


# CRISPR-CAS9-MEDIATED GENOME EDITING OF THE EXTRACELLULAR MATRIX GENE *TNC* IN HUMAN FIBROBLASTS

**GIULIO BRUNETTI**

**Thesis submitted to the University of Nottingham for the degree of Master of Research**

**December 2016**

**Giulio Brunetti**

**University of Nottingham, MRes in Advanced Genomic and Proteomic Science**

**CRISPR-Cas9-mediated genome editing of the extracellular matrix gene *TNC* in human fibroblasts**

## Abstract

Mammalian cells exist in complex three-dimensional (3D) microenvironments, in which diverse signalling molecules and support structures dynamically interact forming a complex network or extracellular matrix (ECM). Beyond providing structural support to cells, the matrix conveys environmental signals that direct many aspects of normal cell behaviour, including shape, differentiation, survival and migration. Tenascin-C (TN-C), an ECM glycoprotein absent in most healthy adult tissues, but persistently expressed during chronic inflammatory diseases and in cancer, has been recently shown to be rapidly induced during the inflammatory response to infection. TN-C is required for an effective pro-inflammatory response to bacterial lipopolysaccharide (LPS) *in vivo*. It sustains the synthesis of the pro-inflammatory cytokine TNF-$\alpha$ by regulating the expression of the microRNA miR-155 in primary murine macrophages upon LPS stimulation. These data highlight a novel, regulatory function of the ECM molecule TN-C that controls inflammatory gene expression by regulating microRNA levels in the mouse. However, we still do not know whether this occurs in human cells and how. Here, I have applied the CRISPR-Cas9-mediated genome editing technology to knock-out *TNC* gene expression in human fibroblast cells, which are a natural and abundant source of ECM. Once validated, these will be used to generate a cell-free fibroblast-derived matrix on which primary human macrophages will be cultured in the presence or absence of infection. This system will allow to investigate whether human TN-C modulates macrophage gene expression by regulating microRNA levels. The HEK 293 cell line was used to design and optimise a CRISPR-Cas9 system that can efficiently knock-out the *TNC* gene.

Three different pairs of oligonucleotides were designed to guide the Cas9 endonuclease to the target gene sequence. The oligonucleotides were cloned into a mammalian expression vector and delivered to HEK 293 cells together with Cas9. PCR amplification and sequencing, in combination with fluorescence in situ hybridization (FISH), were used to validate the genome edit. One oligonucleotide pair edited the *TNC* gene in HEK 293 cells by introducing a premature stop codon at the translation start site. Subsequently, three different approaches were tested to deliver the oligonucleotides and Cas9 to the difficult-to-transfect human fibroblasts. I was able to transfect them with a transfection efficiency of 40-50%. Single cell sorting of transfected cells was performed to generate single-cell clones and a number of single-fibroblast clones are currently growing. Their expansion will allow validation of the genome edit. Together this work has generated a CRISPR-Cas9 system for the specific knock-out of the *TNC* gene in mammalian cells and single-fibroblast clones whose *TNC* expression may be ablated.

## Acknowledgements

*First and foremost, I would like to thank my supervisor Dr. Anna M. Piccinini, School of Pharmacy, University of Nottingham, for her continuous support during my MRes experience. She explained me how to approach several and different research problems. Only with her incessant and generous help, expert knowledge and friendly guidance I could successfully accomplish my goal. I am very grateful to her, because I found a great supervisor, but most important, a friend.*

*I am also thankful to Prof. David M. Heery, School of Pharmacy, University of Nottingham, for his continuous support, kind suggestions and supervision, and for providing all the facilities for my research work.*

*I also would like to express my happiness on the wonderful company found in this laboratory. This is a great group that allowed me to work properly and supported me during this period.*

*I would like to thank my girlfriend Ilaria, with whom I shared this great experience. Her words of appreciation and encouragement were fundamental for me. She was a warm hug in the toughest moments.*

*I dedicate this work and experience to my family. They supported my in each possible way during this year, and I am grateful to them for it.*

# List of contents

**Abstract..... 2**

**Acknowledgments..... 4**

**List of contents..... 5**

**List of figures..... 8**

**List of tables..... 10**

**Abbreviations..... 11**

**1 Introduction..... 13**

1.1 The extracellular matrix (ECM)..... 13

1.1.1 Structure and function of the ECM..... 13

1.1.2 Regulation of the ECM..... 16

1.1.3 Matricellular proteins..... 17

1.2 Tenascin-C (TN-C)..... 18

1.2.1 TN-C in disease..... 21

1.2.2 TN-C and the innate immune response..... 22

1.2.3 Regulation of microRNA expression: does TN-C play a role? ..... 24

1.3 CRISPR-Cas9 genome editing technology.....25

1.3.1 CRISPR-Cas9: a historical overview.....25

1.3.2 The CRISPR-Cas9 nickase: a specific genome editing tool.....27

1.3.3 Applications of CRISPR-Cas9.....28

1.4 Aims.....30

**2 Materials and methods.....31**

2.1 Reagents and sources of materials.....31

2.1.1 Reagents and sources of materials.....31

2.1.2 Molecular biology reagents.....31

2.1.3 Protein Chemistry, Western blotting reagents and equipment.....31

2.1.4 Tissue culture reagents.....32

2.1.5 Composition of solutions and buffers.....32

2.2 Cell culture.....34

2.2.1 Cell lines.....34

2.2.2 Cell culture.....34

2.2.3 Cells cryopreservation.....35

2.2.4 Cell stimulation.....35

2.2.5 Cell lysis.....35

2.3 CRISPR-Cas9 gRNA design and synthesis.....35
2.4 Cloning of gRNAs into CRISPR-Cas9 vector.....36
2.4.1 Annealing of oligonucleotides, restriction digestion and ligation.....36
2.4.2 Agarose gel electrophoresis and plasmid DNA purification from agarose gel.....36
2.4.3 LB agarose plates preparation and bacterial transformation....38
2.4.4 Glycerol stock.....38
2.4.5 Minipreps of plasmid DNA and sequencing.....39
2.4.6 Midipreps of plasmid DNA.....40
2.5 Transfection of cell lines.....41
2.5.1 Transfection of HEK 293 cells.....41
2.5.2 Transfection of BJ fibroblasts.....41
2.5.2.1 Cationic lipid-based transfection.....42
2.5.2.2 Non-lipid-based chemical transfection.....43
2.5.2.3 Electroporation.....44
2.6 Selection of transfected cells.....45
2.6.1 Fluorescence Activated Cell Sorting (FACS) .....45
2.7 Validation of genome edit.....46
2.7.1 Gene copy number estimation by Fluorescence in Situ Hybridization (FISH) .....46
2.7.1.1 Cell preparation.....46
2.7.1.2 Hybridization procedure.....46
2.7.2 Polymerase Chain Reaction (PCR) amplification, gel electrophoresis and sequencing.....47
2.7.2.1 Design and validation of PCR primers.....47
2.7.2.2 PCR amplification and sequencing.....51
2.8 Protein Analysis.....52
2.8.1 Tris-glycine Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE).....52
2.8.2 Western blotting).....52
2.9 Gene expression quantification.....53
2.9.1 RNA extraction and purification.....53
2.9.2 Reverse transcription.....54
2.9.3 Quantitative real time polymerase chain reaction.....54
2.10 Preparation of cell-derived matrices.....54
2.10.1 Matrix decellularization.....55
2.10.2 ECM extraction.....55
2.10.3 Preparation of coverslips.....56
2.10.3 Immunofluorescence (IF) analysis of cell-derived matrices.....56
**3 Results.....58**

**Generation of human *TNC* knock-out fibroblasts by CRISPR-Cas9-mediated genome editing**.....58
3.1 Introduction.....58
3.2 Analysis of TN-C expression in human BJ fibroblasts and HEK 293 cells.....58
3.2.1 BJ fibroblasts synthesise and secrete TN-C into the extracellular space.....59
3.2.2 HEK 293 cells express *TNC* mRNA.....62
3.3 Design of a CRISPR-Cas9 system to knock-out the *TNC* gene.....64
3.3.1 Design of gRNAs.....64
3.3.2 Synthesis of gRNAs.....69
3.3.3 Cloning of gRNAs.....69
3.4 Delivering Cas9n and gRNAs.....72
3.4.1 Transfection of Cas9n and gRNAs into HEK 293 cells.....72
3.4.2 Transfection of Cas9n and gRNAs into BJ fibroblasts.....75
3.5 Validation of *TNC* gene edit.....78
3.5.1 Estimation of gene copy number in HEK 293 cells and BJ fibroblasts by FISH.....78
3.5.2 Validation of *TNC* gene edit in HEK 293 cells by PCR amplification and sequencing.....83
3.5.2.1 Design of gene edit screening primers.....83
3.5.2.2 Validation of gene edit screening primers.....86
3.5.2.3 PCR amplification and sequencing.....87
3.5.3 Validation of *TNC* gene knock-out in HEK 293 cells by RT-qPCR.....92
3.5.4 Validation of *TNC* gene knock-out in HEK 293 cells by immunoblot analysis.....95
3.6 Generation of single-BJ cell clones with *TNC* knock-out expression....95
3.6.1 Selection of transfected cells by Fluorescence Activated Cell Sorting (FACS).....95
3.6.3 Validation of *TNC* gene edit in a bulk population of $GFP^+$ BJ fibroblasts.....97
3.6.4 Expansion of single-BJ cell clones.....101
**4 Discussion.....103**
4.1 General discussion.....103
4.2 Limitations and future work.....104
**5 References.....107**

## List of figures

**Figure 1.1** Model of the molecular structure of the basal lamina
**Figure 1.2** Fibroblast and the surrounding ECM
**Figure 1.3** Schematic representation of the multiple functions of the ECM at the cell-matrix interface
**Figure 1.4** Genomic context of *TNC*
**Figure 1.5** Schematic representation of the *TNC* gene
**Figure 1.6** Schematic representation of *TNC* gene promoters
**Figure 1.7** Schematic representation of the multidomain structure of TN-C
**Figure 1.8** Structure of the TN-C hexabrachion
**Figure 1.9** Schematic representation of CRISPR-Cas9 nickase system
**Figure 1.10** CRISPR-Cas9n-mediated genome editing
**Figure 3.1** BJ fibroblast express TN-C and deposit an extracellular matrix that contains TN-C
**Figure 3.2** HEK 293 cells express *TNC* at the transcriptional level
**Figure 3.3** Design of dual nickase CRISPR system to knockout *TNC* gene
**Figure 3.4** Schematic representation of selected targets
**Figure 3.5** Synthesis of gRNA for cloning
**Figure 3.6** Restriction digest of pSpCas9n(BB)-2A-GFP plasmid with *BbSI* restriction enzyme
**Figure 3.7** Cloning of gRNAs
**Figure 3.8** Delivery of CRISPR-Cas9n and gRNAs into HEK 293 cells by calcium phosphate transfection
**Figure 3.9** Transient expression of reporter GFP in BJ fibroblasts following transfection of pSpCas9n(BB)-2A-GFP plasmid containing gRNA A1 by electroporation
**Figure 3.10** Effect of the total amount of plasmid DNA on BJ fibroblast viability following electroporation
**Figure 3.11** Estimation of target chromosome 9 copy number in HEK 293 cells by fluorescence in situ hybridization (FISH)
**Figure 3.12** Estimation of target chromosome 9 copy number in HEK 293 cells by fluorescence in situ hybridization (FISH)
**Figure 3.13** Estimation of target chromosome 9 copy number in BJ fibroblast by fluorescence in situ hybridization (FISH)
**Figure 3.14** Estimation of target chromosome 9 copy number in BJ fibroblast by fluorescence in situ hybridization (FISH)
**Figure 3.15** Design of primers for validation of *TNC* gene edit by PCR amplification

**Figure 3.16** Optimization of PCR protocols using primers designed above for the validation of *TNC* gene edit
**Figure 3.17** Validation of CRISPR-Cas9n-gRNA pair A-mediated *TNC* gene edit in HEK 293 cells by PCR amplification and sequencing
**Figure 3.18** Validation of CRISPR-Cas9n-gRNA pair B-mediated *TNC* gene edit in HEK 293 cells by PCR amplification and sequencing
**Figure 3.19** Validation of CRISPR-Cas9n-gRNA-mediated *TNC* gene edit in HEK 293 cells by RT-qPCR
**Figure 3.20** Validation of CRISPR-Cas9n-gRNA-mediated *TNC* gene edit in HEK 293 cells by Western blotting
**Figure 3.21** Fluorescence activated cell sorting (FACS) of BJ fibroblast 48 hours after transfection
**Figure 3.22** Single cell sorting of $GFP^+$ BJ fibroblasts 48 hours after transfection
**Figure 3.23** Validation of CRISPR-Cas9n-gRNA pair A-mediated *TNC* gene edit in BJ fibroblasts by PCR amplification and sequencing
**Figure 3.24** Validation of CRISPR-Cas9n-gRNA pair B-mediated *TNC* gene edit in BJ fibroblasts by PCR amplification and sequencing
**Figure 3.25** Establishment of clonal cell lines

## List of tables

**Table 1** Details of antibodies used for Western blotting
**Table 2** Details of CRISPR-Cas9n gRNAs synthesized
**Table 3** Details of plasmid DNA preparations transfected into human HEK 293 cells
**Table 4** Details of lipid-based transfection reactions
**Table 5** Details of non-lipid based chemical transfection reactions
**Table 6** Details of primers for CRISPR-Cas9n gene edit validation
**Table 7** Touchdown PCR protocol used to amplify the *TNC* genomic region edited by the CRISPR-Cas9n-gRNA pair A system using primers 1Fw and 1Rv
**Table 8** Touchdown PCR protocol used to amplify the *TNC* genomic region edited by the CRISPR-Cas9n-gRNA pair A system using primers 2Fw and 2Rv
**Table 9** Touchdown PCR protocol used to amplify the *TNC* genomic region edited by the CRISPR-Cas9n-gRNA pair B system using primers 1Fw and 1Rv
**Table 10** Touchdown PCR protocol used to amplify the *TNC* genomic region edited by the CRISPR-Cas9n-gRNA pair B system using primers 2Fw and 2Rv
**Table 11** Touchdown PCR protocol used to amplify the *TNC* genomic region edited by the CRISPR-Cas9n-gRNA pair B system using primers 3Fw and 3Rv
**Table 12** Composition of 5% SDS-PAGE gels
**Table 13** Details of antibodies used for Western blotting
**Table 14** Plasmid DNA concentration of minipreps
**Table 15** Plasmid DNA concentration of midipreps
**Table 16** Details of primary antibodies used for IF analysis

## Abbreviations

| | |
|---|---|
| **AAV** | Adeno-associated virus |
| **bp** | Base pair |
| **CRISPR** | Clustered Regularly Interspaced Short Palindromic Repeats |
| **CRISPR-Cas9** | CRISPR-associated protein-9 nuclease |
| **CT** | Cycle threshold |
| **DMEM** | Dulbecco's Modified Eagle Medium |
| **DSB** | Double strand break |
| **ECL** | Enhanced chemiluminescence |
| **ECM** | Extracellular matrix |
| **EMEM** | Eagle's Minimun Essental Medium |
| **FACS** | Fluorescence Activated Cell Sorting |
| **FBS** | Fetal Bovine Serum |
| **FISH** | Fluorescence In Situ Hybridization |
| **FPKM** | Fragments per kilobase per million |
| **Fw** | Forward |
| **HEK** | Human Embryonic Kidney |
| **GFP** | Green Fluorescent Protein |
| **gRNA** | Guide RNA |
| **Kb** | Kilobase |
| **kDa** | Kilodalton |
| **LPS** | Lipopolysaccharide |
| **M** | Molar |
| **mA** | Milli ampere |
| **mg** | Milligram |
| **miR** | MicroRNA |
| **mL** | Milliliter |
| **mM** | Millimolar |
| **MMP** | Metalloproteinase |
| **ng** | Nanogram |
| **ORF** | Open reading frame |
| **PAM** | Protospacer adjacent motif |
| **PBS** | Phosphate Buffer Saline |
| **PBST** | Phosphate Buffer Saline (Tween-20) |
| **PCR** | Polymerase Chain Reaction |
| **qPCR** | Quantitative PCR |
| **RNAi** | RNA interference |

| | |
|---|---|
| **RT-qPCR** | Real Time qPCR |
| **Rv** | Reverse |
| **SDS** | Sodium Dodecyl Sulfate |
| **Sec** | Seconds |
| **TAE** | Tris-Acetate-EDTA |
| **TEMED** | Tetramethylethylenediamine |
| **TN-C** | Tenascin-C |
| **TNC** | Tenascin-C gene |
| **TSS** | Transcription start site |
| **°C** | Degree Centigrade |
| **µL** | Microliter |
| **µM** | Micromolar |
| **µg** | Microgram |

# 1 Introduction

## 1.1 The extracellular matrix (ECM)

### 1.1.1 Structure and function of the ECM

The extracellular matrix (ECM) is a complex three-dimensional (3D) network of molecules with a highly regulated and tissue-specific composition that is located around and between cells. Two major classes of macromolecules compose the ECM: polysaccharides and fibrous proteins. The first class of macromolecules includes polysaccharides that are covalently bound to protein ligands, namely proteoglycans, and the polysaccharide hyaluronic acid, which lacks a protein core. Hyaluronic acid molecules form a gel that help resist compressive forces on the matrix and allows rapid and easy diffusion of nutrients, metabolites and hormones. By virtue of their net negative charge, proteoglycans attracts positively charged sodium ions ($Na^+$), which attracts water molecules via osmosis, keeping the ECM and resident cells hydrated. The second class of macromolecules includes collagens, which give structural support to the cells; elastin, which confers elasticity to tissues; laminin, which form web-like structures that resist tensile forces in the basal lamina; and fibronectin, a glycoprotein that, by binding collagen and cell surface integrins, facilitates cell movement through the ECM.

The molecular composition of the ECM and its architecture vary from one anatomical site to the other. However, two main types of ECM can be distinguished: the basal lamina and the interstitial matrix. The basal lamina is a thin, flexible mat of specialised ECM that divides the epithelium (e.g. skin, respiratory and gastrointestinal tracts), mesothelium (e.g. pleural and peritoneal cavities) and endothelium (e.g. blood and lymph vessels) from underlying or surrounding connective tissue. It is made mainly of type IV collagen, laminin, nidogen, and the heparan sulfate proteoglycan perlecan (Fig. 1.1).

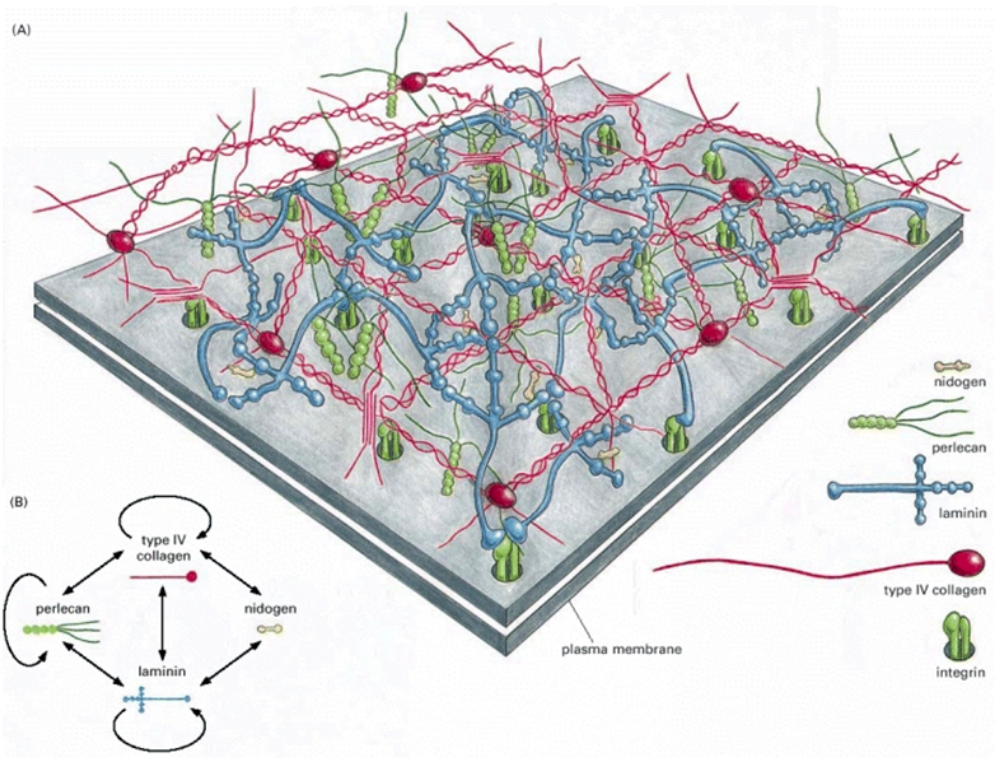



**Figure 1.1. Model of the molecular structure of the basal lamina.** (A) Specific interactions (B) between type IV collagen, laminin, nidogen, and perlecan form the basal lamina. (B) Arrows indicate direct binding. Transmembrane laminin receptors (integrins are shown) in the plasma membrane are thought to organize the assembly of the basal lamina. Adapted rom H. Colognato and P.D. Yurchenco, Dev. Dynamics 218:213–234, 2000.

The interstitial matrix is present in the intercellular spaces and its major components include polysaccharides and, among fibrous proteins, type I collagen, which is the most abundant (Fig. 1.2)(Frantz et al., 2010).

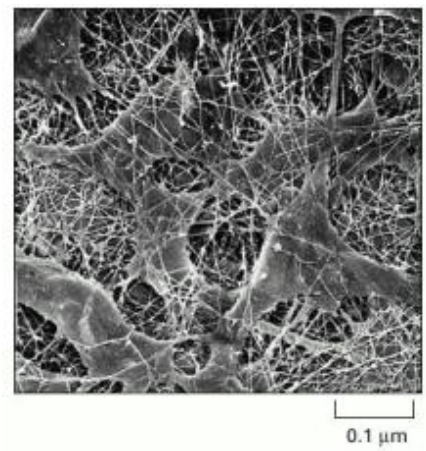



**Figure 1.2. Fibroblasts and the surrounding ECM.** Scanning electron micrograph of a rat cornea. The extracellular matrix surrounding the fibroblasts is composed largely of collagen. Glycoproteins, polysaccharides and proteoglycans were removed by enzyme and acid treatment. Adapted from T. Nishida et al., Invest. Ophthalmol. Vis. Sci. 29:1887–1890, 1988.

Until the early 1990s, the ECM was considered a scaffold, which provided mechanical strength and support to tissues and organs, or a substrate that cells used to migrate. The last two decades have seen a deep change in the way we regard the ECM and so has the research around it. From being entirely focused on the structural function of the ECM, research in this field is now oriented towards the intracellular signalling cascades generated or modulated by the ECM.

The ECM plays an active role in tissue development, where cell-cell and cell-ECM communication direct organization of cells into tissues and organogenesis, homeostasis, repair and remodelling. In the tissue, cells make close contact with the ECM, which profoundly affects their behaviour by regulating fundamental processes such as cell adhesion, migration, proliferation, differentiation, survival and immune system signalling (Hynes, 2009, Frantz et al., 2010) (Fig 1.3).

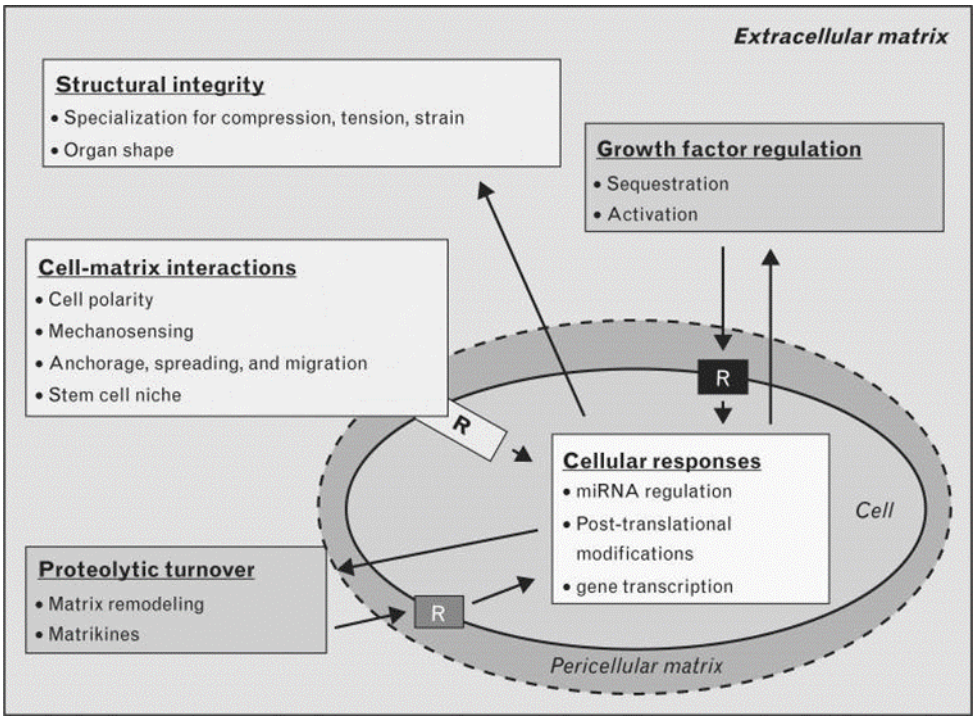



**Figure 1.3 Schematic representation of the multiple functions of the ECM at the cell-matrix interface.** The cartoon shows the various functions of the ECM and exemplifies those that result in feedback loops. Adapted from Chiovaro *et al*. (2015).

The ECM can exert these regulatory functions on cells directly by stimulating cell surface receptor-mediated signalling cascades, but also indirectly by regulating the activity of the soluble mediators found in the cellular microenvironment. For example, diverse molecular components of the ECM can directly activate cell surface receptors such as integrins. Increased deposition of collagen, which enhances ECM stiffness, can also induce integrin clustering and signal transduction (Kim et al., 2011). Moreover, the ECM functions as a reservoir of the soluble mediators found in the cellular microenvironment. It sequester and/or activate growth factors, including the transforming growth factors-β (TGF-β) superfamily (Doyle et al., 2012). Specifically, the ECM binds growth factors through its proteoglycan molecules, which possess many negatively charged glycosaminoglycan (GAG) chains such as heparan sulphate GAGs (Hynes, 2009). In this way, growth factors can be presented to their counterpart receptors on cells or can be released from the ECM following proteolytic cleavage and function as soluble traditional ligands (Mohammadi et al., 2005).

### 1.1.2 Regulation of the ECM

When the ECM becomes dysregulated and/or disorganised, it can lead to the development of disease or contribute to the progression of established disorders. Indeed, a fine balance of synthesis and turnover determines the composition of the ECM and thus the maintenance of its physiological functions (Visse and Nagase, 2003, Verrecchia and Mauviel, 2007) . When the stability of this equilibrium is upset, cellular and tissue changes occur that can lead to diseases, including fibrosis, atherosclerosis, osteoarthritis and cancer (Lu et al., 2012). For this reason, the integrity and structure of the ECM must be kept under strict control.

Several regulatory mechanisms are in place to control its synthesis, maintenance and remodelling. Among the best characterized are signalling pathways that orchestrate ECM synthesis and are initiated by growth factors such as transforming growth factor beta (TGF-β) and bone morphogenetic

proteins (BMPs). Transcriptional regulation, which involves the activity of specific transcription factors such as members of the SOX family or the nuclear factor kappa B (NF-kB), and proteolytic cleavage, involving enzymes such as matrix metalloproteinases (MMPs), are also among the most studied regulatory mechanisms of the ECM. Moreover, microRNAs, which are small non-coding RNAs that control gene expression at the post-transcriptional level (Ambros, 2004), have recently emerged as regulators of the ECM. Some microRNAs directly target the expression of ECM molecules, while others exert their control over the ECM indirectly, by modulating the expression of genes that control the synthesis and degradation of ECM molecules (Piccinini and Midwood, 2014).

### 1.1.3 Matricellular proteins

In the human genome there are hundreds human ECM protein-coding genes. The vast majority of them is highly conserved and undergoes alternative splicing, generating several isoforms, some of which are the predominant forms in disease (Giblin and Midwood, 2015).

As described above, the major constitutive components of the ECM are collagens, elastin, fibronectin, hyaluronic acid and proteoglycans, which together form a macromolecular network that provides structural support for tissues and delivers signals that regulate cell behavior. In addition to these components, there is a subset of ECM molecules that is not constitutively expressed. They are produced at high levels during development, but are present at low levels or absent in most healthy tissues in adulthood. The expression of these proteins is specifically induced upon tissue injury. These molecules have been termed ‘matricellular proteins’ and include the CCN family (CCN1-6), galectins, fibulins, osteopontin, periostin, secreted protein acidic and rich in cysteine (SPARC), small leucine rich proteoglycans, tenascin-C and thrombospondins-1 and -2 (Bornstein and Sage, 2002).

Matricellular proteins show structural diversity and different pattern of expression. However, they share a number of defining common features. They possess a multimodular structure, which allows them to bind to several binding partners, including cell surface receptors, other ECM molecules, hormones, growth factors and cytokines, and thus regulate cell behavior. Furthermore, thanks to their specific pattern of expression, they control cell functions in a highly context- and cell type-specific manner. Finally, mice with genetic ablations of these genes display grossly normal phenotypes, but exhibit abnormal tissue repair following tissue injury (Roberts, 2011, Kyriakides et al., 2009).

### 1.2 Tenascin-C (TN-C)

Tenascin-C (TN-C) is an adhesion modulatory glycoprotein highly expressed during development. In healthy adults, it is present at low levels only at restricted sites, including tendons, lymph nodes and spleen. Its expression is rapidly and transiently induced in the cellular microenvironment of damaged and inflamed tissue, independently of the location of the insult, (Brellier et al., 2009, Tucker and Chiquet-Ehrismann, 2009). Moreover, TN-C expression becomes dysregulated in chronic inflammatory diseases such as rheumatoid arthritis and in cancer (Midwood and Orend, 2009).

The human *TNC* gene (gene ID: 3371) is on chromosome 9q33 (Fig. 1.4) and it comprises 29 exons. 9 of these exons, each coding for a fibronectin type III domain, can be alternatively spliced (Gherzi et al., 1995, Sriramarao and Bourdon, 1993). Exon 1 is a non-coding exon, separated by an intron >20 kb long, and followed by exon 2, which contains the translation start codon ATG. The gene contains a single transcription start site (TSS) localized to the first exon and, 220 bp upstream to the TSS, a region with high promoter activity that contains a classical TATA box at -20 to -26 bp (Fig. 1.5) (Gherzi et al., 1995). Upstream promoter sequences are shown in Figure 1.6.

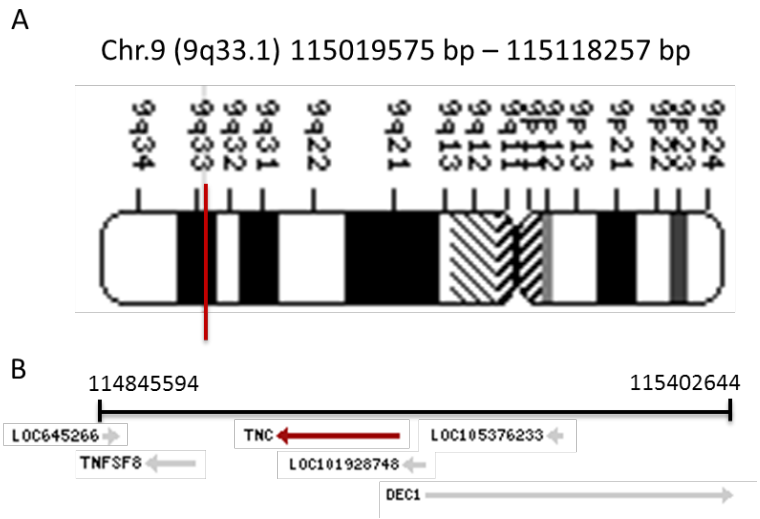



**Figure 1.4 Genomic context of *TNC*.** (A) Chromosomal location of *TNC* gene. Red bar shows position of the *TNC* gene on an idiogram of chromosome 9 that was captured from the NCBI database (http://www.ncbi.nlm.nih.gov/mapview/). (B) Mapping of *TNC* gene and local order on genomic context of the chromosome 9.

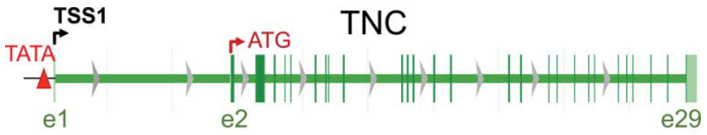



**Figure 1.5 Schematic representation of the *TNC* gene.** TATA and red triangle indicate a TATA box in the promoter of the *TNC* gene; TSS1 indicates the transcription start site; ATG indicates the translation start codon; e1, e2 and e29 indicate the non-coding exon 1, the first coding exon and the last coding exon, respectively. Adapted from Chiovaro *et al*. (2015).

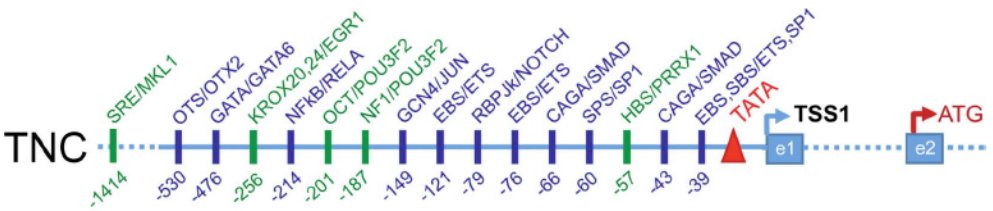



**Figure 1.6 Schematic representation of the gene promoters.** The transcription start site (TSS) and the translation start codon ATG are highlighted. The upstream promoter sequences are shown as horizontal light blue lines, on which the experimentally confirmed transcription factor binding sites are marked by vertical blue and green bars for the human and mouse promoters, respectively. Adapted from Chiovaro *et al*. (2015).

The transcriptional regulation of the *TNC* gene varies depending on the cell type and the nature of the stimulus, which can be mechanical stress, a cytokine, a growth factor or a hormone (Chiovaro et al., 2015). Recently, the *TNC* gene was shown to be regulated at the posttranscriptional level by the microRNA miR-335 in cancer (Oskarsson et al., 2011, Tavazoie et al., 2008).

The TN-C protein possesses a multimodular structure, which includes an N-terminal assembly domain is followed by 14.5 epidermal growth factor (EGF)-like repeats, 17 fibronectin type III-like (FNIII) domains and, at the C-terminus, a fibrinogen-like globe (FBG) domain. The *TNC* gene undergoes alternative splicing, which can generate up to 20 different isoforms. The pattern of expression and function of these isoforms is still unclear. However, the two most represented isoforms have a molecular weight of 180 and 330 kDa, depending on the number of alternatively spliced regions of the FNIII domain (Fig. 1.7) (Giblin and Midwood, 2015).

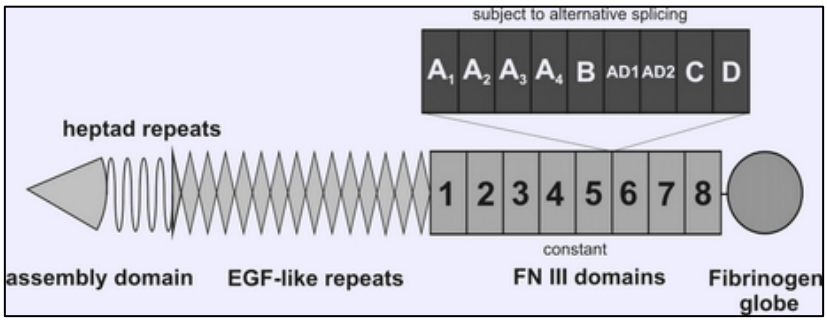



**Figure 2.7 Schematic representation of the multidomain structure of TN-C.** The cartoon shows the constitutively expressed TN-C domains starting from the N-terminal assembly domain to the C-terminal fibrinogen globe. Additional FN III domains generated by alternative splicing are shown. Adapted from Chiovaro *et al*. 2015.

The TN-C monomers assemble themselves into a 6-armed hexabrachion structure forming a huge protein of about 1.5 million Dalton. Assembly of the hexabrachion occurs via the N-terminal assembly domain, which contains a chaperone-like sequence that forms a coiled coil structure and interchanges disulphide bonds necessary for oligomerization (Fig. 1.8).

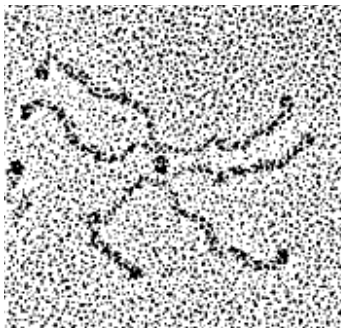


**Figure 3.8 Structure of the TN-C hexabrachion.** Electron micrograph of the TN-C hexabrachion. Adapted from Erickson, 1984.

TN-C is subjected to post-translational modifications, including N- and O-glycosylations, most of which occur in the FN III domains (Giblin and Midwood, 2015, Gulcher et al., 1989) .

After its discovery, TN-C was reported as an anti-adhesive ECM protein. This was observed during initial studies in which cells plated on a TN-C substratum showed a proliferative phenotype, but decreased capacity to adhere (Chiquet-Ehrismann, 1991, Chiquet-Ehrismann et al., 1986). TN-C can interact with several ECM proteins, including versican, aggrecan, brevican and neurocan. For this ability, TN-C plays a fundamental role in the formation and structural organization of the ECM (Day et al., 2004, Lundell et al., 2004) . The multimodular structure of the protein, allows TN-C to interact with several cell receptors such as integrins and proteins such as growth factors and cytokines. In this way, TN-C can modulate diverse cellular functions, ranging from cell survival, proliferation, and activation to cell signaling, migration and adhesion (Midwood and Orend, 2009).

### 1.2.1 TN-C in disease

The expression of TN-C is highly controlled during both development and adulthood. Highly expressed during development, TN-C is absent or present at low levels in restricted anatomical sites in healthy adults (Midwood et al., 2009, Chiquet-Ehrismann et al., 2014). However, its expression is switched on upon tissue injury, where TN-C plays a fundamental role in wound repair. For

instance, in the injured skin, keratinocytes and fibroblasts are the main producers of TN-C, which plays a role during the various phases of wound healing, including inflammation, re-epithelization and new tissue formation (Latijnhouwers et al., 1997, Chiquet-Ehrismann et al., 1994) . Several studies have shown significant correlation of TN-C expression with sterile tissue injury. For example, low levels of TN-C are detected in the joints of healthy individuals, while high levels of TN-C are found in the damaged synovium (specialized connective tissue of the joint), synovial fluid and cartilage of patients with rheumatoid arthritis (Goh et al., 2010, McCachren and Lightner, 1992). Accordingly, *TNC* knock-out mice are protected from joint destruction in a model of antigen-induced arthritis, showing less cell infiltration and bone erosion compared to wild type mice (Midwood et al., 2009). Moreover, up-regulation of TN-C expression has been associated with fibrotic diseases (Van Eyken et al., 1992, Kaarteenaho-Wiik et al., 1998), artherosclerosis (Wallner et al., 1999), asthma (Karjalainen et al., 2003), multiple sclerosis (Gutowski et al., 1999), colitis (Salas et al., 2003), diabetes (Catalan et al., 2012) and cancer (Orend and Chiquet-Ehrismann, 2006). Further on, TN-C levels in serum from patients with inflammatory bowel disease (Riedl et al., 2001), chronic kidney disease (Liabeuf et al., 2011) and myocardial infarction (Fujimoto et al., 2009) can be correlated with disease severity. Therefore, TN-C has been proposed as a potential diagnostic marker.

### 1.2.2 TN-C and the innate immune response

Our immune system has an innate ability to respond to and protect us from foreign invaders like pathogens, but also sterile tissue injury or any form of danger to the body. This immunological theory has recently been developed after the danger model of immunity was introduced by Polly Matzinger, who proposed that our immune system is designed to combat danger rather than discriminate self from non-self (Matzinger, 1994). Central to this theory are pattern-recognition receptors (PRRs) and their ability to recognize pathogen-associated molecular patterns (PAMPs) as well as damage-associated molecular patterns (DAMPs) or danger signals. Stimulation of PRRs leads to

pro-inflammatory gene synthesis and activation of antigen-presenting cells that initiate adaptive immune responses (Gordon, 2002, Medzhitov and Janeway, 2002, Beutler, 2007). DAMPs are endogenous molecules that are normally hidden from the immune system, but are secreted upon necrosis or up-regulated upon tissue injury. DAMPs include a very diverse set of molecules, including proteins, proteoglycans, nucleic acids and fatty acids. Among them are ECM fragments, which result from tissue destruction, and ECM molecules, including TN-C, whose expression is induced upon tissue injury.

During sterile tissue damage, TN-C can activate the pattern-recognition receptor Toll-like receptor 4 (TLR4) through its C-terminal fibrinogen globe domain. In this way, TN-C induces the synthesis of pro-inflammatory cytokines in primary human macrophages and synovial fibroblasts from patients with rheumatoid arthritis (Piccinini et al., 2016, Midwood et al., 2009) and promotes the synthesis of inflammatory mediators at the transcriptional level in a model of antigen-induced arthritis (Ruhmann et al., 2012).

Recent studies have shown that TN-C expression is rapidly induced in response to infection, before the appearance of any obvious signs of tissue damage (Midwood et al., 2009, Lotze and Tracey, 2005, Goh et al., 2010, Piccinini and Midwood, 2012). *TNC* knock-out mice mount a significantly reduced inflammatory response when exposed to bacterial components and primary murine macrophages from these mice express low levels of the pro-inflammatory microRNA miR-155 that results in lower TNF-$\alpha$ synthesis (Piccinini and Midwood, 2012).

Together, these studies indicate that TN-C can alert the immune system of tissue destruction, which exposes the body to greater risks of infection, but can also contribute to the inflammatory response necessary to kill invading pathogens and reestablish tissue homeostasis upon infection. And TN-C appears to do so via two distinct mechanisms, transcriptional and posttranscriptional control of inflammatory gene expression, respectively.

### 1.2.3 Regulation of microRNA expression: does TN-C play a role?

In the last ten years, microRNAs have emerged as key regulators of ECM gene expression (Rutnam et al., 2013). However, more recent findings suggest that not only can microRNAs regulate the biology of the ECM, but also the ECM can affect the expression of specific microRNAs. Notably, altered matrix has been linked with oncogenic microRNA expression (Merline et al., 2011, Mouw et al., 2014). Thus, an intriguing interplay between the ECM and microRNAs is emerging.

Our group has found that TN-C is rapidly induced *in vitro* and *in vivo* in response to lipopolysaccharide (LPS), a component of the outer membrane of gram-negative bacteria (Goh et al., 2010, Piccinini and Midwood, 2012). Mice that lack the *TNC* gene respond very differently to LPS than the corresponding wild type mice, showing significantly lower circulating TNF-$\alpha$ levels when challenged with LPS. Moreover, *in vitro*, stimulation of bone marrow-derived macrophages from *TNC* knock-out mice with LPS results in significantly lower levels of TNF-$\alpha$ compared to wild type cells (Piccinini and Midwood, 2012). With the help of a microRNA-microarray of wild type and *TNC* knock-out bone marrow-derived macrophages, following LPS stimulation, Piccinini *et al*. showed that TN-C regulates TNF-$\alpha$ levels by controlling the expression of the microRNA miR-155 (Piccinini and Midwood, 2012).

These data highlight a novel, regulatory function of the ECM molecule TN-C that control inflammatory gene expression by regulating microRNA levels in the mouse. However, we still do not know whether this occurs in human cells. Moreover, the molecular mechanism by which tenascin-C regulates LPS-induced miR-155 expression is still unclear.

Preliminary data from our group shows that aberrant miR-155 expression in *TNC* knock-out bone marrow-derived macrophages is rescued upon exposure of cells to TN-C-pre-coated plates, but not soluble protein. These results

indicate that TN-C needs to be assembled in a pericellular or extracellular matrix in order to regulate microRNA expression.

To investigate whether and how human TN-C assembled in the ECM regulates microRNA expression, a system that recapitulates the in-vivo-like cell microenvironment and, ideally, allows to knock-out *TNC* expression is needed. Cell-derived matrices (CDM), using human cells that naturally secrete abundant ECM, are cell culture systems that represent the native ECM microenvironments better than conventional 2D or artificially-created 3D cell culture models (Fitzpatrick and McDevitt, 2015). CRISPR-Cas9-mediated genome editing is a new and affordable technology that allows to permanently knock-out target gene expression in mammalian cells. The combination of these two technologies would allow us to engineer an in-vivo-like ECM lacking TN-C that can be directly compared to the wild type ECM and used a substrate (or cellular microenvironment) for the culture of cells whose microRNA expression is of interest.

## 1.3 CRISPR-Cas9 genome editing technology

### 1.3.1 CRISPR-Cas9: a historical overview

The CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) -Cas9 system is a weapon of the adaptive immune system of most archaea and many bacteria that protects against viruses (Sorek et al., 2008, Deveau et al., 2010, Horvath and Barrangou, 2010, Karginov and Hannon, 2010, Koonin and Makarova, 2009). For several years, short repeats interspaced with unique spacers have been found in bacteria and archaeal genomes. However, their biological function has been elucidated only recently. It has also been found how putative operons located close to CRISPR sequences are responsible for the coding of several CRISPR-associated (Cas) proteins. Computational analysis of these proteins has revealed that they contain domains

characteristic of nucleases, RNA binding proteins, helicases and polymerases (Jansen et al., 2002).

Initially, this system was interpreted as a novel pathway for the repair of errors in the genome, but this theory was disproved (Makarova et al., 2002). Later, some of the unique CRISPR sequences that function as spacers have been found to be almost identical to fragments of viral or plasmid genes. This led to the idea that the CRISPR-Cas9 system was a defense mechanism against external genetic elements (Mojica et al., 2005, Bolotin et al., 2005). After a computational re-analysis of the Cas proteins (Makarova et al., 2006, Haft et al., 2005) and a study of their structure and functions, an analogy between the CRISPR-Cas9 system described above and the RNAi (RNA interference) of eukaryotes has been found (Carthew and Sontheimer, 2009). However, unlike RNAi, the CRISPR–Cas9 system agglomerates a small piece of DNA derived from foreign nucleic acid into the CRISPR locus of the host genome as the first step in the series of consecutive events that leads to acquired immunity against the invading phages (Makarova et al., 2006).

The hypothesis that the CRISPR-Cas9 system had a role in the defense against foreign nucleic acid was validated with the phage resistance given by the integration of a short phage-specific sequence into a CRISPR-Cas9 locus of the bacterium *Streptococcus thermophiles*. A double validation was carried out by altering the similarity between the CRISPR insert and the target phage sequence with a single mismatch (Barrangou et al., 2007). Only recently it has been discovered how the CRISPR-Cas9 system possesses a lower stringency for the correlation between spacer and target sequence (Garrett et al., 2011).

The CRISPR–Cas9 system mediates an immune response to invading phages through three stages: adaptation, expression and interference. This process can be split in two main steps: a highly conserved 'information processing' step, which includes the adaptation stage, and an 'executive' step, which includes the expression and interference stages. The proteins involved in the information processing step are highly conserved. On the other hand, the proteins involved in the executive step vary greatly between different

organisms (Deveau et al., 2010, Horvath and Barrangou, 2010, Karginov and Hannon, 2010, van der Oost et al., 2009, Al-Attar et al., 2011). During the stage of adaptation, a short sequence of the invading DNA (the protospacer) is acquired and inserted into the CRISPR array as a spacer. This involves cleavage of the protospacer by Cas proteins and repeat duplication, following recombination in the CRISPR array. During the expression stage, the pre-crRNA is transcribed and undergoes processing to generate mature crRNAs. Then, in the interference stage, crRNA-directed cleavage of foreign nucleic acid is operated by Cas proteins at sites complementary to the crRNA spacer sequence.

There are three types of CRISPR-Cas9 systems (I, II and III) and each of these uses distinct molecular mechanisms for interference: type I and III systems use a large complex of Cas proteins for crRNA-guided targeting, whereas the type II system requires only the Cas9 protein for RNA-guided DNA recognition and cleavage, a property that has proved to be extremely useful for genome engineering applications (Sorek et al., 2008).

### 1.3.2 The CRISPR-Cas9 nickase: a specific genome editing tool

The CRISPR-Cas9 nickase (CRISPR-Cas9n) system uses two main components: a guide RNA (gRNA) and a CRISPR-associated endonuclease Cas9 nickase (Cas9n) (Hubmacher and Apte, 2013).

The gRNA is essential to identify the target DNA. It is a 20 bp long oligonucleotide RNA sequence that binds the genomic target to be modified via Watson-Crick base pairing. A crucial and limiting requirement of Cas9n for recognition and binding to the target DNA is the presence of a PAM sequence. The exact sequence of the PAM varies among Cas9 species. For example, 5' NGG 3' is the base structure of the PAM sequence for *Streptococcus pyogenes* Cas9, where N can be any nucleotide (Hubmacher and Apte, 2013).

The Cas9n is a Cas9 mutant (D10A) that contains a single active catalytic domain in contrast to wild type Cas9, which contains two. As a consequence,

Cas9n can only generate a single nick to the double strand of the DNA. For this reason, the system is called nickase. Therefore, to generate a double strand break (DSB), Cas9n requires two nickases targeting opposite strands. Having two gRNAs in the system dramatically increases the specificity of the cleavage, thus reducing the off-targets. CRISPR-Cas9n is currently used to modify the genome of a huge variety of target organisms, from bacteria to human cells (Hubmacher and Apte, 2013).

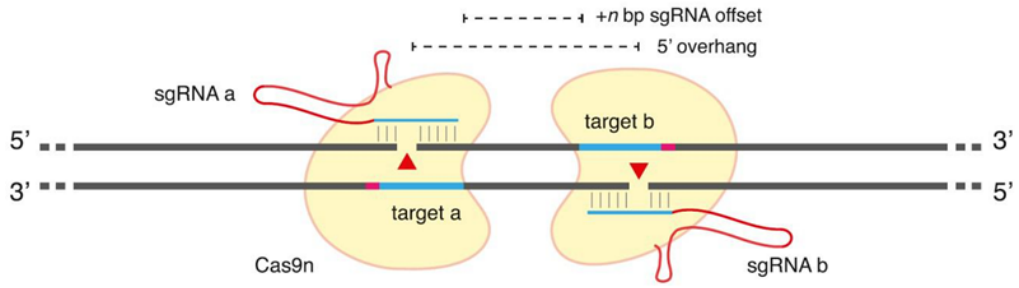



**Figure 1.9 Schematic representation of the CRISPR-Cas9 nickase system.** Two Cas9 nickase endonucleases are driven to the target DNA by two gRNAs (shown in blue). One binds to the top and the other binds to the bottom DNA strand at the PAM sequence (shown in red) and each creates a single nick, which results in a double strand break. Adapted from F. Ann Ran *et al*. 2013.

### 1.3.3 Applications of CRISPR-Cas9

The CRISPR-Cas9 system has become a popular tool for its capacity to directly target and edit the genes of eukaryotic systems. The first application of CRISPR-Cas9 was to knock-out target genes in various cell types and organisms. Now, thanks to modifications to the Cas9 enzyme, CRISPR is used to selectively activate or repress target genes, purify specific regions of DNA, image DNA in live cells, and even to conduct genome-wide screens. Moreover, CRISPR-Cas9 can be used to precisely alter the coding region of a gene to "fix" or create a mutation, introduce a protein tag, or create a new restriction site, using the high-fidelity homology-directed repair (HDR) pathway, which requires a donor template.

To generate a knock-out, specific locations within complex mammalian genomes can be targeted by Cas9 or Cas9n that cause a DSB or a double nick, respectively. These breaks can be repaired by endogenous DNA repair mechanisms through a repair pathway known as non-homologous end-joining (NHEJ) (Fig. 1.10). This error-prone repair pathway causes insertion/deletions (InDels) that can disrupt the gene by shifting its open reading frame (ORF) or generating premature stop codons. The InDels resulting from NHEJ are random and differ from cell to cell. Thus, the exact genomic changes that result must be validated by screening single-cell clones, to which the CRISPR-Cas9 and the gRNA have been delivered.

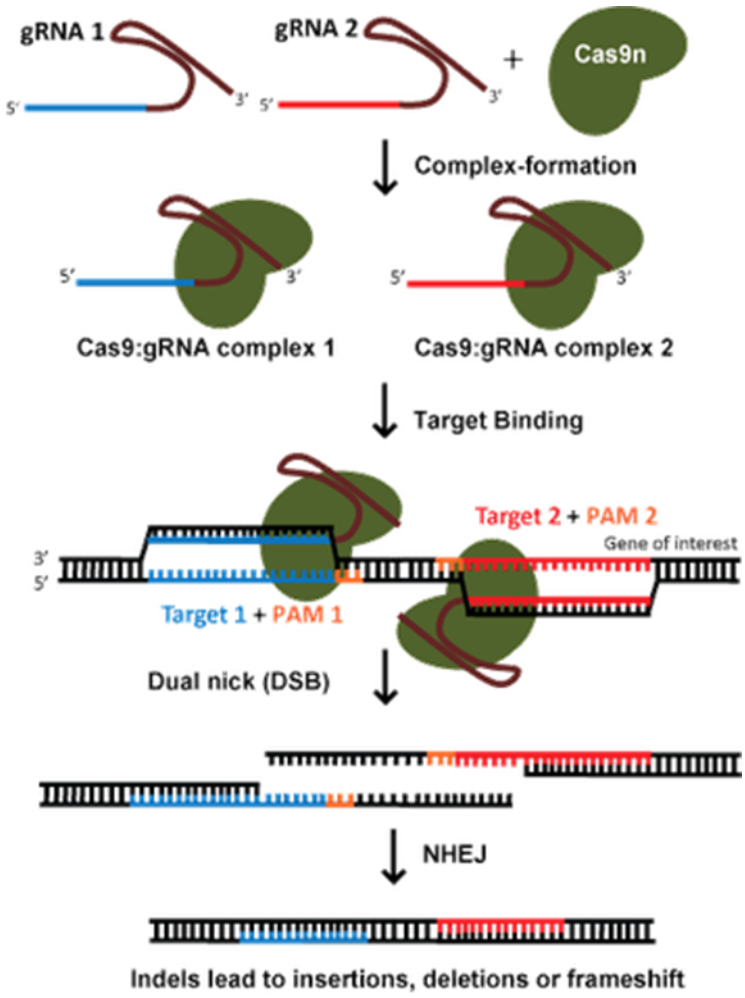



**Figure 1.10 CRISPR-Cas9n-mediated genome editing.** The CRISPR-Cas9n-mediated genome editing process from the initial Cas9n-gRNA complexes formation to the final NHEJ repair is illustrated step by step. Taken from https://www.addgene.org/CRISPR/guide/.

### 1.4 Aims

Recently, it has been shown that expression of the ECM protein TN-C is rapidly induced during the inflammatory response to infection *in vivo*. Specifically, TN-C sustains the synthesis of specific LPS-induced pro-inflammatory genes by regulating microRNA expression in primary murine macrophages. However, it is not known whether this occurs in human cells and how.

The aim of my project is to engineer human *TNC* knock-out fibroblasts to generate cell-derived matrices, which lack TN-C, as a substrate for the culture of primary human macrophages.

This aim will be achieved by: 1) designing and delivering a CRISPR-Cas9 system that specifically targets the *TNC* gene in the HEK 293 cell line; 2) validating the genome edit in HEK 293 cells; and 3) applying the validated CRISPR-Cas9 system to human fibroblasts and generating single-cell clones with knocked-out *TNC* expression.

This system will allow us to investigate whether human TN-C modulates macrophage gene expression by regulating microRNA levels upon infection. Furthermore, the knock-out cells might also be used to study the regulatory function of extracellular TN-C in other cell types of interest and conditions other than infection.

# 2 Materials and methods

## 2.1 Reagents and sources of materials

### 2.1.1 General reagents

All general laboratory reagents were of analytical grade and purchased from Sigma-Aldrich, unless otherwise stated below. Ultrapure water was used for preparing all of the solutions and buffers. The pH of the prepared buffers was adjusted with 5 M sodium hydroxide (NaOH), 0.5 M NaOH or concentrated hydrochloric acid (HCl) using a microprocessor-based pH meter (pH 211, Hanna Instruments).

### 2.1.2 Molecular biology reagents

Taq DNA polymerase, Taq DNA polymerase buffer, MgCl solution, dNTPs and DNA ladder (1 Kb plus) were purchased from New England Biolabs (NEB). Oligonucleotide guide RNAs (gRNAs) for CRISPR-Cas9 genome editing and oligonucleotide primers for polymerase chain reaction (PCR) amplification were purchased from Sigma-Aldrich. Mini-prep and Midi-prep kits were purchased from Macherey-Nagel and Qiagen, respectively. DNA gel extraction kit was purchased from MonarchTM. Restriction digestion, kinase, phosphatase and ligase enzymes and their relevant buffers were purchased from NEB. TRIzol Reagent was purchased from Thermo Fisher Scientific.

### 2.1.3 Protein Chemistry, Western blotting reagents and equipment

PageRulerTM Plus Pre-stained protein ladder (10 to 250 kDa) was purchased from Thermo Scientific. SDS-PAGE electrophoresis apparatus, wet transfer apparatus and 0.45 µm nitrocellulose membranes were all from Bio-Rad. Amersham ECL Western blotting Detection ReagentsTM were purchased from GE Healthcare Life Science.

| ID | Ms X Human Tenascin N-Terminal Region | Polyclonal rabbit anti-Mouse immunoglobulins/HRP |
|---|---|---|
| **Company** | Merck Millipore | Dako |
| **Antigen** | Tenascin-C | Rabbit immunoglobulins |
| **Specificity** | Human | Mouse |
| **Dilution** | 1:1,000 | 1:2,000 |
| **Antibody diluent** | 2 % BSA/PBST | 5 % Milk/PBST |

**Table 1** Details of antibodies used for Western blotting.

### 2.1.4 Tissue culture reagents

All reagents used were of tissue culture grade and sterile. Reagents were purchased from the following sources: 100 X penicillin-streptomycin (10,000 U/mL) and 0.4 % Trypan Blue Solution from Sigma; Dulbecco's modified Eagle's medium (DMEM) with high Glucose 4.5 g/l, L-Glutamine and Sodium Pyruvate, MEM Eagle with Earle's BSS (EMEM) without L-glutamine, 100 X L-glutamine (200 mM in 0.85 % NaCl solution), Opti-MEM medium and 1 X Trypsin-EDTA (200 mg/L Versene EDTA, 170.000 U Trypsin) from Lonza; heat-inactivated fetal bovine serum (FBS) from Gibco; Dimethyl Sulfoxide (DMSO, ≥99.7 %) from Fisher BioReagents.

### 2.1.5 Composition of solutions and buffers

Ampicillin solution: 1 g of ampicillin, 5 mL of $H_2O$, 5 mL of ethanol.

LB medium: 10 g of bacteriological tryptone, 5 g of bacteriological yeast extract, 10 g of NaCl, $H_2O$ up to 1 L, pH 7.0.

2 x Laemmli sample buffer: 4 % β-mercaptoethanol, 100 mM Tris-HCl, pH 6.8, 2% (w/v) SDS, 20% glycerol and 0.012% bromophenol blue.

6 x Laemmli sample buffer: 9% β-mercaptoethanol, 375 mM Tris-HCl pH 6.8, 9% (w/v) SDS, 50% glycerol and 0.03% bromophenol blue.

10 x TAE buffer: 400mM Tris-acetate, 10 mM EDTA, $H_2O$ up to 1L.

10 x Tris-glycine running buffer: 30 g of Tris-base, 144 g of glycine, and 10 g of SDS, $H_2O$ up to 1 L.

10 x Transfer Buffer: 192 mM glycine, 25 mM Tris-base, $H_2O$ up to 2 L.

2 x HBS (250 mL): 0.5 g D-Glucose, 2.5 g HEPES, 0.18 g KCl, 4.0 g NaCl, 0.05 g $Na_2HPO_4$, $H_2O$ to 250 mL, pH up to 7.1 with NaOH.

4 x Tris-Cl/SDS: 0.5 M Tris-Cl, 0.4 % SDS, pH 6.8.

4 x Tris-Cl/SDS: 1.5 M Tris-Cl, 0.4 % SDS, pH 8.8.

Ponceau S solution: 1 % Ponceau S in 1 % acetic acid.

10 x PBS (phosphate-buffered saline): 1.37 M NaCl, 27 mM KCl, 100 mM $Na_2HPO_4$, 18 mM $KH_2PO_4$.

1 x PBST**:** 1 x PBS, 0.05 % Tween20.

Cryopreservation solution: 5% DMSO, 95% cell growth medium, supplemented with 10 % FBS and 1 x L-glutamine.

Fixative solution: 75% methanol, 25% acetic acid.

20 x SSC (sodium saline citrate): 175.3 g NaCl, 88.2 sodium citrate, pH 7.0, $H_2O$ up to 1 L.

Pretreatment solution: 50 mL of 2 x SSC, 0.5% NP-40, pH 7.0.

Denaturation solution: 50 mL of 70% formamide, 1 x SSC, pH 7.0. Store at -20°C.

Washing buffer: 100 mL of 0.5 x SSC, 0.1 % NP-40. Store at 4°C.

## 2.2 Cell culture

### 2.2.1 Cell lines

Cell lines used in these studies: BJ human skin/foreskin fibroblasts from ATCC and human embryonic kidney 293 (HEK 293) cells kindly provided by Dr Hilary Collins.

### 2.2.2 Cell culture

BJ cells were cultured in EMEM medium, supplemented with 10 % (v/v) FBS, L-glutamine (1 X) and penicillin-streptomycin (1 X). Cultures were grown in T75 flasks, 10 cm culture dishes, or 6, 12, 24 or 96-wells plates in an incubator maintained at 37°C with a humid atmosphere supplied with 5 % $CO_2$. The complete growth medium was changed every 2-3 days. Cells were passaged upon ~80 % confluence by trypsinisation after two washes with sterile 1 X PBS. 1 X trypsin/EDTA was added to the cells that were kept at 37$^{o}$C until detached. Complete growth medium was added to neutralise the trypsin-EDTA activity and the cell suspension was centrifuged at 1,500 rpm for 5 minutes. The cell pellet was resuspended in complete growth medium. For routine subcultivation, the cell suspension was aliquoted to new culture vessels in a ratio of 1:2 to 1:9 as recommended by ATCC. For experiments, cells were counted using an improved Neubauer BS.748 haemocytometer (depth 0.1 mm and 1/400 mm2) purchased from Hawksley and a light microscope (Nikon ECLIPSE TS100) and plated at the cell densities stated below.

HEK 293 cell were cultured in DMEM medium supplemented with 10 % (v/v) FBS and penicillin-streptomycin (1 X). Cultures were grown in T75 flasks, 10 cm culture dishes, or 6, 12, 24 or 96-wells plates in an incubator maintained at 37°C with a humid atmosphere supplied with 5 % CO2. The complete growth medium was changed every 2-3 days. Cells were passaged upon ~80 % confluence by gentle pipetting and then the cell suspension was centrifuged

at 1,500 rpm for 5 minutes. The cell pellet was resuspended in complete growth medium. For routine subcultivation, the cell suspension was aliquoted to new culture vessels in a ratio of 1:2 to 1:10 as recommended by ATCC. For experiments, cells were counted using an improved Neubauer haemocytometer and plated at the cell densities stated below.

### 2.2.3 Cells cryopreservation

After two washes with 1 x PBS, adherent cells were dislodged using trypsin/EDTA and resuspended in complete growth medium. The cell suspension was centrifuged at 1,500 rpm and the supernatant discarded. The cell pellet was resuspended in cryopreservation solution at a cell density of 0.75 x $10^6$ cells/mL. The cell suspension was transferred to 1.5 mL cryogenic vials and frozen using a Mr. Frosty freezing container (Thermo Scientific) placed at -80 °C overnight before storage in liquid nitrogen.

### 2.2.4 Cell stimulation

Cells were stimulated for 24 hours with 100 ng/mL of lipopolysaccharide (LPS) from E. coli, Serotype EH100 (Ra) from Enzo Life Sciences.

### 2.2.5 Cell lysis

BJ and HEK 293 cells were seeded in 10 cm culture dishes as described above and grown to ~80% confluence. The growth medium was aspirated and collected. The cells were washed twice with 1 x PBS. 600 µL of 1 X Laemmli sample buffer was added to each dish, which was scraped thoroughly to ensure all cell material was collected. Each sample was collected in a 1.5 ml Eppendorf tube, boiled for 10 minutes, centrifuged at 14,000 rpm for 10 minutes and frozen at -20 °C.

## 2.3 CRISPR-Cas9 gRNA design and synthesis

Three pairs of CRISPR-Cas9n guide RNAs (gRNAs) were designed based on the *TNC* gene sequence (gene ID: 3371). The CRISPR design software

http://crispr.mit.edu/ was used to design the CRISPR-Cas9n gRNAs listed in Table 2. Overhang sequences required for cloning of gRNAs into pSpCas9n(BB)-2A-GFP (PX461) plasmid and an extra G were added at the 5' end of each guide. gRNAs were synthesised by Source Bioscience.

| ID | Sequence | Position in the *TNC* gene |
|---|---|---|
| **GB TNC A1 5'-3' top** | CACCGCTGAGTCATGGCCCCATGG | 27237-27257 |
| **GB TNC A1 3'-5' bottom** | AAACCCATGGGGGCCATGACTCAGC | |
| **GB TNC A2 5'-3' top** | CACCGTTGCCCTCGCTACCGAAGGT | 27286-27306 |
| **GB TNC A2 3'-5' bottom** | AAACACCTTCGGTAGCGAGGGCAAC | |
| **GB TNC B1 5'-3' top** | CACCGGACCCCACCTTCGGTAGCG | 27303-27323 |
| **GB TNC B1 3'-5' bottom** | AAACCGCTACCGAAGGTGGGGTCC | |
| **GB TNC B2 5'-3' top** | CACCGATCCGGCACAAGCGACAGAG | 27324-27344 |
| **GB TNC B2 3'-5' bottom** | AAACCTCTGTCGCTTGTGCCGGATC | |
| **GB TNC C1 5'-3' top** | CACCGCGTGGTTAAACACCACTGGC | 27377-27397 |
| **GB TNC C1 3'-5' bottom** | AAACGCCAGTGGTGTTTAACCACGC | |
| **GB TNC C2 5'-3' top** | CACCGTTACAACATCAAGCTGCCAG | 27407-27427 |
| **GB TNC C2 3'-5' bottom** | AAACCTGGCAGCTTGATGTTGTAAC | |

**Table 2** Details of CRISPR-Cas9n gRNAs synthesized.

## 2.4 Cloning of gRNAs into CRISPR-Cas9 vector

### 2.4.1 Annealing of oligonucleotides, restriction digestion and ligation

Annealing of oligonucleotides

To express Cas9n and the gRNAs, a mammalian expression vector (pSpCas9n(BB)-2A-GFP; PX461; 9289 bp) containing a constitutive CBh Cas9 promoter, a U6 promoter for the gRNA and a GFP reporter gene was used. The lyophilized gRNAs were resuspended in $H_2O$ to a concentration of 100 µM. In a PCR tube, 2 µL of 100 µM gRNA was added to 2 µL of 10 x T4 DNA ligase buffer, 10 Units (1 µL) of polynucleotide kinase and $H_2O$ to a total volume of 20 µL. The mixture was briefly vortexed and then incubated at 37°C for 60 minutes and at 65°C for 20 minutes to allow the phosphorylation of the

oligonucleotides and to inactivate the enzyme, respectively. Once complete, 5 µl of each complementary top and bottom phosphorylated oligonucleotide were mixed together and made up to a total volume of 50 µL with $H_2O$. Using a thermocycler (2720 Thermal Cycler from Applied Biosystem), the mixtures were heated at 95°C for 3 minutes. The heat was then reduced by 1°C each minute until a temperature of 26°C was reached, followed by 5 minutes at 25°C. The products were stored at -20°C until required.

Restriction digestion

DNA was digested with the *BbsI* restriction enzyme using the manufacturer's recommended buffer at the recommended temperature for 2 hours. Typical restriction enzyme digestions consisted of: 2 µl of 10 x restriction enzyme buffer, 1 µg plasmid DNA (pSpCas9n(BB)-2A-GFP), 10 Units *BbsI* and $H_2O$ to a total volume of 20 µl (digestion was done by Mr Jonathan Whitchurch).

1 µg of fully digested plasmid was dephosphorylated at the 5' end by incubation with 1 µl of thermo-sensitive alkaline phosphatase (TSAP; 1 Unit) for 1 hour at 37°C to prevent re-ligation without the insert. TSAP was denatured at 74°C for 20 minutes. The products were stored at -20°C until required.

Ligation

Plasmids were ligated to linear DNA fragments after restriction enzyme digestion to produce new circular constructs containing a genetic region of interest. A typical ligation reaction consisted of: 100 ng of dephosphorylated and fully digested plasmid and oligonucleotide inserts (in a molar ratio of 3:1 with respect to the plasmid), 2 µL of 10 x T4 DNA ligase buffer, 3 Units of T4 DNA ligase and $H_2O$ to a total volume of 20 µL. The ligation mixture was vortexed, centrifuges and allowed to ligate overnight at 4°C. The products were stored at -20°C until required.

### 2.4.2 Agarose gel electrophoresis and plasmid DNA purification from agarose gel

Plasmid DNA was separated in a 0.6 % agarose gel (0.3 g of agarose, 50 ml of 1x TAE buffer and 2µL of ethidium bromide). The gel was run for 2 hours at 80 V using a PowerPac Universal system from Bio-Rad. The extraction from the gel was done with the DNA Gel Extraction Kit from Monarch following the manufacturer's instructions. The DNA fragment was excised from the agarose gel and transferred to a clean and weight-known microfuge tube. The gel slices were weighted and an equal volume of binding buffer was added. The gel slices were completely dissolved in a water bath at 60°C. The mixture was loaded into a spin column and centrifuged for 1 minute at 13,000 rpm. 500 µL of wash solution I was added and the mixture centrifuged for 1 minute at 13,000 rpm. The flow-through solution was discarded and centrifugation was repeated. After 1 additional minute of centrifugation to remove the residual wash solution I, the DNA was eluted by adding 40 µL of elution buffer and stored at 4°C.

### 2.4.3 LB agarose plates preparation and bacterial transformation

LB agar was melted and allowed to cool to 50°C. After adding ampicillin solution to a final concentration of 0.1 mg/mL, the LB agar was poured into plastic Petri dishes and was allowed to set. 50 µL of E. coli DH5-α competent cells, purchased from Stratagene, were mixed in a 1.5 ml Eppendorf tube with 50 ng of the annealed plasmid DNA. To increase the permeability of the bacterial cell membrane a heat shock was used. The mixture of cells and plasmids was left for 30 minutes on ice and then heat shocked for 2 minutes at 42°C. After that, it was placed on ice for 5 minutes and LB medium was added to make up a total volume of 1 ml. The mixture was incubated for 1 hour at 37°C, 220 rpm and then centrifuged at 7,000 rpm for 2 minutes. The pellet was resuspended and the bacterial suspension was plated on the prepared LB agar dish containing ampicillin and incubated overnight at 37°C.

### 2.4.4 Glycerol stock

500 µL of transformed E. coli DH5-α overnight culture was added to 500 µL of 100 % glycerol in a 2 mL sterile screw-top tube and gently mixed. The glycerol stocks were frozen at -80°C.

### 2.4.5 Minipreps of plasmid DNA and sequencing

The isolation of the plasmid was carried out by using the Nucleospin® plasmid kits from Macherey-Nagel and the procedure was executed by following the manufacturing protocol. First of all, 10 mL of a transformed E. coli LB culture with O.D.600 between 0.6 and 0.8 was centrifuged at 11,000 x g and pelleted. 500 µL of Buffer A1 were used to completely resuspend the cells and 500 µL of Buffer A2 were then added to the mixture. After inverting the tubes 8 times, the samples were incubated at room temperature for 5 minutes. 600 µL of Buffer A3 were added and then the mixture was centrifuged at 11,000 x g for 10 minutes at RT. The supernatant was decanted using a Nucleospin plasmid/plasmid column with a silica membrane and then centrifuged at 11,000 x g for 1 minute. After that, 500 µL of Buffer AW, pre-heated at 50°C, were added and centrifuged at 11,000 x g for 1 minute. 600 µL of Buffer A4, supplemented with ethanol, were added and centrifuged at 11,000 x g for 1 minute. The silica membrane was dried with two centrifugations at 11,000 x g for 1 minute. The DNA was eluted with 50 µL of Buffer EA pre-heated at 70°C.

The quality and the concentration of the DNA was determined using a NanoDrop ND-1000 UV-Vis Spectrophotometer. A ratio of sample absorbance at 260 and 280 nm was used to assess the purity of DNA; ~1.8 was accepted as pure. A ratio of sample absorbance at 260 and 230 nm was also used as a measurement of DNA purity, and a range of 1.8 - 2.2 was considered as pure nucleotides. A lower value for either of these ratios was considered as an indication of the presence of contaminants in the sample.

DNA was sequenced by Source Bioscience using an Applied BioSystems 3730 DNA Analyzer and PeakTrace™ software. 5 μL of plasmid DNA (100 ng/μL) and 5 μL of primer (3.2 pmol/μL) were necessary for successful sequencing.

### 2.4.6 Midipreps of plasmid DNA

Extraction and purification of larger amounts of plasmid DNA was carried out by using the QIAGEN Plasmid Midi Kit.

100 mL of transformed E. coli culture was prepared as described above and was centrifuged at 6,000 x g. 4 mL of Buffer P1 were used to completely resuspend the cell pellet and 4 mL of Buffer P2 were then added to the mixture. After inverting the tubes 6 times, the samples were incubated at RT for 5 minutes. 4 mL of Buffer P3 were added and the mixture immediately mixed by inverting the tube 6 times. The mixture was incubated on ice for 15 minutes and then centrifuged at 9,500 rpm for 90 minutes at 4°C. The supernatant was removed promptly and centrifuged at 6,000 rpm for 30 minutes at 4°C. The supernatant was applied on a QIAGEN-tip 100 (pre-equilibrated with 4 mL of buffer QBT) and allowed to enter the resin by gravity flow. The QIAGEN-tip 100 was washed twice with 10 mL of buffer QC. The DNA was eluted in 5 fractions of 1 mL of buffer QF each and its concentration measured using a NanoDrop spectrophotometer. DNA fractions with low to no DNA and A260/A230 ratio <1.6 were discarded to reduce chemical contamination and increase DNA quality and concentration. The rest of the DNA fractions were precipitated by adding 3.5 mL of isopropanol and centrifuged immediately at 10,000 rpm for 80 min at 4°C. The DNA pellet was washed with 2 mL of 100 % and 70 % ethanol and centrifuged at 9,000 rpm for 30 minutes each time. The DNA pellet was air-dried and the DNA resuspended in 50 μL of $H_2O$, pH 8.0.

## 2.5 Transfection of cell lines

### 2.5.1 Transfection of HEK 293 cells

HEK 293 cells were transfected by calcium phosphate co-precipitation. At day one, 140,000 cells per well were seeded in 6-well plates in DMEM medium, supplemented with 10 % FBS and 1 x penicillin/streptomycin. The cells were incubated overnight at 37°C, in 5 % $CO_2$ atmosphere. At day two or when the cells reached ~80 % confluence, the medium was aspirated from the plates and replaced with 2 mL of complete growth medium. Separately, one 1.5 mL Eppendorf tube per each sample was filled with 200 µL of 2 x HBS. Another 1.5 mL Eppendorf tube for each sample was prepared containing 1 µg of each plasmid DNA as shown in Table 3, 25 µL /well of 2M $CaCl_2$, and $H_20$ up to 200 µL/well. The DNA/$CaCl_2$ mixture was added to the HBS solution while vortexing. The solution was allowed to precipitate for 20 minutes and then added drop-wise to the cells and the plate was gently rocked to allow for even cell distribution. After 16 hours the supernatant was removed and, after a single wash with PBS, replaced with complete growth media. Cell viability and transfection's efficiency were evaluated 24 hours later by light and fluorescence microscopy. Genomic DNA was extracted 48 hours after transfection.

| | **Tube 1** | **Tube 2** | **Tube 3** | **Tube 4** | **Tube 5** | **Tube 6** |
|---|---|---|---|---|---|---|
| **Plasmid DNA (1 µg each)** | A1+A2 | B1+B2 | C1+C2 | A1 | B1 | C1 |

**Table 3** Details of plasmid DNA preparations transfected into human HEK 293 cells.

### 2.5.2 Transfection of BJ fibroblasts

Three different approaches were used to transfect the difficult-to-transfect BJ cells: cationic lipid-based transfection, non-lipid-based chemical transfection and electroporation.

#### 2.5.2.1 Cationic lipid-based transfection

At day one, a specific number of cells per well (Table 4) was seeded in 24-well plates in 0.5 ml per well of EMEM medium, supplemented with 10 % FBS and 1 x L-glutamine, without antibiotics. Cells were incubated overnight at 37°C, in 5 % $CO_2$ atmosphere. At day two or when the cells reached 80 % confluence, for each well to be transfected, the specific transfection reagent was diluted with the appropriate medium (Table 4) and, in the case of the NanoJuice transfection reagent, incubated at room temperature for 5 minutes. The plasmid DNA was added to the diluted transfection reagent and the mixture was incubated for 5 to 15 minutes (Table 4) to allow for DNA-lipid complex formation. After that, the mixture was added drop-wise to the cells and the plate was rocked for 30 seconds. The cells were then incubated at 37°C, in 5 % $CO_2$ atmosphere. In the case of NanoJuice, the medium with the transfection mixture was removed 8 hours after transfection and was replaced with complete growth medium. Cell viability and transfection's efficiency were analysed by light and fluorescence microscopy at 24, 48 and 72 hours after transfection.

| **Transfection reagent** | **NanoJuice (Merck Millipore)** | **Lipofectamine 2000 (ThermoFisher)** |
|---|---|---|
| **Number of cells/well** | $4 \times 10^4$ | $4 \times 10^4$ |
| **Plate type** | 24-well | 24-well |
| **Amount of medium (mL/well)** | 0.5 | 0.5 |
| **Amount of DNA (µg/well)** | 0.8 | 0.5 |
| **Amount of reagent (µL/well)** | NanoJuice reagent: 1<br>NanoJuice booster: 2 | 4 |
| **Medium for transfection reaction** | Serum-free EMEM medium | Opti-MEM medium |
| **Mixture incubation time (minutes)** | 15 | 5 |

**Table 4** Details of lipid-based transfection reactions.

### 2.5.2.2 Non-lipid-based chemical transfection

At day one, a specific amount of cells per well (Table 5) was seeded in a 24-well plate in 0.5 mL of EMEM medium, supplemented with 10 % FBS and 1 x L-Glu. The day of the transfection, the transfection reagents were allowed to equilibrate at room temperature and diluted in the appropriate medium in a sterile tube (Table 5). The DNA was added to the diluted transfection reagent, and after a specific incubation time (Table 5), the mixture of transfection reagent/DNA was added dropwise to the cells and the plates rocked for 30 seconds. The cells were incubated for 72 hours at 37°C, in 5% $CO_2$ atmosphere. The cell viability and transfection's efficiency were analysed by light and fluorescence microscopy at 24, 48 and 72 hours after transfection.

| | ***Trans*IT-X2 (Mirus)** | ***Trans*IT-2020 (Mirus)** | **FuGene 6 (Promega)** | **Fugene HD (Promega)** | **GeneJuice (Merck Millipore)** |
|---|---|---|---|---|---|
| Number of cells/well | $5 \times 10^4$ | $5 \times 10^4$ | $4 \times 10^4$ | $3.6 \times 10^3$ | $2 \times 10^4$, $4 \times 10^4$, $8 \times 10^4$ |
| Plate type | 24-well | 24-well | 24-well | 96-well | 24-well |
| Amount of medium (mL/well) | 0.5 | 0.5 | 0.5 | 0.1 | 0.5 |
| Transfection reagent/DNA ratio(µg/µL) | 3 | 3 | 6, 3, 1.5 | 3, 0.25 | 3 |
| Amount of DNA (µg/well) | 0.5 | 0.5 | 1 | 0.1 | 1 |
| Amount of reagent (µL/well) | 1.5 | 1.5 | 6, 3, 1.5 | 0.3, 0.25 | 3 |
| Mixture incubation time (minutes) | 20 | 20 | 15 | 15 | 10 |
| Medium for transfection reaction | Opti-MEM reduced-Serum Medium | Opti-MEM reduced-Serum Medium | Serum and P/S-free EMEM medium | P/S-free EMEM medium | Serum-free EMEM medium |

**Table 5** Details of non-lipid based chemical transfection reactions.

### 2.5.2.3 Electroporation

Transfection of BJ cells by electroporation was carried out using the Neon® Transfection System from Invitrogen and the 4D-Nucleofector™ System from Lonza with the Neon® Transfection System 10 μL Kit and the Cell Line Nucleofector™ Kit/Solution SF, respectively.

Neon® Transfection System: at the day of transfection, the cells were washed twice using 1 x PBS without $Ca^{2+}$ and $Mg^{2+}$, detached by trypsinization and counted as described above. Cells were centrifuged for 5 minutes at 400 x g and the cell pellet was resuspended in Resuspension Buffer R (Invitrogen) at a final cell density of 5 x $10^6$ cells/mL. A 24- or 6-well plate was pre-filled with 0.5 or 2 mL of EMEM medium, supplemented with 10% FBS and 1 x L-glutamine (without antibiotics), and pre-warmed at 37°C. A Neon Tube with 3 mL Electrolytic Buffer E (Invitrogen) was set into the Neon Pipette Station. The DNA (0,5 – 2μg) was added to 10 μL of cell suspension. The cell/DNA mixture was loaded into the Neon Pipette tip and positioned in the Neon transfection device. After delivering a single electric pulse (1,650 v, 20 ms), the sample containing the cells and the DNA was immediately placed into the previously prepared plate containing pre-warmed medium. The cells were incubated at 37°C, in 5% $CO_2$ atmosphere and cell viability and transfection's efficiency were analysed by light and fluorescence microscopy at 24, 48 and 72 hours after transfection.

4D-Nucleofector™ System: at the day of transfection, the cells were washed twice with 1 x PBS, detached by trypsinization and counted as described above. Cells were centrifuged for 10 minutes at 100 x g and the cell pellet was resuspended in 20 μL Buffer SF (Lonza). A 6-well plate was filled with 0.5 mL/well of EMEM medium, supplemented with 10% FBS and 1 x L-glutamine (without antibiotics), and pre-warmed at 37°C. For each reaction, 3 x $10^5$ cells were resuspended in 20 μL of Buffer SF and 1 μg of plasmid DNA or 1 μg of pmaxGFP (positive control, plasmid DNA; Lonza) were added. Each cell/DNA mixture was transferred to a transfection cuvette and an electric pulse was

delivered as recommended by Lonza (program FF-120). After 10 minutes of incubation at RT, 80 µL of pre-warmed medium were added to each cuvette and the cell suspension was transferred into the pre-warmed 6-well plate and incubated for 24 hours at 37°C, in 5% $CO_2$ atmosphere. Cell viability and transfection's efficiency were analysed by light and fluorescence microscopy at 24, 48 and 72 hours after electroporation and quantified by Fluorescence Activated Cell Sorting (FACS). Positive cells, expressing the green fluorescent protein (GFP) reporter gene were selected by FACS 72 hours after electroporation.

### 2.6 Selection of transfected cells

#### 2.6.1 Fluorescence Activated Cell Sorting (FACS)

BJ cells that were transfected with plasmid DNA containing Cas9n and gRNA pair A, Cas9n and gRNA pair B and Cas9n and a single gRNA by electroporation using the Neon system were sorted on the basis of their GFP expression using the Beckman Coulter MoFlo XDP (Flow Cytometry Facility, University of Nottingham). For sorting, BJ cells were washed with 1 X PBS and detached by trypsinisation 72 hours after transfection. The cells from each well of a 6-well plate were resuspended in 500 µL of complete growth medium and transferred into a 5 mL FACS tube. Single $GFP^+$ sorted cells were each seeded into a single well of 96-well plates containing 200 µL/well of complete growth medium and BJ-conditioned medium (1:1) and were incubated at 37°C in a 5% $CO_2$ atmosphere. BJ-conditioned medium was collected from wild type, non-transfected BJ cells after three days of subculture, and was centrifuged at 3,000 rpm and filtered through a 0.2µm pore size filter before use. The remaining bulk $GFP^+$ cells were seeded in a 6-well plate and were incubated in complete growth medium at 37°C in a 5% $CO_2$ atmosphere.

## 2.7 Validation of genome edit

### 2.7.1 Gene copy number estimation by Fluorescence in Situ Hybridization (FISH)

#### 2.7.1.1 Cell preparation

100,000 BJ and HEK 293 cells were harvested into a 15 mL sterile centrifuge tube, and centrifuged at 1,000 x g for 5 minutes. The cell pellet was resuspended in 0.2 mL of growth medium and 10 mL of a 1:1 mixture of 0.4 % KCl and 0.8 % sodium citrate were added drop-wise to the cell suspension. After gently mixing by inverting the tube 4 times, the mixture of cells and saline solution was incubated for 20 minutes in a 37°C water bath. 0.1 mL of fixative solution was added to the cell suspension that was mixed and centrifuged at 1,000 x g for 5 minutes. The cell pellet was resuspended in 5 mL of fixative solution added dropwise. The cell suspension was mixed and incubated at RT for 20 minutes. The samples were centrifuged at 3,000 x g to pellet the cells and the wash with the fixative solution was repeated 3 times. The cell pellet was resuspended in 1 mL of fixative solution and 10 µL of cell suspension was dropped on a glass slide at RT. The slides were baked overnight at 56°C.

#### 2.7.1.2 Hybridization procedure

VividFISH CEP Kit by GeneCopoeia was used to carry out the FISH experiment. For the hybridization procedure, 50 mL Coplin glass slide jars were used. A slide jar with 50 mL of pretreatment solution was pre-warmed at 37°C in a water bath. The slides, with the fixed cells, were incubated in the pre-warmed pretreatment solution for 30 minutes at 37°C. The slides were dehydrated in 70 %, 90 % and 100 % ethanol, each time for 1 minute. 2 µL of the FISH probes (VividFISH CEP9; human chromosome 9) were diluted to 10 µL using hybridization solution. The pre-treated slides were incubated in a slide jar at 73°C with 50 mL of denaturation solution in a water bath for 5 minutes. The slides were dehydrated in 3 solutions with growing ratio of ethanol in $H_2O$ (70

%, 90 % and 100 %) each for 1 minute. 10 µL of the FISH probes were added to each slide, and the area with the cells was covered with a 22 x 22 mm cover slip, allowing the solution to slowly spread and fill under the cover slip. The perimeter of the cover slips were sealed with rubber cement and they were incubated in the pre-warmed moisture chamber at 42°C for 24 hours. The slides with removed rubber cement were putted into a pre-warmed slide jar with 2 X SSC, allowing the cover slips to float away. The slides were incubated with gentle shaking in a pre-warmed slide jar with 50 mL of 0.5 X SCC + 0.1 % NP40 at 73°C for 5 minutes. Next, the slides were transferred into a slide jar with 2 X SSC and 0.1% NP-40 at room temperature for 1 minute, rinsed with $H_2O$ and air dried. 15 µL of AntiFade w/DAPI I were loaded on the slides, which were covered with 22 x 22 mm cover slips. The slides were analysed by fluorescence microscopy using a 100X oil immersion objective lens (LEICA DM2000).

### 2.7.2 Polymerase Chain Reaction (PCR) amplification, gel electrophoresis and sequencing

#### 2.7.2.1 Design and validation of PCR primers

Two to three pairs of primers flanking the genomic region to be edited were designed for each CRISPR-Cas9n system of gRNAs (Table 6). We designed a 20 nucleotides long oligonucleotides with desired offset, with a GC content between 50% and 65%, a Tm between 55°C and 80°C and a G or C at the 5' and 3' ends.

| Primer | Sequence 5' - 3' | Amplicon length (bp) | GC content (%) | Tm (°C) | Base pairs position in TNC gene |
|---|---|---|---|---|---|
| **Primers for screening of Cas9n-gRNA pair A** | | | | | |
| **1 Fw** | GATACCCTAGAGCCCTAGAG | 180 | 55 | 56.1 | 27187-27207 |
| **1 Rv** | CTGGCAGGGTGGCGTTCACC | 180 | 70 | 74.4 | 27347-27367 |

| | | | | | |
|---|---|---|---|---|---|
| **2 Fw** | CTCTGTGCTTCTAAATCCCC | 232 | 50 | 60.3 | 27149-27169 |
| **2 Rv** | CTGGCTGGTTCTCTTCTGGC | 232 | 60 | 66.3 | 27362-27382 |
| | **Primers for screening of Cas9n-gRNA pair B** | | | | |
| **1 Fw** | GGGCCATGACTCAGCTGTTG | 177 | 60 | 68.2 | 27244-27264 |
| **1 Rv** | CCACTGGCAGCTTGATGTTG | 177 | 55 | 66.9 | 27401-27421 |
| **2 Rv** | CCACTGGCTGACTCCAGATC | 271 | 60 | 65 | 27187-27207 |
| **3 Fw** | CCAAGGGACCAGAGCTACAG | 781 | 60 | 63.8 | 26801-26821 |
| **3 Rv** | CCTCCAGTCTGCTCAGCAGC | 781 | 65 | 67.8 | 27602-27622 |

**Table 6** Details of primers for CRISPR-Cas9n gene edit validation.

The primer 1Fw for the CRISPR-Cas9n gRNA pair A system was also used as primer 2Fw for the validation of CRISPR-Cas9n gRNA pair B system.

The reaction was assembled completely on ice and then quickly transferred to the thermocycler preheated to the denaturation temperature of 95°C. The polymerase chain reaction (PCR) was performed using 2.5 µL 10 x Standard Taq (Mg-free) Reaction Buffer, 1.5 µL of 25 mM $MgCl_2$, 0.5 µL of 10 mM dNTPs, 0.5 µL of 10 mM Forward Primer, 0.5 µL of 10 mM Reverse Primer, 0.125 µL of Taq DNA polymerase and 120 ng of target DNA to amplify. The total volume of the reaction was 25 µL for each sample to amplify. The 2720 Thermal Cycler from Applied Biosystem was used to carry out this PCR analysis.

A touchdown PCR protocol was designed for each pair of primers (Table 7-11).

| Temperature (°C) | Time | Cycles |
|---|---|---|
| **95** | 5 min | 1 |
| **95** | 1 min | 27 |
| **75** | 30 sec | 27 |

| | | |
|---|---|---|
| **68** | 30 sec | 27 |
| **95** | 1 min | 13 |
| **52** | 30 sec | 13 |
| **68** | 30 sec | 13 |
| **68** | 5 min | 1 |
| **4** | ∞ | 1 |

**Table 7** Touchdown PCR protocol used to amplify the *TNC* genomic region edited by the CRISPR-Cas9n-gRNA pair A system using primers 1Fw and 1Rv.

| Temperature (°C) | Time | Cycles |
|---|---|---|
| **95** | 5 min | 1 |
| **95** | 1 min | 18 |
| **70** | 30 sec | 18 |
| **68** | 30 sec | 18 |
| **95** | 1 min | 22 |
| **54** | 30 sec | 22 |
| **68** | 30 sec | 22 |
| **68** | 5 min | 1 |
| **4** | ∞ | 1 |

**Table 8** Touchdown PCR protocol used to amplify the *TNC* genomic region edited by the CRISPR-Cas9n-gRNA pair B system using primers 2Fw and 2Rv.

| Temperature (°C) | Time | Cycles |
|---|---|---|
| **95** | 5 min | 1 |
| **95** | 1 min | 18 |
| **77** | 30 sec | 18 |
| **68** | 30 sec | 18 |
| **95** | 1 min | 22 |
| **61** | 30 sec | 22 |
| **68** | 30 sec | 22 |
| **68** | 5 min | 1 |
| **4** | ∞ | 1 |

**Table 9** Touchdown PCR protocol used to amplify the *TNC* genomic region edited by the CRISPR-Cas9n-gRNA pair B using primers 1Fw and 1Rv.

| Temperature (°C) | Time | Cycles |
|---|---|---|
| **95** | 5 min | 1 |
| **95** | 1 min | 18 |
| **66** | 30 sec | 18 |
| **68** | 30 sec | 18 |
| **95** | 1 min | 22 |
| **51** | 30 sec | 22 |
| **68** | 30 sec | 22 |

| | | |
|---|---|---|
| **68** | 5 min | 1 |
| **4** | ∞ | 1 |

**Table 10** Touchdown PCR protocol used to amplify the *TNC* genomic region edited by the CRISPR-Cas9n-gRNA pair B system using primers 2Fw and 2Rv.

| Temperature (°C) | Time | Cycles |
|---|---|---|
| **95** | 5 min | 1 |
| **95** | 1 min | 18 |
| **78** | 30 sec | 18 |
| **68** | 30 sec | 18 |
| **95** | 1 min | 22 |
| **60** | 30 sec | 22 |
| **68** | 30 sec | 22 |
| **68** | 5 min | 1 |
| **4** | ∞ | 1 |

**Table 11** Touchdown PCR protocol used to amplify the *TNC* genomic region edited by the CRISPR-Cas9n-gRNA pair B system using primers 3Fw and 3Rv.

#### 2.7.2.2 PCR amplification and sequencing

PCR samples were run on a 1.5 % agarose gel and DNA bands were cut. The DNA was extracted from the gel and purified as described above. PCR products were sequenced by Source Bioscience as described above.

## 2.8 Protein Analysis

### 2.8.1 Tris-glycine Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

50 µL of HEK 293 and 7.5 µL of BJ conditioned media and cell lysates and 6 X Laemmli sample buffer were mixed and loaded into 5% SDS-PAGE gels (Table 12). Each sample was boiled for 5 minutes at 95°C before loading. Samples were run in 1 X Tris-glycine running buffer at 10 mA/gel for 20 minutes and at 20 mA/gel for 2 hours using a PowerPac Universal from BIO RAD.

| Component | Separating gel (mL) | Stacking gel (mL) |
|---|---|---|
| 30 % acrylamide/0.8 % bisacrylamide | 2.50 | 0.65 |
| 4 X Tris-Cl/SDS, pH 8.8 | 3.75 | - |
| 4 X Tris-Cl/SDS, pH 6.8 | - | 1.25 |
| $H_2O$ | 8.75 | 3.05 |
| 10 % w/v ammonium persulfate | 0.05 | 0.025 |
| N,N,N',N'- Tetramethylethylenediamine (TEMED) | 0.01 | 0.005 |

**Table 12** Composition of 5 % SDS-PAGE gels.

### 2.8.2 Western blotting

Protein samples separated by SDS-PAGE were transferred to a nitrocellulose membrane. The membrane was first hydrated with water and then equilibrated in 1 X, ice-cold transfer buffer for 10 minutes with shaking. Protein transfer was carried out using a PowerPac Universal fromBIO RAD, at 350 mA for 90 minutes in ice-cold 1 X transfer buffer. After electroblotting, protein transfer was confirmed by visualization of the pre-stained protein ladder, and the membrane was stained with 1 % Ponceau S solution and imaged using an image analyser (Fujifilm LAS-4000) to check equal loading of

protein samples. After destaining with $H_20$, the membrane was blocked with 20 mL of 5 % BSA/PBST on a shaker overnight at 4°C. The membrane was incubated with 5 mL of 2 % BSA/PBST containing the primary antibody (Table 1) for 3 hours on a slow shaker at room temperature. After 3 washes of 10 minutes each with 1 X PBST with shaking, the membrane was incubated in 5 mL of 5 % dry milk/PBST containing the secondary antibody (Table 1) for 1 hour at room temperature. After 3 washes of 10 minutes each with 1 X PBST on a shaker, the membrane was incubated for 2 minutes in 1.6 mL of freshly prepared ECL detection solution. The membrane was covered in Saran wrap, drained from the liquid excess and the chemiluminescent signal detected using an image analyser (Fujifilm LAS-4000).

### 2.9 Gene expression quantification

#### 2.9.1 RNA extraction and purification

For total RNA extraction from HEK 293 and BJ cell, the medium was removed from ~80% confluent 10 cm culture dishes and 1 mL of TRIzol Reagent was added. The cells were lysed directly in the culture dish using a cell scraper. Cell lysates were transferred to RNase-free 1,5 mL tubes and incubated for 10 minutes at room temperature. 0.2 mL of chloroform was added to the samples and mixed vigorously by shaking the tubes for 15 seconds. After an incubation of 3 minutes at room temperature, the samples were centrifuged at 10,000 rpm for 15 minutes at 4°C. The top, aqueous phase was transferred into a RNase-free 1.5 mL tube, where 0.5 mL of 100 % isopropanol was previously added, and incubated at room temperature for 10 minutes. After a 10 minutes centrifugation at 12,000 x g at 4°C the RNA pellet was washed with 1 mL of 75 % ethanol. The mixture was centrifuged at 7,500 x g for 5 minutes at 4°C and the pellet air-dried for 10 minutes. 30 µL RNase-free $H_2O$ was used to re-suspend the RNA, which was incubated for 15 minutes at 60°C before measuring the RNA concentration and purity using a NanoDrop spectrophotometer. RNA was stored at -80$^{o}$C until further analysis.

### 2.9.2 Reverse Transcription

Genomic DNA removal from RNA samples and reverse transcription polymerase chain reaction (RT-PCR) were carried out using the QuantiTect Reverse Transcription Kit from QIAGEN, following the manufacturer's protocol. A reverse-transcription master mix was prepared by mixing 1 μL of quantitative reverse transcriptase, 4 μL of 5 x quantitative RT buffer and 1 μL RT primer mix. The genomic DNA elimination reaction (2 μL of 7 x gDNA wipeout buffer, 0.5 μg of template RNA and RNase-free $H_2O$ to 14 μL) was carried out at 42°C for 2 minutes and immediately stopped on ice. The master mix was added to the treated RNA solution to make up a total volume of 20 μL. The reverse transcription reaction was carried out at 42°C for 15 minutes followed by inactivation of the reverse transcriptase at 95°C for 3 minutes. For long-term storage, the reactions were stored at -20°C.

### 2.9.3 Quantitative real time polymerase chain reaction

Quantitative real-time PCR was performed with TaqMan primer sets containing two unlabelled primers and one 6-FAM dye-labelled TaqMan MGB probe for human TN-C (Hs01115665_m1) and HPRT1 (Hs02800695_m1) from Applied Biosystems. cDNA samples (reverse transcription reactions) were diluted 1:3 with autoclaved, ultrapure $H_2O$. Each reaction was performed in triplicate, in 10 μL total volume, by mixing 5 μL of 5 X Universal PCR Master Mix (Applied Biosystems), 0.5 μL of 20 X TaqMan primer set, 1.5 μL of $H_2O$ and 3 μL of diluted cDNA. The real-time PCR reactions were run in a Corbett Rotor-gene machine using the following program: hold at 95$^{o}$C for 10 minutes and 40 cycles at 95$^{o}$C for 15 seconds and 60$^{o}$C for 1 minute. Relative expression was calculated by the change-in-threshold (ΔΔCT) method with HPRT1 as endogenous control for gene expression and was normalized to results obtained with untreated cells.

## 2.10 Preparation of cell-derived matrices

For preparation of cell-derived matrices, a modified version of Martin

Humphries's protocol was used. 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 or 1 mL in a 12-well plate. Cells were cultured at 37˚C (5 % $CO_2$). When cells reached 100 % confluence, the growth medium was changed to fresh medium supplemented with 50 µg/ml L-ascorbic acid (L-AA; Sigma). The L-AA containing medium was changed every day to promote collagen synthesis and deposition of a thick matrix. Cells were cultured for 14 days in the presence of L-AA.

**2.10.1 Matrix decellularization**

Cells were washed twice with 1 x PBS. 6 mL/dish of pre-warmed extraction buffer (20mM $NH_4OH$, 0.5% Triton X-100 in PBS) were gently added. Cells were lysed for 2 min until no intact cells were visible by phase. The extraction buffer was removed partially and washed with 1 x PBS. The gentle wash with PBS was repeated twice. Residual DNA was digested with 10 µg/ml DNase I (Roche) in 1 x PBS containing $Ca2^+$ and $Mg2^+$ for 30 min at 37˚C. DNase I was then removed and the matrix washed twice with 1 x PBS and used immediately.

**2.10.3 ECM extraction**

For total ECM molecule extraction, 300 µl of 1 x Laemmli sample buffer were added to each 10 cm dish after decellularization. The dish was scraped off with a cell scraper and the lysate was collected into a 1.5 ml mini-centrifuge tube. Additional 300 µl of 1 x Laemmli sample buffer were added and the solution was collected as above. The lysates were vortexed and boiled for 10 min.

For extraction of soluble and insoluble ECM molecules, 1 ml of modified RIPA buffer was applied to each 10 cm dish, which was incubated at room temperature for 5 min on a rocker. The buffer was aspirated and the solution was collected in a 1.5 ml tube. Immediately after, 1 ml of urea lysis buffer was added directly to each dish, which was incubated at room temperature for 5 min on a rocker. The plate was scraped off with a cell scraper and the extract

was collected in a 1.5 ml tube. Both extracts were dialyzed using Slide-A-Lyzer G2 Dialysis Cassette, 2K MWCO (Thermo Scientific). The total protein concentrations of samples was quantified using the Bradford assay. Extracts were also analysed by Western blotting.

### 2.10.4 Preparation of coverslips

In order to prepare cell-derived matrices on glass coverslips for immunofluorescence analysis, glass coverslips were boiled 3 times in ultrapure $H_2O$ and autoclaved. They were then transferred to 24-well plate. After a wash with 1 x PBS, coverslips were incubated with 0.2 % sterile gelatin (Sigma) for 1 hour at 37˚C. After that, they were washed 3 times with 1 x PBS. 1 % sterile-filtered glutaraldehyde in PBS was then added for 30 minutes at room temperature. Coverslips were then washed 3 times with 1 x PBS. 1 M sterile-filtered glycine in PBS was added for 20 minutes at room temperature. Coverslips were then washed 3 times with 1 x PBS and incubated for 30 minutes at 37˚C with complete growth medium. The medium was removed and replaced with fresh growth media. A BJ cell suspension was then added as described above.

### 2.10.5 Immunofluorescence (IF) analysis of cell-derived matrices

Cell-derived matrices were fixed with 3 % paraformaldehyde in PBS at 4˚C for 15 minutes. Paraformaldehyde was removed and matrices were incubated with blocking solution containing bovine serum albumin and goat serum at room temperature for 1 hour. Coverslips were then incubated with primary antibody or no antibody in blocking solution at room temperature for 1 hour (Table 16). After 4 washes with Tris-buffered saline (TBS), coverslips were incubated with the appropriate secondary antibody and DAPI (Invitrogen) in blocking buffer at room temperature for 1 hour. Coverslips were washed 4 times with TBS and mounted onto microscope slides using 10 µl of Prolong Gold antifade agent (Thermo Scientific). Coverslips were sealed on the slides by applying nail polish at the edges and examined using an inverted

fluorescence microscope (EVOS® Cell Imaging System, Thermo Scientific).

| Antibody | Use | Dilution | Provider |
|---|---|---|---|
| Rabbit anti-human fibronectin (F3648) | IF | 1:200 | Sigma |
| Mouse anti-human tenascin-C (MAB1908) | IF | 1:200 | Chemicon |

**Table 16** Details of primary antibodies used for IF analysis.

# 3 Results

## Generation of human *TNC* knock-out fibroblasts by CRISPR-Cas9-mediated genome editing

### 3.1 Introduction

To knockout the *TNC* gene in human BJ fibroblasts we have chosen the CRISPR-Cas9 nickase (CRISPR-Cas9n) technology. Our objective is to create a double strand break (DSB) in the first expressed exon near the N-terminus where mutations are more likely to disrupt the ability of the gene to code for the protein. Specifically, our target DNA sequence is located in proximity of the translation start codon ATG, which is found in exon 2. We expect that the DSB is repaired by the error-prone non-homologous end joining (NHEJ) repair pathway, which causes insertion/deletions (InDels) that can disrupt the gene by shifting its open reading frame (ORF) or generating premature stop codons. The *S. pyogenes* endonuclease Cas9n mutant D10A, which contains only one catalytic domain, will be used. Cas9n requires two 20 bp long oligonucleotides, known as guide RNAs (gRNAs), which direct the nuclease to two specific sequences on the opposite strands of the target DNA, allowing Cas9n to generate a DSB. Compared to the wild type Cas9, this mutant operates with increased specificity and reduced off-target cleavage. Before using this technology to knock-out the *TNC* gene in human BJ fibroblasts, I will deliver and test the designed CRISPR-Cas9n system in the HEK 293 cell line.

### 3.2 Analysis of TN-C expression in human BJ fibroblasts and HEK 293 cells

In animals, the predominant connective tissue cells are fibroblasts, which synthesise and secrete abundant extracellular matrix molecules, including TN-C. For this ability, fibroblasts can be used to prepare in vivo-like cell-derived matrices, which can then be employed as a substrate for the culture of other cells and to study the function of the matrix and its molecules. In order to

study how TN-C assembled in the extracellular matrix signals to cells and influences their gene expression by regulating microRNA levels during inflammation, we aim to generate cell-derived matrices lacking TN-C by engineering *TNC* knock-out fibroblasts. BJ fibroblasts, a cell line derived from normal human skin/foreskin, were chosen for their ability to secrete abundant extracellular matrix and their long lifespan compared to primary fibroblasts or other normal human fibroblast cell lines. However, as fibroblasts grow slowly and are hard to transfect, HEK 293 cells, which are easy to grow and transfect, were used in the first place to design and optimise CRISPR-Cas9-mediated edit of the *TNC* gene.

### 3.2.1 BJ fibroblasts synthesise and secrete TN-C into the extracellular space

To analyse TN-C expression at the mRNA and protein level in BJ fibroblasts, cells were cultured in 10 cm culture dishes to confluence and were stimulated with 100 ng/mL of LPS or left unstimulated for 24 hours. Cells were lysed to extract and purify total RNA or to extract total proteins. *TNC* mRNA expression was measured by RT-qPCR and normalised using the housekeeping gene *HPRT1*, an endogenous control whose expression level does not vary between samples, as it has been extensively tested in our lab. TN-C protein in cell lysate and conditioned medium was analysed by Western blotting. To analyse TN-C secretion and deposition in the ECM synthesized by BJ fibroblasts, cells were cultured in 10 cm culture dishes or on sterile glass coverslips in 12-well plates to confluence and for 14 additional days during which cells were supplemented daily with fresh complete media by Dr. Anna Piccinini. TN-C protein was analysed in total cell extracts (cell and matrix) and matrix extracts after decellularization by Western blotting by Miss Er Pey Ling. Moreover, TN-C protein was visualised in the extracellular matrix deposited by BJ fibroblasts by immunofluorescence by Miss Er Pey Ling.

BJ fibroblasts in the steady-state express *TNC* at the mRNA level and LPS stimulation of cells further increases its expression by ~2.5 fold (Fig. 3.1A). LPS

was previously shown to transcriptionally induce the expression of *TNC* in myeloid cells (Goh et al., 2010). However, similar levels of TN-C protein were detected in the conditioned medium and cell lysate of unstimulated and LPS stimulated cells for 24 hours (Fig. 3.1B). These results, together with the kinetic of *TNC* mRNA and TN-C protein expression in myeloid cells that peaks at 4 and 24 hours respectively (Goh et al., 2010), indicate that TN-C protein levels should be measured at later time points to confirm whether the induction of TNC mRNA leads to an enhancement in TN-C protein in BJ fibroblasts.

BJ fibroblasts secrete TN-C protein into the extracellular space and higher levels of TN-C were detected in the insoluble fraction compared to the soluble fraction of matrix extracts (Fig. 3.1C). This suggests that most of the TN-C that is secreted by BJ fibroblasts is assembled into the matrix meshwork where TN-C is known to bind to other matrix molecules. Immunofluorescence analysis confirmed secretion and deposition of TN-C in the extracellular matrix and showed a macromolecular organization for TN-C that is different to that of the abundantly expressed matrix molecule fibronectin (Fig. 3.1D).

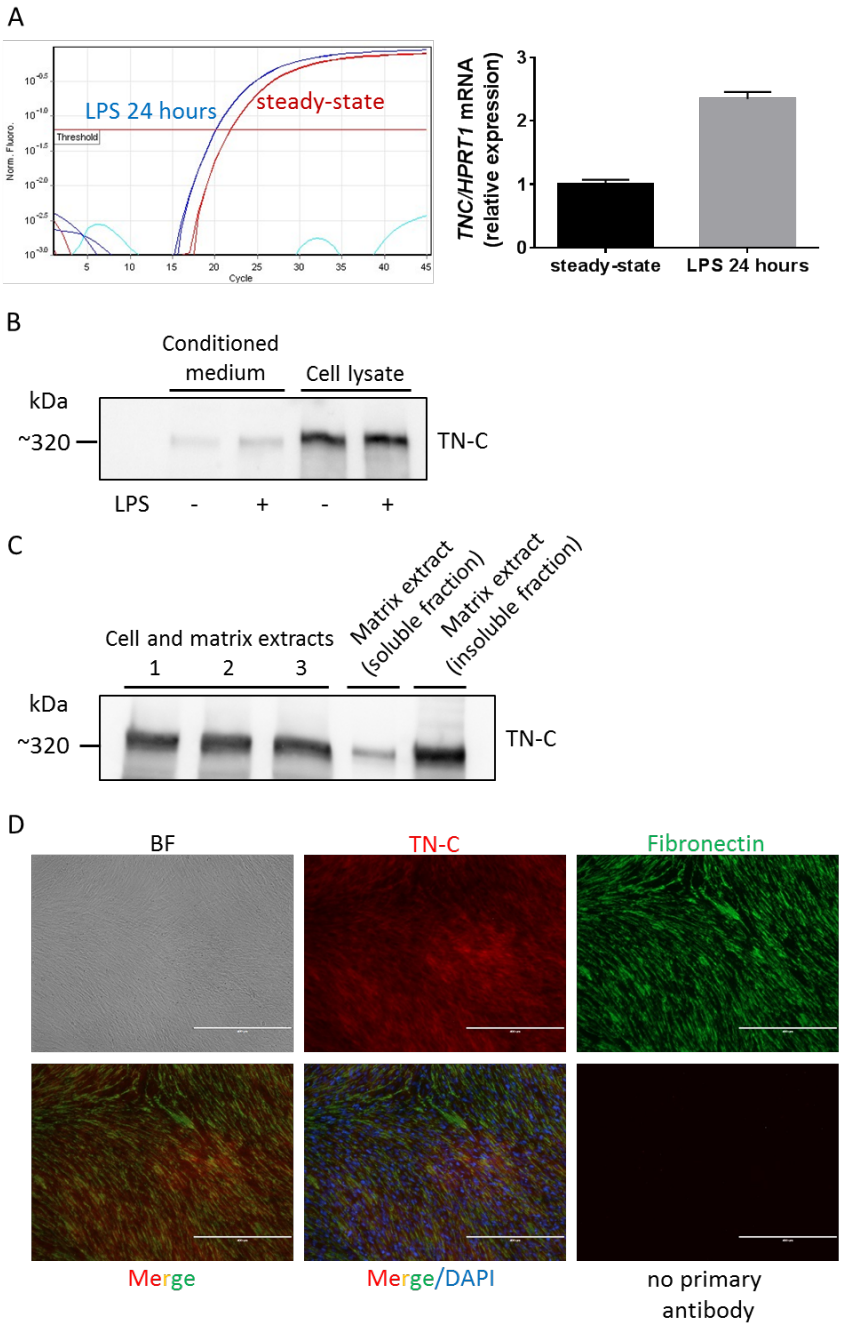



**Figure 3.1 BJ fibroblasts express TN-C and deposit an extracellular matrix that contains TN-C.** (A) Quantitative RT-PCR analysis of *TNC* mRNA in BJ fibroblasts at steady-state or stimulated with 100 ng/ml LPS for 24 hours. Left: amplification plot and threshold cycles; right: results are presented relative to the abundance of *TNC* mRNA at steady-state. Data are means ± SD of three replicates. (B) BJ cells were treated for 24 hours with 100 ng/ml LPS and conditioned medium and cell lysate were analyzed by Western blotting for the presence of TN-C. (C) Western blotting analysis of TN-C in three Laemmli sample buffer BJ cells and matrix extracts and soluble and insoluble decellularized, BJ-derived matrix extracts. (D) TN-C or fibronectin single staining, TN-C/fibronectin or no primary antibody with or without DAPI double-immunofluorescence staining of BJ cells and matrix. (BF: bright field; scale bar 400μm).

### 3.2.2 HEK 293 cells express *TNC* mRNA

To ensure that HEK 293 cells can be used as a cellular system to design and optimise CRISPR-Cas9-mediated editing of the *TNC* gene, expression of *TNC* mRNA in these cells was investigated.

The Human Protein Atlas database was searched for *TNC* expression in HEK 293 cells. Moreover, cells were cultured in T75 flasks to ~70% confluence and were stimulated with 100 ng/mL of LPS or left unstimulated for 24 hours. Cells were lysed to extract and purify total RNA and *TNC* mRNA expression was measured by RT-qPCR.

RNA-seq analysis of HEK 293 cell line deposited in the Human Protein Atlas showed that *TNC* is expressed in these cells at the transcriptional level (Fig. 3.2A). RT-qPCR performed in our lab confirmed that HEK 293 cells express *TNC* mRNA and showed that the levels of expression are significantly lower compared to BJ fibroblasts (Fig. 3.2B). These findings confirm that HEK 293 cells are suitable to test the CRISPR-Cas9-mediated genome editing approach.

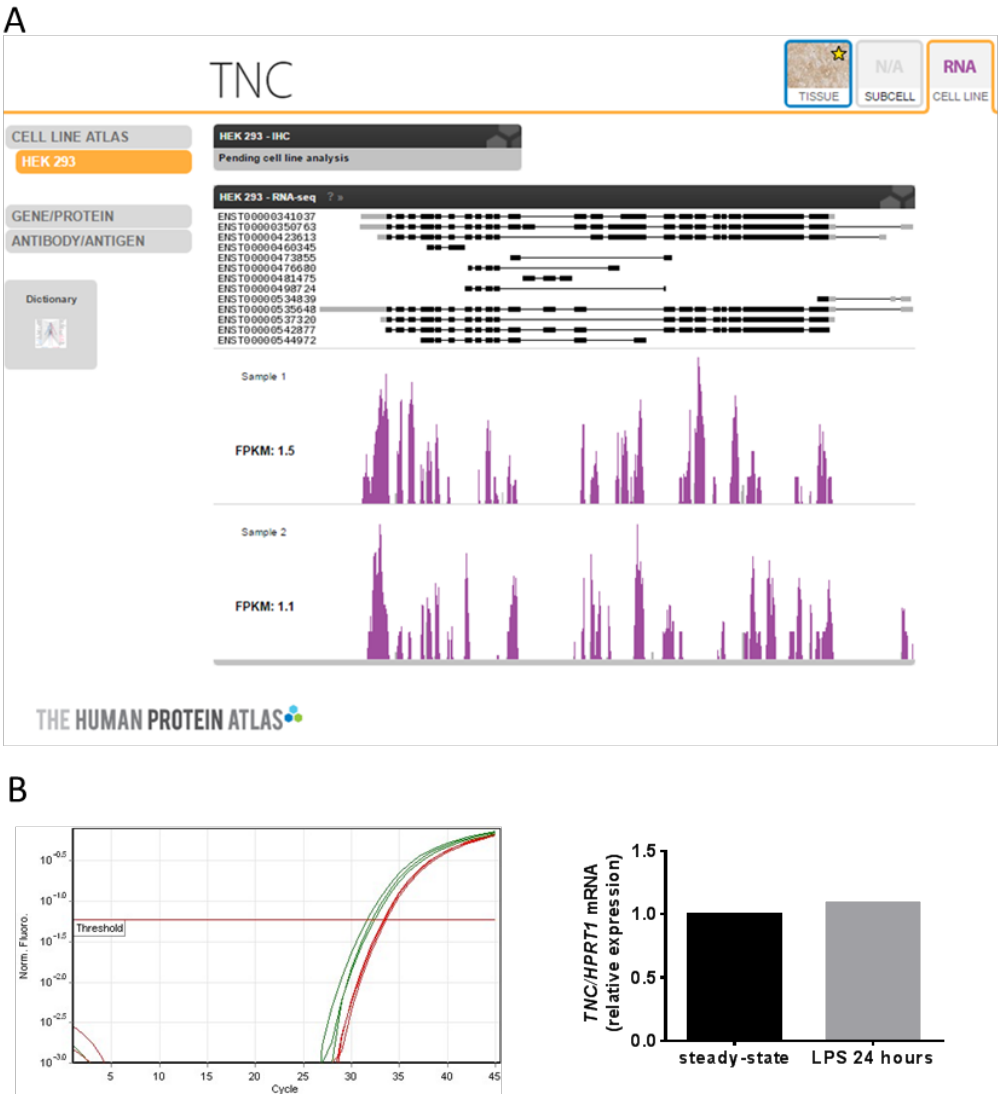



**Figure 3.2 HEK 293 cells express *TNC* at the transcriptional level.** (A) RNA-seq analysis of HEK 293 cell line from the Human Protein Atlas (http://www.proteinatlas.org/). On top the transcript layout for all coding and non-coding transcripts of the gene is shown. Untranslated regions (UTRs) of protein-coding transcripts are in gray. Gene abundance is quantified by the FPKM (Fragments Per Kilobase of exon per Million fragments mapped), which is calculated as the sum of FPKMs of all protein-coding transcripts. The plots show the mapped read coverage along all exons and introns of the gene. (B) RT-qPCR analysis of *TNC* mRNA in HEK 293 cells at steady-state or 24 hours after stimulation with 100 ng/ml LPS. Left: amplification plot and threshold cycles (LPS 24 hours in green and steady-state in red); right: results are presented relative to the abundance of *TNC* mRNA at steady-state.

## 3.3 Design of a CRISPR-Cas9 system to knock-out the *TNC* gene

### 3.3.1 Design of gRNAs

To knock-out the *TNC* gene using the CRISPR-Cas9n system, gRNAs were designed to bind in close proximity on the top and bottom DNA strands near the translation start codon ATG, which is found in exon 2 of the *TNC* gene. The design was carried out using the CRISPR design software (http://crispr.mit.edu/). For this, a 250 bp *TNC* coding sequence within exon 2, including the ATG, was identified and submitted (Fig. 3.3A). The software generates a list of all the possible gRNAs located exactly next to a DNA sequence encoding the species-specific protospacer adjacent motif (PAM) NGG at their 3' end and scans them for potential offtargets as well as target specificity. Single and pair gRNAs are then ranked based on an aggregate score and pairs with a score > 50 and 0 off-targets are considered candidate targeting sequences. Because it is still unclear which properties of the gRNA are important for gRNA specificity, three different pairs of gRNAs were designed in the region of interest to maximise the chances of success of the system (Fig. 3.3B-D).

The first pair of gRNAs (pair A) consists of the gRNAs A1 and A2, which have an offset of 28 bp, 0 offtargets and an aggregate score of 60 (Fig. 3.3B). The second pair of gRNAs (pair B) consists of the gRNAs B1 and B2, which have an offset of 11 bp, 0 offtargets and an aggregate score of 81 (Fig. 3.3BC). The third pair of gRNAs (pair C) consists of the gRNAs C1 and C2, which have an offset of 1 bp, 0 offtargets and an aggregate score of 66 (Fig. 3.3D). A schematic overview of the three gRNA pairs in the context of the locus of interest is shown in Figure 3.4.

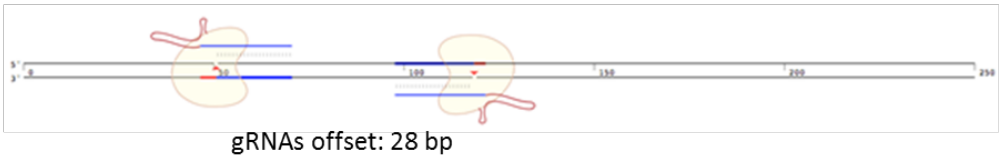
A

```
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
```

B

Pair of gRNAs: A (A1 and A2)

```
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
```



gRNAs offset: 28 bp

```
5′ CTGAGTCATGGCCCCCATGG  3′
3′ GACTCAGTACCGGGGGTACC  5′
```

gRNA A1 aggregate score: 63

```
5′ TTGCCCTCGCTACCGAAGGT  3′
3′ AACGGGAGCGATGGCTTCCA  5′
```

gRNA A2 aggregate score: 97

gRNA A1

63

**quality**
high
**cuts after position**
50 in query
**sequence**
CTGAGTCATGGCCCCCATGGTGG
**# offtargets**
218
**# genic offtargets**
27
Export off-targets to .csv

gRNA A2

97

**quality**
high
**cuts after position**
118 in query
**sequence**
TTGCCCTCGCTACCGAAGGTGGG
**# offtargets**
32
**# genic offtargets**
10
Export off-targets to .csv

Pair score, A1 & A2

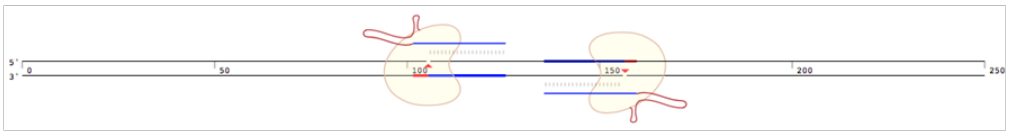
60

**quality**
high
**# offtarget pairs of A & B**
0
**# genic OT pairs of A & B**
0

C

Pair of gRNAs: B (B1 and B2)

```
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
```




gRNAs offset: 11 bp

```
5′ GGACCCCACCTTCGGTAGCG  3′
3′ CCTGGGGTGGAAGCCATCGC  5′
```

gRNA B1 aggregate score: 93

```
5′ ATCCGGCACAAGCGACAGAG  3′
3′ TAGGCCGTGTTCGCTGTCTC  5′
```

gRNA B2 aggregate score: 87

gRNA B1

93

**quality**
high
**cuts after position**
105 in query
**sequence**
GGACCCCACCTTCGGTAGCGAGG
**# offtargets**
56
**# genic offtargets**
10
Export off-targets to .csv

gRNA B2

87

**quality**
high
**cuts after position**
156 in query
**sequence**
ATCCGGCACAAGCGACAGAGTGG
**# offtargets**
65
**# genic offtargets**
9
Export off-targets to .csv

Pair score, B1 & B2

81

**quality**
no
**# offtarget pairs of A & B**
0
**# genic OT pairs of A & B**
0

D

Pair of gRNAs: C (C1 and C2)

```
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
```




gRNAs offset: 1 bp

```
5′ CGTGGTTAAACACCACTGGC  3′
3′ GCACCAATTTGTGGTGACCG  5′
```

gRNA C1 aggregate score: 88

```
5′ TTACAACATCAAGCTGCCAG  3′
3′ AATGTTGTAGTTCGACGGTC  5′
```

gRNA C2 aggregate score: 75

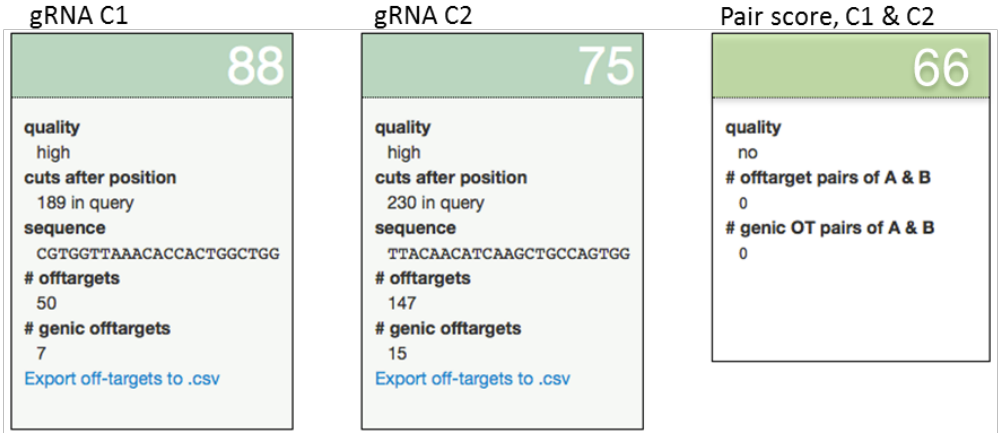



**Figure 3.3 Design of dual nickase CRISPR system to knockout *TNC* gene.** (A) The 250 bp *TNC* coding sequence (within exon 2; obtained from the NCBI gene database) submitted to the CRISPR design tool (http://crispr.mit.edu) for double nickase analysis. The translation start codon ATG is highlighted in green. (B-D) The top 3 pairs of guide oligonucleotides (gRNAs) near the start codon with the highest target specificity and the lowest to no potential genome-wide offtargets as indicated by the aggregate score for each individual guide and pair of guides. PAM sequences are highlighted in green, gRNAs in light orange and blue and ATG in red. Triangles indicate cleavage sites.

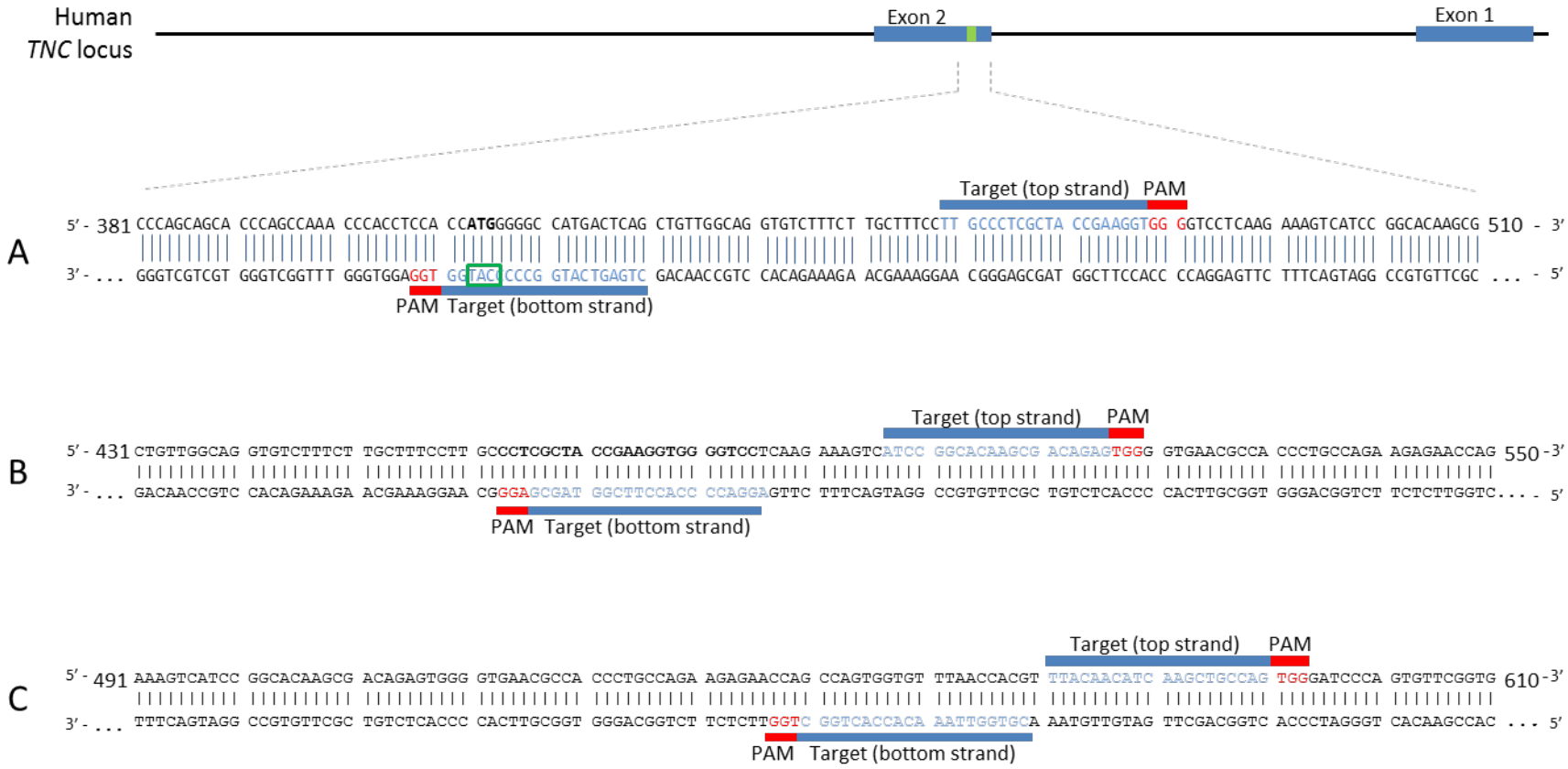



**Figure 3.4 Schematic representation of selected targets.** 20 nt long targets (in blue) followed by 5'-NGG (PAM sequence in red) at their 3' ends in the top and the bottom strand of the human *TNC* gene. Green outline highlights the translation start codon ATG in the bottom strand of genomic DNA.

### 3.3.2 Synthesis of gRNAs

To transiently express Cas9n and gRNAs in HEK 293 cells and BJ fibroblasts, we have chosen a mammalian expression plasmid (pSpCas9n(BB)-2A-GFP; PX461) containing a constitutive CBh Cas9 promoter, a U6 promoter for the gRNA and a GFP reporter gene. In order to clone the gRNAs into the plasmid, we added specific overhangs for ligation into the pSpCas9n(BB)-2A-GFP plasmid and a G for optimal U6 transcription at the 5' end of the gRNA sequence (Fig. 3.5). gRNA were synthesized by Sigma.

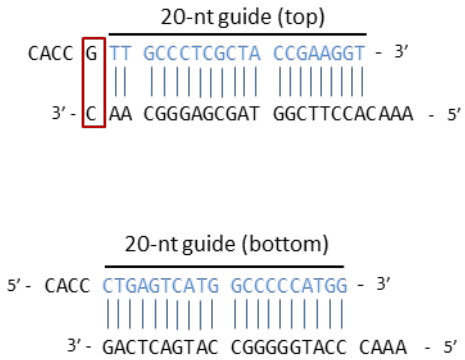



**Figure 3.5 Synthesis of gRNA for cloning.** The gRNAs for the top and bottom strands of the human *TNC* gene (in blue) include overhangs for ligation into the pSpCas9n(BB)-2A-GFP plasmid and a G-C pair (brown outline) at the 5' end of the gRNA for optimal U6 transcription. One gRNA is shown as an example.

### 3.3.3 Cloning of gRNAs

The cloning of each single gRNA into the PX461 mammalian expression plasmid was performed by digesting the vector with the *BbsI* restriction enzyme. The digested plasmid was then agarose gel purified to remove the 22 nucleotide-excised fragment and to separate the digested plasmid from any uncut plasmid (Fig. 3.6).

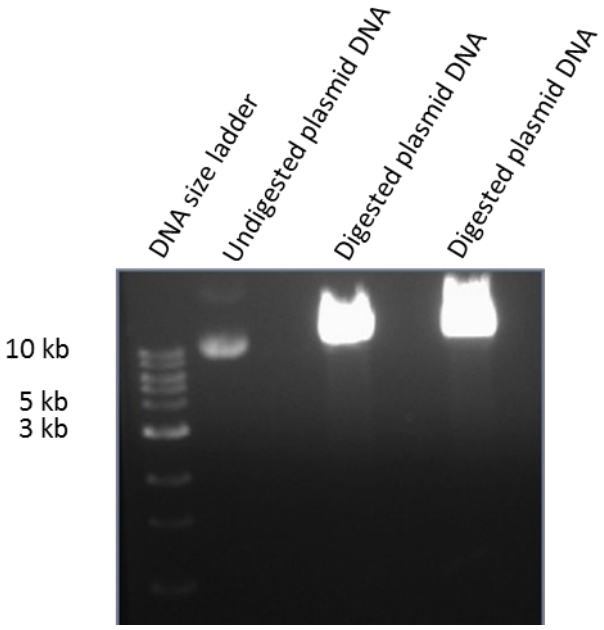



**Figure 3.6 Restriction digest of pSpCas9n(BB)-2A-GFP plasmid with *BbsI* restriction enzyme.** 5μL of undigested and digested plasmid DNA run on a 0.6% w/v agarose gel stained with ethidium bromide.

Each of the six gRNAs was annealed and phosphorylated, and then ligated into the digested PX461 plasmid (Fig. 3.7).

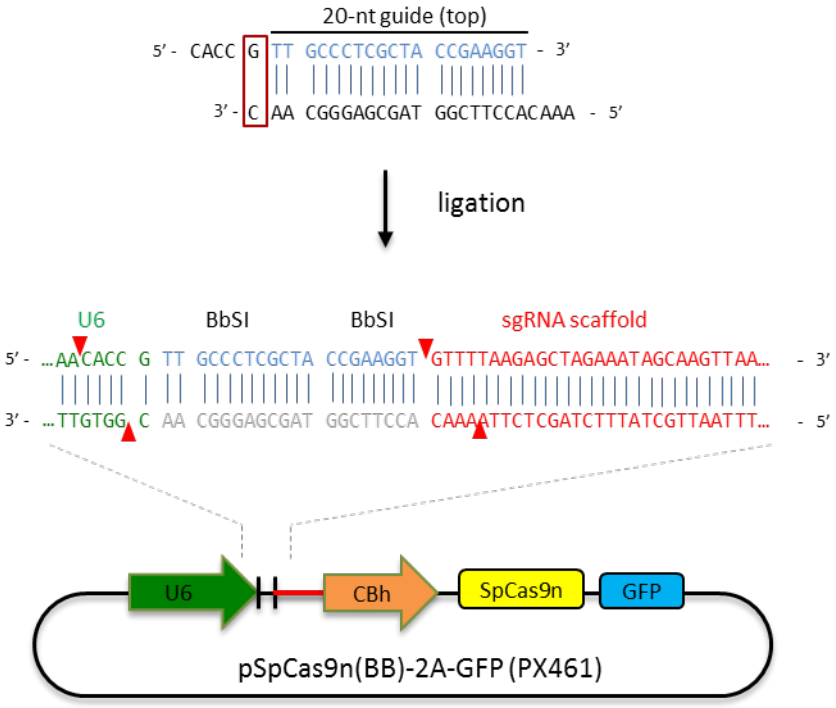



**Figure 3.7 Cloning of gRNAs.** Cloning of annealed gRNAs into plasmid by digestion of pSpCas9n(BB)-2A-GFP plasmid with *BbsI*. Insertion of a single gRNA is shown as an example.

The resulting ligation products were transformed into E. coli DH5α strain by heat shock and selected on ampicillin LB agar plates. One single fresh colony from each plate was miniprepped. Purified plasmid DNA concentration and purity was analysed (Table 14) and minipreps were sent for sequencing (Source Bioscience) with the human U6 forward primer to confirm correct insertion of the desired gRNA. Sequencing showed that all of the six gRNAs were successfully cloned into PX461-gRNAs. Thus, midipreps were done to obtain enough plasmid DNA of high purity and low endotoxin contamination for transfection into mammalian cells (Table 15).

| ID | ng/µL | A260/A280 | A260/A230 |
|---|---|---|---|
| A1 | 591.5 | 1.88 | 2.20 |
| A2 | 543.3 | 1.87 | 2.22 |
| B1 | 478.1 | 1.89 | 2.23 |
| B2 | 432.2 | 1.89 | 2.19 |
| C1 | 800.7 | 1.88 | 2.10 |
| C2 | 768.3 | 1.88 | 2.20 |

**Table 14** Plasmid DNA concentration of minipreps. The ID refers to the specific gRNA cloned into the plasmid.

| ID | µg/µL | A260/A280 | A260/A230 |
|---|---|---|---|
| A1 | 2.4 | 1.89 | 2.24 |
| A2 | 2.6 | 1.88 | 2.21 |
| B1 | 2.1 | 1.90 | 2.25 |
| B2 | 2.2 | 1.88 | 2.23 |
| C1 | 2.1 | 1.88 | 2.25 |
| C2 | 2.4 | 1.88 | 2.23 |

**Table 15** Plasmid DNA concentration of midipreps. The ID refers to the specific gRNA cloned into the plasmid.

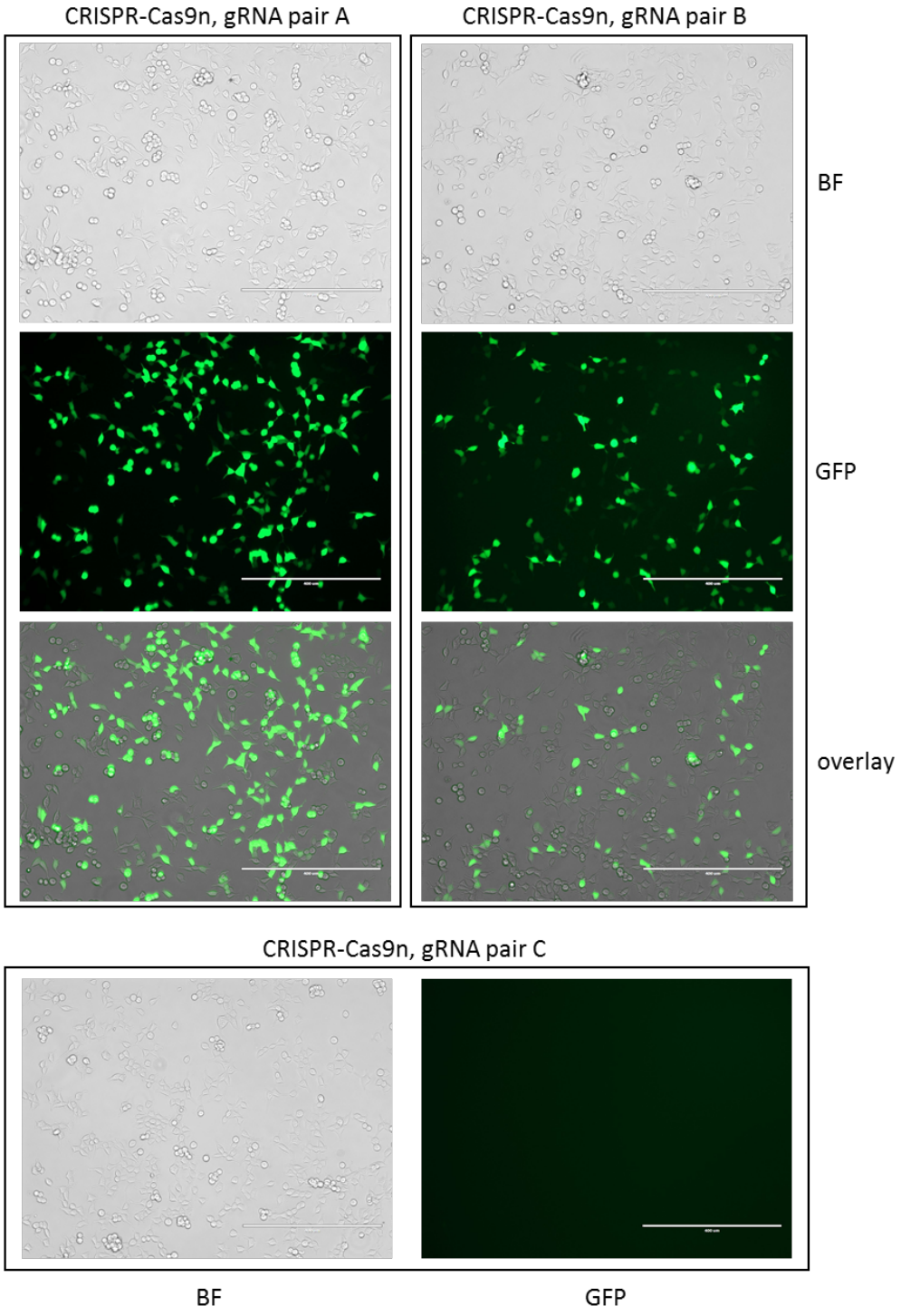
## 3.4 Delivering Cas9n and gRNAs

### 3.4.1 Transfection of Cas9n and gRNAs into HEK 293 cells

The purified plasmid DNA samples (Table 3.2), each containing Cas9n and a gRNA, were transfected into HEK 293 cells using the calcium phosphate method. For this, cells were seeded in 6-well plates and, when they reached ~80 % confluence, they were left non-transfected or transfected with 1 μg of each gRNA either in pairs or individually as a negative control for the Cas9n activity. A mock-transfection with all the reaction steps and reagents, apart from the plasmid DNA, was performed to control for the toxicity of the transfection method on the cells. Transfection of a GFP-plasmid control (kindly provided by Dr Hilary Collins) was also included in the experiment as a positive control for the transfection reaction.

Cas9n and gRNA pair A were transfected into HEK 293 cells with the highest transfection efficiency (~80 %), which was estimated by light microscopy combined with fluorescence microscopy analysis of the reporter green fluorescent protein (GFP) expression (Fig. 3.8A). gRNA pair B were successfully transfected into the cells, but with lower efficiency compared to gRNA pair A (Fig. 3.8A). The third gRNA pair C and single gRNA C1 and C2 were consistently not transfected into the cells (Fig. 3.8A). The quality of the purified DNA is a critical factor for transfection. However, the sequencing results and spectrophotometric analysis of the plasmids containing these gRNAs were similar to those of the other gRNAs that were successfully transfected. A possible explanation of the failure to transfect these guides may be the loss of plasmid integrity. Non transfected and mock-transfected cells showed no GFP fluorescence while cells transfected with the single gRNA A1 or a GFP-plasmid control showed GFP fluorescence as expected (Fig. 3.8B). These results indicate that the plasmid DNA containing the gRNA pairs A and B can be transfected into mammalian cells and the Cas9n activity at the target DNA sequence can be validated.

A


CRISPR-Cas9n, gRNA pair A
CRISPR-Cas9n, gRNA pair B
BF
GFP
overlay
CRISPR-Cas9n, gRNA pair C
BF
GFP

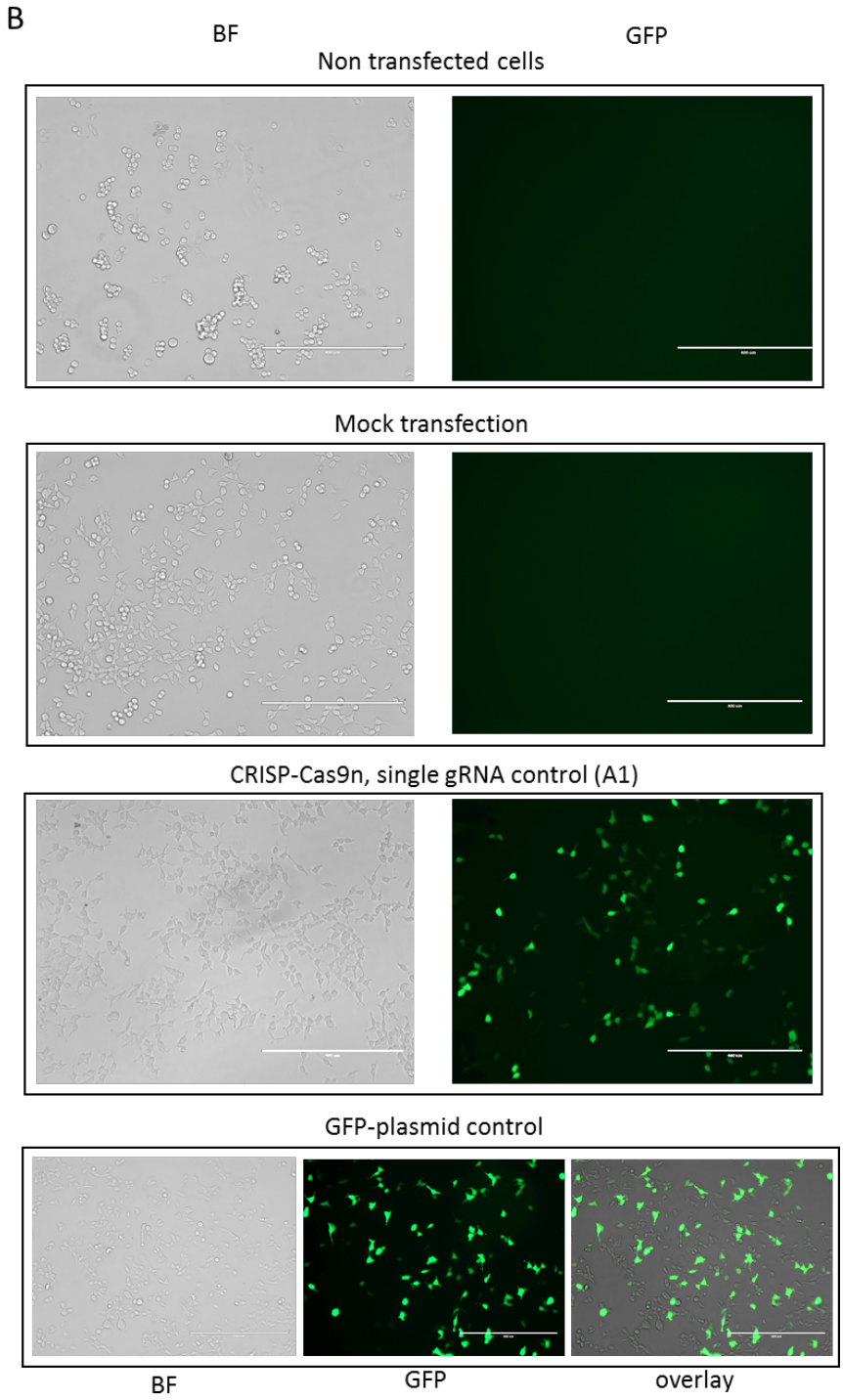



**Figure 3.8 Delivery of CRISPR-Cas9n and gRNAs into HEK 293 cells by calcium phosphate transfection.** (A) Expression of reporter green fluorescent protein (GFP) in HEK 293 cells 24 hours after transfection with pSpCas9n(BB)-2A-GFP plasmid containing specfic gRNAs. (B) Expression of GFP in non or mock-transfected cells and in cells transfected with a single gRNA (negative controls), and cells transfected with a GFP-plasmid (positive control). One representative image for each condition in one of three experiments is shown. BF: bright field; GFP: green fluorescence protein. Scale bar: 400 µm.

### 3.4.2 Transfection of Cas9n and gRNAs into BJ fibroblasts

The transfection of plasmid DNA into fibroblasts, including BJ fibroblasts, is known to represent a great challenge. The proliferation rate of these cells is significantly lower than other cell lines like HEK 293 and this makes them harder to transfect with high efficiency. However, transfection of BJ cells can be achieved as it has been reported before (Chiquet-Ehrismann et al., 2014, Avci-Adali et al., 2014).

To transfect BJ fibroblasts, we tested three different approaches: cationic lipid-based reagents, non-lipid-based chemical reagents and electroporation. All of the methods and conditions tested are described in detail in section 2.5 of the materials and methods. Briefly, the transfection efficiency obtained with the cationic lipids and non-lipid chemical reagents was very low, with 10-20% of total cells being transfected, which is insufficient to subsequently perform single transfected cell sorting and expansion of single-cell clones.

For electroporation of BJ fibroblasts, we used two different systems: the Neon® Transfection system from Invitrogen and the 4D-Nucleofector™ system from Lonza. Both Invitrogen and Lonza had established and published a protocol for the electroporation of BJ fibroblasts that we followed. For electroporation with the Neon system, a cell density of 5 x $10^6$ cells/mL was used and two different amounts of each plasmid DNA were tested (500ng and 1μg). A single electric pulse of 1,650 v for 20 ms could transfect 35-50% of the cells as observed by light and fluorescence microscopy 24 hours after electroporation (Fig. 3.9).

The mammalian expression vector used in this study allows transient expression of Cas9n and gRNAs. To verify this and estimate the kinetic of expression of Cas9 and gRNAs in BJ fibroblasts, the expression of the reporter GFP was followed for 9 days. Expression peaked at day 1-2 after transfection and started to gradually decline at day 3 until it completely switched off at day 8-9 (Fig. 3.9).

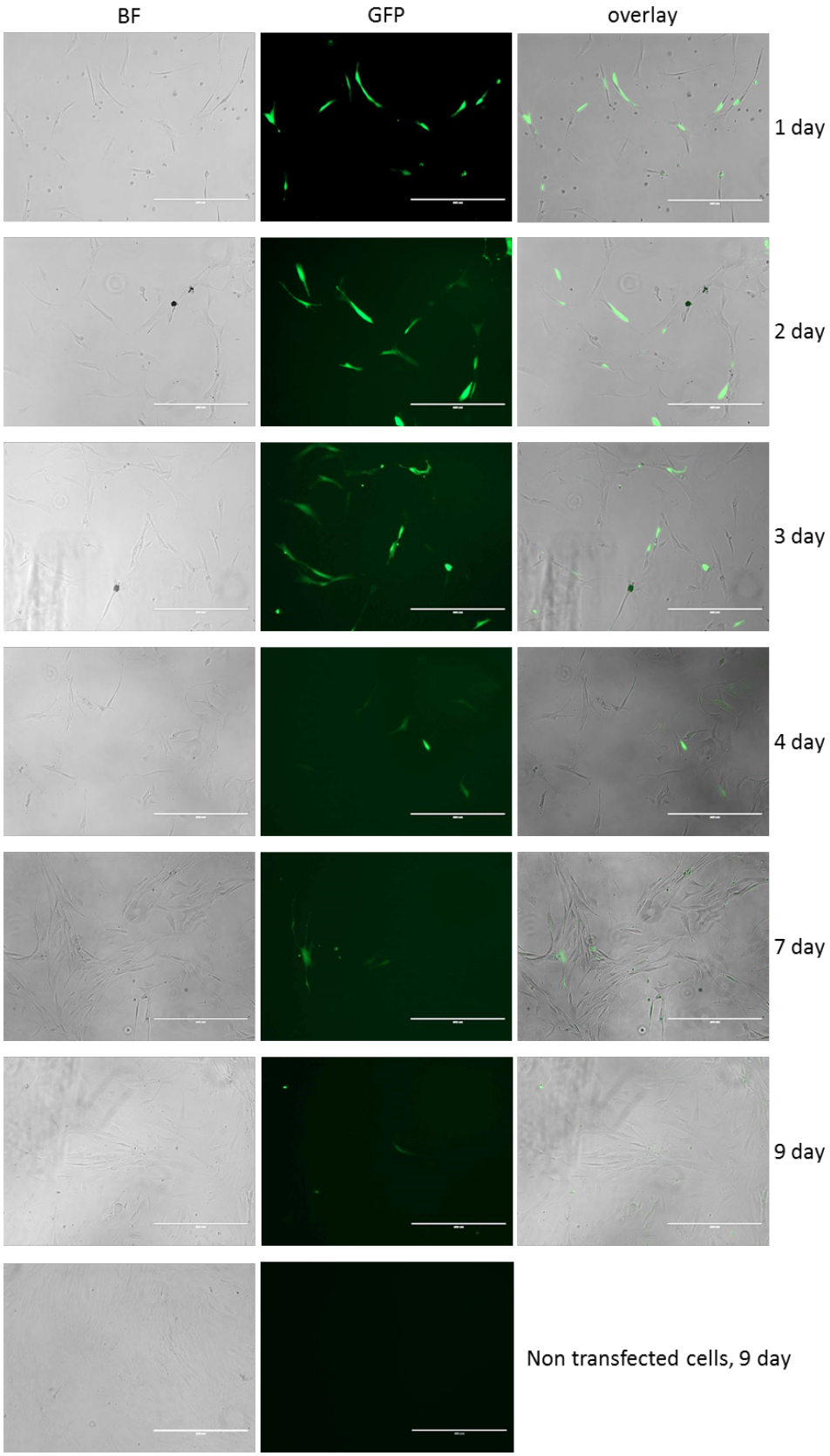



**Figure 3.9 Transient expression of reporter GFP in BJ fibroblasts following transfection of pSpCas9n(BB)-2A-GFP plasmid containing gRNA A1 by electroporation.** Expression of reporter GFP in BJ fibroblasts at 1, 2, 3, 4, 7 and 9 days after transfection. One representative image for each condition is shown. BF: bright field; GFP: green fluorescence protein. Scale bar: 400 µm.

To determine the amount of plasmid DNA that gives the highest transfection efficiency and the lowest cytotoxicity, two amounts of each plasmid DNA were tested: 500ng and 1μg of each plasmid DNA for the delivery of 1μg and 2μg of plasmid DNA in total. 1μg was the maximum amount of total plasmid DNA tolerated by the cells (Fig. 3.10).

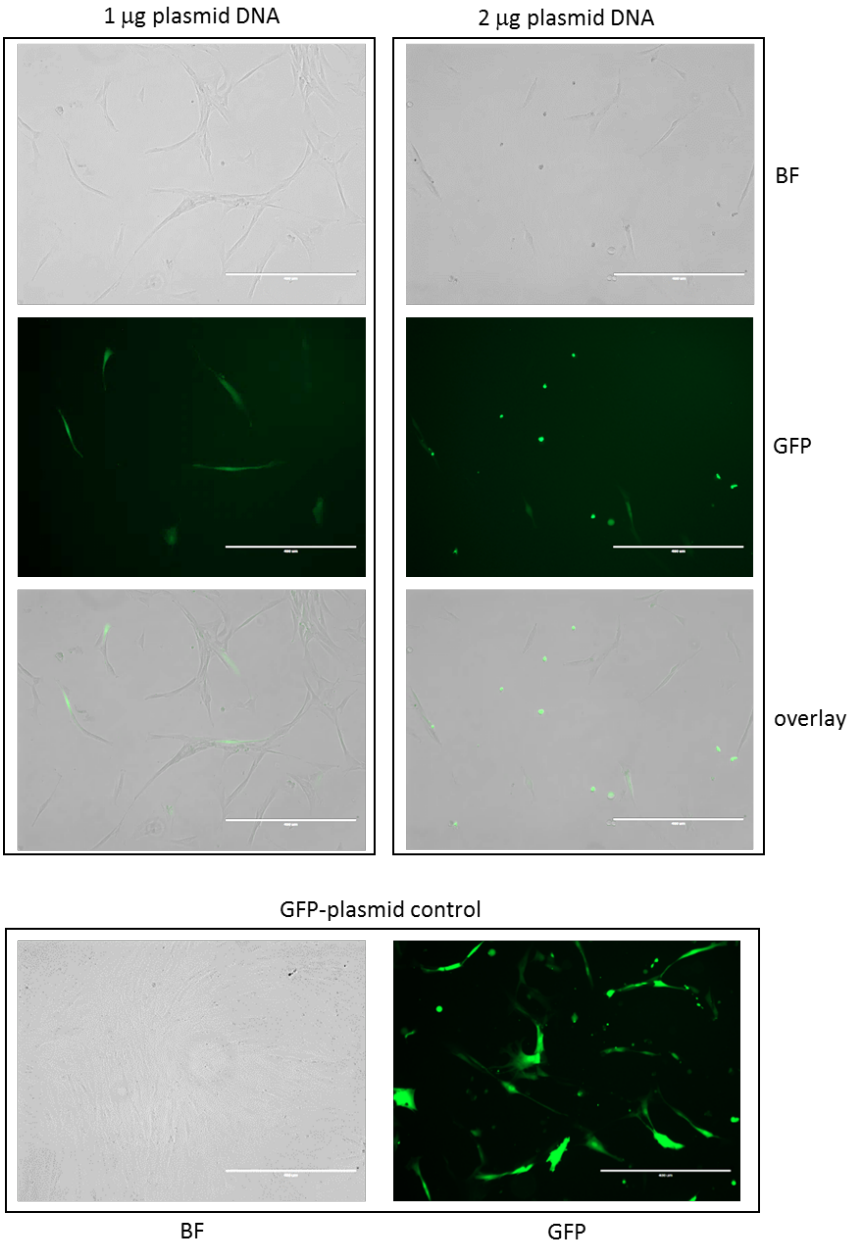



**Figure 3.10 Effect of the total amount of plasmid DNA on BJ fibroblasts viability following electroporation.** Expression of reporter GFP in BJ fibroblasts 24 hours after transfection of 1 μg total plasmid DNA (500 ng of pSpCas9n(BB)-2A-GFP containing gRNA A1 plus 500 ng of pSpCas9n(BB)-2A-GFP containing gRNA A2) and 2 μg total plasmid DNA (1 μg of pSpCas9n(BB)-2A-GFP containing gRNA A1 plus 1 μg of pSpCas9n(BB)-2A-GFP containing gRNA A2). One representative image for each condition is shown. BF: bright field; GFP: green fluorescence protein. Scale bar: 400 μm.

Even though transfection with the Neon electroporation system significantly increased the transfection efficiency of BJ fibroblasts, it also resulted in high cell death as observed by light microscopy analysis of transfected cells compared to non transfected cells. Thus, I tested the 4D-Nucleofector system from Lonza, which was delivered to our laboratory in June for a demo.

For electroporation with the 4D-Nucleofector system, 3 x $10^5$ cells were transfected with 1 µg of plasmid DNA or pmaxGFP positive control (Lonza). The transfection efficiency estimated by light and fluorescence microscopy analysis was comparable to that obtained with the Neon system, but no cell death was observed. Thus, electroporation using the 4D-Nucleofector system was the method of choice for the subsequent transfection of BJ fibroblasts with Cas9n and the gRNA pairs A and B followed by single cell sorting, single-cell clone expansion and genome edit validation.

### 3.5 Validation of *TNC* gene edit

To facilitate validation of CRISPR-Cas9n-mediated *TNC* gene edit in mammalian cells, fluorescence in situ hybridization (FISH) was performed to estimate the target gene copy number in HEK 293 cells and BJ fibroblasts by determining the target chromosome 9 copy number. *TNC* gene edit was then validated at the gene, RNA and protein level.

#### 3.5.1 Estimation of gene copy number in HEK 293 cells and BJ fibroblasts by FISH

Cell lines can carry more than 2 copies of a gene, challenging complete gene knock-out by the CRISPR-Cas9 system. The target gene copy number in HEK 293 cells and BJ fibroblasts was estimated by determining the number of copies of the chromosome 9, which harbours the *TNC* gene, by FISH.

Cell were fixed on microscope slides and probes containing fluorophore-labelled DNA specific for human chromosome 9 were hybridized to the

nuclear DNA of the cells. DAPI was used to visualize single nuclei. Chromosome enumeration was done using a fluorescence microscope and a 100X objective. Three slides for each cell type were screened. All of the BJ cells and > 97% of the HEK 293 cells that could be analysed were diploid (Fig. 3.11-3.14). < 3% of the HEK 293 cells carried three to four copies of chromosome 9.

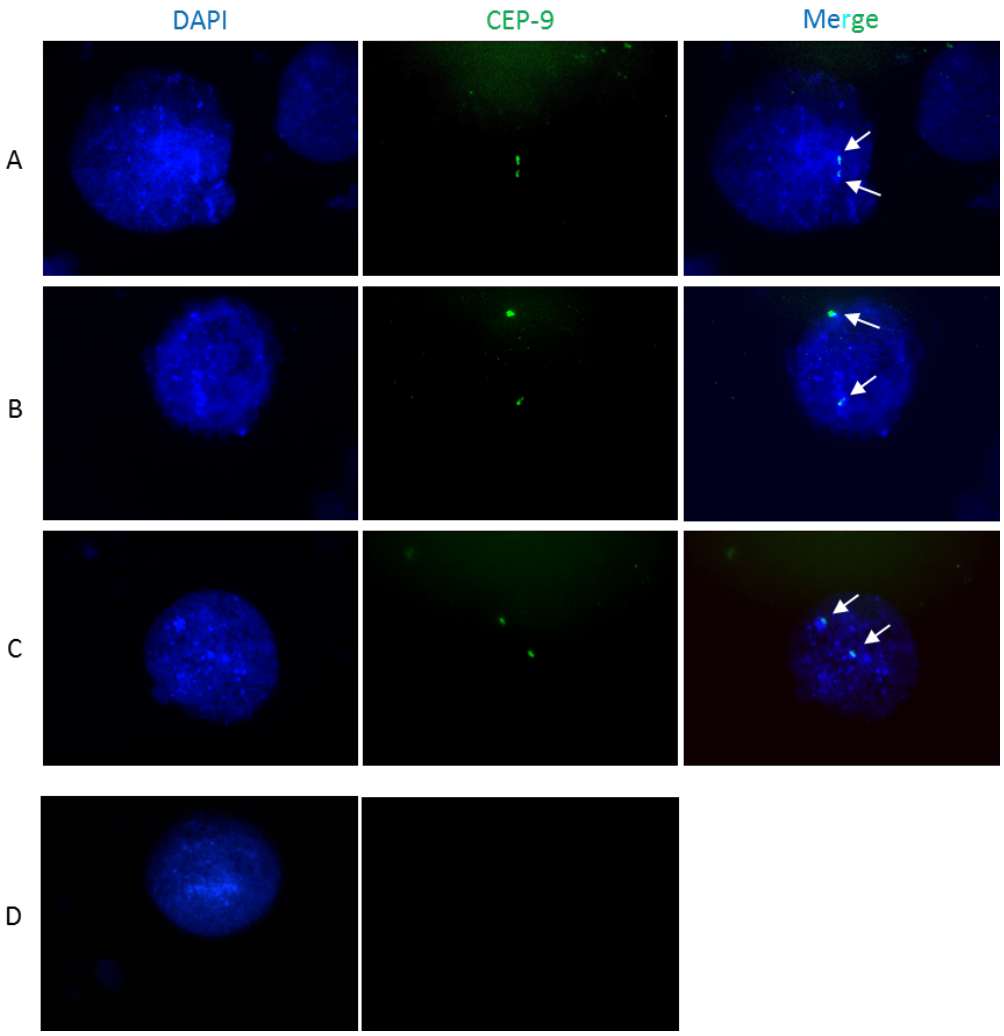



**Figure 3.11 Estimation of target chromosome 9 copy number in HEK 293 cells by fluorescence in situ hybridization (FISH).** Images were captured by fluorescence microscopy using a 100X oil immersion objective. White arrows indicate copies of chromosome 9. Three representative images from three experiments (A-C) and a negative control without CEP-9 (Chromosome Enumeration Probe-9) probe (D) are shown.

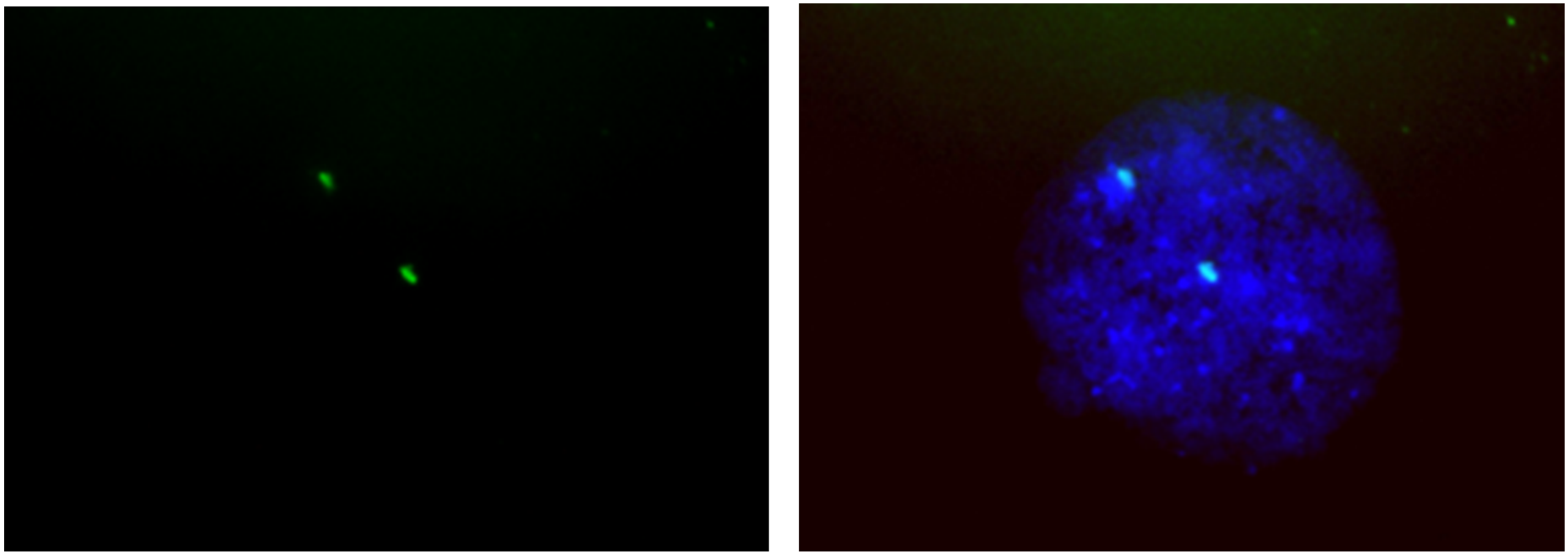


**Figure 3.12 Estimation of target chromosome 9 copy number in HEK 293 cells by fluorescence in situ hybridization (FISH).** Images were captured by fluorescence microscopy using a 100X oil immersion objective. White arrows indicate copies of chromosome 9. Large representative image is shown.

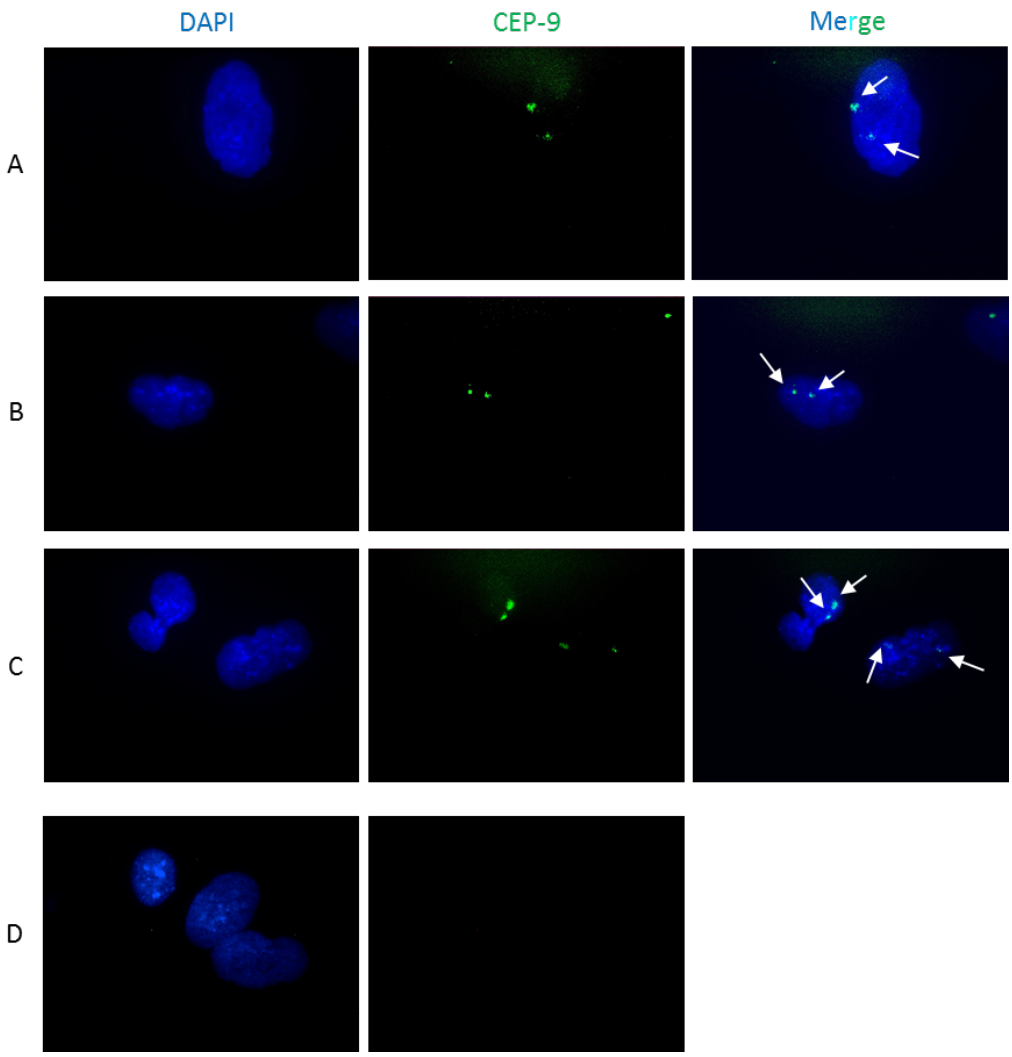



**Figure 3.13 Estimation of target chromosome 9 copy number in BJ fibroblasts by fluorescence in situ hybridization (FISH).** Images were captured by fluorescence microscopy using a 100X oil immersion objective. White arrows indicate copies of chromosome 9. Three representative images from three experiments (A-C) and a negative control without CEP-9 probe (D) are shown.

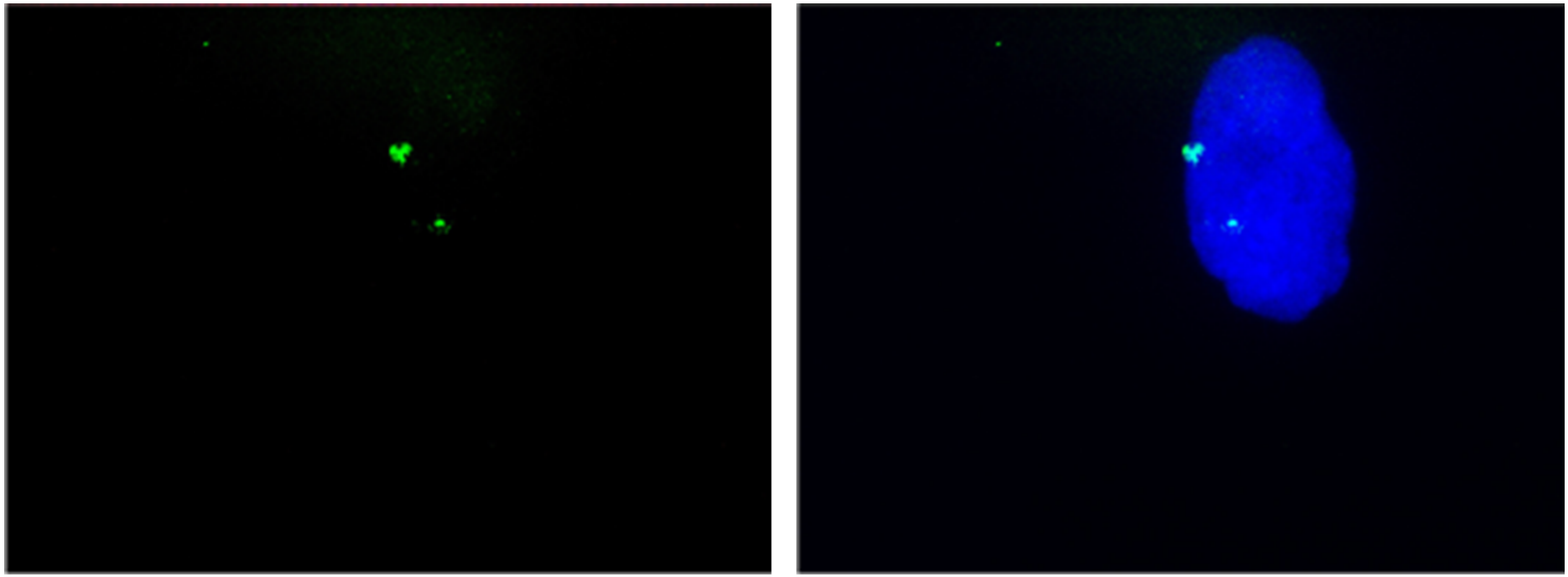


**Figure 3.14 Estimation of target chromosome 9 copy number in BJ fibroblasts by fluorescence in situ hybridization (FISH).** Images were captured by fluorescence microscopy using a 100X oil immersion objective. White arrows indicate copies of chromosome 9. Large representative image is shown.

### 3.5.2 Validation of *TNC* gene edit in HEK 293 cells by PCR amplification and sequencing

#### 3.5.2.1 Design of gene edit screening primers

To determine if the *TNC* gene in HEK 293 cells was edited, following transfection with the CRISPR-Cas9n and gRNA pairs A and B respectively (Fig. 3.8), PCR amplification and sequencing of the targeted DNA region were done. For this, 5 pairs of primers flanking the regions to be edited by Cas9n, under the direction of the gRNA pair A and B respectively, were designed.

As shown in table 6 (section 2.7.2.1 of materials and methods), each pair of primers, forward and reverse, was designed to amplify an area of exon 2 of the *TNC* gene that starts at least ~30 bp upstream of the PAM sequence of the top strand gRNA and ends at least ~35 bp downstream of the PAM sequence of the bottom strand gRNA (Fig. 3.15). The length of the amplicons generated with these primers ranged from 177 bp to 781 bp (Table 6).

A

```
26881 TGTCTCCTTA AATTTACAGA AGTTGAATAG TGAGTGAAAT CATTCCTTCA AGTGACAAAG
26941 TTCTTTTAGA GACCATTGTA TGTTATTTCA AGGGTATATA ATGAAAAATT ATCAGGAGAG
27001 CTTCAAATAC AGGGATATAA AATCAAAATT CACTTGGAAT TGTAGATATT CAATATATTC
27061 AATAGTAGAT ACTCATAATT TTTTTGTTCA TCTTTGCTTG TAGAAAATAA TTCTTTCAAG
27121 AAGATCAGGG ACAACTGATT TGAAGTCTAC TCTGTGCTTC TAAATCCCCA ATTCTGCTGA
27181 AAGTGAGATA CCCTAGAGCC CTAGAGCCCC AGCAGCACCC AGCCAAACCC ACCTCCACCA
27241 TGGGGGCCAT GACTCAGCTG TTGGCAGGTG TCTTTCTTGC TTTCCTTGCC CTCGCTACCG
27301 AAGGTGGCGT CCTCAAGAAA GTCATCCGGC ACAAGCGACA GAGTGGGGTG AACGCCACCC
27361 TGCCAGAAGA GAACCAGCCA GTGGTGTTTA ACCACGTTTA CAACATCAAG CTGCCAGTGG
27421 GATCCCAGTG TTCGGTGGAT CTGGAGTCAG CCAGTGGGGA GAAAGACCTG GCACCGCCTT
27481 CAGAGCCCAG CGAAAGCTTT CAGGAGCACA CAGTGGATGG GGAAAACCAG ATTGTCTTCA
27541 CACATCGCAT CAACATCCCC CGCCGGGCCT GTGGCTGTGC CGCAGCCCCT GATGTTAAGG
27601 AGCTGCTGAG CAGACTGGAG GAGCTGGAGA ACCTGGTGTC TTCCCTGAGG GAGCAATGTA
27661 CTGCAGGAGC AGGCTGCTGT CTCCAGCCTG CCACAGGTAT GAGCTGGTGG CCCGCCCAGT
27721 CCCACACTGT GCAAACATGG AGAAATGAAG AGCATAGCAG CATCCATCAC TTGCTTCCAA
27781 AGTGGGCTCA CTAGAAAACT AGCTTCCCAG AGTGTTCATT ACAAAGAGAT TTTATGTTCA
27841 AATAAGTGTG GAGACTGCAG CCATCTCTCT CCTGCTTGGA GGAGTCACGG CGACAACTGC
```

GATACCCTAG AGCCCTAGAG 1Fw
GGTGAACGCC ACCCTGCCAG 1Rv

B

```
26881 TGTCTCCTTA AATTTACAGA AGTTGAATAG TGAGTGAAAT CATTCCTTCA AGTGACAAAG
26941 TTCTTTTAGA GACCATTGTA TGTTATTTCA AGGGTATATA ATGAAAAATT ATCAGGAGAG
27001 CTTCAAATAC AGGGATATAA AATCAAAATT CACTTGGAAT TGTAGATATT CAATATATTC
27061 AATAGTAGAT ACTCATAATT TTTTTGTTCA TCTTTGCTTG TAGAAAATAA TTCTTTCAAG
27121 AAGATCAGGG ACAACTGATT TGAAGTCTAC TCTGTGCTTC TAAATCCCCA ATTCTGCTGA
27181 AAGTGAGATA CCCTAGAGCC CTAGAGCCCC AGCAGCACCC AGCCAAACCC ACCTCCACCA
27241 TGGGGGCCAT GACTCAGCTG TTGGCAGGTG TCTTTCTTGC TTTCCTTGCC CTCGCTACCG
27301 AAGGTGGCGT CCTCAAGAAA GTCATCCGGC ACAAGCGACA GAGTGGGGTG AACGCCACCC
27361 TGCCAGAAGA GAACCAGCCA GTGGTGTTTA ACCACGTTTA CAACATCAAG CTGCCAGTGG
27421 GATCCCAGTG TTCGGTGGAT CTGGAGTCAG CCAGTGGGGA GAAAGACCTG GCACCGCCTT
27481 CAGAGCCCAG CGAAAGCTTT CAGGAGCACA CAGTGGATGG GGAAAACCAG ATTGTCTTCA
27541 CACATCGCAT CAACATCCCC CGCCGGGCCT GTGGCTGTGC CGCAGCCCCT GATGTTAAGG
27601 AGCTGCTGAG CAGACTGGAG GAGCTGGAGA ACCTGGTGTC TTCCCTGAGG GAGCAATGTA
27661 CTGCAGGAGC AGGCTGCTGT CTCCAGCCTG CCACAGGTAT GAGCTGGTGG CCCGCCCAGT
27721 CCCACACTGT GCAAACATGG AGAAATGAAG AGCATAGCAG CATCCATCAC TTGCTTCCAA
27781 AGTGGGCTCA CTAGAAAACT AGCTTCCCAG AGTGTTCATT ACAAAGAGAT TTTATGTTCA
27841 AATAAGTGTG GAGACTGCAG CCATCTCTCT CCTGCTTGGA GGAGTCACGG CGACAACTGC
```

CTCTGTGCTT CTAAATCCCC 2Fw
GCCAGAAGAG AACCAGCCAG 2Rv

C

```
26881 TGTCTCCTTA AATTTACAGA AGTTGAATAG TGAGTGAAAT CATTCCTTCA AGTGACAAAG
26941 TTCTTTTAGA GACCATTGTA TGTTATTTCA AGGGTATATA ATGAAAAATT ATCAGGAGAG
27001 CTTCAAATAC AGGGATATAA AATCAAAATT CACTTGGAAT TGTAGATATT CAATATATTC
27061 AATAGTAGAT ACTCATAATT TTTTTGTTCA TCTTTGCTTG TAGAAAATAA TTCTTTCAAG
27121 AAGATCAGGG ACAACTGATT TGAAGTCTAC TCTGTGCTTC TAAATCCCCA ATTCTGCTGA
27181 AAGTGAGATA CCCTAGAGCC CTAGAGCCCC AGCAGCACCC AGCCAAACCC ACCTCCACCA
27241 TGGGGGCCAT GACTCAGCTG TTGGCAGGTG TCTTTCTTGC TTTCCTTGCC CTCGCTACCG
27301 AAGGTGGGGT CCTCAAGAAA GTCATCCGGC ACAAGCGACA GAGTGGCGGTG AACGCCACCC
27361 TGCCAGAAGA GAACCAGCCA GTGGTGTTTA ACCACGTTTA CAACATCAAG CTGCCAGTGG
27421 GATCCCAGTG TTCGGTGGAT CTGGAGTCAG CCAGTGGGGA GAAAGACCTG GCACCGCCTT
27481 CAGAGCCCAG CGAAAGCTTT CAGGAGCACA CAGTGGATGG GGAAAACCAG ATTGTCTTCA
27541 CACATCGCAT CAACATCCCC CGCCGGGCCT GTGGCTGTGC CGCAGCCCCT GATGTTAAGG
27601 AGCTGCTGAG CAGACTGGAG GAGCTGGAGA ACCTGGTGTC TTCCCTGAGG GAGCAATGTA
27661 CTGCAGGAGC AGGCTGCTGT CTCCAGCCTG CCACAGGTAT GAGCTGGTGG CCCGCCCAGT
27721 CCCACACTGT GCAAACATGG AGAAATGAAG AGCATAGCAG CATCCATCAC TTGCTTCCAA
27781 AGTGGGCTCA CTAGAAAACT AGCTTCCCAG AGTGTTCATT ACAAAGAGAT TTTATGTTCA
27841 AATAAGTGTG GAGACTGCAG CCATCTCTCT CCTGCTTGGA GGAGTCACGG CGACAACTGC
```

GGGCCATGAC TCAGCTGTTG 1Fw
CAACATCAAG CTGCCAGTGG 1Rv

D

```
26881 TGTCTCCTTA AATTTACAGA AGTTGAATAG TGAGTGAAAT CATTCCTTCA AGTGACAAAG
26941 TTCTTTTAGA GACCATTGTA TGTTATTTCA AGGGTATATA ATGAAAAATT ATCAGGAGAG
27001 CTTCAAATAC AGGGATATAA AATCAAAATT CACTTGGAAT TGTAGATATT CAATATATTC
27061 AATAGTAGAT ACTCATAATT TTTTTGTTCA TCTTTGCTTG TAGAAAATAA TTCTTTCAAG
27121 AAGATCAGGG ACAACTGATT TGAAGTCTAC TCTGTGCTTC TAAATCCCCA ATTCTGCTGA
27181 AAGTGAGATA CCCTAGAGCC CTAGAGCCCC AGCAGCACCC AGCCAAACCC ACCTCCACCA
27241 TGGGGGCCAT GACTCAGCTG TTGGCAGGTG TCTTTCTTGC TTTCCTTGCC CTCGCTACCG
27301 AAGGTGGGGT CCTCAAGAAA GTCATCCGGC ACAAGCGACA GAGTGGGGTG AACGCCACCC
27361 TGCCAGAAGA GAACCAGCCA GTGGTGTTTA ACCACGTTTA CAACATCAAG CTGCCAGTGG
27421 GATCCCAGTG TTCGGTGGAT CTGGAGTCAG CCAGTGGGGA GAAAGACCTG GCACCGCCTT
27481 CAGAGCCCAG CGAAAGCTTT CAGGAGCACA CAGTGGATGG GGAAAACCAG ATTGTCTTCA
27541 CACATCGCAT CAACATCCCC CGCCGGGCCT GTGGCTGTGC CGCAGCCCCT GATGTTAAGG
27601 AGCTGCTGAG CAGACTGGAG GAGCTGGAGA ACCTGGTGTC TTCCCTGAGG GAGCAATGTA
27661 CTGCAGGAGC AGGCTGCTGT CTCCAGCCTG CCACAGGTAT GAGCTGGTGG CCCGCCCAGT
27721 CCCACACTGT GCAAACATGG AGAAATGAAG AGCATAGCAG CATCCATCAC TTGCTTCCAA
27781 AGTGGGCTCA CTAGAAAACT AGCTTCCCAG AGTGTTCATT ACAAAGAGAT TTTATGTTCA
27841 AATAAGTGTG GAGACTGCAG CCATCTCTCT CCTGCTTGGA GGAGTCACGG CGACAACTGC

      GATACCCTAG AGCCCTAGAG 2Fw
      GATCTGGAGT CAGCCAGTGG 2Rv
```

E

```
26641 TATAACACTA TAATAATTCA CTCAGAAGAC ATAGTAAATA CTGAAGGAAT TAATGAATAC
26701 ACAAATAACT TCTACATCAT TCTTTTAACC CTGTCATTAG CAACTATGAT CTTCACTATA
26761 CAGAAAGTGA AATAAGGCAA AGAAAAATTA CATAACGTAT CCAAGGGACC AGAGCTACAG
26821 CCAAGATACA AACCTATGTC TATTTTATTT CAATATTTTA CCTTTTTCTC TCTATTTTTC
26881 TGTCTCCTTA AATTTACAGA AGTTGAATAG TGAGTGAAAT CATTCCTTCA AGTGACAAAG
26941 TTCTTTTAGA GACCATTGTA TGTTATTTCA AGGGTATATA ATGAAAAATT ATCAGGAGAG
27001 CTTCAAATAC AGGGATATAA AATCAAAATT CACTTGGAAT TGTAGATATT CAATATATTC
27061 AATAGTAGAT ACTCATAATT TTTTTGTTCA TCTTTGCTTG TAGAAAATAA TTCTTTCAAG
27121 AAGATCAGGG ACAACTGATT TGAAGTCTAC TCTGTGCTTC TAAATCCCCA ATTCTGCTGA
27181 AAGTGAGATA CCCTAGAGCC CTAGAGCCCC AGCAGCACCC AGCCAAACCC ACCTCCACCA
27241 TGGGGGCCAT GACTCAGCTG TTGGCAGGTG TCTTTCTTGC TTTCCTTGCC CTCGCTACCG
27301 AAGGTGGGGT CCTCAAGAAA GTCATCCGGC ACAAGCGACA GAGTGGGGTG AACGCCACCC
27361 TGCCAGAAGA GAACCAGCCA GTGGTGTTTA ACCACGTTTA CAACATCAAG CTGCCAGTGG
27421 GATCCCAGTG TTCGGTGGAT CTGGAGTCAG CCAGTGGGGA GAAAGACCTG GCACCGCCTT
27481 CAGAGCCCAG CGAAAGCTTT CAGGAGCACA CAGTGGATGG GGAAAACCAG ATTGTCTTCA
27541 CACATCGCAT CAACATCCCC CGCCGGGCCT GTGGCTGTGC CGCAGCCCCT GATGTTAAGG
27601 AGCTGCTGAG CAGACTGGAG GAGCTGGAGA ACCTGGTGTC TTCCCTGAGG GAGCAATGTA
27661 CTGCAGGAGC AGGCTGCTGT CTCCAGCCTG CCACAGGTAT GAGCTGGTGG CCCGCCCAGT
27721 CCCACACTGT GCAAACATGG AGAAATGAAG AGCATAGCAG CATCCATCAC TTGCTTCCAA

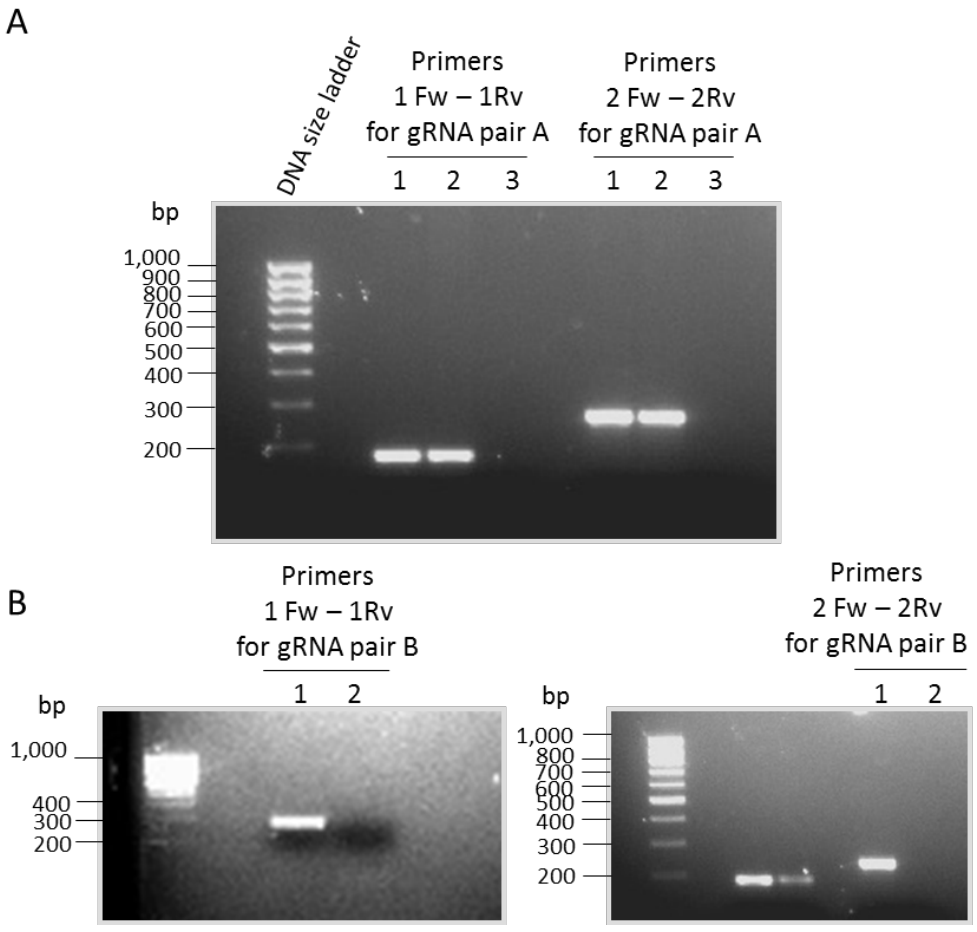
      CCAAGGGACC AGAGCTACAG 3Fw
      GCTGCTGAGC AGACTGGAGG 3RV
```

**Figure 3.15 Design of primers for validation of *TNC* gene edit by PCR amplification.** (A-E) The sequences of two sets of forward and reverse primers for validation of CRISPR-Cas9n-gRNA pair A (A, B) and CRISPR-Cas9n-gRNA pair B (C-E) gene edit are shown in lavender and dark red, respectively, in the context of the *TNC* gene (26881bp – 27901bp and 26641bp – 27721bp). Exon 2 is shown in green, start codon ATG in red, PAM sequences in blue and gRNA pairs are underlined.

### 3.5.2.2 Validation of gene edit screening primers

For each pair of primers a touchdown PCR protocol was designed taking into account the melting temperature of each primer (Table X-X; section 2.7.2.1 of materials and methods). The validation of the primers in the PCR reaction was done using genomic DNA from non-transfected cells or from mock-transfected cells as template DNA or with a non-template negative control.

A clear single band of the expected size was obtained by PCR amplification of HEK 293 genomic DNA with all of the designed primers. No unspecific binding of the primers to the DNA and amplification was detected and the efficiency of amplification was high (Fig. 3.16). These results indicate that these primers and PCR protocols can be used to validate the genome edit of interest.



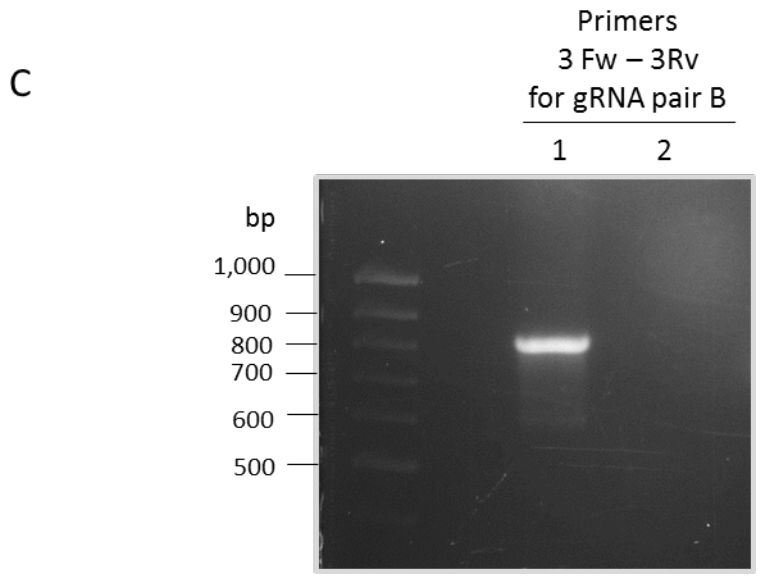



**Figure 3.16 Optimization of PCR protocols using primers designed above for the validation of *TNC* gene edit.** (A-C) PCR products obtained using two (A) or three (B and C) sets of forward and reverse primers for validation of CRISPR-Cas9n-gRNA pair A (A) and CRISPR-Cas9n-gRNA pair B (B and C), run on a 1.5% w/v agarose gel stained with ethidium bromide. (A) The PCR products in lanes 1 and 2 were amplified from genomic DNA (gDNA) from not transfected HEK 293 cells and mock-transfected cells, respectively. In lane 3, the product of a non-template control PCR is shown. (B and C) The PCR products in lane 1 of each gel were amplified from genomic DNA (gDNA) from not transfected HEK 293 cells. In lane 2, the product of a non-template control PCR is shown.

### 3.5.2.3 PCR amplification and sequencing

The primers validated above were used to amplify the region of exon 2 of the *TNC* gene targeted by the CRISPR-Cas9n system in HEK 293 cells. The PCR products were then run on agarose gel, the DNA was extracted from the gel and purified for subsequent sequencing.

Genomic DNA was extracted and purified from HEK 293 cells that were not transfected, mock-transfected or transfected with Cas9n and gRNA pair A, Cas9n and gRNA pair B or Cas9n and a single gRNA (Fig. 3.8). Each genomic DNA sample was amplified by PCR with two pairs of primers for each gRNA pair used to edit the genome. A non-template control, in which water was added to the reaction instead of genomic DNA, was also included. The amplification products run on agarose gel appeared as sharp, single bands of

the expected size (Fig. 3.17A and 3.18A). Compared to non-transfected, mock-transfected and Cas9n/single gRNA transfected samples, the PCR amplification products of Cas9n-gRNA pair A and Cas9n-gRNA pair B transfected samples were visible on the gel as a band with slightly different size and signal intensity (Fig. 3.17A and 3.18A). This could be the result of an insertion in the target DNA sequence.

The fragments of DNA run on the agarose gel were cut, extracted from the gel, purified and sent to sequence with the forward primer used in the PCR (Source Bioscience). Analysis of the resulting electropherograms showed that transfection of cells with Cas9n and the gRNA pair A did not generate any changes to the analysed DNA sequence (Fig. 3.17B). However, transfection of Cas9n and the gRNA pair B into HEK 293 cells edited the target DNA sequence (Fig. 3.18B). The analysed sequence of the mutant contained a premature stop codon (TAA) likely in place of the translation start codon ATG and an area located 40 bp downstream of the ATG that was completed altered (Fig. 3.18B).

Notably, the genomic DNA analysed here was extracted from all of the cells subjected to transfection in each sample, including those not successfully transfected and those successfully transfected. Potentially, the latter include cells mutated by CRISPR-Cas9n as well as non-mutated, wild type cells. Indeed, the electropherogram of the DNA sequence targeted by Cas9n and gRNA pair B clearly showed three signals in mutated regions (Fig. 3.18B). Analysis of these three signals indicated that they correspond to the wild type gene and to two distinct mutants, respectively. The fact that not all of the cells were transfected explains the presence of the wild type signals. The detection of two mutants suggests that Cas9n has edited each of the two copy of the *TNC* gene differently and this is in agreement with the results of the FISH analysis conducted above (Fig. 3.11-3.12).

Transfection of HEK 293 cells with Cas9n and a single gRNA did not cause any change in the analysed DNA sequence (Fig. 3.18C), which corresponded

exactly to that of the wild type gene. This was expected given that Cas9n requires two gRNAs to cause a DSB. Similarly, the analysed DNA sequence of mock-transfected cells was identical to that of the wild type gene.

These results indicate that the CRISPR-Cas9n-gRNA pair B designed here successfully edited our target gene at the genomic level in HEK 293 cells and that this system can be used to edit the *TNC* gene in other mammalian cells, including BJ fibroblasts.

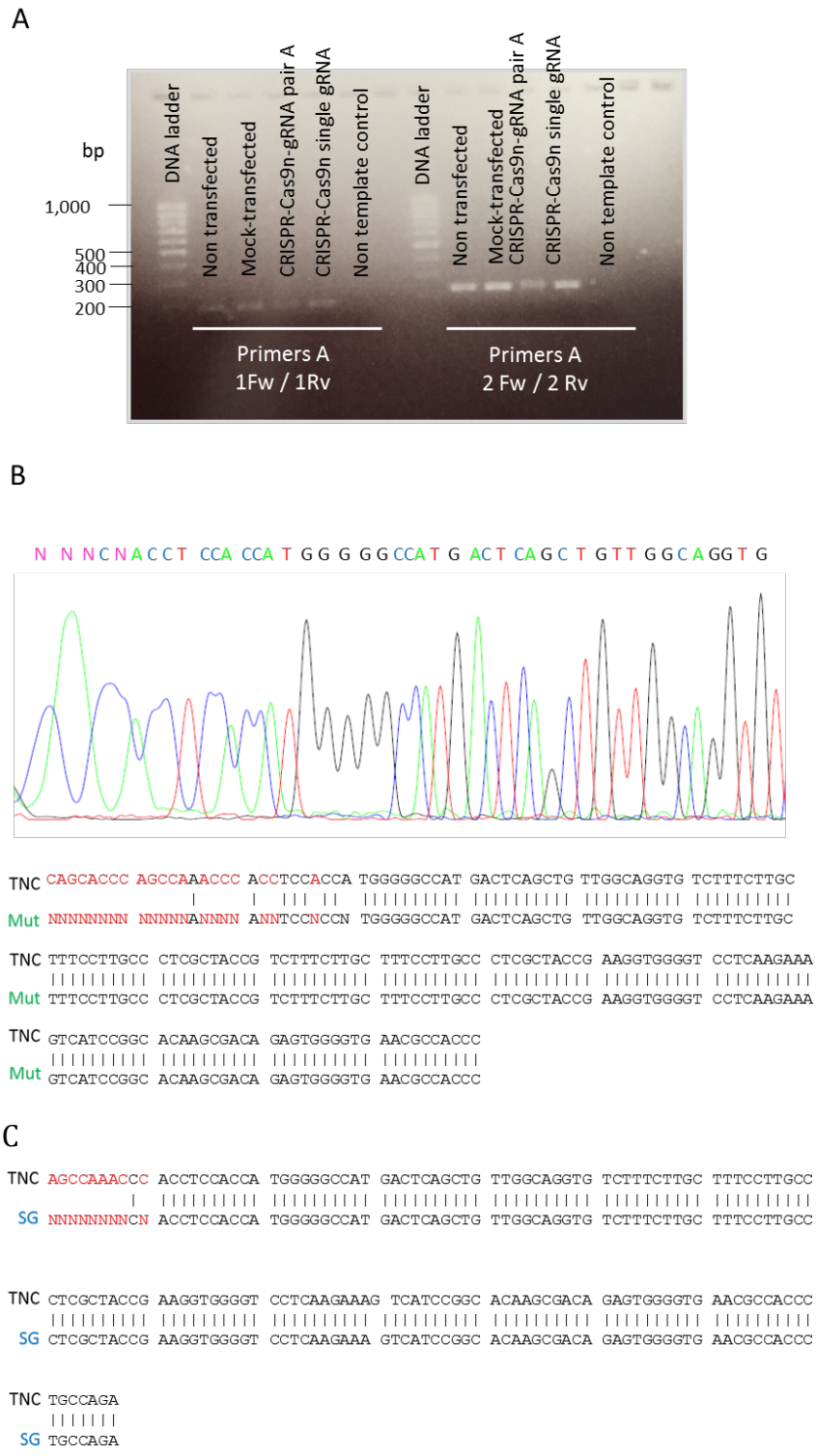



**Figure 3.17 Validation of CRISPR-Cas9n-gRNA pair A-mediated *TNC* gene edit in HEK 293 cells by PCR amplification and sequencing.** (A) PCR amplicons obtained using two sets of forward and reverse primers for validation of CRISPR-Cas9n-gRNA pair A and run on a 1,5% agarose w/v gel stained with ethidium bromide. (B) Sequencing electropherogram of CRISPR-Cas9n-gRNA pair A amplicon showing part of 45-bp segment of the TNC gene. 1Fw was the sequencing primer. Pairwise sequence alignment of PCR amplicon and wild type *TNC* gene is shown below the electropherogram. (C) Pairwise sequence alignment of PCR amplicon of gDNA from HEK-293 transfected with CRISPR-Cas9n single guide control and wild type *TNC*.

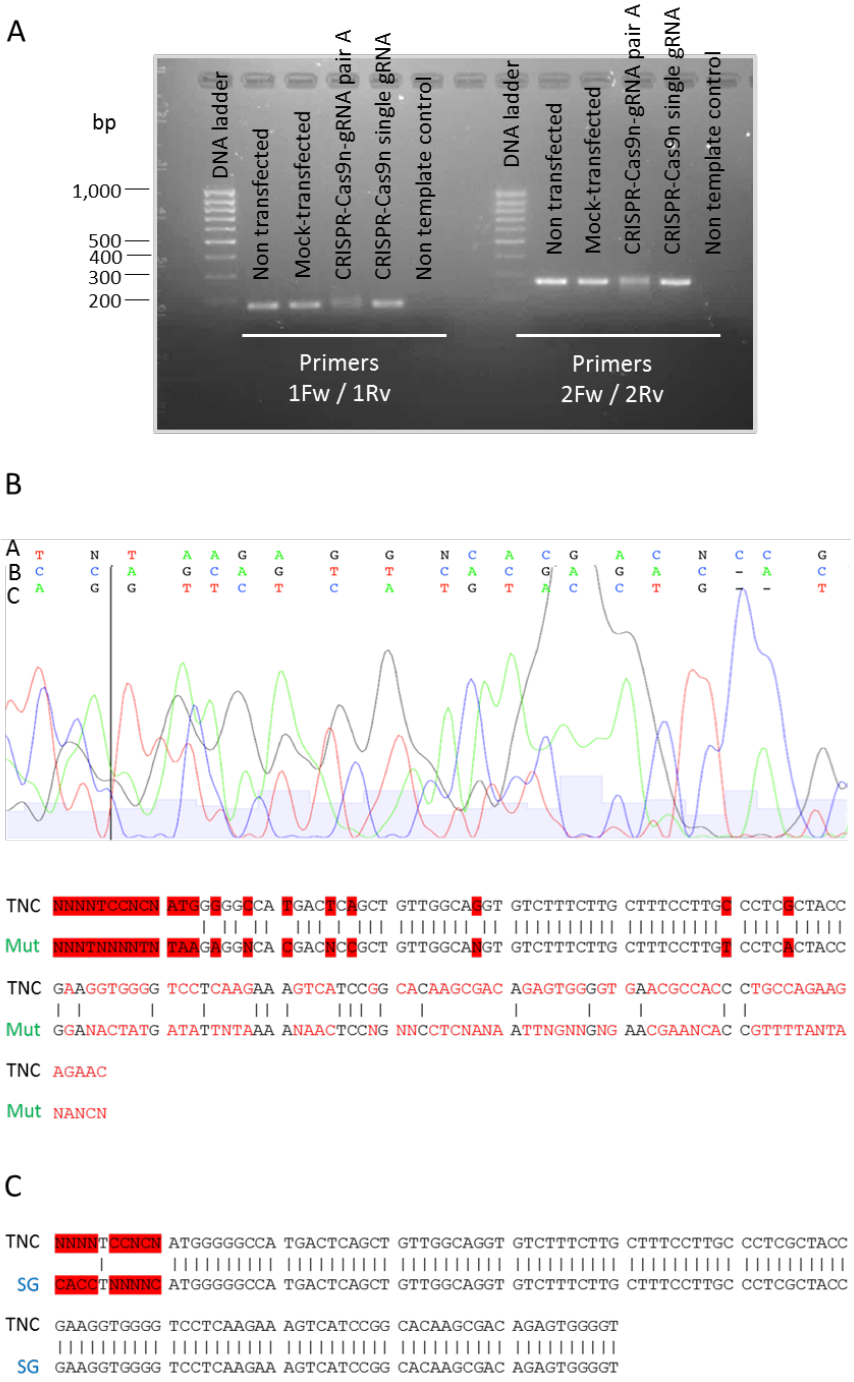



**Figure 3.18 Validation of CRISPR-Cas9n-gRNA pair B-mediated *TNC* gene edit in HEK 293 cells by PCR amplification and sequencing.** (A) PCR amplicons obtained using two sets of forward and reverse primers for validation of CRISPR-Cas9n-gRNA pair B and run on a 1,5% agarose w/v gel stained with ethidium bromide. (B) Sequencing electropherogram of CRISPR-Cas9n-gRNA pair B amplicon showing part of 20 bp segment of the TNC gene. 1Fw was the sequencing primer. Three different sequences are identified in the elecropherogram: the wild type *TNC* gene (B) and two different mutant sequences (A and C), suggesting modification of two alleles. Pairwise sequence alignment of mutant sequence A and wild type *TNC* gene is shown below the electropherogram. (C) Pairwise sequence alignment of PCR amplicon of gDNA from HEK-293 transfected with CRISPR-Cas9n single guide control and wild type *TNC*.

### 3.5.3 Validation of *TNC* gene knock-out in HEK 293 cells by RT-qPCR

To verify whether the CRISPR-Cas9n-mediated genome edit of the *TNC* gene validated above resulted into a knock-out of *TNC* mRNA expression, HEK 293 cells were transfected as described above followed by total RNA extraction and RT-qPCR analysis.

After extraction and purification, the RNA concentration and purity was checked by NanoDrop spectroscopy analysis. All of the samples had A260/A280 and A260/A230 ratios ≥ 2. 500 ng of RNA per sample was subjected to genomic DNA removal followed by reverse transcription with random primers to synthesize the cDNA. cDNA was dilute 1:3 with water and 3 μL of each cDNA was used as template in the RT-qPCR reaction with TaqMan primers and fluorescent probe specific for human *TNC*. The HPRT1 gene was chosen as endogenous control.

Unexpectedly, expression of *TNC* mRNA was found to be induced in HEK 293 cells transfected with Cas9n and gRNA pair A, Cas9n and gRNA pair B and Cas9n and a single gRNA. Transfection of Cas9n and one gRNA induced the expression the most (Fig. 3.19).

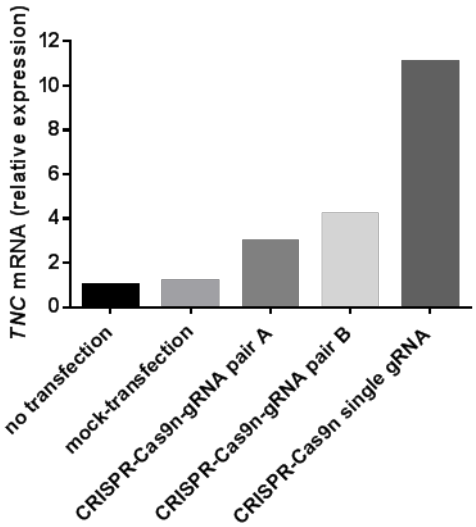



**Figure 3.19 Validation of CRISPR-Cas9n-mediated *TNC* gene edit in HEK 293 cells by RT-qPCR.** Quantitative RT-PCR analysis of TNC mRNA in HEK 293 cells 72 hours after transfection. Results are presented relative to the abundance of TNC mRNA in non transfected cells.

Given that in mock-transfected cells the expression of *TNC* mRNA did not change compared to non-transfected cells, it is possible that bacterial products that may contaminate the plasmid DNA preparations induce the expression of *TNC* mRNA in this cells. For instance, the transcription of the *TNC* gene has been shown to be induced by several types of bacterial products in human myeloid cells through activation of Toll-like receptors (Goh et al., 2010).

This experimental approach did not allow us to validate knock-out of *TNC* mRNA expression in HEK 293 cells.

### 3.5.4 Validation of *TNC* gene knock-out in HEK 293 cells by immunoblot analysis

The effective knock-out of the *TNC* gene in HEK 293 was analysed at the protein level by immunoblot analysis. As TN-C is a secreted protein, the cell conditioned media was analysed. The major isoform of the TN-C protein that is expressed in mammalian cells is expected to run in acrylamide gel under reducing conditions as a band at ~320 kDa.

A protein band of ~320 kDa like in the positive control was detected in every protein sample tested. The relative amount of the protein band was the same in each sample as confirmed by the Ponceau S staining of the membrane which showed equal loading of the protein samples (Fig. 3.20A).

Side-by-side immunoblot analysis of TN-C in HEK 293 conditioned media and fresh, complete media supplemented with 10 % FBS showed that the anti-TN-C antibody used cross-reacted with bovine TN-C found in serum (Fig. 3.20B). Thus, the band detected in the conditioned media most likely corresponds to bovine TN-C rather than human TN-C secreted by HEK 293 cells. This result is not entirely unexpected and is in line with the very low expression of *TNC* in HEK 293 cells at the transcriptional level (Fig. 3.2).

These data indicate that either HEK 293 cells do not secrete TN-C protein or that the immunoblot analysis conducted here is not sensitive enough to detect the low expression levels of TN-C in these cells.

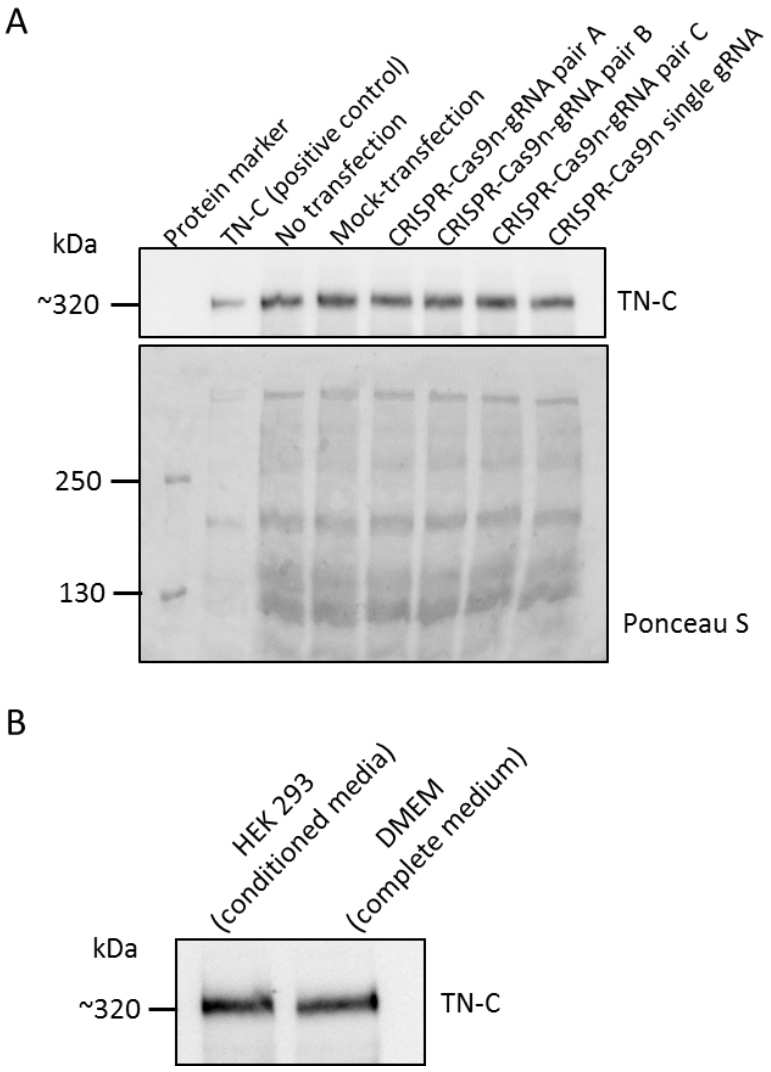
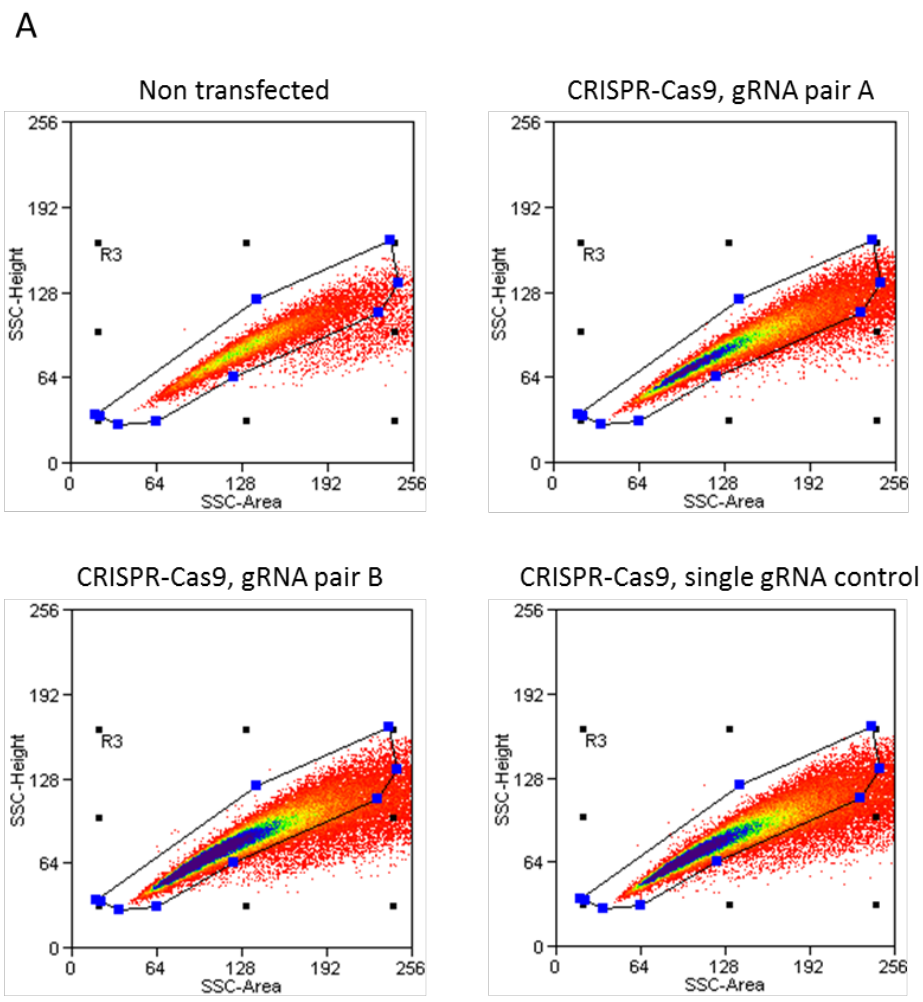



**Figure 3.20 Validation of CRISPR-Cas9n-mediated *TNC* gene edit in HEK 293 cells by Western blotting.** (A) Top: Western blotting analysis of TN-C in 40μl of HEK 293 conditioned media 72 hours after transfection and 15μl of BJ fibroblast conditioned media (positive control); bottom: Ponceau S staining of nitrocellulose membrane showing equal loading. (B) Western blotting analysis of TN-C in 40μl of HEK 293 conditioned media and equal volume of DMEM medium supplemented with 10% FBS, L-Glutamine and antibiotics.

### 3.6 Generation of single-BJ cell clones with *TNC* knock-out expression

#### 3.6.1 Selection of transfected cells by Fluorescence Activated Cell Sorting (FACS)

BJ fibroblasts were transfected with Cas9n and gRNA pair A, Cas9n and gRNA pair B and Cas9n and a single gRNA by electroporation using the 4D-Nucleofector™ system and seeded in 6-well plates as described above. 48 hours after transfection, cells were trypsinized and subjected to single-cell sorting (Flow Cytometry Facility, University of Nottingham) based on their GFP expression (Fig. 3.21B). 42 to 48.9% of the transfected cells were $GFP^+$ (Fig. 3.21B).



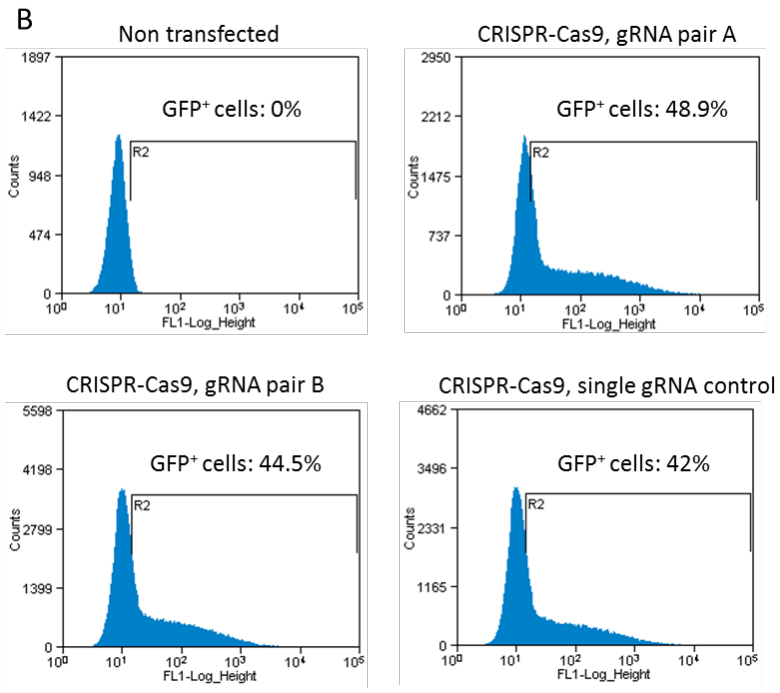



**Figure 3.21 Fluorescence activated cell sorting (FACS) of BJ fibroblasts 48 hours after transfection.** (A): FCS-SSC dot plots of BJ fibroblasts that were not transfected or transfected with pSpCas9n(BB)-2A-GFP plasmids (500 ng each) containing the indicated gRNAs. (B) Histograms show GFP$^+$ cells.

Single GFP+ cells were seeded each in a single well of 96-well plates. Six 96-well plates in total were prepared, two for each transfection with different gRNAs (Fig. 3.22).

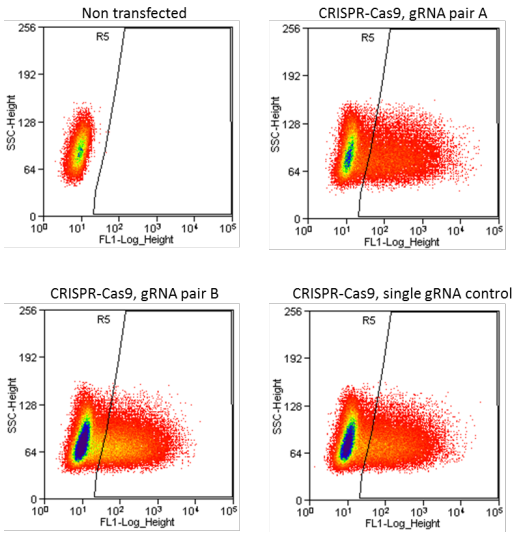



**Figure 3.22 Single cell sorting of GFP$^+$ BJ fibroblasts 48 hours after transfection.** Dot plots of BJ fibroblasts that were not transfected or transfected with pSpCas9n(BB)-2A-GFP plasmids (500 ng each) containing the indicated gRNAs. Gates include single GFP$^+$ cells sorted and collected in 96-well plates (one cell/well).

In addition, ~5 x $10^3$ $GFP^+$ cells were sorted in a FACS tube for subsequent expansion of $GFP^+$ cells and preliminary genome edit validation as performed in the case of HEK 293 cells.

### 3.6.3 Validation of *TNC* gene edit in a bulk population of $GFP^+$ BJ fibroblasts

The $GFP^+$ cells that were sorted in a FACS tube were expanded. A validation of the genome edit was done as for HEK 293 cells at the genomic level. Notably, these cells were all successfully transfected, but included a mixture of wild type cells that have not been edited by Cas9n and cells that have potentially been mutated.

The primers validated above were used to amplify the region of exon 2 of the *TNC* gene targeted by the CRISPR-Cas9n system in $GFP^+$ BJ cells. The PCR products were then run on agarose gel, the DNA was extracted from the gel and purified for subsequent sequencing.

Genomic DNA was extracted and purified from $GFP^+$ BJ cells that were transfected with Cas9n and gRNA pair A, Cas9n and gRNA pair B and Cas9n and a single gRNA. Each genomic DNA sample was amplified by PCR with two pairs of primers for each gRNA pair used to edit the genome. A non-template control, in which water was added to the reaction instead of genomic DNA, was also included. The amplification products run on agarose gel appeared as sharp, single bands of the expected size and no apparent differences in size or signal intensity of the bands could be seen between the different samples (Fig. 3.23A and 3.24A).

The fragments of DNA run on the agarose gel were cut, extracted from the gel, purified and sent to sequence with the forward primer used in the PCR (Source Bioscience). Analysis of the resulting electropherograms showed no changes to the analysed DNA sequence in the genome of $GFP^+$ BJ cells that were transfected with Cas9n and the gRNA pair A or Cas9n and the gRNA pair B (Fig. 3.23B and 3.24B).

Transfection of $GFP^+$ BJ cells with Cas9n and a single gRNA did not cause any change in the analysed DNA sequence (Fig. 3.24C), which corresponded exactly to that of the wild type gene. This was expected given that Cas9n requires two gRNAs to cause a DSB.

These results suggest that either the designed CRISPR-Cas9n system did not edit the genome of BJ cells at the target site or only a low percentage of mutants was present in the bulk of cells that was below the detection threshold. Another possibility is that DSB were generated and repaired efficiently and with high-fidelity in BJ fibroblasts. Validation of *TNC* gene edit in single-BJ cell clones is therefore imperative to conclusively determine whether the designed CRISPR-Cas9n system mutated the *TNC* gene in BJ cells as it did in HEK 293 cells.

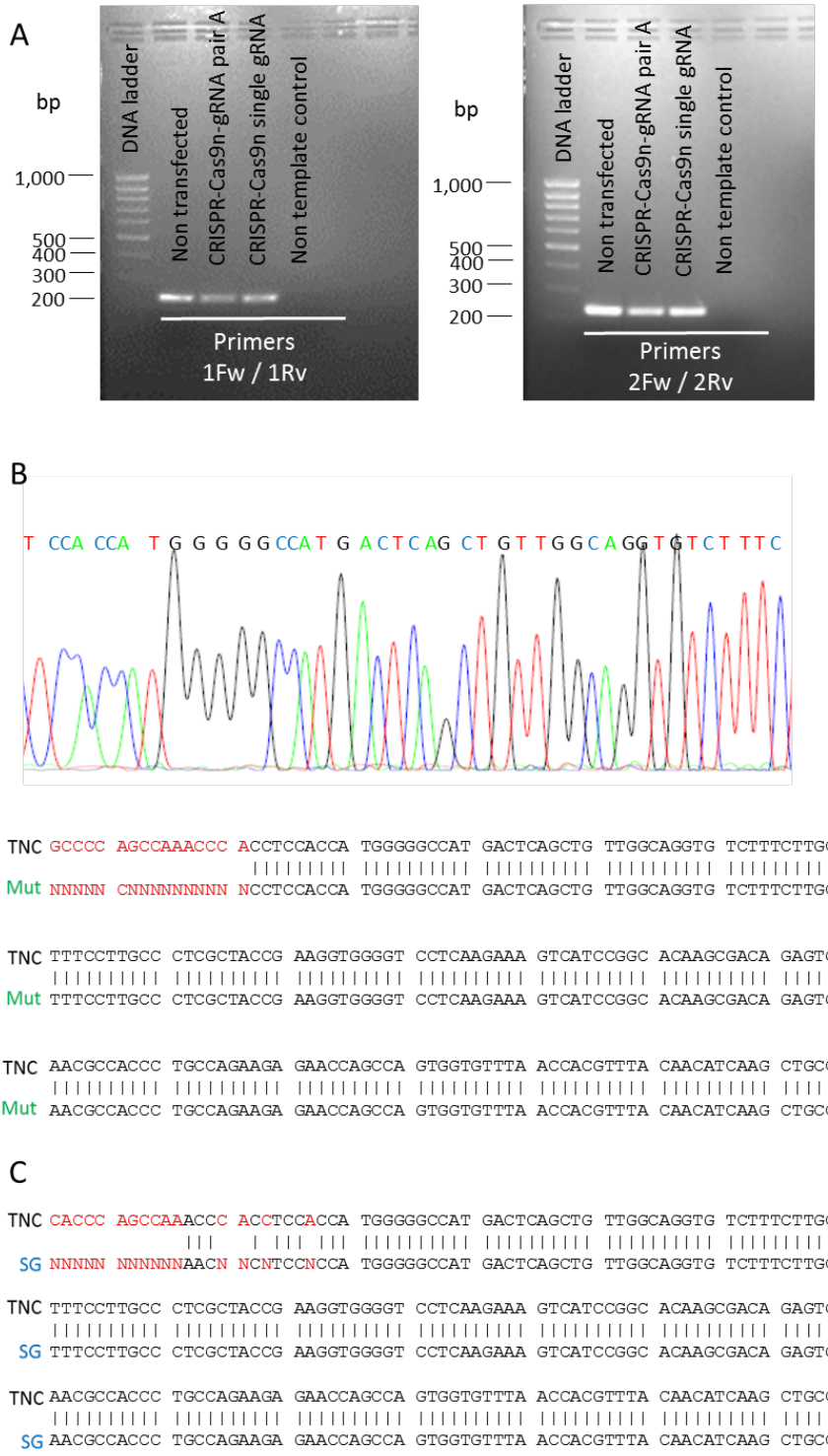



**Figure 3.23 Validation of CRISPR-Cas9n-gRNA pair A-mediated *TNC* gene edit in BJ fibroblasts by PCR amplification and sequencing.** (A) PCR amplicons obtained using two sets of forward and reverse primers for validation of CRISPR-Cas9n-gRNA pair A and run on a 1,5% agarose w/v gel stained with ethidium bromide. (B) Sequencing electropherogram of CRISPR-Cas9n-gRNA pair A amplicon showing part of 40 bp segment of the TNC gene. 1Fw was the sequencing primer. Pairwise sequence alignment of PCR amplicon and wild type *TNC* gene is shown below the electropherogram.

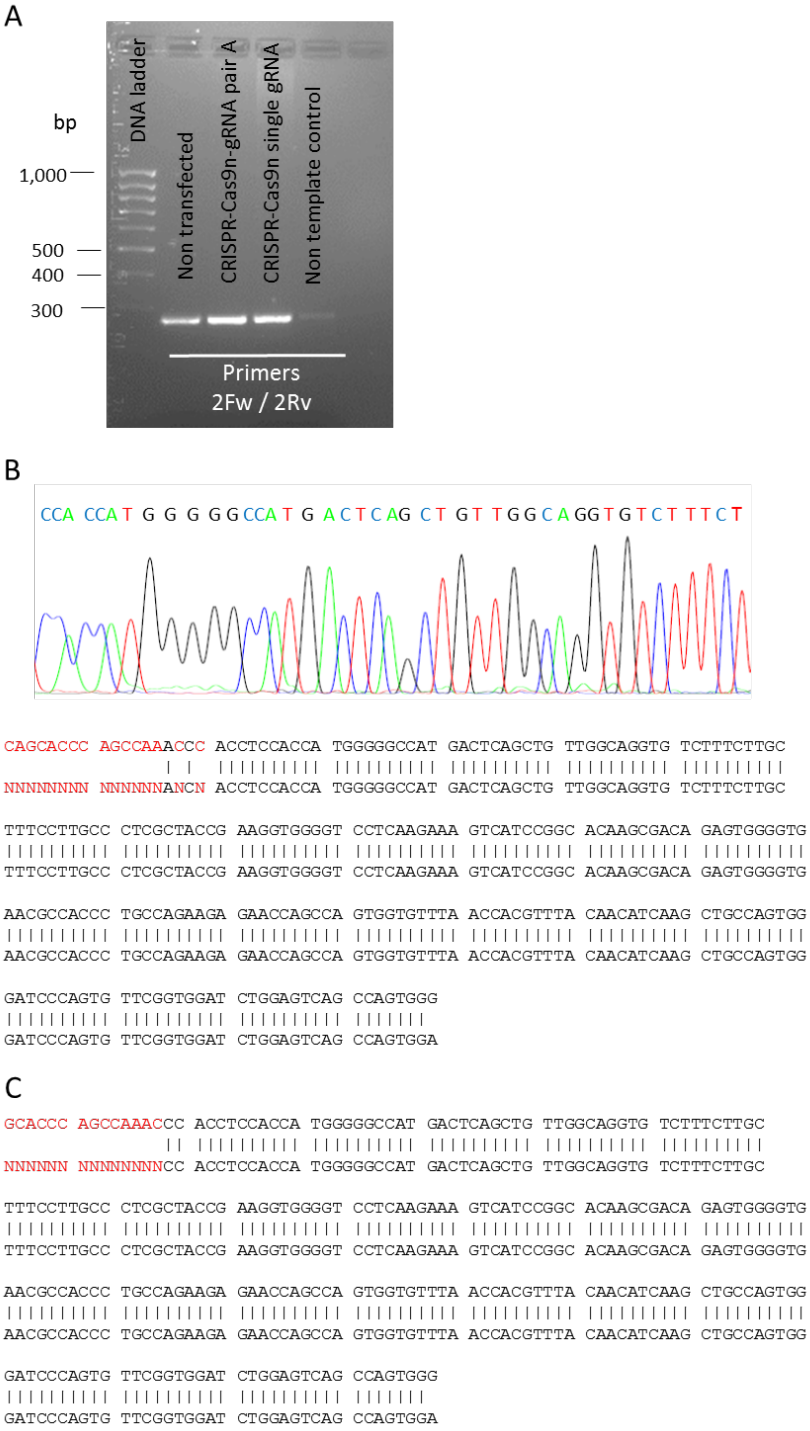



**Figure 3.24 Validation of CRISPR-Cas9n-gRNA pair B-mediated *TNC* gene edit in GFP$^+$ BJ fibroblasts by PCR amplification and sequencing.** (A) PCR amplicons obtained using one set of forward and reverse primers for validation of CRISPR-Cas9n-gRNA pair B and run on a 1,5% agarose w/v gel stained with ethidium bromide. (B) Sequencing electropherogram of CRISPR-Cas9n-gRNA pair B amplicon showing part of XXX-bp segment of the TNC gene. 1Fw was the sequencing primer. Pairwise sequence alignment of PCR amplicon and wild type *TNC* gene is shown below the electropherogram. (C) Pairwise sequence alignment of PCR amplicon of gDNA from BJ fibroblasts transfected with CRISPR-Cas9n single guide control and wild type *TNC*.

### 3.6.4 Expansion of single-BJ cell clones

To obtain single-cell clones, each GFP$^+$ BJ fibroblast was sorted into a single well of 96-well plate. Normal stromal cells, including fibroblasts, are difficult to grow as single cells in the absence of cell-cell contacts and soluble factors secreted by the cells. Therefore, single cells were grown in 200 µL of complete growth medium supplemented with BJ-conditioned medium in a ratio of 1:1.

The majority of single cells died, independently of the plasmid DNA transfected. The cells that survived showed different growth speed and reached ~80% confluency in the 96-well plate after 2 months in culture (Fig. 3.25).

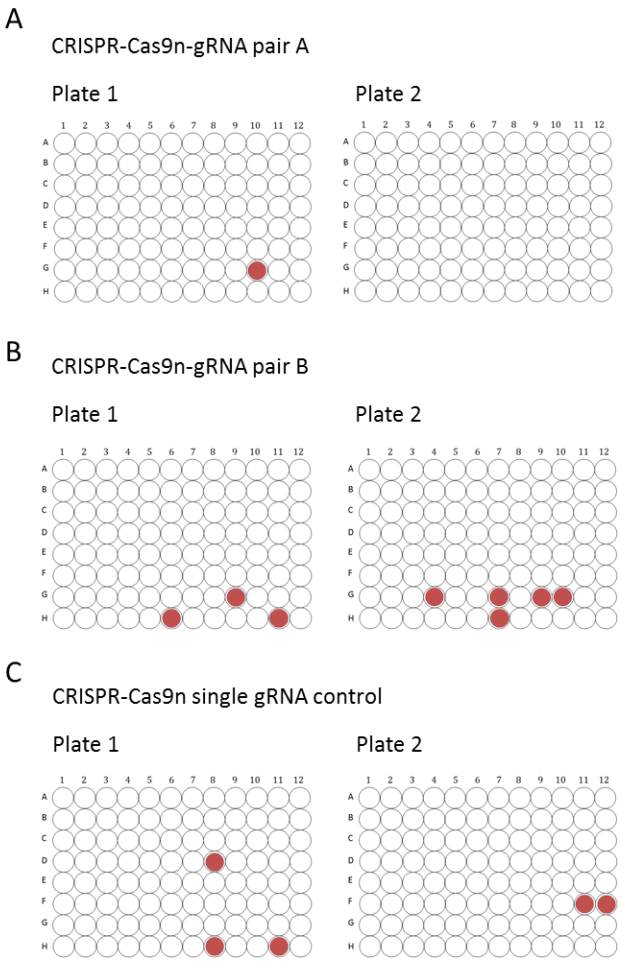



**Figure 3.25 Establishment of clonal cell lines.** (A-C) Schematic representation of single GFP$^+$ BJ fibroblasts seeded in 96-well plates. Cells transfected with CRISPR-Cas9n-gRNA pair A (A), CRISPR-Cas9n-gRNA pair B (B) and CRISPR-Cas9n single gRNA control (C) are shown. Dark red wells represent single cells that are currently expanding.

At this point, the single-BJ cell clones were trypsinised and seeded in 48- or 24-well plates, depending on the number of cells, and cultured in complete growth medium without BJ-conditioned media. These clones are now being cultured in 6-well plates and, when they will reach ~80% confluency, they will be passaged and cultured in 10 cm dishes to generate enough cells for the propagation and validation of single-BJ cell clones.

# 4 Discussion

## 4.1 General discussion

TN-C is an ECM protein whose expression is rapidly induced during the inflammatory response to infection, where it influences the synthesis of a subset of LPS-induced pro-inflammatory genes by regulating the expression of the microRNA miR-155 in primary murine macrophages (Piccinini and Midwood, 2012, Goh et al., 2010). However, it is not known whether this occurs in human cells and how. Here, I have designed, delivered and validated a CRISPR-Cas9n system that specifically targets the *TNC* gene in HEK 293 cells. I have applied this system to human fibroblasts to generate single-cell clones with knocked-out *TNC* expression. Following validation, positive clones will be used to generate cell-derived matrices, which lack TN-C, as a substrate for the culture of primary human macrophages. This system will be applied to investigate whether human TNC modulates macrophage gene expression by regulating microRNA levels upon infection.

Since the discovery that exogenous DNA can be taken up by cells and integrated into the genome, an increasing number of genome engineering approaches has been established. Some of these, RNA interference (RNAi) in particular, have become a routine method in many laboratories. The popularity of RNAi (Fire et al., 1998) is now being echoed by CRISPR-Cas9, a new genome engineering technology that was chosen in this project to knock-out *TNC* expression in human fibroblasts. This choice was made as: 1) CRISPR-Cas9 allows permanent gene knock-out in contrast to RNAi, which provides only transient gene knock-down (Alic et al., 2012); 2) CRISPR-Cas9 off-target effects can be predicted and significantly reduced by using mutant Cas9 proteins in contrast to the unpredictable RNAi off-target effects; 3) compared to other recent technologies such as zinc-finger nucleases (ZFNs) (Urnov et al., 2005) and transcription-activator like effector nucleases (TALENs) (Mussolino et al., 2011), CRISPR-Cas9 is considerably less expensive and time-consuming.

The properties of the gRNAs that can increase the specificity and efficiency of the CRISPR-Cas9n system are unclear. Therefore, here we designed three different pairs of gRNAs to maximise our chances of targeting the *TNC* gene. The gRNA pairs differed in the gRNA offset, their position in the *TNC* locus relative to the translation start codon ATG and their aggregate score. The gRNA pair that could edit the target gene in HEK 293 cells possesses an intermediate offset of 11 bp, is located downstream of the ATG at intermediate distance compared to the other pairs and has the highest agglomerate score, which suggests high specificity and no off-target effects.

### 4.2 Limitations and future work

Limitations for the experimental approaches in this study have been discussed here based on the data analysis and discussion described in the previous sections.

HEK 293 cells were used in the first part of this project to design and optimise CRISPR-Cas9n-mediated edit of the *TNC* gene as matrix producing cells, including fibroblasts, grow slowly and are hard to transfect. HEK 293 cells grow easily and are amenable to transfection, thus have been very useful to validate the genome editing ability of the designed CRISPR-Cas9n systems at the genomic level. However, as HEK 293 cells express very low levels of *TNC* mRNA and do not secrete TN-C protein, they have not allowed us to confirm genome editing at the protein level.

This work also showed that validation of genome edit at the protein level by Western blot analysis can be performed only in matrix extracts or total cells extracts (including the ECM), but not in conditioned medium. This is most likely due to anti-TN-C antibody cross-reactivity with a bovine protein of the same size as TN-C that is found in serum. An alternative approach that could be used in future is proteomics analysis.

To validate CRISPR-Cas9n-mediated *TNC* gene edit in HEK 293 cells and fibroblasts, fluorescence in situ hybridization (FISH) was used in combination

with PCR amplification and sequencing. This technique estimates the target gene copy number in cells of interest by determining the target chromosome copy number. However, one caveat for this application of FISH is that the resulting gene copy number may be underestimated as events such as gene duplication or translocation are not taken into account. Nevertheless, this tool has been useful as: 1) it confirmed that the population of BJ fibroblasts used in the lab is diploid, as stated by ATCC; and 2) it facilitated the validation of the genome edit in HEK 293 cells. More than 97% of HEK 293 cells were estimated to carry 2 copies of the gene and this was in agreement with the PCR amplification and sequencing data.

Finally, validation of genome edit in single-fibroblasts clones has not been completed within the duration of this project due to the low growth rate of human BJ fibroblasts, which were chosen for their ability to secrete and deposit high amounts of ECM, including TN-C. Other fibroblast cell line could be tested, however slow growth is a characteristic of the majority of human stromal cells that synthesize and secrete a matrix.

Future work will include:

1. Use of adeno-associated virus (AVV) or lentiviral transduction, which are useful for efficient delivery of Cas9 and gRNAs into difficult-to-transfect cells like fibroblasts;
2. PCR amplify and sequence wider regions (up to ~500/700 bp) of the *TNC* gene around the target site for genome edit validation. This is to sequence areas of the targeted gene that have not been mutated, both upstream and downstream of the edited region, in addition to the mutated area.
3. Include transfection of cells with the empty backbone plasmid DNA as an additional control. This is to confirm whether increased *TNC* mRNA expression, following transfection with Cas9 and gRNAs, is due to unspecific and unpredicted activity of Cas9 or to the induction of *TNC* transcription by exogenous DNA or contaminants of bacterial origin

through activation of NF-kB, which is known to promote *TNC* synthesis.

4. Validation of single-fibroblast cell clones that are currently growing at the gene, mRNA and protein level. This will be done using PCR amplification and sequencing, RT-qPCR or Northern blot analysis and Western blot analysis of fibroblast-derived matrices, respectively.
5. Production and application of cell-free fibroblast-derived matrices using wild type and *TNC* knock-out fibroblasts as a substrate for the culture and screening of primary human macrophages.

## References


AL-ATTAR, S., WESTRA, E. R., VAN DER OOST, J. & BROUNS, S. J. 2011. Clustered regularly interspaced short palindromic repeats (CRISPRs): the hallmark of an ingenious antiviral defense mechanism in prokaryotes. *Biol Chem,* 392**,** 277-89.

ALIC, N., HODDINOTT, M. P., FOLEY, A., SLACK, C., PIPER, M. D. & PARTRIDGE, L. 2012. Detrimental effects of RNAi: a cautionary note on its use in Drosophila ageing studies. *PLoS One,* 7**,** e45367.

AMBROS, V. 2004. The functions of animal microRNAs. *Nature,* 431**,** 350-5.

AVCI-ADALI, M., BEHRING, A., KELLER, T., KRAJEWSKI, S., SCHLENSAK, C. & WENDEL, H. P. 2014. Optimized conditions for successful transfection of human endothelial cells with in vitro synthesized and modified mRNA for induction of protein expression. *J Biol Eng,* 8**,** 8.

BARRANGOU, R., FREMAUX, C., DEVEAU, H., RICHARDS, M., BOYAVAL, P., MOINEAU, S., ROMERO, D. A. & HORVATH, P. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science,* 315**,** 1709-12.

BEUTLER, B. 2007. Neo-ligands for innate immune receptors and the etiology of sterile inflammatory disease. *Immunol Rev,* 220**,** 113-28.

BOLOTIN, A., QUINQUIS, B., SOROKIN, A. & EHRLICH, S. D. 2005. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology,* 151**,** 2551-61.

BORNSTEIN, P. & SAGE, E. H. 2002. Matricellular proteins: extracellular modulators of cell function. *Curr Opin Cell Biol,* 14**,** 608-16.

BRELLIER, F., TUCKER, R. P. & CHIQUET-EHRISMANN, R. 2009. Tenascins and their implications in diseases and tissue mechanics. *Scand J Med Sci Sports,* 19**,** 511-9.

CARTHEW, R. W. & SONTHEIMER, E. J. 2009. Origins and Mechanisms of miRNAs and siRNAs. *Cell,* 136**,** 642-55.

CATALAN, V., GOMEZ-AMBROSI, J., RODRIGUEZ, A., RAMIREZ, B., ROTELLAR, F., VALENTI, V., SILVA, C., GIL, M. J., SALVADOR, J. & FRUHBECK, G. 2012. Increased tenascin C and Toll-like receptor 4 levels in visceral adipose tissue as a link between inflammation and extracellular matrix remodeling in obesity. *J Clin Endocrinol Metab,* 97**,** E1880-9.

CHIOVARO, F., CHIQUET-EHRISMANN, R. & CHIQUET, M. 2015. Transcriptional regulation of tenascin genes. *Cell Adh Migr,* 9**,** 34-47.

CHIQUET-EHRISMANN, R. 1991. Anti-adhesive molecules of the extracellular matrix. *Curr Opin Cell Biol,* 3**,** 800-4.

CHIQUET-EHRISMANN, R., MACKIE, E. J., PEARSON, C. A. & SAKAKURA, T. 1986. Tenascin: an extracellular matrix protein involved in tissue interactions during fetal development and oncogenesis. *Cell,* 47**,** 131-9.

CHIQUET-EHRISMANN, R., OREND, G., CHIQUET, M., TUCKER, R. P. & MIDWOOD, K. S. 2014. Tenascins in stem cell niches. *Matrix Biol,* 37**,** 112-23.

CHIQUET-EHRISMANN, R., TANNHEIMER, M., KOCH, M., BRUNNER, A., SPRING, J., MARTIN, D., BAUMGARTNER, S. & CHIQUET, M. 1994. Tenascin-C expression by fibroblasts is elevated in stressed collagen gels. *J Cell Biol,* 127**,** 2093-101.

DAY, J. M., OLIN, A. I., MURDOCH, A. D., CANFIELD, A., SASAKI, T., TIMPL, R., HARDINGHAM, T. E. & ASPBERG, A. 2004. Alternative splicing in the aggrecan G3 domain influences binding interactions with tenascin-C and other extracellular matrix proteins. *J Biol Chem,* 279**,** 12511-8.

DEVEAU, H., GARNEAU, J. E. & MOINEAU, S. 2010. CRISPR/Cas system and its role in phage-bacteria interactions. *Annu Rev Microbiol,* 64**,** 475-93.

DOYLE, J. J., GERBER, E. E. & DIETZ, H. C. 2012. Matrix-dependent perturbation of TGFbeta signaling and disease. *FEBS Lett,* 586**,** 2003-15.

FIRE, A., XU, S., MONTGOMERY, M. K., KOSTAS, S. A., DRIVER, S. E. & MELLO, C. C. 1998. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature,* 391**,** 806-11.

FITZPATRICK, L. E. & MCDEVITT, T. C. 2015. Cell-derived matrices for tissue engineering and regenerative medicine applications. *Biomater Sci,* 3**,** 12-24.

FRANTZ, C., STEWART, K. M. & WEAVER, V. M. 2010. The extracellular matrix at a glance. *J Cell Sci,* 123**,** 4195-200.

FUJIMOTO, N., ONISHI, K., SATO, A., TERASAKI, F., TSUKADA, B., NOZATO, T., YAMADA, T., IMANAKA-YOSHIDA, K., YOSHIDA, T., ITO, M. & HIROE, M. 2009. Incremental prognostic values of serum tenascin-C levels with blood B-type natriuretic peptide testing at discharge in patients with dilated cardiomyopathy and decompensated heart failure. *J Card Fail,* 15**,** 898-905.

GARRETT, R. A., SHAH, S. A., VESTERGAARD, G., DENG, L., GUDBERGSDOTTIR, S., KENCHAPPA, C. S., ERDMANN, S. & SHE, Q. 2011. CRISPR-based immune systems of the Sulfolobales: complexity and diversity. *Biochem Soc Trans,* 39**,** 51-7.

GHERZI, R., CARNEMOLLA, B., SIRI, A., PONASSI, M., BALZA, E. & ZARDI, L. 1995. Human tenascin gene. Structure of the 5'-region, identification, and characterization of the transcription regulatory sequences. *J Biol Chem,* 270**,** 3429-34.

GIBLIN, S. P. & MIDWOOD, K. S. 2015. Tenascin-C: Form versus function. *Cell Adh Migr,* 9**,** 48-82.

GOH, F. G., PICCININI, A. M., KRAUSGRUBER, T., UDALOVA, I. A. & MIDWOOD, K. S. 2010. Transcriptional regulation of the endogenous danger signal tenascin-C: a novel autocrine loop in inflammation. *J Immunol,* 184**,** 2655-62.

GORDON, S. 2002. Pattern recognition receptors: doubling up for the innate immune response. *Cell,* 111**,** 927-30.

GULCHER, J. R., NIES, D. E., MARTON, L. S. & STEFANSSON, K. 1989. An alternatively spliced region of the human hexabrachion contains a repeat of potential N-glycosylation sites. *Proc Natl Acad Sci U S A,* 86**,** 1588-92.
GUTOWSKI, N. J., NEWCOMBE, J. & CUZNER, M. L. 1999. Tenascin-R and C in multiple sclerosis lesions: relevance to extracellular matrix remodelling. *Neuropathol Appl Neurobiol,* 25**,** 207-14.
HAFT, D. H., SELENGUT, J., MONGODIN, E. F. & NELSON, K. E. 2005. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput Biol,* 1**,** e60.
HORVATH, P. & BARRANGOU, R. 2010. CRISPR/Cas, the immune system of bacteria and archaea. *Science,* 327**,** 167-70.
HUBMACHER, D. & APTE, S. S. 2013. The biology of the extracellular matrix: novel insights. *Curr Opin Rheumatol,* 25**,** 65-70.
HYNES, R. O. 2009. The extracellular matrix: not just pretty fibrils. *Science,* 326**,** 1216-9.
JANSEN, R., EMBDEN, J. D., GAASTRA, W. & SCHOULS, L. M. 2002. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol,* 43**,** 1565-75.
KAARTEENAHO-WIIK, R., MERTANIEMI, P., SAJANTI, E., SOINI, Y. & PAAKKO, P. 1998. Tenascin is increased in epithelial lining fluid in fibrotic lung disorders. *Lung,* 176**,** 371-80.
KARGINOV, F. V. & HANNON, G. J. 2010. The CRISPR system: small RNA-guided defense in bacteria and archaea. *Mol Cell,* 37**,** 7-19.
KARJALAINEN, E. M., LINDQVIST, A., LAITINEN, L. A., KAVA, T., ALTRAJA, A., HALME, M. & LAITINEN, A. 2003. Airway inflammation and basement membrane tenascin in newly diagnosed atopic and nonatopic asthma. *Respir Med,* 97**,** 1045-51.
KIM, S. H., TURNBULL, J. & GUIMOND, S. 2011. Extracellular matrix and cell signalling: the dynamic cooperation of integrin, proteoglycan and growth factor receptor. *J Endocrinol,* 209**,** 139-51.
KOONIN, E. V. & MAKAROVA, K. S. 2009. CRISPR-Cas: an adaptive immunity system in prokaryotes. *F1000 Biol Rep,* 1**,** 95.
KYRIAKIDES, T. R., WULSIN, D., SKOKOS, E. A., FLECKMAN, P., PIRRONE, A., SHIPLEY, J. M., SENIOR, R. M. & BORNSTEIN, P. 2009. Mice that lack matrix metalloproteinase-9 display delayed wound healing associated with delayed reepithelization and disordered collagen fibrillogenesis. *Matrix Biol,* 28**,** 65-73.
LATIJNHOUWERS, M., BERGERS, M., PONEC, M., DIJKMAN, H., ANDRIESSEN, M. & SCHALKWIJK, J. 1997. Human epidermal keratinocytes are a source of tenascin-C during wound healing. *J Invest Dermatol,* 108**,** 776-83.
LIABEUF, S., BARRETO, D. V., KRETSCHMER, A., BARRETO, F. C., RENARD, C., ANDREJAK, M., CHOUKROUN, G. & MASSY, Z. 2011. High circulating levels of large splice variants of tenascin-C is associated with mortality and cardiovascular disease in chronic kidney disease patients. *Atherosclerosis,* 215**,** 116-24.

LOTZE, M. T. & TRACEY, K. J. 2005. High-mobility group box 1 protein (HMGB1): nuclear weapon in the immune arsenal. *Nat Rev Immunol,* 5**,** 331-42.
LU, P., WEAVER, V. M. & WERB, Z. 2012. The extracellular matrix: a dynamic niche in cancer progression. *J Cell Biol,* 196**,** 395-406.
LUNDELL, A., OLIN, A. I., MORGELIN, M., AL-KARADAGHI, S., ASPBERG, A. & LOGAN, D. T. 2004. Structural basis for interactions between tenascins and lectican C-type lectin domains: evidence for a crosslinking role for tenascins. *Structure,* 12**,** 1495-506.
MAKAROVA, K. S., ARAVIND, L., GRISHIN, N. V., ROGOZIN, I. B. & KOONIN, E. V. 2002. A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis. *Nucleic Acids Res,* 30**,** 482-96.
MAKAROVA, K. S., GRISHIN, N. V., SHABALINA, S. A., WOLF, Y. I. & KOONIN, E. V. 2006. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol Direct,* 1**,** 7.
MATZINGER, P. 1994. Tolerance, danger, and the extended family. *Annu Rev Immunol,* 12**,** 991-1045.
MCCACHREN, S. S. & LIGHTNER, V. A. 1992. Expression of human tenascin in synovitis and its regulation by interleukin-1. *Arthritis Rheum,* 35**,** 1185-96.
MEDZHITOV, R. & JANEWAY, C. A., JR. 2002. Decoding the patterns of self and nonself by the innate immune system. *Science,* 296**,** 298-300.
MERLINE, R., MORETH, K., BECKMANN, J., NASTASE, M. V., ZENG-BROUWERS, J., TRALHAO, J. G., LEMARCHAND, P., PFEILSCHIFTER, J., SCHAEFER, R. M., IOZZO, R. V. & SCHAEFER, L. 2011. Signaling by the matrix proteoglycan decorin controls inflammation and cancer through PDCD4 and MicroRNA-21. *Sci Signal,* 4**,** ra75.
MIDWOOD, K., SACRE, S., PICCININI, A. M., INGLIS, J., TREBAUL, A., CHAN, E., DREXLER, S., SOFAT, N., KASHIWAGI, M., OREND, G., BRENNAN, F. & FOXWELL, B. 2009. Tenascin-C is an endogenous activator of Toll-like receptor 4 that is essential for maintaining inflammation in arthritic joint disease. *Nat Med,* 15**,** 774-80.
MIDWOOD, K. S. & OREND, G. 2009. The role of tenascin-C in tissue injury and tumorigenesis. *J Cell Commun Signal,* 3**,** 287-310.
MOHAMMADI, M., OLSEN, S. K. & GOETZ, R. 2005. A protein canyon in the FGF-FGF receptor dimer selects from an a la carte menu of heparan sulfate motifs. *Curr Opin Struct Biol,* 15**,** 506-16.
MOJICA, F. J., DIEZ-VILLASENOR, C., GARCIA-MARTINEZ, J. & SORIA, E. 2005. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol,* 60**,** 174-82.
MOUW, J. K., OU, G. & WEAVER, V. M. 2014. Extracellular matrix assembly: a multiscale deconstruction. *Nat Rev Mol Cell Biol,* 15**,** 771-85.
MUSSOLINO, C., MORBITZER, R., LUTGE, F., DANNEMANN, N., LAHAYE, T. & CATHOMEN, T. 2011. A novel TALE nuclease scaffold enables

high genome editing activity in combination with low toxicity. *Nucleic Acids Res,* 39**,** 9283-93.
OREND, G. & CHIQUET-EHRISMANN, R. 2006. Tenascin-C induced signaling in cancer. *Cancer Lett,* 244**,** 143-63.
OSKARSSON, T., ACHARYYA, S., ZHANG, X. H., VANHARANTA, S., TAVAZOIE, S. F., MORRIS, P. G., DOWNEY, R. J., MANOVA-TODOROVA, K., BROGI, E. & MASSAGUE, J. 2011. Breast cancer cells produce tenascin C as a metastatic niche component to colonize the lungs. *Nat Med,* 17**,** 867-74.
PICCININI, A. M. & MIDWOOD, K. S. 2012. Endogenous control of immunity against infection: tenascin-C regulates TLR4-mediated inflammation via microRNA-155. *Cell Rep,* 2**,** 914-26.
PICCININI, A. M. & MIDWOOD, K. S. 2014. Illustrating the interplay between the extracellular matrix and microRNAs. *Int J Exp Pathol,* 95**,** 158-80.
PICCININI, A. M., ZULIANI-ALVAREZ, L., LIM, J. M. & MIDWOOD, K. S. 2016. Distinct microenvironmental cues stimulate divergent TLR4-mediated signaling pathways in macrophages. *Sci Signal,* 9**,** ra86.
RIEDL, S., TANDARA, A., REINSHAGEN, M., HINZ, U., FAISSNER, A., BODENMULLER, H., BUHR, H. J., HERFARTH, C. & MOLLER, P. 2001. Serum tenascin-C is an indicator of inflammatory bowel disease activity. *Int J Colorectal Dis,* 16**,** 285-91.
ROBERTS, D. D. 2011. Emerging functions of matricellular proteins. *Cell Mol Life Sci,* 68**,** 3133-6.
RUHMANN, M., PICCININI, A. M., KONG, P. L. & MIDWOOD, K. S. 2012. Endogenous activation of adaptive immunity: tenascin-C drives interleukin-17 synthesis in murine arthritic joint disease. *Arthritis Rheum,* 64**,** 2179-90.
RUTNAM, Z. J., WIGHT, T. N. & YANG, B. B. 2013. miRNAs regulate expression and function of extracellular matrix molecules. *Matrix Biol,* 32**,** 74-85.
SALAS, A., FERNANDEZ-BANARES, F., CASALOTS, J., GONZALEZ, C., TARROCH, X., FORCADA, P. & GONZALEZ, G. 2003. Subepithelial myofibroblasts and tenascin expression in microscopic colitis. *Histopathology,* 43**,** 48-54.
SOREK, R., KUNIN, V. & HUGENHOLTZ, P. 2008. CRISPR--a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nat Rev Microbiol,* 6**,** 181-6.
SRIRAMARAO, P. & BOURDON, M. A. 1993. A novel tenascin type III repeat is part of a complex of tenascin mRNA alternative splices. *Nucleic Acids Res,* 21**,** 163-8.
TAVAZOIE, S. F., ALARCON, C., OSKARSSON, T., PADUA, D., WANG, Q., BOS, P. D., GERALD, W. L. & MASSAGUE, J. 2008. Endogenous human microRNAs that suppress breast cancer metastasis. *Nature,* 451**,** 147-52.
TUCKER, R. P. & CHIQUET-EHRISMANN, R. 2009. The regulation of tenascin expression by tissue microenvironments. *Biochim Biophys Acta,* 1793**,** 888-92.

URNOV, F. D., MILLER, J. C., LEE, Y. L., BEAUSEJOUR, C. M., ROCK, J. M., AUGUSTUS, S., JAMIESON, A. C., PORTEUS, M. H., GREGORY, P. D. & HOLMES, M. C. 2005. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature,* 435**,** 646-51.

VAN DER OOST, J., JORE, M. M., WESTRA, E. R., LUNDGREN, M. & BROUNS, S. J. 2009. CRISPR-based adaptive and heritable immunity in prokaryotes. *Trends Biochem Sci,* 34**,** 401-7.

VAN EYKEN, P., GEERTS, A., DE BLESER, P., LAZOU, J. M., VRIJSEN, R., SCIOT, R., WISSE, E. & DESMET, V. J. 1992. Localization and cellular source of the extracellular matrix protein tenascin in normal and fibrotic rat liver. *Hepatology,* 15**,** 909-16.

VERRECCHIA, F. & MAUVIEL, A. 2007. Transforming growth factor-beta and fibrosis. *World J Gastroenterol,* 13**,** 3056-62.

VISSE, R. & NAGASE, H. 2003. Matrix metalloproteinases and tissue inhibitors of metalloproteinases: structure, function, and biochemistry. *Circ Res,* 92**,** 827-39.

WALLNER, K., LI, C., SHAH, P. K., FISHBEIN, M. C., FORRESTER, J. S., KAUL, S. & SHARIFI, B. G. 1999. Tenascin-C is expressed in macrophage-rich human coronary atherosclerotic plaque. *Circulation,* 99**,** 1284-9.