENST00000504154	19.69478467	13.21229733	0.575932	0.042434
MYBL2	ENST00000217026	7.831147333	15.43780667	-0.97917	0.042434
ARF4	ENST00000486310	23.33136433	43.63845567	-0.90333	0.042434
GINS2	ENST00000253462	3.756146	5.936635333	-0.66039	0.042434
CD44-AS1	TCONS_00005072	9.496160667	13.125652	-0.46697	0.042434

1B14	TCONS_00027856	101.27345	130.1542103	-0.36197	0.042434
NCS1	ENST00000372398	16.02104567	19.72212867	-0.29985	0.042434
DKK1	ENST00000373970	285.801229	242.6014407	0.236424	0.042434
HEXIM1	ENST00000332499	16.558945	13.356303	0.31009	0.042434
ARPC3	ENST00000228825	51.235462	41.04509367	0.319933	0.042434
LTA4H	ENST00000228740	16.84235	13.274071	0.343483	0.042434
THY1	ENST00000284240	118.4060387	85.54394767	0.469005	0.042434
SLIT2	ENST00000504154	19.69478467	13.21229733	0.575932	0.042434
GTSE1	ENST00000454366	5.912902333	9.118908	-0.62499	0.04423
OLFM2	ENST00000264833	3.809542	8.782147333	-1.20496	0.04423
GTSE1	ENST00000454366	5.912902333	9.118908	-0.62499	0.04423
SOCS3	ENST00000330871	5.788808333	7.988723333	-0.4647	0.04423
SH3BGRL3	ENST00000270792	210.7969613	160.5335747	0.392979	0.04423
CYB5D2	TCONS_00011037	2.473111333	0.298904333	3.048571	0.04423
RTN4	ENST00000337526	24.94967467	0.015870333	10.61847	0.04423
SGCD	ENST00000435422	5.14818	7.503138667	-0.54343	0.044455
CRTAP	ENST00000320954	32.87897233	29.00426533	0.1809	0.044455
NCAPG	ENST00000251496	6.536203667	11.833014	-0.85629	0.045109
NTM	ENST00000470371	7.938304	6.055331333	0.390625	0.045109
MYBL1	ENST00000522677	4.812037	8.570806333	-0.83278	0.045143
VGLL3	ENST00000398399	8.188533	10.74641267	-0.39218	0.045143
MT1E	ENST00000306061	3.479061	8.064820333	-1.21294	0.045413
PFN2	ENST00000452853	3.674418333	8.44783	-1.20106	0.045413
PRRX2	ENST00000372469	32.82476667	28.30644533	0.213654	0.045413
MT-ND3	ENST00000361227	71.74133567	35.76980567	1.004062	0.045413
RFC3	ENST00000380071	5.465439667	9.964194333	-0.86642	0.045622
TINAGL1	ENST00000271064	0.230192	3.684407	-4.00052	0.045622
CHRM2	TCONS_00024288	0.527966	3.309175667	-2.64795	0.045622
NHLRC3	ENST00000379599	0.564437	2.864450333	-2.34337	0.045622
RFC3	ENST00000380071	5.465439667	9.964194333	-0.86642	0.045622
SLC8A1	ENST00000406785	8.425230667	10.64086933	-0.33683	0.045622
ADAMTS1	ENST00000284984	90.24699433	64.50590667	0.484448	0.045622
KRT81	TCONS_00006854	18.89250167	6.722913	1.490655	0.045622
ARHGEF2	TCONS_00002579	6.321931667	1.989934	1.667645	0.045622
AC013451.1	TCONS_00008112	3.603819667	5.491475333	-0.60767	0.045669
DPP7	ENST00000491807	5.191917333	2.708207	0.938929	0.045669
BUB3	ENST00000368865	6.098169	8.175317667	-0.4229	0.045757
GDI2	ENST00000380181	24.71292367	2.057105	3.586578	0.045757
STAC	ENST00000273183	5.394262	1.667225667	1.693976	0.047718
UBASH3B	ENST00000284273	7.515005667	12.10276233	-0.68749	0.047732
FBXW5	ENST00000487794	4.974275	1.843772667	1.431825	0.047732
CCND2	ENST00000261254	23.15497967	13.421145	0.786815	0.048522
SCD	ENST00000370355	15.444299	6.956150333	1.150713	0.048522
CYB5D2	ENST00000301391	0.386306333	6.457658667	-4.06319	0.074224
H2BC17	ENST00000616182	82.735316	137.1078897	-0.72874	0.07547
JUN	ENST00000371222	27.81082233	36.46738333	-0.39096	0.075689
FOS	ENST00000555347	3.361864667	1.454536333	1.208702	0.075811

LAMB2	ENST00000305544	116.9459587	104.5488843	0.161664	0.076561
RPS6KA2	ENST00000265678	18.45143367	14.16877533	0.381018	0.084446
CDK6	ENST00000424848	16.339913	26.595735	-0.70279	0.089288
SPIN4	ENST00000374884	3.034630333	4.895684667	-0.68999	0.090052
MMP1	ENST00000315274	118.5888643	235.4767457	-0.98962	0.106613
MT-CYB	ENST00000361789	254.0102537	154.7837523	0.714633	0.111477
CCND1	ENST00000227507	665.3713783	755.7168987	-0.18369	0.115192
H2BC14	ENST00000621112	35.59838767	47.92624567	-0.429	0.125603
GDF15	ENST00000252809	41.543204	24.71776267	0.749064	0.136457
CDK2	ENST00000266970	6.181659333	11.114636	-0.84639	0.13894
CCNA2	ENST00000618014	0.175954667	2.727542667	-3.95433	0.147479
ANAPC11	ENST00000579978	0.715421333	3.472501	-2.27911	0.148316
H2AC14	ENST00000333151	85.32386	132.1408237	-0.63106	0.176915
H2AC4	ENST00000615868	40.33872067	57.294337	-0.50623	0.250145
H4C1	ENST00000617569	44.08044567	58.21684267	-0.4013	0.254144
UBE2C	ENST00000356455	15.260974	24.343668	-0.6737	0.256277
SCD5	ENST00000319540	27.486379	30.64283567	-0.15683	0.258339
CDKN2C	ENST00000371761	2.647852333	0.77913	1.764887	0.275955
UFM1	ENST00000239878	20.57287533	23.17754367	-0.17198	0.283921
H3C1	ENST00000613854	58.99547967	80.59188833	-0.45003	0.285422
UBE2S	ENST00000264552	15.089822	21.161464	-0.48786	0.290766
H2BC13	ENST00000377401	37.887587	54.08668767	-0.51355	0.307284
IGFBP4	ENST00000269593	362.620524	405.5375267	-0.16137	0.331858
H2BC11	ENST00000339812	52.34360233	90.148046	-0.78428	0.332956
H2BC12	ENST00000356950	55.115012	78.36138433	-0.5077	0.37565
H2BC3	ENST00000615966	36.27053567	51.14742267	-0.49586	0.378574
MAPK1	ENST00000215832	29.25100633	31.229117	-0.09441	0.435515
H2AC20	ENST00000331380	102.80425	126.587158	-0.30023	0.43988
CHMP5	ENST00000223500	16.112825	13.42038533	0.263783	0.443167
H2AC7	ENST00000341023	34.37885033	51.95480467	-0.59574	0.463653
H3C15	ENST00000403683	35.134604	49.28841933	-0.48836	0.473228
ANTXR1	ENST00000303714	32.434005	39.36736533	-0.27949	0.473228
CDT1	ENST00000301019	7.234608	10.45539233	-0.53126	0.477309
H2AZ1	ENST00000296417	34.11189	49.21174733	-0.52873	0.491872
ZNHIT1	ENST00000485387	6.158212	2.116575333	1.54078	0.535762
H2AC18	ENST00000369159	163.8735097	199.1688437	-0.28141	0.536715
SCOC	ENST00000394201	17.872443	20.852708	-0.2225	0.569023
CDKN2D	ENST00000393599	3.779732333	5.687747333	-0.58957	0.57392
LAMB1	ENST00000222399	22.00279833	0.560837	5.293962	0.57392
CEBPB	ENST00000303004	54.121035	49.477773	0.129409	0.621187
UBB	ENST00000302182	41.82481	18.621229	1.167411	0.655749
STAG1	ENST00000383202	12.27562533	14.67007133	-0.25708	0.675729
UBA52	ENST00000442744	25.138375	20.159956	0.318399	0.693343
DYNLT3	ENST00000378578	10.66756467	11.921318	-0.16031	0.705122
IGFBP3	ENST00000381083	85.402497	228.8877003	-1.42229	0.711935
IGFBP7	ENST00000295666	182.2930657	213.9906363	-0.23129	0.73677
CDC16	ENST00000356221	1.293377667	4.252031	-1.71701	0.740562

H3-3A	ENST00000366815	13.63940433	8.903672333	0.615308	0.793229
SLC10A3	ENST00000369649	1.468555333	3.507119333	-1.25589	0.80248
FAM214B	ENST00000378557	8.805070333	9.834882333	-0.15957	0.80248
ZBTB7A	ENST00000322357	10.19261967	10.658548	-0.06449	0.80248
KLC1	ENST00000389744	5.138877333	3.739073333	0.458772	0.80248
NFKB1	ENST00000226574	1.007541	2.193786667	-1.12258	0.80248
H2BC15	ENST00000612898	10.419884	14.76046767	-0.5024	0.80248
STAT3	ENST00000264657	22.80284867	21.89814967	0.058405	0.80248
RPS27A	ENST00000272317	31.939839	18.35681333	0.799041	0.80248
TIMP1	ENST00000218388	391.094177	354.1385497	0.143202	0.80248
HMGB1	ENST00000339872	4.550266667	0.390317333	3.543232	0.80248
STAT1	ENST00000361099	1.366189	2.572162667	-0.91282	0.80248
TOLLIP	ENST00000317204	11.83123833	8.584935333	0.462722	0.803124
H2BC5	ENST00000289316	6.021840333	6.813114333	-0.17811	0.803917
ZC3H4	ENST00000601973	3.950988667	2.756689333	0.519277	0.809106
GLB1	ENST00000307363	19.263877	15.18789833	0.342976	0.826358
H2BC21	ENST00000369155	14.989951	15.916514	-0.08653	0.83544
SLC16A3	ENST00000582743	21.027877	16.11675967	0.383741	0.838194
H2AJ	ENST00000544848	15.139337	12.02903133	0.331782	0.840339
UBE2E1	ENST00000346855	0.496841333	1.496138	-1.59039	0.856396
FZR1	ENST00000441788	2.628733	2.193158333	0.261358	0.856882
UBC	ENST00000339647	52.005069	54.49474967	-0.06746	0.87127
CDKN1B	ENST00000228872	13.82271333	13.35666833	0.049481	0.874404
H2AC6	ENST00000602637	2.174131333	1.734414	0.325991	0.875575
CDKN1A	ENST00000244741	91.52600333	132.7973683	-0.53697	0.875737
TIMP2	ENST00000262768	1.585787667	6.356878333	-2.00312	0.87591
ARHGAP35	ENST00000404338	16.87323367	18.86782567	-0.16119	0.883306
ADPGK	ENST00000562621	1.525796667	2.110077	-0.46773	0.883742
POFUT2	ENST00000612472	4.823609333	7.453908667	-0.62788	0.890639
TAF13	ENST00000338366	33.44970333	31.675052	0.078646	0.902692
ANAPC7	ENST00000455511	4.694503333	4.950608667	-0.07663	0.906694
CDK4	ENST00000257904	20.91852833	21.53343067	-0.0418	0.912629
H2AZ2	ENST00000308153	7.594718	8.772123333	-0.20793	0.916005
BCL2L2	ENST00000557579	1.716984667	1.17517	0.547008	0.949118
ANAPC5	ENST00000541887	3.574641333	3.212155667	0.154257	0.965578
H2BC4	ENST00000314332	42.5131	54.94546433	-0.37009	0.966388
H2BC9	ENST00000619466	4.360334	5.280552333	-0.27625	0.978839
RHNO1	ENST00000623153	5.828676	7.329760667	-0.3306	0.982127
ARF4	ENST00000489843	2.966427333	4.149986667	-0.48438	0.982127
IGFBP5	ENST00000233813	623.4568887	624.863342	-0.00325	0.985028
STC1	ENST00000290271	18.891791	18.78578867	0.008118	0.995171
DDA1	ENST00000359866	2.106256333	2.037208	0.048088	0.995709

11 PUBLICATIONS

Immunology
The Journal of cells, molecules, systems and technologies

British Society for immunology

IMMUNOLOGY REVIEW ARTICLE

A complex interplay between the extracellular matrix and the innate immune response to microbial pathogens

Hannah Tomlin and Anna M. Piccinini

School of Pharmacy, University of Nottingham, Nottingham, UK

doi:10.1111/imm.12972
Received 29 March 2018; revised 26 April 2018; accepted 1 June 2018.
Correspondence: Anna M. Piccinini, Pharmacy School Building, East Drive, University Park, Nottingham NG7 2RD, UK.
Email: anna.piccinini@nottingham.ac.uk
Senior author: Anna M. Piccinini

Summary

The role of the host extracellular matrix (ECM) in infection tends to be neglected. However, the complex interactions between invading pathogens, host tissues and immune cells occur in the context of the ECM. On the pathogen side, a variety of surface and secreted molecules, including microbial surface components recognizing adhesive matrix molecules and tissue-degrading enzymes, are employed that interact with different ECM proteins to effectively establish an infection at specific sites. Microbial pathogens can also hijack or misuse host proteolytic systems to modify the ECM, evade immune responses or process biologically active mole- cules such as cell surface receptors and cytokines that direct cell beha- viour and immune defence. On the host side, the ECM composition and three-dimensional ultrastructure undergo significant modifications, which have a profound impact on the specific signals that the ECM conveys to immune cells at the forefront of infection. Unexpectedly, activated immune cells participate in the remodelling of the local ECM by synthe- sizing ECM glycoproteins, proteoglycans and collagen molecules. The close interplay between the ECM and the innate immune response to microbial pathogens ultimately affects the outcome of infection. This review explores and discusses recent data that implicate an active role for the ECM in the immune response to infection, encompassing antimicro- bial activities, microbial recognition, macrophage activation, phagocytosis, leucocyte population balance, and transcriptional and post-transcriptional regulation of inflammatory networks, and may foster novel antimicrobial approaches.

Keywords: extracellular matrix; immunity; infection.

Important questions to emerge from these studies include (i) does the ECM play an active role in infection rather than simply providing a scaffold for bacterial adhe- sion or being a barrier to breach; (ii) which ECM–patho- gen interactions significantly impact the ability of pathogens to colonize and invade host tissues, and/or bypass host defences, and influence how host cells respond to pathogens; and (iii) can we interfere with these interactions to develop new antimicrobial strategies or improve existing ones. Here, we briefly review the ECM and its interactions with microbial pathogens, and discuss the evidence for a direct implication of the ECM in the immune response to infection.

The ECM: an immunological perspective

Most infections and the resulting immune responses are tissue-specific. Each tissue has distinct ECM signatures that arise from complex and dynamic combinations of up to ~ 300 different proteins in varying concentrations and geometries within the three-dimensional extracellular space. This ever increasing number of proteins that con- tribute to matrices (the 'matrisome') include approxi- mately 43 collagen subunits, 36 proteoglycans and 200 glycoproteins.[6] ECM molecular multiplicity and complex- ity are amplified by post-translational modifications, which can unveil cryptic epitopes or generate new ones capable of triggering immune reactions, and alternative splicing, which generates several isoforms with distinct functions.[7,8] For instance, laminin *a*4 and *a*5 isoforms in basement membranes contribute to immune cell subtype selectivity during leucocyte recruitment to sites of inflammation.[9,10]

Interstitial ECMs, loose fibril-like matrices that fill the tissue stroma, and basement membranes, laminar sheets that anchor cell layers to underlying tissues, are basic forms of ECM. Additionally, specialized reticular fibre networks combining features of these two ECM structures are found in secondary lymphoid organs.[11,12]

The ECM is a highly dynamic yet strictly regulated tis- sue component. Its composition and normal function are determined and maintained by a fine balance between ECM synthesis, orchestrated by cytokines such as trans- forming growth factor-*b* (TGF-*b*), and turnover, accom- plished by matrix metalloproteinases (MMPs), a disintegrin and metalloproteinases (ADAMs) and a disin- tegrin and metalloproteinases with thrombospondin motifs (ADAMTSs), whose activity is controlled by tissue inhibitors of metalloproteinases (TIMPs).[13] During injury and infection, these enzymes are produced by activated immune cells such as monocytes/macrophages and pro- mote immune cell migration into sites of infection and tissue damage, and affect their ability to mount inflam- matory responses.[14]

The central dogma of matrix biology describes how the ECM provides structural support for cells and contributes to the unique structure of a tissue. However, this is only one function of the ECM. Host immune responses are carried out in the context of the ECM, so when immune cells contact the ECM, they receive vital instructions for their survival, proliferation, differentiation and activation, in addition to support for adhesion and guidance for migration. Several mechanisms are responsible for the communication between the ECM and cells. One such mechanism is signalling through cell surface adhesion molecules and receptors, including integrins and dis- coidin-domain receptors (reviewed in refs 15,16). Others involve binding, storage, activation and release of secreted molecules with potent immunomodulatory activity, including cytokines, chemokines and growth factors.[17,18] For instance, TGF-*b* has been shown *in vitro* and *in vivo* to affect activation, proliferation and differentiation of most immune cell types, modulating nearly all stages of the immune response.[19] Moreover, bioactive ECM frag- ments generated by tissue destruction (e.g. hyaluronan and heparan sulphate oligosaccharides) and ECM mole- cules whose expression is specifically induced upon tissue injury (e.g. fibronectin isoforms containing extra domain A, versican and biglycan) promote inflammation by inducing pro-inflammatory gene expression or exhibiting chemoattractant properties (reviewed in refs 20,21). These molecules form a class of endogenous damage-associated molecular patterns,[22] which, by activating pattern recog- nition receptors (PRRs) such as toll-like receptors (TLRs) in immune cells (e.g. macrophages and dendritic cells) and non-immune cells (e.g. fibroblasts and epithelial cells) alert the immune system to tissue damage and infection, which initiates not only pathogen clearance, but also tissue repair.[23] Hence, the ECM and the immune sys- tem are intertwined: signals from the ECM help to coor- dinate immune responses and, in turn, immune cells promote ECM repair and regeneration through the release of cytokines such as tumour necrosis factor (TNF), interferon-*c* (IFN-*c*) and TGF-*b*, which regulate the expression of many ECM molecules.

ECM–pathogen interactions

A tale of adhesion and colonization of the host

Tissue adherence represents the first essential mechanism that bacteria adopt to colonize the host. Failure to do so results in the organism being removed by physiological cleansing systems at sites of entry. The surface of host cells and the ECM are negatively charged and microbes employ a number of physico-chemical forces to overcome these repulsive forces and establish interactions between a microbial ligand or 'adhesin' and a complementary mole- cule or 'receptor' on the host tissue. These interactions are highly specific, contributing to tissue tropism, species specificity and genetic specificity within a species.

Bacterial adhesins are components of capsules, cell walls, pili or fimbriae, and those that bind to ECM components are called 'microbial surface components recognizing adhesive matrix molecules'. Host receptors are usually glycoproteins found on the cell membrane and ECM components.[24]

The binding of microbial pathogens to specific ECM proteins has been extensively reviewed elsewhere[1,3,25,26] and will be briefly discussed here. The first report of a host ECM–pathogen interaction was published in 1978 by Kuusela *et al.*[2] who studied the binding of *Staphylococcus aureus* to fibronectin and, a few years later, revealed two separate binding sites on fibronectin, in the N-terminus and C-terminus, respectively.[27] Other groups further characterized the interaction of this ubiquitous and promiscuous ECM glycoprotein with *Staphylococcus aur- eus* and *Streptococcus pyogenes*, which has a remarkably large number of fibronectin binding adhesins.[28–33] Several other ECM proteins such as laminin, collagen, heparan and chondroitin sulphate, vitronectin, thrombospondin, elastin, bone sialoprotein and tenascin-C, have since been implicated in specific interactions with pathogenic bacte- ria[26,34] (Table 1).

Early studies were mostly limited by the use of binding assays, involving bacterial cells or recombinant adhesins and ECM molecules, and inhibition experiments to demonstrate these interactions. Cutting-edge technologies, including live cell imaging and particle-tracking methods, have now started to reveal the implications of these inter- actions in pathogenesis. For example, Niddam *et al.* showed that the Lyme disease spirochaete *Borrelia burgdorferi* exploits fibronectin to interact with vascular surfaces under physiological shear stress. Specifically, it recruits and induces polymerization of soluble plasma fibronectin that strengthens and stabilizes bacterial inter- actions with endothelia by a catch–bond mechanism.[35]

Breaking down barriers to invade the host

Having colonized the host, most pathogens need to invade tissues to cause disease. This requires the break-down of primary and/or secondary defences of the host, involving the crossing of basement membranes and inter-stitial matrices. Pathogens have developed distinct ways to modify the ECM.[1,36] They can directly degrade ECM components using 'invasins' or bacterial tissue-degrading enzymes such as hyaluronidases and collagenases, causing local tissue damage. Specific examples of these interac- tions are reported in Table 2 and reviewed elsewhere.[26,36] Degradation of the ECM not only facilitates the spread of pathogens, but can also favour tissue necrosis, bacterial toxin diffusion and host cell adhesion, migration and sur- vival. Pathogens can also indirectly modify the ECM by altering the synthesis and turnover of ECM components by host cells in response to their presence (discussed later in this review). Moreover, pathogens can hijack or misuse host proteolytic systems. For instance, *Staphylococcus aur-eus*,[37] *Haemophilus influenzae*[38] and *Pseudomonas aerugi-nosa*,[39] among other common pathogenic bacteria, manipulate the plasminogen–plasmin system thereby degrading laminin and fibronectin and activating MMP zymogens that not only degrade all types of ECM pro- teins, but also process many biologically active molecules (e.g. cell surface receptors and cytokines) that direct cell behaviour and host defence.

Strategies to evade the host immune response?

Although numerous pathogens undoubtedly exploit ECM components and ECM-associated molecules to adhere to and degrade tissues for efficient host colonization and invasion, it is not always clear whether pathogens interact with these ECM molecules also to evade immune responses. For instance, *Borrelia burgdorferi* can avoid antibody-mediated clearance and this may be partly explained by its specific interaction with decorin.[36] The *Haemophilus influenzae* surface protein E binds to plas- minogen, which is converted to plasmin, and uses plas- min for complement evasion and innate immune escape.[37] Several other pathogenic bacteria take advantage of the host plasminogen system to facilitate their own spread and invasion through tissues.[40,41] Notably, *Heli- cobacter pylori* benefits from the complement regulatory property and the plasminogen-binding ability of vit- ronectin to protect itself from innate immune responses.[42] Furthermore, Group B streptococci and other Gram-positive bacteria secrete hyaluronidases, whose activity allows immune evasion besides tissue invasion. Specifically, their hyaluronidases process pro-inflamma-tory hyaluronan fragments into disaccharides, which block TLR2/4 signalling triggered by host-derived hyaluronan fragments and pathogenic ligands, including lipopolysaccharide (LPS), thereby evading immune detec-tion.[43]

Altered ECM dynamics in infection and its implications

The ECM undergoes significant alterations upon infection that promote or inhibit the establishment of infection and the host response to it. Below is an overview of the changes in ECM synthesis, degradation and post-transla- tional modification during infection and a discussion of their implications in pathogenesis.

ECM synthesis and deposition

Signal transduction pathways activated upon pathogen entry and recognition, and by mediators of inflammation and tissue repair during infection all contribute to ECM

Table 1. Examples of specific host ECM–pathogen interactions that facilitate microbial adhesion to tissues and pathogenesis (* indicates that animal models of infection were used in the study)

ECM molecule	Pathogen	Adhesin	Effect of interaction	Disease
Fibronectin	Enterohaemorrhagic *Escherichia coli* O157:H7	Lpf fimbriae (LpfA1 major subunit)	Colonization of GI tract	Acute diarrhoea; bloody diarrhoea; haemolytic uraemic syndrome	130
	Streptococcus pyogenes (group A streptococci, GAS)	Protein F1 (functional upstream domain)	Fibronectin links F1 to integrin receptors, helping bacterial uptake	Tonsillopharyngitis; necrotizing fasciitis; myositis; streptococcal toxic shock syndrome	131–133
		Protein F2 (C-terminal domains)	Host cell adhesion and internalization		33
		SfbII (C-terminal domain)	Host cell adhesion and internalization		134
		Fba	Host cell adhesion and internalization		31*
	Staphylococcus aureus	MntC	Mucosal colonization	Nosocomial infections, septicaemia, osteomyelitis, endocarditis, etc.	135
	Salmonella typhimurium	MisL (N-terminal non-conserved region)	Intestinal colonization	Gastroenteritis	136*
		ShdA			137
Plasma fibronectin	*Borrelia burgdorferi*	BBK32	Binding to fibronectin 13FnIII repeat module and intestinal colonization	Lyme disease	35
Fibronectin N-terminal proteolytic fragments (30 and 70 kDa)	*Streptococcus pyogenes*	Protein F2(C-terminal domains)	Fibronectin polymerization and colonization of vascular surfaces	Tonsillopharyngitis; necrotizing fasciitis; myositis; streptococcal toxic shock syndrome	33 130
Laminin	Enterohaemorrhagic*Escherichia coli* O157:H7	Lpf fimbriae (LpfA1 major subunit)	Eukaryotic cell adhesion and internalization	Acute diarrhoea; bloody diarrhoea; haemolytic uraemic syndrome	130
	Enterohaemorrhagic*Escherichia coli* O157:H7	Lpf fimbriae (LpfA1 major subunit)	Colonization of GI tract	Acute diarrhoea; bloody diarrhoea; haemolytic uraemic syndrome	135
	Staphylococcus aureus	MntC	Colonization of GI tract	Nosocomial infections, septicaemia, osteomyelitis, endocarditis, etc.	138
	Streptococcus gallolyticus (*gallolyticus* endocarditis isolates)	FimB, gtf and pilB	Mucosal colonization	Infective endocarditis	139
Collagen I	*Borrelia burgdorferi*	ErpX	Adhesion and invasion of endothelial cells	Lyme disease	138

Table 1 *(Continued)*

ECM molecule	Pathogen	Adhesin	Effect of interaction	Disease	References
Collagen II	*Enterococcus faecium*	Pilus subunits EmpA and EmpB	Intestinal colonization and persistence	UTIs, bacteraemia, and infective endocarditis	138
Collagen IV	*Streptococcus gallolyticus* (*gallolyticus* endocarditis isolates)	FimB, gtf and pilB	Adherence to host tissue and biofilm formation	Infective endocarditis	130
	Enterohaemorrhagic*Escherichia coli* O157: H7	Lpf fimbriae (LpfA1 major subunit)	Adhesion and invasion of endothelial cells	Acute diarrhoea; bloody diarrhoea; haemolytic uraemic syndrome	135
	Staphylococcus aureus	MntC			138
	Streptococcus gallolyticus (*gallolyticus* endocarditis isolates)	FimB, gtf and pilB	Colonization of GI tract	Nosocomial infections, septicaemia, osteomyelitis, endocarditis, etc.	136*
Collagen V	*Salmonella typhimurium*	MisL (N-terminal non-conserved region)	Mucosal colonization	Infective endocarditis	141
Collagen VI	*Enterococcus faecium*	EcbA	Adhesion and invasion of endothelial cells	Gastroenteritis	142*
Tenascin-C	*Legionella pneumophila*	Mip	Intestinal colonization	UTIs, bacteraemia, and infective endocarditis	138
Vitronectin	*Streptococcus gallolyticus* (*gallolyticus* endocarditis isolates)	FimB, gtf and pilB	Host tissue adhesion and biofilm formation	Legionellosis	138
	Streptococcus gallolyticus (*gallolyticus* endocarditis isolates)	FimB, gtf and pilB	Adhesion to lung tissue and bacterial dissemination	Infective endocarditis	143*
	Yersinia enterocolitica	YadA	Adhesion and invasion of endothelial cells	Infective endocarditis	46
Thrombospondin	*Trypanosoma cruzi*	TcCRT	Adhesion and invasion of endothelial cells	Enteric and systemic diseases	144*
Decorin	*Borrelia burgdorferi*	DbpA and DbpB	Adhesion to host cells and tissue; improved bacterial survival	Chagas disease	141
Nidogen 1 and 2	*Enterococcus faecium*	SgrA	Enhancement of cellular infection	Lyme disease	145,146
Soluble and immobilized fibrinogen (a- and b-chains)	*Staphylococcus aureus*	ClfA, ClfB	Specific localization to decorin-rich niches in the tunica adventitia and myocardial connective tissue; persistence of infection	UTIs, bacteraemia, and infective endocarditis	
			Host tissue adhesion and biofilm formation	Nosocomial infections, septicaemia, osteomyelitis, endocarditis, etc.	
			Colonization of biomaterial implants; bacterial spread		

synthesis and deposition. This involves structural ECM proteins (e.g. collagens, laminins and proteoglycans) as well as non-structural ECM components termed 'matricel- lular proteins', which are normally absent or scarcely expressed in healthy tissues (e.g. osteopontin, throm- bospondins, galectins, tenascins). For instance, rhinovirus activates TLR3 and TLR7/8 signalling in airway smooth muscle cells, which leads to increased deposition of fibro- nectin, perlecan and collagen IV, contributing to airway remodelling and facilitating the migration of airway smooth muscle cells to the infection site.[44] Interleukin-33 (IL-33) signalling induced by *Staphylococcus aureus* enhances fibronectin and collagen IIIa expression and deposition, accelerating wound repair.[45] Systems biology approaches elucidating the ECM interactome network reg- ulated by *Trypanosoma cruzi* and its gp83 ligand, which mediates trypanosome attachment and entry, have shown that activation of gp83 receptors in the cell via extracellular signal-regulated kinase 1/2 results in up-regulation of lami- nin $c1$ and thrombospondin expression to facilitate trypo- mastigote recruitment, enhancing cellular infection.[46] In mice, *Citrobacter rodentium* infection induces osteopontin and fibronectin expression through integrin-linked kinase activation, facilitating bacterial colonization of the intes- tine.[47] High expression levels of osteopontin have also been shown to be induced in murine acute and chronic coxsack- ievirus B3-myocarditis together with those of MMP-3, TIMP1, urokinase-type plasminogen activator and TGF-b_1 and, in turn, procollagen-1a mRNA expression and fibro- sis. Accordingly, osteopontin-null mice are protected from viral myocarditis, and inhibition of osteopontin transcrip- tion by vitamin D decreases cardiac fibrosis in wild-type animals.[48] Viral myocarditis also results in up-regulation of tenascin-C before immune cell infiltration occurs and until scar tissue is formed.[49] Elevated tenascin-C levels have also been reported in patients with sepsis, parapneu- monic infection, tuberculosis and *Staphylococcus aureus* infection.[50–54] Furthermore, the lungs of mice infected with Gram-negative bacteria show accumulation of versi- can, a chondroitin sulphate proteoglycan with multiple cytokine, chemokine, adhesion molecule and growth fac- tor-binding domains, that is implicated in the innate immune response.[55] Similarly, when primary human lung fibroblasts are treated with the viral mimetic Poly I:C, they deposit a higher-order structured ECM, rich in versican and hyaluronan, to which T cells avidly adhere and cease migration, an effect reversed by hyaluronidase treatment or versican antibody during matrix formation.[56] Impor- tantly, besides expression levels, infection can also alter the spatial distribution of ECM components as it emerges from influenza-infected lungs, which show distinct regions that are enriched with either fibronectin or collagen.[57] Hence, specific infectious diseases seem to generate distinct compositional changes of the ECM, inevitably influencing its biophysical structure and presentation of bioactive compounds that impact bacterial colonization and invasion, immune cell response and tissue repair.

ECM degradation

The catabolic machinery that breaks down and remodels the ECM is also altered upon infection, affecting the supportive, barrier and biological functions of the ECM. Tissue degrading enzymes such as MMPs play a crucial role in regulating immune cell recruitment: they cleave the basement membrane ECM; expose cryptic pro-migratory sites of ECM components (e.g. $c2$ chain of laminin 5); target non-ECM proteins such as adhesion molecules (e.g. E-cadherin); activate, deactivate or regulate the bioavailability of chemokines (e.g. monocyte cheomattrac- tant protein-1 and IL-8) and cytokines (e.g. IL-1b and TNF-a); and shed cell surface receptors associated with cell migration (e.g. CD44), modulating inflammation (re- viewed in refs 58–60). With the exception of neutrophils, tissue-degrading enzymes are not stored, but require *de novo* synthesis that is strictly regulated and can be induced by pro-inflammatory cytokines (e.g. TNF-a and IL-1) and also by bacterial products (e.g. LPS and chlamydial heat-shock proteins).[61]

Several pathogens, including *Mycobacterium bovis*, *Mycobacterium tuberculosis*, *Streptococcus pyogenes* and *Helicobacter pylori*, induce expression and activity of a number of MMPs, including MMP-1, MMP-2, MMP-7, MMP-9 and MMP-13.[62–64] In particular, the extensive production of MMP-2 and MMP-9 during mycobacterial infection is regulated by macrophage-derived and T-cell-derived cytokines and causes ECM breakdown. This may be necessary for cell recruitment and granuloma formation, both protective immune responses to mycobacteria. However, dysregulation of MMP production at late stages of the infection could contribute to tissue damage and, by compromising tissue integrity, may facilitate bacterial dissemination and persistence of infection.[62] Systemic *Escherichia coli* infection, acute Lyme neuroborreliosis and pneumococcal meningitis can all lead to secretion of high amounts of MMP-9. In meningitis, MMP-9 has been suggested to contribute to blood–brain barrier destruction and neuronal injury.[65–67] Similarly, gastrin-dependent induction of MMP-7 upon *Helicobacter pylori* infection has been impli- cated in the development of gastric cancer through the release of heparin-binding epidermal growth factor.[64] *Chlamydia trachomatis*, which causes trachoma and blind- ness, up-regulates MMP-7, MMP-9, MMP-12 and TIMP-1 expression, while it down-regulates MMP-10 and SPARC (secreted protein, acidic, cysteine-rich)-like 1, a matricellular protein that regulates decorin production and collagen assembly.[68] In this study the expression pattern of these ECM-modifying enzymes was correlated with the clinical scarring grade and inflammation. Interestingly, a dual function has been shown for MMP-12 in viral myocarditis

Table 2. Examples of specific ECM–pathogen interactions that facilitate host invasion through direct degradation of ECM components (# indicates that *ex vivo* mammalian tissue degradation models, were used in the study)

ECM molecule	Pathogen	Microbial enzyme	Effect of ECM cleavage	Disease	References
Laminin	*Psedomonas aeruginosa*	Elastase; alkaline protease	Tissue invasion and necrosis	Necrotizing pneumonia, septic shock, UTI, skin and soft-tissue infections	147#
	Clostridium difficile	Cwp84	Tissue integrity loss; facilitation of toxin diffusion	Pseudomembranous colitis and nosocomial diarrhoea	148
Collagen I	*Porphyromonas gingivalis*	Gingipains in *P. gingivalis* supernatant	Tissue degradation	Periodontal disease	149
	Vibrio parahaemolyticus	Metalloprotease VppC	Tissue damage	Acute gastroenteritis	150
Collagen I, II, III, IV, V and VI	*Clostridium histolyticum*	Class I and II collagenases (ColG, ColH)	Necrotic tissue degradation; promote keratinocyte migration	Gas gangrene, infective endocarditis	151,152
Collagen IV	*Streptococcus gordonii*	Serine protease	Basement membrane breakdown	Infective endocarditis	153
Fibronectin	*Clostridium difficile*	Cwp84	Tissue integrity loss; facilitation of toxin diffusion	Pseudomembranous colitis and nosocomial diarrhoea	148
	Porphyromonas gingivalis	Gingipains in *P. gingivalis* supernatant; HRgpA and RgpB gingipains	Cleavage and inactivation of cell-binding region of fibronectin; gingival fibroblast detachment and death; tissue destruction	Periodontal disease	149,154#
Vitronectin	*Clostridium difficile*	Cwp84	Tissue integrity loss; facilitation of toxin diffusion	Pseudomembranous colitis and nosocomial diarrhoea	148
Tenascin-C (large isoforms)	*Porphyromonas gingivalis*	HRgpA, RgpB and Kgp gingipains	Enhanced anti-adhesive activity of tenascin-C; gingival fibroblasts detachment, apoptosis and tissue destruction	Periodontal disease	154#

Cwp84, putative cell surface-associated cysteine protease; HRgpA and RgpB, arginine-gingipains; Kgp, lysine-gingipains; UTI, urinary tract infection.

caused by coxsackievirus type B3. While intracellular MMP-12 causes IFN-*a* secretion and host protection, extracellular MMP-12 cleaves the IFN-*a* receptor 2 binding site of IFN-*a*, preventing an uncontrolled immune response.[69] In the same model of viral myocarditis, MMP-9 exerts a protective role by inactivating IFN-*b*/*c*. Indeed, MMP-9-null mice display higher viral load, infiltration of $CD3^+$ cells and tissue damage.[70] However, murine MMP-9 can also enhance susceptibility to infection and increase morbidity and mortality. This is the case in *Francisella tularensis* pulmonary infection where MMP-9 generates pro-inflammatory ECM-derived peptides (i.e. Pro-Gly-Pro peptide from collagen I), enhancing neutrophil infiltration to lungs.[71]

In addition to inducing MMP expression, pathogens can also activate pro-MMPs by secreting their own activating enzymes, as in the case of *Pseudomonas aeruginosa*, which activates pro-MMP-2 using LasB, a thermolysin-like metalloprotease.[39] The role of dysregulated tissue-degrading enzymes, individually or in combination, is not well understood in every infectious disease. How- ever, it is clear that the MMP/TIMP system changes the ECM composition and biophysics and its presentation of bioactive molecules, generating environmental cues that are detected and processed by immune cells into sig- nalling events that direct their behaviour and response to infection.

ECM post-translational modification

Post-translational modifications (PTMs) add complexity to the ECM. Some are generated by proteolytic cleavage,

while others are generated by citrullination of arginine, glycosylation, cross-linking, hydroxylation of prolines, nitrosylation of tyrosines and aspartate isomerization (reviewed in ref. 7). Given the ability of pathogens to hijack host enzymes or secrete enzymes targeting host molecules, it is tempting to ask whether PTMs in the ECM are altered in infection. Indeed, the *Porphyromonas gingivalis* enzyme peptidylarginine deiminase can convert arginine residues to citrulline in mammalian ECM proteins, including fibrinogen, collagen II, fibronectin and tenascin-C.[72–76] Although arginine is positively charged at neutral pH, citrulline is uncharged, increasing protein hydrophobicity and, hence, altering protein three-dimensional structure and function. Notably, *Porphyromonas gingivalis* periodontal infection has been linked to rheumatoid arthritis, an autoimmune disease of the joints characterized by high levels of citrullinated proteins and anti-citrullinated ECM protein antibodies.[77] *Trypanosoma cruzi*, the aetiological agent of Chagas disease, which features extensive inflammation and fibrosis of the heart, has been reported to increase the expression of lysyl oxidase (LOX).[78] This enzyme carries out cross-linking of collagen fibres thereby altering matrix stiffness, which has been linked to cancer metastasis.[79] By irreversibly altering collagen structure and function, LOX has been proposed to cause dysfunction of cardiomyocytes and, in turn, of the heart in Chagas disease. Similarly, dengue virus infection has been shown to suppress the expression of cartilage-associated protein (CRTAP), alongside the protein associated to tight junctions (PATJ). CRTAP associates with the proteoglycan leprecan, which has collagen prolyl 3-hydroxylase activity, and cyclophilin B in the endoplasmic reticulum. This trimeric complex is required for proline 3-hydroxylation of collagen and, hence, collagen assembly. As the levels of CRTAP mRNA negatively correlate with viral replication, the authors of the study have speculated that CRTAP, together with PATJ, restrict dengue infection by influencing cell–cell adhesion.[80]

While evidence supporting a link between specific PTMs of ECM proteins and diseases such as arthritis and cancer is accumulating, the implications of altered PTMs of ECM components during infection are only beginning to emerge. For example, collagen fibre cross-linking mediated by LOX can significantly alter tissue structure and ECM mechanics. Mammoto *et al.* recently showed that LOX-dependent changes in ECM mechanics control vascular permeability and pulmonary oedema. *In vivo*, mouse lungs treated with LPS, which contributes to pulmonary oedema and acute respiratory distress syndrome in patients with sepsis,[81] become much stiffer than untreated lungs and exhibit enhanced vascular permeability. Increased LOX expression and LOX and LOX1 protein isoform activity control alveolar architecture and vascular permeability, which are restored by LOX activity inhibition.[82]

Macrophages: not just destroyers of the ECM

In bacterial and viral infection, activated macrophages [e.g. M(IFN-γ), M(LPS)] have long been implicated in ECM destabilization and destruction through the secretion of tissue-degrading enzymes, including MMP-9. Similarly, parasitic infections involve activated macrophages [i.e. M(IL-4)], which release proteases such as MMP-1 and MMP-12.[83] However, macrophages can also produce a number of ECM proteins. Increasing evidence of this and its implications are discussed below.

Fibronectin is the first ECM glycoprotein reported to be produced by human macrophages and IFN-γ-stimulated mouse peritoneal macrophages.[84,85] Moreover, bacterial components induce tenascin-C expression in human monocyte-derived and mouse bone-marrow-derived macrophages.[53,86] Lipopolysaccharide from *Porphyromonas gingivalis* induces thrombospondin-1 production in THP-1 cells,[87] while LPS from *Escherichia coli* and IFN-γ up-regulate galectin-1 and galectin-3, including five galectin-3 truncated forms, in primary human macrophages.[88] In mice with viral myocarditis, macrophages infiltrating the heart are the main producers of osteopontin.[48] Notably, in murine macrophages, LPS causes the formation of chromosomal loops in the osteopontin promoter by bridging nuclear factor-κB (NF-κB) and activator protein-1 together, leading to osteopontin transcription,[89] which is negatively regulated by GSK3b.[90]

Macrophage expression of proteoglycans has also been demonstrated. Treatment of bone marrow-derived and alveolar macrophages with *Escherichia coli* LPS results in rapid induction of versican and hyaluronan synthase 1, and simultaneous inhibition of the major hyaluronan-degrading enzymes (hyaluronidases 1/2).[91] Serglycin, decorin and biglycan are also secreted by LPS-activated mouse peritoneal macrophages.[92–94] By using inhibition, chromatin immunoprecipitation and NF-κB reporter gene assays, some of these studies have demonstrated that pathogenic activation of the NF-κB signalling pathway downstream of TLR4 leads to ECM molecule transcription.[86,89,93–95] Importantly, in the case of tenascin-C, biglycan and decorin, which can activate TLR4, this can promote autocrine loops of inflammation.[86,93,94] Hence, the infected cellular microenvironment influences TLR function and, in turn, TLR activation affects the microenvironment.

Intriguingly, there is mounting evidence that macrophages can synthesize collagen molecules. In *Drosophila*, phagocytes (haemocytes) produce functional collagen IV, which controls key signalling events in the germline stem cell niche.[96] The first reports of collagen synthesis by mammalian macrophages were published in the 1990s. The first showed collagen I synthesis in mouse peritoneal

macrophages and the second demonstrated collagen VIII synthesis and secretion in human macrophages, which was decreased by IFN-*c* and LPS treatment.[97,98] Collagen VIII is a short-chain, non-fibrillar collagen that forms unique hexagonal lattice structures and possesses both structural and signalling properties. Later, human macrophages were shown to secrete collagen VI, which forms beaded filaments with a multidomain structure that interact with ECMs and cell surface receptors to anchor interstitial structures and cells within tissues. Expression of collagen VI was decreased by IFN-*c* and LPS, but increased upon IL-4 treatment.[99] Recently, the expression of all 28 collagen-encoding mRNAs was quantified in steady-state and LPS-activated primary human macrophages and compared with that of human dermal fibroblasts, an abundant matrix source.[95] Steady-state macrophages expressed basal levels (lower than those in dermal fibroblasts) of collagen mRNAs with the exception of collagen III, X, XI, XVI, XX and XXVI. However, LPS specifically increased the expression of fibril-associated collagens with interrupted triple helices (FACITs; collagen VII, XII, XV, XIX and XXI), the collagenase-resistant col- lagen V, the collagen-containing von Willebrand factor collagen XXVIII and collagen I, IV, XVIII, XXV and

XXVII. Collagen II, VIII, IX, XIII, XIV, XXIII and XXIV expression was not increased. LPS also down-regulated the expression of collagen VI, confirming previous studies,[99] and collagen XIII and XVII.[95] By secreting collagens, depending upon their mode of activation, macrophages may contribute to the ECM, and therefore to tissue stabi- lization and repair, and to cell–cell and cell–matrix inter- actions (e.g. *in vitro*, monocytes adhere strongly to collagen VI[99]). Furthermore, macrophages may bind to their secreted collagen molecules as they express several receptors known to interact with collagen (e.g. integrins and proteoglycans). However, collagen secretion by macrophages and its potential role in macrophage adhe- sion and the immune response *in vivo* remain to be clari- fied.

The ECM: an integral part of the innate immune response to infection?

Antimicrobial activity of the ECM

The innate immune system employs endogenous peptides like *a*-defensin and LL-37, which bind to heparin and dermatan sulphate glycosaminoglycans and have antimicrobial properties. During the inflammatory response to infection, tissue-degrading enzymes generate bioactive ECM fragments. A number of studies showed that heparin-binding peptides derived from laminin isoforms, vitronectin, thrombospondin and fibronectin, exert antimicrobial activities against Gram-positive and Gram-negative bacteria, and the fungus *Candida albicans*.[100–102]

Certain ECM proteins are also found in biological fluids. One such protein is tenascin-C, which has been found in human breast milk (2·2–671 ▮g/ml) where it acts as an innate broad-spectrum human immunodefi- ciency virus 1 (HIV-1) -neutralizing protein.[103] Tenascin- C directly captures HIV-1 virions by binding to the HIV- 1 envelope gp120 protein at a CD4-inducible epitope that overlaps the chemokine co-receptor binding site. Accordingly, tenascin-C depletion abolishes the HIV-1-neutralizing activity of milk.[103]

Direct antimicrobial activity has also been reported for the ECM-associated protein MMP-12, which is abun- dantly expressed in mature tissue macrophages and mobi- lized to macrophage phagolysosomes after the ingestion of bacteria. Inside phagolysosomes, MMP-12 adheres to bacterial cell walls and disrupts cellular membranes, kill- ing the bacteria. Notably, the C-terminal domain of MMP-12, but not its catalytic domain, contains a four- amino-acid sequence on an exposed *b* loop of the protein that is unique in nature and confers antimicrobial activ- ity.[104] Together, these data may help the search for safe, endogenous antimicrobial molecules from complex bio- logical mixtures. Furthermore, binding epitopes could serve as templates for *de novo* synthesis of novel antimi- crobial molecules.

ECM-mediated recognition of microbial pathogens, macrophage activation and phagocytosis

The innate immune system employs highly conserved receptors, namely PRRs, to recognize conserved motifs in microbial pathogens, called pathogen-associated molecu- lar patterns. Mindin, a member of the mindin-F-spondin family of secreted ECM proteins, has emerged as a unique pattern recognition molecule in the ECM for microbial pathogens and has been proposed to function as an inte- gral part of the innate immune response.[105] He *et al.*, showed that genetic ablation of mindin confers resistance to LPS-induced shock and systemic *Salmonella typhimur- ium* and *Streptococcus pneumoniae* infections *in vivo*. Moreover, mindin-null mice feature impaired bacterial clearance in lungs infected with Gram-positive group B streptococcus or *Haemophilus influenzae*. *In vitro*, macro- phages and mast cells lacking mindin display impaired TNF-*a* and IL-6 production and defective phagocytosis. By using recombinant mindin, the authors showed that this ECM protein recognizes carbohydrate moieties of Gram-positive and Gram-negative bacterial components and, by binding to them, it agglutinates bacteria. As glu- cose inhibited not only mindin binding to pathogens but also macrophage activation by LPS, the authors con- cluded that mindin-mediated carbohydrate recognition of microbial pathogens is a secondary stimulation necessary for the activation of macrophages and mast cells[105] (Fig. 1). Notably, mindin is not a universal, but is a

specific pattern recognition molecule as it only recognizes and opsonizes certain bacteria. Following this, the innate immune function of mindin has been extended to include promoting influenza virus clearance from the nasal cavity by allowing efficient macrophage activation.[106] Further- more, the proteoglycan lumican, whose core protein con- tains tandem repeats of leucine-rich motifs similarly to PRRs, interacts with CD14 on the surface of macrophages and neutrophils, promoting CD14-TLR4-mediated responses to LPS (Fig. 1). Hence, lumican-null mice are hyporesponsive to LPS-induced septic shock.[107] In a *Pseu- domonas aeruginosa* model of lung infection, mortality of lumican-deficient mice is increased as animals fail to clear bacteria from tissues. This study showed that CD14- mediated phagocytosis of *Escherichia coli* and *Pseu- domonas aeruginosa* bacteria by macrophages is impaired in the absence of lumican and identified Tyr-20 as a vital residue for CD14 binding and phagocytosis.[108] Finally, infection of the cornea with *Pseudomonas aeruginosa* readily increases lumican expression before inflammatory cell infiltration, and lumican-null mice display poor reso- lution of bacterial keratitis and sustained production of pro-inflammatory cytokines.[109] Another ECM protein that senses a number of microbial pathogens is galectin-3, which binds to carbohydrate structures on glycoproteins and glycolipids (e.g. *N*-acetyl-$_D$-lactosamine and LPS) from (myco)bacteria, protozoan parasites and yeast.[110] Although the *in vivo* function of galectin-3 during infec- tion has not been fully investigated, upon microorganism recognition, it contributes to macrophage-mediated phagocytosis, at least *in vitro*[111] (Fig. 1). Whether these ECM components exert the same essential immune func- tions in humans represents an outstanding, important question.

Leucocyte population balance in infection: emerging roles for ECM components

Severe infection demands large numbers of leucocytes that are compensated by the emergency myelopoiesis response, which is initiated by activated PRRs and cytoki- nes, including IL-6, granulocyte–macrophage colony-sti- mulating factor and granulocyte colony-stimulating factor. This protects the host from systemic infection by quickly generating the required leucocyte population. Kanayama *et al.*[112] have recently found that osteopontin skewed the balance of myeloid and lymphoid cell popula- tions during systemic infection with *Candida albicans.* Specifically, the authors showed that osteopontin limited the supply of neutrophils and Ly6C$^+$ monocytes–macro- phages by enhancing the apoptosis of common myeloid and granulocyte–macrophage progenitors through a down- regulation of the expression of the apoptosis inhi- bitor survivin. This resulted in greater fungal load in kid- neys and significantly higher mortality of wild-type mice with systemic fungal infection compared with osteopon- tin-deficient littermates. The detrimental effect of osteo- pontin was observed early, 24 hr after infection.[112]

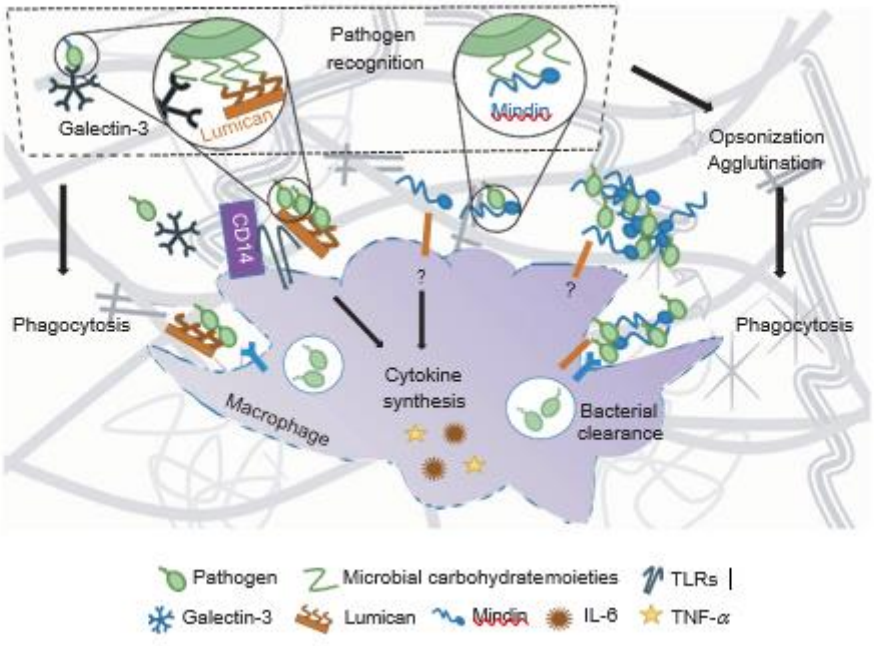


Figure 1. Mindin, lumican and galectin-3: three extracellular sentinels. Mindin, lumican and galectin-3 recognize and bind to sugar moieties found in the cell wall of several types of bacteria. All of them promote phagocytosis of bacteria by macrophages. Mindin binds to bacteria, causes their opsonization and agglutination, and facilitates their phagocytosis by macrophages. Mindin also induces the synthesis of pro-inflammatory cytokines by these cells. Lumican instead interacts with CD14 on the surface of macrophages, promoting CD14-TLR4-mediated responses to lipopolysaccharide (LPS) and CD14-mediated phagocytosis.

However, it is unclear if this effect is maintained, exacerbated or reversed at later stages of infection.

Dendritic cells (DCs) are the main antigen-presenting cells and, upon capturing microbial pathogens, mature and migrate to lymphoid tissues where they activate naive T cells. Distinct classes of microbes elicit lineage-specific responses from the effector T-cell repertoire. Type 1 helper T (Th1) cells are involved in infection by intracel- lular bacteria and viruses, Th2 cells in parasitic infection, and Th17 cells in infection by extracellular pathogens and facultative and obligate intracellular bacteria and fungi. Regulatory T cells hold the inflammatory response in check. Specific ECM proteins have been shown to con- tribute to the T-cell polarizing function of DCs without affecting DC development. For instance, generation of Th17 cells by *Escherichia coli* LPS or *Mycobacterium tuber- culosis* stimulated bone marrow-derived DCs from tenas- cin-C-null mice is significantly impaired.[113] Furthermore, the expression of galectin-3 in DCs controls the magni- tude of T-cell priming *in vitro* and *in vivo* during hel- minthic infection with *Schistosoma mansoni*. Galectin-3- deficient mice have significantly fewer T cells in their spleen and higher cellular and humoral Th1 responses.[114] Although this study shows that galectin-3 expression by DCs modulates the proliferation and cytokine release by T cells, it does not explain the mechanism responsible for the biased Th1 response. Efficient T-cell priming by DCs has also been shown to depend on mindin.[115] When DCs from mindin-null mice are activated with bacterial com- ponents, including LPS from *Salmonella typhosa* or *Escherichia coli* and lipoteichoic acid from *Staphylococcus aureus*, $CD4^+$ T-cell priming is 60–70% lower than that of wild-type mice. Investigation of this demonstrated that DCs interact with mindin via integrins $\alpha_4\beta_1$ and $\alpha_5\beta_1$, leading to up-regulated expression of the Rho GTPases Rac1/2, which are known to regulate DC priming of T cells.[115] As DCs link innate to adaptive immunity and unbalanced effector T-cell populations lead to pathologi- cal inflammation, understanding how the ECM can regu- late T-cell responses is crucial.

Transcriptional and post-transcriptional roles for ECM molecules in host defence signalling pathways

Pathogen recognition via PRRs initiates inflammatory signalling pathways that are tightly regulated to allow microbial clearance with minimal damage to the host. Recent research implicates a role for ECM and ECM-associated proteins in regulating inflammatory networks at the transcriptional and post-transcriptional levels during the immune response to infection.

An elegant study by Marchant *et al*.[69] found a transcriptional role for MMP-12 (macrophage elastase) in immunity against viral infection. During coxsackievirus type B3 and respiratory syncytial virus infections, MMP-12-null mice display increased viral load, mortality and lower levels of IFN-α, which is essential for viral immunity. Mechanistically, secreted MMP-12 is taken up by virus-infected cells and traffics to the nucleus, where it binds to the *NFKBIA* promoter, driving its transcription, which is essential for optimal IFN-α secretion and host protection. Additionally, MMP-12 regulates specific sub- strates by two distinct mechanisms: (i) through DNA binding of gene exons (e.g. exons encoding PSME3, the immunoproteasome cap protein, and SPARC-like protein 1, which decreases their mRNA and protein levels in MMP-12-null mice); and (ii) extracellularly, through sub- strate protein cleavage (e.g. IFN-α receptor 2 binding site). Hence, MMP-12 clears systemic IFN-α and, accord- ingly, selective inhibition of extracellular MMP-12 in infected wild-type mice elevates systemic IFN-α levels and reduces viral replication.[69]

At the post-transcriptional level, a role for decorin and tenascin-C has been found in regulating microRNAs in LPS-induced sepsis.[53,93,116] Merline *et al.* detected increased decorin levels in patients with sepsis and mice with LPS-induced sepsis, and elevated IL-10 amounts in decorin-null mice. They demonstrated that, in the presence of LPS, decorin reduces the levels of the anti-inflammatory cytokine IL-10 through two mechanisms. First, it activates extracellular signal-regulated kinase and p38 mitogen-activated protein kinases downstream of TLR2 and TLR4, thereby inducing the expression of the pro-inflammatory modulator programmed cell death 4 (PDCD4), which translationally represses IL-10. Second, decorin inhibits TGF-β_1 signalling, leading to lower levels of the microRNA miR-21, which represses PDCD4 expression and so decreases IL-10 levels.[93] Soon after, tenascin-C was shown to orchestrate the secretion of specific cytokine subsets in response to LPS. Specifically, tenascin-C regulates the biosynthesis of the LPS-responsive miRNA miR-155, allowing optimal TNF-α production in macrophages and effective immune response to LPS *in vivo*.[53] Hence, specific ECM and ECM-associated proteins possess previously unknown transcriptional and post-transcriptional regulatory activities, which fine-tune the innate immune response to infection.

The ECM in infectious granulomas

Tuberculosis, syphilis, toxoplasmosis, infectious mononucleosis and measles are a few examples of infectious diseases characterized by the formation of granulomas. Traditionally considered host-protective structures, infectious granulomas are compact, organized immune cell clusters that are generated in response to specific pathogens. Granulomas contain large numbers of mature macrophages, which can fuse into multinucleated giant cells or differentiate into foam cells. Neutrophils, den- dritic cells, B, T and natural killer cells, and fibroblasts are also found in granulomas, which are surrounded by

epithelial cells. Granulomas also contain ECM proteins, whose expression and function in infectious granulomatous disease pathogenesis has only recently been investigated.

In humans, osteopontin has been detected in granulo- mas of diverse aetiology;[117] in those caused by *Paracoccidioides brasiliensis*, it localizes in ECM, macrophages and multinucleated giant cells at the centre of lesions in the early phase of infection.[118] In tuberculosis granulomas, osteopontin is markedly expressed within lymphocytes, macrophages, epithelioid cells and multinucleated giant cells, but not in the central necrotic core.[54] These granulomas also stain strongly for tenascin-C in the surround- ing fibrotic rings and weakly inside, in the ECM. A diffuse ECM-associated staining of galectin-9 is detected in the granuloma and in epithelioid and multinucleated giant cells. Notably, osteopontin, tenascin-C and galectin-9 are absent in non-infectious Crohn disease granulo- mas.[54] Analysis of human lung biopsies from patients with atypical mycobacteriosis and tuberculosis revealed expression of tenascin-C and precursor proteins of collagens I and III around granulomas. Precursor proteins of collagen I were also found within granulomas that co-localized with myofibroblasts.[119] In mice infected with *Paracoccidioides brasiliensis*, Gonzalez *et al.* detected increased expression of laminin, fibronectin, fibrinogen, collagen I, collagen III, elastic fibres and proteoglycans during granuloma formation.[120] Analysis of their arrangement showed that they were mostly surrounding the granuloma, but sporadically inside it. Initially arranged in

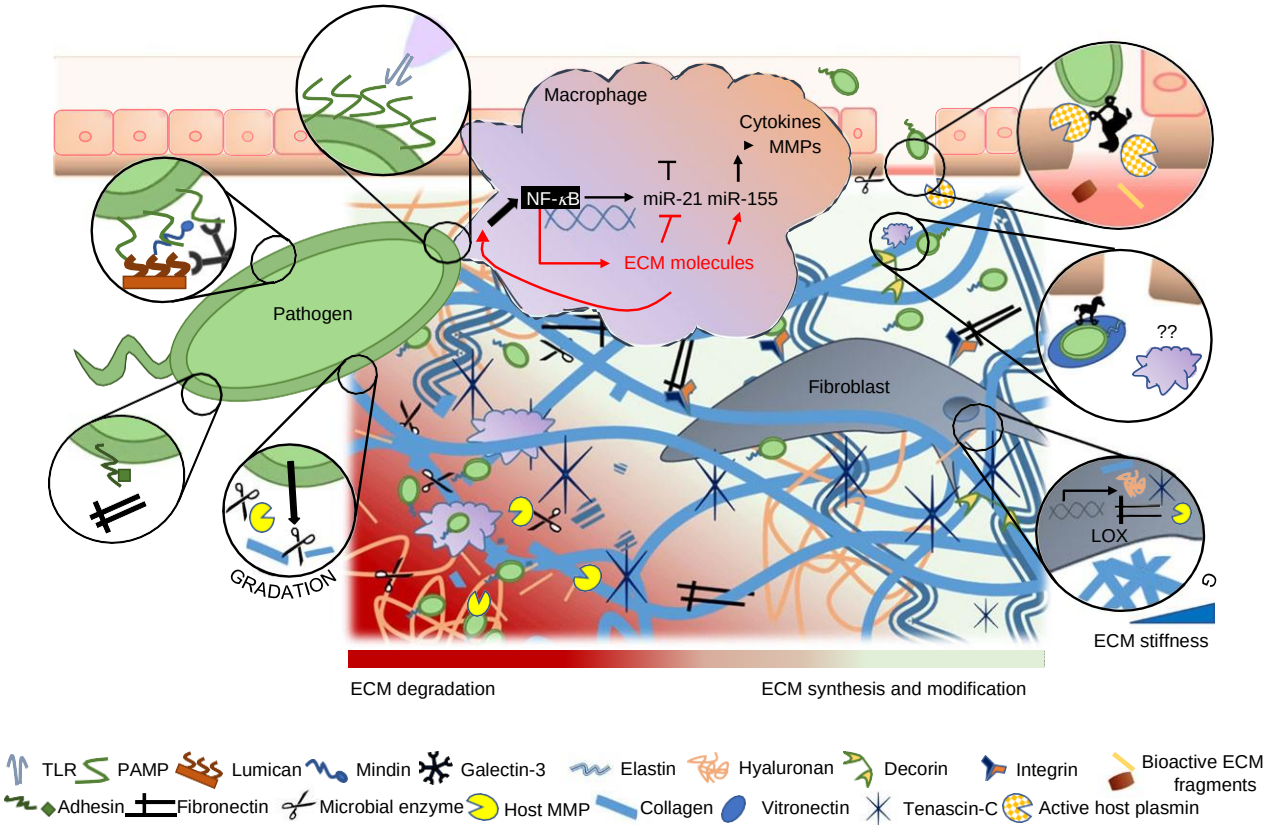


Figure 2. Multiple functions of the extracellular matrix (ECM) in the immune response to infection. Recognition of pathogen-associated molecular patterns by macrophages through pattern recognition receptors (PRRs) and by ECM components is shown. Pathogen binding to ECM molecules such as fibronectin helps host colonization. Degradation of the ECM through microbial tissue-degrading enzymes or host matrix met- alloproteases (MMPs) activated by pathogen facilitates host invasion. To establish the infection, pathogens can also hijack host proteolytic systems such as the plasminogen–plasmin system and evade innate immune responses by binding to ECM components such as vitronectin. Fibroblasts in the interstitial ECM produce and secrete ECM proteins, MMPs and higher levels of lysyl oxidase (LOX), which cross-links collagen fibres, increas- ing ECM stiffness. Pathogen-mediated activation of macrophages triggers inflammatory signalling pathways such as the nuclear factor-jB (NF- jB) pathway, which culminates in the synthesis of cytokines, MMPs and microRNAs. Activated macrophages synthesize also ECM components such as decorin and tenascin-C, which regulate the biosynthesis of miR-21 and miR-155 and generate positive feedback loops that propagate inflammation.

a disorganized manner, ECM fibres later acquired a compact, concentric arrangement, which originated from an anchorage point that may contribute to tissue integrity and enhance distribution of growth factors and cytokines.[120] Notably, during chronic mycobacterial infection, the fibrinolytic system has been shown to limit progressive fibrosis with plasminogen regulating the turnover of ECM proteins within the granuloma.[121]

Recent studies on ECM turnover in tuberculosis granuloma have helped to rewrite tuberculosis immunopathology. Traditionally considered a host-protective structure that 'quarantines' the infecting mycobacteria, tuberculosis granuloma has been implicated in the expansion and dissemination of infection. Specifically, caseous necrosis leading to ECM destruction and bacterial dissemination was thought to be the cornerstone of tuberculosis pathogenesis. Conversely, collagen destruction has now been proposed to initiate caseous necrosis. Elkington *et al.*[122] have first shown that MMP-1-expressing mice develop collagen destruction within granulomas upon infection with *Mycobacterium tuberculosis* in the absence of caseous necrosis. They then demonstrated that proteolytic colla- gen destruction of the lung ECM is the initial pathologi- cal event that reduces the survival of *Mycobacterium tuberculosis*-infected cells, resulting in caseous necrosis and cavitation, and so diverting the immune response in favour of the pathogen. Conversely, intact collagen fibrils increase survival of infected cells.[123,124] In line with this, Parasa *et al.* reported up-regulation of MMP-1, MMP-3, MMP-9 and MMP-12 in a human lung-tissue model and in biopsies from patients with non-cavitary tuberculosis. Global MMP inhibition via marimastat reduced granu- loma formation and bacterial load.[125]

Future work should take into account the hypoxic conditions in infectious granulomas and include unbiased global analyses of ECM and ECM-associated molecules in granulomas at various disease stages and gain- and loss-of-function experiments in three-dimensional systems and/or genetically modified animals to understand the role of the ECM in granuloma formation and function.

Concluding remarks and future perspectives

The ECM and the innate immune response to infection are inextricably linked. Largely ignored or overlooked, the diverse yet specific functions of the ECM in infection influence the establishment and dissemination of microbial pathogens in host tissues as well as the outcome of the immune response to the infection (Fig. 2). In addition to the studies discussed here, ECM immunological research, which is expanding into the areas of prion disease,[126] virus transmission,[127] exosomes in viral pathogenesis[128] and vaccine development,[129] shows the breadth and complexity of ECM activity and regulation in infection.

The ECM undergoes significant changes upon micro- bial invasion, magnifying the complexity of the cellular microenvironment at sites of infection. This presents challenges in setting up appropriate model systems and identifying the ECM and immunological pathways that are directly responsible for the outcome of individual infectious diseases. Unbiased omics, systems biology and genome editing approaches are promising resources for defining these pathways, designing mechanistic studies, and, in the longer-term, elucidating them in the context of the microbiota.

As antimicrobial resistance is of global concern and tackling it is becoming increasingly challenging, investigating the now evident role of the ECM in infection may reveal novel therapeutic strategies or improve existing ones. It may also inform biomaterials and tissue engineer- ing in their effort to prolong the lifespan and integrity of medical devices that are compromised by infection.

Acknowledgements

The authors are supported by an Anne McLaren Fellow- ship (University of Nottingham; awarded to AMP) and a BBSRC-DTP studentship (HT). We thank E.P. Ling for assistance with the literature search and N. Zordan for help with the figures.

Disclosures

None to declare.

References

1 Pizarro-Cerda J, Cossart P. Bacterial adhesion and entry into host cells. *Cell* 2006; 124:715–27.
2 Kuusela P. Fibronectin binds to *Staphylococcus aureus*. *Nature* 1978; 276:718–20.
3 Chagnot C, Listrat A, Astruc T, Desvaux M. Bacterial adhesion to animal tissues: protein determinants for recognition of extracellular matrix components. *Cell Microbiol* 2012; 14:1687–96.
4 Sorokin L. The impact of the extracellular matrix on inflammation. *Nat Rev Immunol* 2010; 10:712–23.
5 Boyd DF, Thomas PG. Towards integrating extracellular matrix and immunological pathways. *Cytokine* 2017; 98:79–86.
6 Hynes RO, Naba A. Overview of the matrisome – an inventory of extracellular matrix constituents and functions. *Cold Spring Harb Perspect Biol* 2012; 4:a004903.
7 Leeming DJ, Bay-Jensen AC, Vassiliadis E, Larsen MR, Henriksen K, Karsdal MA. Post-translational modifications of the extracellular matrix are key events in cancer progression: opportunities for biochemical marker development. *Biomarkers* 2011; 16:193–205.
8 Boyd CD, Pierce RA, Schwarzbauer JE, Doege K, Sandell LJ. Alternate exon usage is a commonly used mechanism for increasing coding diversity within genes coding for extracellular matrix proteins. *Matrix.* 1993; 13:457–69.
9 Kenne E, Soehnlein O, Genove G, Rotzius P, Eriksson EE, Lindbom L. Immune cell recruitment to inflammatory loci is impaired in mice deficient in basement membrane protein laminin α4. *J Leukoc Biol* 2010; 88:523–8.
10 Wu C, Ivars F, Anderson P, Hallmann R, Vestweber D, Nilsson P, *et al.* Endothelial basement membrane laminin α5 selectively inhibits T lymphocyte extravasation into the brain. *Nat Med* 2009; 15:519–27.
11 Lokmic Z, Lammermann T, Sixt M, Cardell S, Hallmann R, Sorokin L. The extracellular matrix of the spleen as a potential organizer of immune cell compartments. *Semin Immunol* 2008; 20:4–13.

12 Sixt M, Kanazawa N, Selg M, Samson T, Roos G, Reinhardt DP, et al. The conduit system transports soluble antigens from the afferent lymph to resident dendritic cells in the T cell area of the lymph node. Immunity 2005; 22:19–29.

13 Lu P, Takai K, Weaver VM, Werb Z. Extracellular matrix degradation and remodeling in development and disease. Cold Spring Harb Perspect Biol 2011; 3:a005058.

14 Chou J, Chan MF, Werb Z. Metalloproteinases: a functional pathway for myeloid cells. Microbiol Spectr 2016; 4:1–9.

15 Leitinger B. Transmembrane collagen receptors. Annu Rev Cell Dev Biol 2011; 27:265–90.

16 Campbell ID, Humphries MJ. Integrin structure, activation, and interactions. Cold Spring Harb Perspect Biol 2011; 3:a004994.

17 Doyle JJ, Gerber EE, Dietz HC. Matrix-dependent perturbation of TGFb signaling and disease. FEBS Lett 2012; 586:2003–15.

18 Yan D, Lin X. Shaping morphogen gradients by proteoglycans. Cold Spring Harb Perspect Biol 2009; 1:a002493.

19 Travis MA, Sheppard D. TGF-b activation and function in immunity. Annu Rev Immunol 2014; 32:51–82.

20 Chen GY, Nunez G. Sterile inflammation: sensing and reacting to damage. Nat Rev Immunol 2010; 10:826–37.

21 Piccinini AM, Midwood KS. DAMPening inflammation by modulating TLR signalling. Mediators Inflamm 2010; 2010:1–21.

22 Matzinger P. The danger model: a renewed sense of self. Science 2002; 296:301–5.

23 Gordon S. Pattern recognition receptors: doubling up for the innate immune response. Cell 2002; 111:927–30.

24 Kline KA, Falker S, Dahlberg S, Normark S, Henriques-Normark B. Bacterial adhesins in host–microbe interactions. Cell Host Microbe 2009; 5:580–92.

25 Westerlund B, Korhonen TK. Bacterial proteins binding to the mammalian extracellular matrix. Mol Microbiol 1993; 9:687–94.

26 Singh B, Fleury C, Jalalvand F, Riesbeck K. Human pathogens utilize host extracellular matrix proteins laminin and collagen for adhesion and invasion of the host. FEMS Microbiol Rev 2012; 36:1122–80.

27 Kuusela P, Vartio T, Vuento M, Myhre EB. Binding sites for streptococci and staphylococci in fibronectin. Infect Immun 1984; 45:433–6.

28 Kreikemeyer B, Klenk M, Podbielski A. The intracellular status of Streptococcus pyogenes: role of extracellular matrix-binding proteins and their regulation. Int J Med Microbiol 2004; 294:177–88.

29 Schwarz-Linek U, Werner JM, Pickford AR, Gurusiddappa S, Kim JH, Pilka ES, et al. Pathogenic bacteria attach to human fibronectin through a tandem b-zipper. Nature 2003; 423:177–81.

30 Joh D, Wann ER, Kreikemeyer B, Speziale P, Hook M. Role of fibronectin-binding MSCRAMMs in bacterial adherence and entry into mammalian cells. Matrix Biol 1999; 18:211–23.

31 Terao Y, Kawabata S, Kunitomo E, Murakami J, Nakagawa I, Hamada S. Fba, a novel fibronectin-binding protein from Streptococcus pyogenes, promotes bacterial entry into epithelial cells, and the fba gene is positively transcribed under the Mga regulator. Mol Microbiol 2001; 42:75–86.

32 Hanski E, Caparon M. Protein-F, a fibronectin-binding protein, is an adhesin of the group a streptococcus Streptococcus pyogenes. Proc Natl Acad Sci U S A 1992; 89:6172– 6.

33 Kreikemeyer B, Oehmcke S, Nakata M, Hoffrogge R, Podbielski A. Streptococcus pyogenes fibronectin-binding protein F2: expression profile, binding characteristics, and impact on eukaryotic cell interactions. J Biol Chem 2004; 279:15850–9.

34 Ljungh A, Moran AP, Wadstrom T. Interactions of bacterial adhesins with extracellular matrix and plasma proteins: pathogenic implications and therapeutic possibilities. FEMS Immunol Med Microbiol 1996; 16:117–26.

35 Niddam AF, Ebady R, Bansal A, Koehler A, Hinz B, Moriarty TJ. Plasma fibronectin stabilizes Borrelia burgdorferi–endothelial interactions under vascular shear stress by a catch-bond mechanism. Proc Natl Acad Sci U S A 2017; 114:E3490–8.

36 Steukers L, Glorieux S, Vandekerckhove AP, Favoreel HW, Nauwynck HJ. Diverse microbial interactions with the basement membrane barrier. Trends Microbiol 2012; 20:147–55.

37 Beaufort N, Wojciechowski P, Sommerhoff CP, Szmyd G, Dubin G, Eick S, et al. The human fibrinolytic system is a target for the staphylococcal metalloprotease aureolysin. Biochem J 2008; 410:157–65.

38 Barthel D, Singh B, Riesbeck K, Zipfel PF. Haemophilus influenzae uses the surface protein E to acquire human plasminogen and to evade innate immunity. J Immunol 2012; 188:379–85.

39 Beaufort N, Seweryn P, de Bentzmann S, Tang A, Kellermann J, Grebenchtchikov N, et al. Activation of human pro-urokinase by unrelated proteases secreted by Pseudomonas aeruginosa. Biochem J 2010; 428:473–82.

40 Bergmann S, Hammerschmidt S. Fibrinolysis and host response in bacterial infections.Thromb Haemost 2007; 98:512–20.

41 Lahteenmaki K, Edelman S, Korhonen TK. Bacterial metastasis: the host plasminogen system in bacterial invasion. Trends Microbiol 2005; 13:79–85.

42 Ringner M, Valkonen KH, Wadstrom T. Binding of vitronectin and plasminogen to Helicobacter pylori. FEMS Immunol Med Microbiol 1994; 9:29–34.

43 Kolar SL, Kyme P, Tseng CW, Soliman A, Kaplan A, Liang J, et al. Group B Strepto- coccus evades host immunity by degrading hyaluronan. Cell Host Microbe 2015; 18:694–704.

44 Kuo C, Lim S, King NJ, Johnston SL, Burgess JK, Black JL, et al. Rhinovirus infection induces extracellular matrix protein deposition in asthmatic and nonasthmatic airway smooth muscle cells. Am J Physiol Lung Cell Mol Physiol 2011; 300:L951–7.

45 Yin H, Li X, Hu S, Liu T, Yuan B, Ni Q, et al. IL-33 promotes Staphylococcus aureus- infected wound healing in mice. Int Immunopharmacol 2013; 17:432–8.

46 Nde PN, Lima MF, Johnson CA, Pratap S, Villalta F. Regulation and use of the extra- cellular matrix by Trypanosoma cruzi during early infection. Front Immunol 2012; 3:337.

47 Assi K, Bergstrom K, Vallance B, Owen D, Salh B. Requirement of epithelial integrin- linked kinase for facilitation of Citrobacter rodentium-induced colitis. BMC Gastroen- terol 2013; 13:137.

48 Szalay G, Sauter M, Haberland M, Zuegel U, Steinmeyer A, Kandolf R, et al. Osteo- pontin: a fibrosis-related marker molecule in cardiac remodeling of enterovirus myocarditis in the susceptible host. Circ Res 2009; 104:851–9.

49 Imanaka-Yoshida K, Hiroe M, Yasutomi Y, Toyozaki T, Tsuchiya T, Noda N, et al. Tenascin-C is a useful marker for disease activity in myocarditis. J Pathol 2002; 197:388–94.

50 Schenk S, Muser J, Vollmer G, Chiquet-Ehrismann R. Tenascin-C in serum: a ques- tionable tumor marker. Int J Cancer 1995; 61:443–9.

51 Kaarteenaho-Wiik R, Lakari E, Soini Y, Pollanen R, Kinnula VL, Paakko P. Tenascin expression and distribution in pleural inflammatory and fibrotic diseases. J Histochem Cytochem 2000; 48:1257–68.

52 Paallysaho T, Tervo K, Kivela T, Virtanen I, Tarkkanen A, Tervo T. Cellular fibronec- tin and tenascin in an orbital nylon prosthesis removed because of infection caused by Staphylococcus aureus. Graefes Arch Clin Exp Ophthalmol 1993; 231:61–5.

53 Piccinini AM, Midwood KS. Endogenous control of immunity against infection: tenascin-C regulates TLR4-mediated inflammation via microRNA-155. Cell Rep 2012; 2:914–26.

54 Hasibuan FM, Shiratori B, Senoputra MA, Chagan-Yasutan H, Koesoemadinata RC, Apriani L, et al. Evaluation of matricellular proteins in systemic and local immune response to Mycobacterium tuberculosis infection. Microbiol Immunol 2015; 59:623–32.

55 Wight TN, Frevert CW, Debley JS, Reeves SR, Parks WC, Ziegler SF. Interplay of extra- cellular matrix and leukocytes in lung inflammation. Cell Immunol 2017; 312:1–14.

56 Evanko SP, Potter-Perigo S, Bollyky PL, Nepom GT, Wight TN. Hyaluronan and ver- sican in the control of human T-lymphocyte adhesion and migration. Matrix Biol 2012; 31:90–100.

57 Overstreet MG, Gaylo A, Angermann BR, Hughson A, Hyun YM, Lambert K, et al. Inflammation-induced interstitial migration of effector CD4+ T cells is dependent on integrin aV. Nat Immunol 2013; 14:949–58.

58 Sternlicht MD, Werb Z. How matrix metalloproteinases regulate cell behavior. Annu Rev Cell Dev Biol 2001; 17:463–516.

59 McCawley LJ, Matrisian LM. Matrix metalloproteinases: they're not just for matrix anymore!. Curr Opin Cell Biol 2001; 13:534–40.

60 Van Lint P, Libert C. Chemokine and cytokine processing by matrix metalloproteinases and its effect on leukocyte migration and inflammation. J Leukoc Biol 2007; 82:1375–81.

61 Gaffney J, Solomonov I, Zehorai E, Sagi I. Multilevel regulation of matrix metallopro- teinases in tissue homeostasis indicates their molecular specificity in vivo. Matrix Biol 2015; 44–46:191–9.

62 Quiding-Jarbrink M, Smith DA, Bancroft GJ. Production of matrix metalloproteinases in response to mycobacterial infection. Infect Immun 2001; 69:5661–70.

63 Sakurai A, Okahashi N, Maruyama F, Ooshima T, Hamada S, Nakagawa I. Streptococ- cus pyogenes degrades extracellular matrix in chondrocytes via MMP-13. Biochem Bio- phys Res Commun 2008; 373:450–4.

64 Yin Y, Grabowska AM, Clarke PA, Whelband E, Robinson K, Argent RH, et al. Heli- cobacter pylori potentiates epithelial:mesenchymal transition in gastric cancer: links to soluble HB-EGF, gastrin and matrix metalloproteinase-7. Gut 2010; 59:1037–45.

65 Kirchner A, Koedel U, Fingerle V, Paul R, Wilske B, Pfister HW. Upregulation of matrix metalloproteinase-9 in the cerebrospinal fluid of patients with acute Lyme neu- roborreliosis. J Neurol Neurosurg Psychiatry 2000; 68:368–71.

66 Leib SL, Leppert D, Clements J, Tauber MG. Matrix metalloproteinases contribute to brain damage in experimental pneumococcal meningitis. Infect Immun 2000; 68:615– 20.

67 Paemen L, Jansen PM, Proost P, Van Damme J, Opdenakker G, Hack E, et al. Induc- tion of gelatinase B and MCP-2 in baboons during sublethal and lethal bacteraemia. Cytokine 1997; 9:412–5.

68 Hu VH, Weiss HA, Ramadhani AM, Tolbert SB, Massae P, Mabey DC, et al. Innate immune responses and modified extracellular matrix regulation characterize bacterial infection and cellular/connective tissue changes in scarring trachoma. Infect Immun 2012; 80:121–30.

69 Marchant DJ, Bellac CL, Moraes TJ, Wadsworth SJ, Dufour A, Butler GS, et al. A new transcriptional role for matrix metalloproteinase-12 in antiviral immunity. Nat Med 2014; 20:493–502.

70 Cheung C, Marchant D, Walker EK, Luo Z, Zhang J, Yanagawa B, et al. Ablation of matrix metalloproteinase-9 increases severity of viral myocarditis in mice. Circulation 2008; 117:1574–82.

71 Malik M, Bakshi CS, McCabe K, Catlett SV, Shah A, Singh R, et al. Matrix metalloproteinase 9 activity enhances host susceptibility to pulmonary infection with type A and B strains of Francisella tularensis. J Immunol 2007; 178:1013–20.

72 Wegner N, Wait R, Sroka A, Eick S, Nguyen KA, Lundberg K, et al. Peptidylarginine deiminase from Porphyromonas gingivalis citrullinates human fibrinogen and a-enolase: implications for autoimmunity in rheumatoid arthritis. Arthritis Rheum 2010; 62:2662–72.

73 van der Woude D, Rantapaa-Dahlqvist S, Ioan-Facsinay A, Onnekink C, Schwarte CM, Verpoort KN, et al. Epitope spreading of the anti-citrullinated protein antibody response occurs before disease onset and is associated with the disease course of early arthritis. Ann Rheum Dis 2010; 69:1554–61.

74 Tutturen AE, Fleckenstein B, de Souza GA. Assessing the citrullinome in rheumatoid arthritis synovial fluid with and without enrichment of citrullinated peptides. J Pro- teome Res 2014; 13:2867–73.

75 Chang X, Yamada R, Suzuki A, Kochi Y, Sawada T, Yamamoto K. Citrullination of fibronectin in rheumatoid arthritis synovial tissue. Rheumatology (Oxford) 2005; 44:1374–82.

76 Anzilotti C, Pratesi F, Tommasi C, Migliorini P. Peptidylarginine deiminase 4 and citrullination in health and disease. Autoimmun Rev 2010; 9:158–60.

77 Wegner N, Lundberg K, Kinloch A, Fisher B, Malmstrom V, Feldmann M, et al. Autoimmunity to specific citrullinated proteins gives the first clues to the etiology of rheumatoid arthritis. Immunol Rev 2010; 233:34–54.

78 Soares MB, de Lima RS, Rocha LL, Vasconcelos JF, Rogatto SR, dos Santos RR, et al. Gene expression changes associated with myocarditis and fibrosis in hearts of mice with chronic chagasic cardiomyopathy. J Infect Dis 2010; 202:416–26.

79 Erler JT, Bennewith KL, Nicolau M, Dornhofer N, Kong C, Le QT, et al. Lysyl oxidase is essential for hypoxia-induced metastasis. Nature 2006; 440:1222–6.

80 Afroz S, Giddaluru J, Abbas MM, Khan N. Transcriptome meta-analysis reveals a dysregulation in extra cellular matrix and cell junction associated gene signatures during Dengue virus infection. Sci Rep 2016; 6:33752.

81 Mammoto T, Parikh SM, Mammoto A, Gallagher D, Chan B, Mostoslavsky G, et al. Angiopoietin-1 requires p190 RhoGAP to protect against vascular leakage in vivo. J Biol Chem 2007; 282:23910–8.

82 Mammoto A, Mammoto T, Kanapathipillai M, Wing Yung C, Jiang E, Jiang A, et al. Control of lung vascular permeability and endotoxin-induced pulmonary oedema by changes in extracellular matrix mechanics. Nat Commun 2013; 4:1759.

83 Mosser DM, Edwards JP. Exploring the full spectrum of macrophage activation. Nat Rev Immunol 2008; 8:958–69.

84 Alitalo K, Hovi T, Vaheri A. Fibronectin is produced by human macrophages. J Exp Med 1980; 151:602–13.

85 Cofano F, Comoglio PM, Landolfo S, Tarone G. Mouse immune interferon enhances fibronectin production of elicited macrophages. J Immunol 1984; 133:3102–6.

86 Goh FG, Piccinini AM, Krausgruber T, Udalova IA, Midwood KS. Transcriptional regulation of the endogenous danger signal tenascin-C: a novel autocrine loop in inflammation. J Immunol 2010; 184:2655–62.

87 Gokyu M, Kobayashi H, Nanbara H, Sudo T, Ikeda Y, Suda T, et al. Throm- bospondin-1 production is enhanced by Porphyromonas gingivalis lipopolysaccharide in THP-1 cells. PLoS One 2014; 9:e115107.

88 Novak R, Dabelic S, Dumic J. Galectin-1 and galectin-3 expression profiles in classically and alternatively activated human macrophages. Biochim Biophys Acta 2012; 1820:1383–90.

89 Zhao W, Wang L, Zhang M, Wang P, Zhang L, Yuan C, et al. NF-jB- and AP-1- mediated DNA looping regulates osteopontin transcription in endotoxin-stimulated murine macrophages. J Immunol 2011; 186:3173–9.

90 Song H, Deng B, Zou C, Huai W, Zhao R, Zhao W. GSK3b negatively regulates LPS-induced osteopontin expression via inhibiting its transcription. Scand J Immunol 2015; 81:186–91.

91 Chang MY, Tanino Y, Vidova V, Kinsella MG, Chan CK, Johnson PY, et al. Reprint of: A rapid increase in macrophage-derived versican and hyaluronan in infectious lung disease. Matrix Biol 2014; 35:162–73.

92 Zernichow L, Abrink M, Hallgren J, Grujic M, Pejler G, Kolset SO. Serglycin is the major secreted proteoglycan in macrophages and has a role in the regulation of macrophage tumor necrosis factor-a secretion in response to lipopolysaccharide. J Biol Chem 2006; 281:26792–801.

93 Merline R, Moreth K, Beckmann J, Nastase MV, Zeng-Brouwers J, Tralhao JG, et al. Signaling by the matrix proteoglycan decorin controls inflammation and cancer through PDCD4 and MicroRNA-21. Sci Signal 2011; 4:ra75.

94 Zeng-Brouwers J, Beckmann J, Nastase MV, Iozzo RV, Schaefer L. De novo expression of circulating biglycan evokes an innate inflammatory tissue response via MyD88/ TRIF pathways. Matrix Biol 2014; 35:132–42.

95 Piccinini AM, Zuliani-Alvarez L, Lim JM, Midwood KS. Distinct microenvironmental cues stimulate divergent TLR4-mediated signaling pathways in macrophages. Sci Signal 2016; 9:ra86.

96 Van De Bor V, Zimniak G, Papone L, Cerezo D, Malbouyres M, Juan T, et al. Companion blood cells control ovarian stem cell niche microenvironment and homeosta- sis. Cell Rep 2015; 13:546–60.

97 Vaage J, Lindblad WJ. Production of collagen type I by mouse peritoneal macrophages. J Leukoc Biol 1990; 48:274–80.

98 Weitkamp B, Cullen P, Plenz G, Robenek H, Rauterberg J. Human macrophages syn- thesize type VIII collagen in vitro and in the atherosclerotic plaque. FASEB J 1999; 13:1445–57.

99 Schnoor M, Cullen P, Lorkowski J, Stolle K, Robenek H, Troyer D, et al. Production of type VI collagen by human macrophages: a new dimension in macrophage func- tional heterogeneity. J Immunol 2008; 180:5707–19.

100 Andersson E, Rydengard V, Sonesson A, Morgelin M, Bjorck L, Schmidtchen A. Antimicrobial activities of heparin-binding peptides. Eur J Biochem 2004; 271:1219–26.

101 Malmsten M, Davoudi M, Schmidtchen A. Bacterial killing by heparin-binding pep- tides from PRELP and thrombospondin. Matrix Biol 2006; 25:294–300.

102 Kobayashi N, Yoshida T. Binding sites on laminin receptors as components for antibi- otics. Protein Pept Lett 2007; 14:33–6.

103 Fouda GG, Jaeger FH, Amos JD, Ho C, Kunz EL, Anasti K, et al. Tenascin-C is an innate broad-spectrum, HIV-1-neutralizing protein in breast milk. Proc Natl Acad Sci U S A 2013; 110:18220–5.

104 Houghton AM, Hartzell WO, Robbins CS, Gomis-Ruth FX, Shapiro SD. Macrophage elastase kills bacteria within murine macrophages. Nature 2009; 460:637–41.

105 He YW, Li H, Zhang J, Hsu CL, Lin E, Zhang N, et al. The extracellular matrix protein mindin is a pattern-recognition molecule for microbial pathogens. Nat Immunol 2004; 5:88–97.

106 Jia W, Li H, He YW. Pattern recognition molecule mindin promotes intranasal clearance of influenza viruses. J Immunol 2008; 180:6255–61.

107 Wu F, Vij N, Roberts L, Lopez-Briones S, Joyce S, Chakravarti S. A novel role of the lumican core protein in bacterial lipopolysaccharide-induced innate immune response. J Biol Chem 2007; 282:26409–17.

108 Shao H, Lee S, Gae-Scott S, Nakata C, Chen S, Hamad AR, et al. Extracellular matrix lumican promotes bacterial phagocytosis, and Lum–/– mice show increased Pseudomonas aeruginosa lung infection severity. J Biol Chem 2012; 287:35860–72.

109 Shao H, Scott SG, Nakata C, Hamad AR, Chakravarti S. Extracellular matrix protein lumican promotes clearance and resolution of Pseudomonas aeruginosa keratitis in a mouse model. PLoS One 2013; 8:e54765.

110 van den Berg TK, Honing H, Franke N, van Remoortere A, Schiphorst WE, Liu FT, et al. LacdiNAc-glycans constitute a parasite pattern for galectin-3-mediated immune recognition. J Immunol 2004; 173:1902–7.

111 Sano H, Hsu DK, Apgar JR, Yu L, Sharma BB, Kuwabara I, et al. Critical role of galectin-3 in phagocytosis by macrophages. J Clin Invest 2003; 112:389–97.

112 Kanayama M, Xu SJ, Danzaki K, Gibson JR, Inoue M, Gregory SG, et al. Skewing of the population balance of lymphoid and myeloid cells by secreted and intracellular osteopontin. Nat Immunol 2017; 18:973–84.

113 Ruhmann M, Piccinini AM, Kong PL, Midwood KS. Endogenous activation of adaptive immunity: tenascin-C drives interleukin-17 synthesis in murine arthritic joint disease. Arthritis Rheum 2012; 64:2179–90.

114 Breuilh L, Vanhoutte F, Fontaine J, van Stijn CM, Tillie-Leblond I, Capron M, et al. Galectin-3 modulates immune and inflammatory responses during helminthic infection: impact of galectin-3 deficiency on the functions of dendritic cells. Infect Immun 2007; 75:5148–57.

115 Li H, Oliver T, Jia W, He YW. Efficient dendritic cell priming of T lymphocytes depends on the extracellular matrix protein mindin. EMBO J 2006; 25:4097–107.

116 Piccinini AM, Midwood KS. Illustrating the interplay between the extracellular matrix and microRNAs. Int J Exp Pathol 2014; 95:158–80.

117 Carlson I, Tognazzi K, Manseau EJ, Dvorak HF, Brown LF. Osteopontin is strongly expressed by histiocytes in granulomas of diverse etiology. Lab Invest 1997; 77:103–8.

118 Nishikaku AS, Scavone R, Molina RF, Albe BP, Cunha Cda S, Burger E. Osteopontin involvement in granuloma formation and in the severity of Paracoccidioides brasiliensis infection. Med Mycol 2009; 47:495–507.

119 Kaarteenaho-Wiik R, Sademies O, Paakko P, Risteli J, Soini Y. Extracellular matrix proteins and myofibroblasts in granulomas of sarcoidosis, atypical mycobacteriosis, and tuberculosis of the lung. Hum Pathol 2007; 38:147–53.

120 Gonzalez A, Lenzi HL, Motta EM, Caputo L, Restrepo A, Cano LE. Expression and arrangement of extracellular matrix proteins in the lungs of mice infected with Paracoccidioides brasiliensis conidia. Int J Exp Pathol 2008; 89:106–16.

121 Sato J, Schorey J, Ploplis VA, Haalboom E, Krahule L, Castellino FJ. The fibrinolytic system in dissemination and matrix protein deposition during a mycobacterium infection. Am J Pathol 2003; 163:517–31.

122 Elkington P, Shiomi T, Breen R, Nuttall RK, Ugarte-Gil CA, Walker NF, et al. MMP- 1 drives immunopathology in human tuberculosis and transgenic mice. J Clin Invest 2011; 121:1827–33.

123 Al Shammari B, Shiomi T, Tezera L, Bielecka MK, Workman V, Sathyamoorthy T, et al. The extracellular matrix regulates granuloma necrosis in tuberculosis. J Infect Dis 2015; 212:463–73.

124 Tezera LB, Bielecka MK, Chancellor A, Reichmann MT, Shammari BA, Brace P, et al. Dissection of the host-pathogen interaction in human tuberculosis using a bioengi- neered 3-dimensional model. Elife. 2017; 6:e21283.

125 Parasa VR, Muvva JR, Rose JF, Braian C, Brighenti S, Lerm M. Inhibition of tissue matrix metalloproteinases interferes with Mycobacterium tuberculosis-induced granuloma formation and reduces bacterial load in a human lung tissue model. Front Microbiol 2017; 8:2370.

126 Marbiah MM, Harvey A, West BT, Louzolo A, Banerjee P, Alden J, et al. Identification of a gene regulatory network associated with prion replication. EMBO J 2014; 33:1527–47.

127 Pais-Correia AM, Sachse M, Guadagnini S, Robbiati V, Lasserre R, Gessain A, et al. Biofilm-like extracellular viral assemblies mediate HTLV-1 cell-to-cell transmission at virological synapses. Nat Med 2010; 16:83–9.

128 Kulkarni R, Prasad A. Exosomes derived from HIV-1 infected DCs mediate viral trans-infection via fibronectin and galectin-3. Sci Rep 2017; 7:14787.

129 Sun X, Mei M, Zhang X, Han F, Jia B, Wei X, et al. The extracellular matrix protein mindin as a novel adjuvant elicits stronger immune responses for rBAG1, rSRS4 and rSRS9 antigens of Toxoplasma gondii in BALB/c mice. BMC Infect Dis 2014; 14:429.

130 Farfan MJ, Cantero L, Vidal R, Botkin DJ, Torres AG. Long polar fimbriae of enterohemorrhagic Escherichia coli O157:H7 bind to extracellular matrix proteins. Infect Immun 2011; 79:3744–50.

131 Ensenberger MG, Tomasini-Johansson BR, Sottile J, Ozeri V, Hanski E, Mosher DF. Specific interactions between F1 adhesin of Streptococcus pyogenes and N-terminal modules of fibronectin. J Biol Chem 2001; 276:35606–13.

132 Ozeri V, Tovi A, Burstein I, Natanson-Yaron S, Caparon MG, Yamada KM, et al. A two-domain mechanism for group A streptococcal adherence through protein F to the extracellular matrix. EMBO J 1996; 15:989–98.

133 Cue D, Southern SO, Southern PJ, Prabhakar J, Lorelli W, Smallheer JM, et al. A non-peptide integrin antagonist can inhibit epithelial cell ingestion of Streptococcus pyogenes by blocking formation of integrin a5b1-fibronectin-M1 protein complexes. Proc Natl Acad Sci U S A 2000; 97:2858–63.

134 Kreikemeyer B, Talay SR, Chhatwal GS. Characterization of a novel fibronectin-binding surface protein in group A streptococci. Mol Microbiol 1995; 17:137–45.

135 Salazar N, Castiblanco-Valencia MM, da Silva LB, de Castro I, Monaris D, Masuda HP, et al. Staphylococcus aureus manganese transport protein C (MntC) is an extracel- lular matrix- and plasminogen-binding protein. PLoS One 2014; 9:e112730.

136 Dorsey CW, Laarakker MC, Humphries AD, Weening EH, Baumler AJ. Salmonella enterica serotype Typhimurium MisL is an intestinal colonization factor that binds fibronectin. Mol Microbiol 2005; 57:196–211.

137 Kingsley RA, Keestra AM, de Zoete MR, Baumler AJ. The ShdA adhesin binds to the cationic cradle of the fibronectin 13FnIII repeat module: evidence for molecular mimicry of heparin binding. Mol Microbiol 2004; 52:345–55.

138 Vollmer T, Hinse D, Kleesiek K, Dreier J. Interactions between endocarditis-derived Streptococcus gallolyticus subsp. gallolyticus isolates and human endothelial cells. BMC Microbiol 2010; 10:78.

139 Brissette CA, Verma A, Bowman A, Cooley AE, Stevenson B. The Borrelia burgdorferi outer-surface protein ErpX binds mammalian laminin. Microbiology 2009; 155:863–72.

140 Montealegre MC, Singh KV, Somarajan SR, Yadav P, Chang C, Spencer R, et al. Role of the Emp Pilus subunits of Enterococcus faecium in biofilm formation, adherence to host extracellular matrix components, and experimental infection. Infect Immun 2016; 84:1491–500.

141 Hendrickx AP, van Luit-Asbroek M, Schapendonk CM, van Wamel WJ, Braat JC, Wij- nands LM, et al. SgrA, a nidogen-binding LPXTG surface adhesin implicated in bio- film formation, and EcbA, a collagen binding MSCRAMM, are two novel adhesins of hospital-acquired Enterococcus faecium. Infect Immun 2009; 77:5097–106.

142 Wagner C, Khan AS, Kamphausen T, Schmausser B, Unal C, Lorenz U, et al. Collagen binding protein Mip enables Legionella pneumophila to transmigrate through a barrier of NCI-H292 lung epithelial cells and extracellular matrix. Cell Microbiol 2007; 9:450–62.

143 Muhlenkamp MC, Hallstrom T, Autenrieth IB, Bohn E, Linke D, Rinker J, et al. Vitronec- tin binds to a specific stretch within the head region of Yersinia Adhesin A and thereby modulates Yersinia enterocolitica host interaction. J Innate Immun 2017; 9:33–51.

144 Imai DM, Feng S, Hodzic E, Barthold SW. Dynamics of connective-tissue localization during chronic Borrelia burgdorferi infection. Lab Invest 2013; 93:900–10.

145 Eidhin DN, Perkins S, Francois P, Vaudaux P, Hook M, Foster TJ. Clumping factor B (ClfB), a new surface-located fibrinogen-binding adhesin of Staphylococcus aureus. Mol Microbiol 1998; 30:245–57.

146 McDevitt D, Nanavaty T, House-Pompeo K, Bell E, Turner N, McIntire L, et al. Char- acterization of the interaction between the Staphylococcus aureus clumping factor (ClfA) and fibrinogen. Eur J Biochem 1997; 247:416–24.

147 Heck LW, Morihara K, Abrahamson DR. Degradation of soluble laminin and depletion of tissue-associated basement membrane laminin by Pseudomonas aeruginosa elas- tase and alkaline protease. Infect Immun 1986; 54:149–53.

148 Janoir C, Pechine S, Grosdidier C, Collignon A. Cwp84, a surface-associated protein of Clostridium difficile, is a cysteine protease with degrading activity on extracellular matrix proteins. J Bacteriol 2007; 189:7174–80.

149 Sofat N, Wait R, Robertson SD, Baines DL, Baker EH. Interaction between extracellular matrix molecules and microbial pathogens: evidence for the missing link in autoimmunity with rheumatoid arthritis as a disease model. Front Microbiol 2014; 5:783.

150 Kim SK, Yang JY, Cha J. Cloning and sequence analysis of a novel metalloprotease gene from Vibrio parahaemolyticus 04. Gene 2002; 283:277–86.

151 Shi L, Ermis R, Garcia A, Telgenhoff D, Aust D. Degradation of human collagen iso- forms by Clostridium collagenase and the effects of degradation products on cell migration. Int Wound J 2010; 7:87–95.

152 Watanabe K. Collagenolytic proteases from bacteria. Appl Microbiol Biotechnol 2004;
63:520–6.

153 Juarez ZE, Stinson MW. An extracellular protease of Streptococcus gordonii hydrolyzes type IV collagen and collagen analogues. Infect Immun 1999; 67:271–8.

154 Ruggiero S, Cosgarea R, Potempa J, Potempa B, Eick S, Chiquet M. Cleavage of extra- cellular matrix in periodontitis: gingipains differentially affect cell adhesion activities of fibronectin and tenascin-C. Biochim Biophys Acta 2013; 1832:517–26.