Evaluation of CRISPR-Cas9 approaches to investigate microRNA targeting in human chondrocytes

Noman Mahmood Chaudhry

Exeter College

Supervisors: Prof. Tonia Vincent and Prof. Tudor Fulga

Kennedy Institute of Rheumatology

Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences

University of Oxford

Thesis submitted to the University of Oxford for the degree of Doctor of Philosophy (D.Phil.) in Molecular and Cellular Medicine

Michaelmas 2019

Abstract

Micro RNAs (miRNAs) are small, non-coding, negative regulators of genes that play a major role in a wide variety of biological and pathological processes. The chondrocyte-selective *miR-140* is required for normal cartilage development, as demonstrated by the developmental defects in *miR-140* knockout (KO) mice. However, only a few studies have been conducted to explore the mechanism of action of *miR-140* at the cellular level.

In the present study, we describe a CRISPR-Cas9 system to identify functional *miR-140* targets in isolated primary human osteoarthritic articular chondrocytes. CRISPR-Cas9 editing offers a highly target-specific and stable method for exploring functional miRNA targets, without perturbing other miRNA levels. Developing a systematic approach of double transfection using CRISPR-Cas9, we produced rapid, efficient, and reproduceable genetically engineered populations of primary human articular chondrocytes, without prior need for clonal selection.

Stable knockdown of *miR-140* (> 90 %), with no detectable off-target activity, identified several *miR-140* regulated targets, of which Septin 2 (*SEPT2)*, bone morphogenetic protein 2 (*BMP2),* and fibroblast growth factor 2 (*FGF2)* had previously been reported. The major aggrecanase *ADAMTS-5* was not a *miR-140* target in our study, even though this had been reported previously. A number of novel targets were identified including Agrin (*AGRN*), a gene encoding for a heparan sulfate proteoglycan with chondrogenic potential, two regulators of the retinoic acid pathway, retinoic acid receptor gamma

(*RARG*) and cytochrome P450 26B1 (*CYP26B1*), and two molecules critical for primary cilium function, intraflagellar transport 88 (*IFT88*) and Tau Tubulin Kinase 2 (*TTBK2*).

In order to attempt to differentiate between direct and indirect effects of the miRNA, I adopted a novel CRISPR-Ca9 based technology, in which the miRNA response element (MRE) of the target gene, rather than the miRNA itself, was targeted. I report the challenges associated with this approach in primary human chondrocytes.

Taken together, I describe a highly efficient and reproduceable CRISPR-Cas9-mediated protocol, in which I identified novel *miR-140* targets of disease relevance using primary OA chondrocytes.

Table of contents

Acknowledgements

This has been quite a journey. My PhD experience has taught me invaluable lessons in how to learn from failures, and convert these, through persistence and hard work, into accomplishments. Professor Tonia Vincent has been instrumental in this process. Her trust, guidance, and immense knowledge have provided me with a firm basis to grow as a scientist. I would like to sincerely thank her for not only taking me on as a student during a time of uncertainty, but also for her enormous contribution towards making this a positive story.

I am also grateful towards my co-supervisor Professor Tudor Fulga, who possesses incredible technical insight in the field of genome engineering, and provided support whenever required. Similarly, I would like to thank Professor Chris Murphy who recruited and put his trust in me.

In the lab, Dr Christine Seidl helped me develop enormously as a researcher during my first years. Her thought-provoking discussions helped me become an independently thinking scientist, and her daily support and guidance in the lab were vital in aiding my understanding of the complex mechanisms surrounding CRISPR-Cas9.

I am equally thankful to Dr Hayat Muhammad, not only for his generous support in handling and running stem cell experiments with me, but also for forming a valuable friendship. Our countless pizza parties at the end of long, tiring lab days will never be forgotten.

I'd also like to express my gratitude towards Professor Afsie Sabokbar, Director of Graduate Studies. Afsie fulfils her duties beyond any expectation and is an incredible student supporter, who puts the student's (legitimate) interest above anyone else's. Her constant and unconditional support made a significant difference during adverse times. Thank you, Afsie.

I wish to thank my family from the bottom of my heart, who have supported all of my efforts and academic pursuits throughout my life, and are my main anchor.

My niece Eliza has grown in parallel to my PhD journey. Spending time with her and seeing her grow has been one of the greatest joys during my PhD studies. This happiness only multiplied later with Aleha. Thank you for always having been there and letting me be your kids's favourite uncle, Fahid and Emen.

I would also like to express my gratitude to my aunt Tahira Chaudhry for encouraging me to stride beyond the boundaries of complacency in order to pursue knowledge – a timeless piece of advice that turned out to be invaluable.

Finally, my parents; Arshid Chaudhry and Musrat Naz. Both have provided me with continuous hope and reassurances beyond any expectation not just during my PhD, but throughout my life. With their constant prayers and support, I am what I am now. This work is dedicated to them.

Words cannot express how grateful I am to achieve this landmark. All praise belongs to the Almighty, the Knower of the Unseen and Seen.

Noman Chaudhry Oxford, the 19th of December in 2019

Declaration of Authorship

I, Noman Chaudhry, declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

Thesis title: “Evaluation of CRISPR-Cas9 approaches to investigate miRNA targeting in human chondrocytes”

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. None of this work has been published before submission

Signed: 

Date: 19.12.2019

List of figures

List of tables

List of abbreviations

ACAN	Aggrecan
AGRN	Agrin
ADAMT-5	A disintegrin and metalloproteinase with thrombospondin motifs
BMP	Bone morphogenetic protein
bp	Base pairs
C9orf7	chromosome 9 open reading frame 7
Cas9	CRISPR associated protein 9
CDK	Cyclin-dependent kinase
cDNA	complementary DNA
COL2	Collagen type II
COMP	Cartilage oligomeric matrix protein
CRISPR	Clustered regularly interspaced short palindromic repeats
crRNA	CRISPR RNA
CTGF	Connective tissue growth factor
D1 / 2	Deletion 1 / 2 (barcode)
DMEM	Dulbecco's modified eagle's medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DSB	Double-strand break
ECM	Extracellular matrix
FBS	Fetal bovine serum

FGF	Fibroblast growth factor
Fw	Forward (primer)
GAG	Glycosaminoglycans
gDNA	genomic DNA
h	Hour(s)
hAC(s)	Human articular chondrocyte(s)
HEK293	Human embryonic kidney 293
HDR	Homology-directed repair
HS	Heparan sulfate
HSPG	Heparan sulfate proteoglycan
IL	Interleukin
ITS	Insulin-transferrin-selenium
JNK	c-Jun N-terminal kinase
KD	Knock down
KO	Knock out
M1 / 2	Maintenance 1 / 2 (barcode)
MAPK	Mitogen activated protein kinase
MEK	Mitogen activated protein kinase
MKP3	Mitogen activated protein Kinase, phosphatase 3
min(s)	Minute(s)
miRISC	miRNA-induced silencing complex
mRNA	Messenger RNA
miRNA	Micro RNA

MMP	Matrix metalloproteases
MSCs	Mesenchymal stem cells
NGF	Nerve growth factor
NHEJ	Non-homologous end joining
NFκB	Nuclear factor kappa-light-chain-enhancer of activated B cells
ns	Non-significant
NT	Non-targeting
nt	Nucleotide(s)
OA	Osteoarthritis / osteoarthritic
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCM	Pericellular matrix
PCR	Polymerase chain reaction
P/S	Penicillin/streptomycin
qPCR	Real-Time PCR
RA(R)	Retinoic acid (receptor)
RNA	Ribonucleic acid
RNP	Ribonucleoprotein
Rv	Reverse (primer)
Sept	Septin
SDS	Sodium dodecyl sulfate
sgRNA	single guide RNA
siRNA	Small interfering RNA

SNP	Single nucleotide polymorphism
SOX	Sex-determining region Y-box 9
ssODN(s)	single stranded oligo DNA nucleotide(s)
T_A	Annealing temperature
T3 / 7	T3 /7 (promotor sequence) barcode
T7E1	T7 endonuclease 1
TAK-1	Transforming growth factor β–activated kinase 1
TIMP	Tissue inhibitor of metalloproteinase
TGF	Transforming growth factor
tracrRNA	trans-activating CRISPR RNA
VEGF	Vascular endothelial growth factor
wt	Wild type
ΔCt	Delta Ct
ΔΔCt	Delta delta Ct

"(…) Thou seest not, in the creation of the All-merciful any imperfection, Return thy gaze, seest thou any fissure. Then Return thy gaze, again and again. Thy gaze, Comes back to thee dazzled, aweary."

(The Holy Qur'an 67:4-5)

Introduction

1

Chapter 1. Introduction

1.1 Cartilage Biology

1.1.1 Osteoarthritis

Osteoarthritis (OA) is a prevalent degenerative joint disease especially amongst older adults, that carries morbidity in the form of joint pain and disability. Other important aetiological factors in addition to increasing age are joint injury, obesity and genetic predisposition (Heidari, 2011). The common generalisation of OA as simply a disease of "wear and tear" is somewhat misleading. Though wear and tear indeed may constitute the triggering events leading to OA development, the disease process is not simply the presence of symptoms as a direct result of such attrition. Rather, it but actually entails the persistent, abnormal remodelling of the tissue that is driven by inflammatory responses including those intrinsic to the cartilage that are induced by injury – so-called mechano-inflammation (Vincent, 2019). The disease is primarily characterised by the progressive degradation of articular cartilage, due to mechanical and biochemical changes that occur in response to injury and with age.

OA is characterised by progressive destruction and loss of articular cartilage, thickening of the layer of bone immediately underneath the cartilage (subchondral bone), the formation of osteophytes (bony projections), and chronic low-grade inflammation of the joint. Additionally, pathological findings can be present in periarticular muscles, fat, and nerves which contribute to the disease in an individual (Loeser et al., 2012). The

presence of pathological changes in all tissues of the joint provides a rationale for describing OA as a disease of the whole joint as an organ, i.e. a “joint failure”. In this introduction I will focus on the changes seen within the cartilage of the joint, and their role in OA.

1.1.2 Cartilage

Cartilage is a connective tissue, of which there are three different types: elastic cartilage, hyaline cartilage and fibrocartilage. Each type of cartilage possesses a unique function within the body, reflected by its unique structure. Like all connective tissue types of the body, cartilage comprises relatively few cells embedded in their rigid secreted extracellular matrix (ECM); in the case of cartilage, the cells secreting and embedded in their ECM are chondrocytes (Bhosale and Richardson, 2008). Elastic cartilage is the most cellular of the three types, and an abundance of elastic fibres in the ECM provides elasticity and structural integrity, for example in the external ear. Fibrocartilage is the least cellular, is the strongest and least flexible type of cartilage, contains thick layers of dense collagen fibres, and is found in the intervertebral discs (Chang and Martin, 2018). Hyaline cartilage is the cartilage type found at articular surfaces in the joints, of which chondrocytes form between 1 – 5 % of total volume (Buckwalter and Mankin, 1997).

1.1.3 Articular cartilage structure

Depending on their place and distance from the surface of the cartilage, the function of the chondrocyte varies. There are four morphological zones described within the cartilage: the superficial zone, transitional zone, deep zone, and calcified cartilage (Fig. 1.1). The superficial zone is at the joint surface and is the thinnest of all layers, comprised of flattened chondrocytes covered by a thin layer of synovial fluid, which produce relatively high amounts of collagen, thereby providing tensile and shear strength. Moving down into the transitional zone, spheroid shaped cells are found with a higher matrix:cell ratio and large amounts of aggrecan in the matrix. In the deep zone, there is the highest proteoglycan concentration, lowest cell density, and largest diameter of collagen fibrils (Sophia Fox et al., 2009).

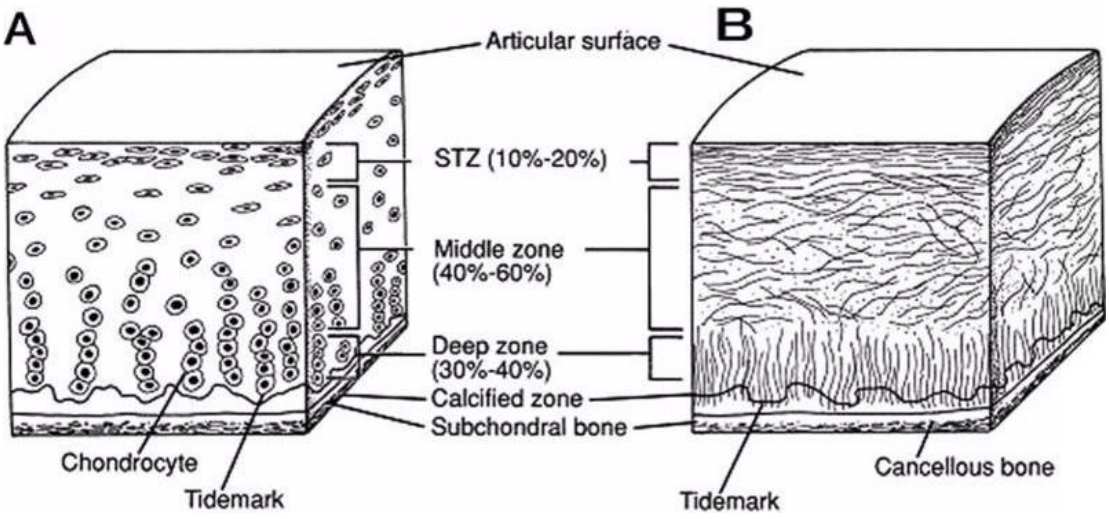


Fig. 1.1: Schematic representation of a cross selection of healthy articular cartilage.

A: cellular organization, B: collagen fibre architecture. Adapted from Sophia Fox et al. (2009). STZ – superficial tangential zone.

1.1.4 Articular cartilage composition

Articular cartilage is highly specialised to meet the function of the joint, facilitating smooth and safe movement between articulating bone surfaces. The ECM of articular cartilage contains high amounts of type II collagen fibres. Collagen type II alpha 1 chain (*Col2a1*) encodes the main component of type II collagen, the pro-alpha1(II) chain. Type II fibres make up 90 – 95 % of all collagens present, which all contribute to 10 – 20 % of the cartilage wet weight (Sophia Fox et al., 2009). There are at least five types of collagen found in articular cartilage which are genetically distinct. Types II, IX and XI are specific to cartilage. The network of these cross-linked collagen fibrils creates the framework and infrastructure of the ECM, providing the cartilage with its form and tensile strength. Type IX collagen molecules make up around 1 % of total collagens, and are found covalently linked to type II collagen fibrils (Eyre, 1991). Type X collagen is predominantly synthesised by hypertrophic chondrocytes (Mayne, 1989). *Col10a1* encodes for the alpha chain of type X collagen and is a specific marker of chondrocyte hypertrophy. Collagen I (*COL1A1* encodes for the alpha I chain) is implicated in the mineralisation of bone (Eyre, 1991).

While collagen type II is the principal collagen found in articular cartilage, the pericellular matrix (PCM), a narrow area of distinct matrix surrounding the chondrocyte (Fig. 1.2), bears collagen type VI (Poole et al., 1987). This helps to anchor the chondrocyte to the matrix and provide a framework along with additional proteins such as anchorin CII. Another molecule present in the PCM is perlecan, which interacts with collagen VI (Wilusz et al., 2012), but also binds to growth factors such as connective tissue growth

factor (CTGF) and fibroblast growth factor 2 (FGF2), thus possibly having a role in both stabilisation and cell signalling. These growth factors are released from the PCM upon injury where they exert a cellular signalling response (Vincent et al., 2002, Vincent et al., 2004, Tang et al., 2018).

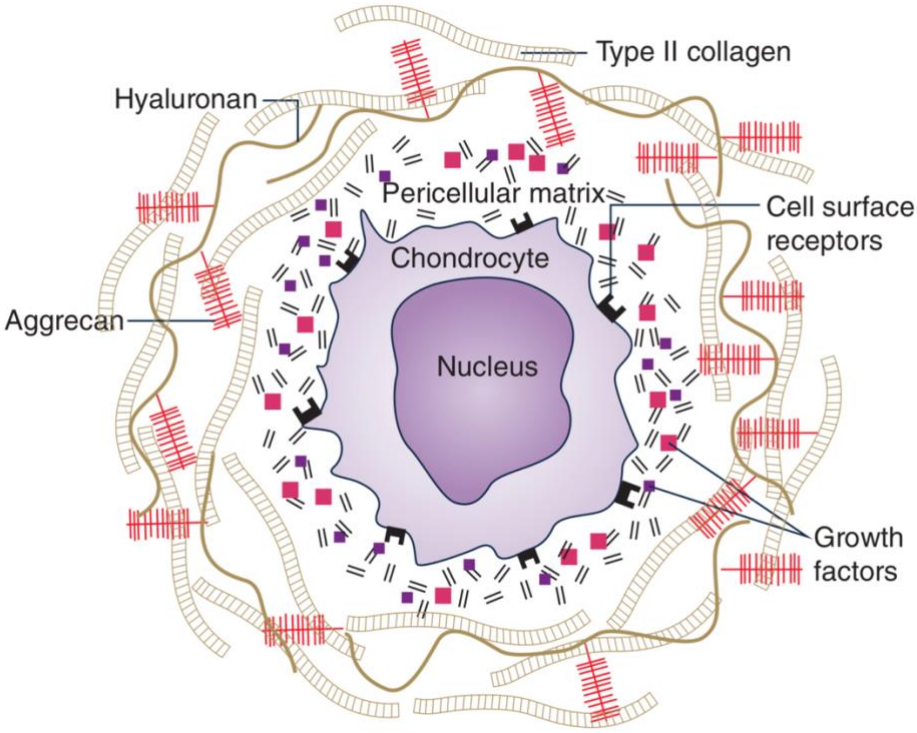


Fig. 1.2: Schematic representation of a Chondrocyte and its immediate surroundings.

The pericellular matrix (PCM) circuits the chondrocyte, and together they form the "chondron". Note that the PCM is lacking fibrillar collagens and aggrecan. Adapted from Vincent and Wann (2019).

1.1.5 Proteoglycans

Another crucial component of the ECM are proteoglycans, principally aggrecan (Fig. 1.2). Aggrecan is the most abundant proteoglycan and is responsible for water retention and hydration of the cartilage. Proteoglycans consist of glycosaminoglycans (GAGs;

chondroitin sulfate, keratan sulfate, heparan sulfate, dermatan sulfate) attached to a core protein which, in the case of aggrecan, is attached by link protein to hyaluronic acid (Kiani et al., 2002). Aggrecan specifically contains a core protein with three globular domains, G1, G2 and G3. GAGs, namely chondroitin sulfate and keratan sulfate, attach at an extended region between G2 and G3, while G1 interacts with hyaluronic acid and link protein, all together forming stable complexes (Hardingham and Fosang, 1995). The GAGs are negatively charged, and therefore become associated with mobile cations. Water thus follows down an osmotic gradient into the ion-concentrated cartilage ECM, and forms 65 – 80 % of cartilage wet weight (Bhosale and Richardson, 2008). These structural features are important, as through facilitating water retention, the cartilage is given its compressive strength, and is able to maintain the necessary matrix hydration that enables efficient pressure absorption and load-dependent deformation of the joint. The coefficient of friction between two layers of hyaline cartilage at the joint surface is three times less than that of ice on ice (Wang and Ateshian, 1997), providing smooth joint movement with very little friction. Though in much lower amounts than aggrecan, smaller proteoglycans such as decorin, biglycan, and fibromodulin aid in the stabilisation of the matrix, as well as having a role in binding growth factors (Yamaguchi et al., 1990, Hildebrand et al., 1994).

Another heparan sulfate proteoglycan in articular cartilage is agrin (encoded by *Agrn*) (Groffen et al., 1998). Agrin is reported to be increasingly lost upon OA progression *in vitro*. Knockdown of Agrin leads to decreased glycosaminoglycan content, reduced cartilage specific ECM molecules and reduced SOX9 (Eldridge et al., 2016). Moreover, it

strongly contributes towards chondrocyte differentiation and cartilage formation *in vivo* (Eldridge et al., 2016).

In addition to collagen and proteoglycan molecules, proteins such as cartilage oligomeric protein (COMP) are found in cartilage, which acts as a marker of turnover and degeneration of the tissue (Tseng et al., 2009).

The pH of the ECM is acidic compared to other tissues at around 6.9, and is necessary to maintain integrity of the matrix infrastructure (Wilkins et al., 2000).

1.1.6 ECM homeostasis

Chondrocytes are responsible for organising and maintaining the ECM. There is little cell-to-cell contact between chondrocytes in the tissue. They respond to mechanical stimuli to achieve homeostasis of the ECM (Sanchez-Adams et al., 2014). The tissue undergoes remodelling as chondrocytes replenish components lost through degradation. Therefore, chondrocytes need to be able to detect and respond to alterations in the surrounding matrix, such as the presence of degraded molecules. Importantly, the matrix itself is a signal transducer for the cells. As joint movement occurs and forces are transduced across the joint, mechanical and physiochemical stresses signal to the chondrocyte to guide activity. Compression of the cartilage efficiently regulates matrix genes through mechanotransduction pathways. Putative mechanosensors include cell-surface integrins (Katsumi et al., 2004), and stretch and cation-sensitive ion channels (Lee et al., 2014, O'Conor et al., 2014, Mobasheri et al., 2010) as well as aforementioned matrix-dependent release of PCM growth factors (Vincent et al., 2002, Vincent et al., 2004, Tang et al., 2018). Studies have confirmed a

critical role for this mechanotransduction *in vivo*. In 1979 it was demonstrated that proteoglycan content of the ECM is rapidly lost when joints are immobilised in canines, but can be restored by re-introducing weight-bearing activities (Palmoski et al., 1979). Similarly, in patients with spinal cord injury, cartilage is seen to undergo atrophy (Vanwanseele et al., 2003).

Another stress that impacts on chondrocyte physiology is hypoxia (Lafont, 2010). Articular cartilage is aneural, avascular and alymphatic, and as such, chondrocytes receive nutrition via diffusion through the cartilage ECM from surrounding tissues. The avascular nature of the tissue means that chondrocytes function in a hypoxic environment compared with other tissues (0.5 – 5 % O_2, depending on zone) (Lund-Olesen, 1970, Brighton and Heppenstall, 1971) and rely on anaerobic respiration. Work by Murphy and Sambaniss (2001) showed promotion of chondrocyte matrix gene expression in hypoxic conditions in bovine articular chondrocytes, while Domm et al. (2002), demonstrated similar effects in human articular chondrocytes. In the state of hypoxia, hypoxia-inducible transcription factors (HIFs, e.g. HIF1-α) are regulated by hydroxylation (Semenza and Wang, 1992), such that in high oxygen environments, HIFs are hydroxylated and degraded (Semenza, 2000).

In contrast, in a hypoxic environment with oxygen levels similar to those of the cartilage ECM, hydroxylation and subsequent degradation is reduced. In the absence of rapid degradation, HIF1-α is able to dimerise with constitutively expressed HIF-1β, translocate to the nucleus, and bind gene promoters (Dengler et al., 2014).

1.1.7 Pathological turnover of extracellular matrix

Degradation of cartilage is a cardinal feature of OA. Matrix-degrading enzymes, specifically metalloproteinases (MMPs), are upregulated in disease and lead to degradation of ECM protein components including collagen and proteoglycans (Troeberg and Nagase, 2012, Sapolsky and Howell, 1982).

The cartilage-degrading MMPs are a family of 23 zinc-dependent proteinases, of which there are four main functional subtypes based on their respective substrates: collagenases, gelatinases, stromelysins, and membrane type enzymes (MT-MMPs). Collagenases cleave collagens and include MMP-1, -8, and -13. Gelatinases include MMP-2 and -9, and cleave collagens type VI and XI, laminin and aggrecan, with MMP-2 also cleaving collagen type II (Nagase et al., 2006). Stromelysins include MMP-3 and have a broad specificity with substrates including proteoglycans and fibronectin (Yasuda, 2006). MT-MMPs such as MMPs -14 to -16 are located at the cell membrane, and can activate other zymogen MMP precursors. MMP-14 also cleaves collagen type II (Nagase et al., 2006). The peptide fragments of ECM proteins released as products of MMP regulated breakdown further upregulate proteinases, resulting in a form of positive feedback loop, enhancing catabolism further (Yasuda, 2006).

In osteoarthritis MMP-13 is thought to be the main collagenase that cleaves type II collagen, and is significantly expressed in osteoarthritic joints whilst almost undetectable in normal, disease-free adult tissues (Li et al., 2017). Its increased expression is observable by immunohistochemistry and qPCR in the initial onset period in rat OA models (Bo et al., 2012). Shen et al. (2013) found a critical role for upregulated

MMP-13 in the development of OA via TGF-β inhibition pathways. Neuhold et al. (2001) expressed a constitutively active *MMP-13* mutant in transgenic mice and found that they exhibited an OA-like pathology. Similarly, *MMP13* knockout mice displayed some reduction in disease that was surgically induced by medial meniscal destabilisation (Little et al., 2009). MMP-13 levels and activity are controlled by transcriptional and post-transcriptional mechanisms (Young et al., 2019) including epigenetic modification (Chernov and Strongin, 2011) and proenzyme activity (Knäuper et al., 1996).

The main aggrecanases in OA are a disintegrin and metalloproteinase with thrombospondin motifs 4 (ADAMTS-4) (Troeberg and Nagase, 2012) and ADAMTS-5 (Hurskainen et al., 1999). Similar to MMPs, members of the ADAMTS family are zinc-dependant metalloproteinases (Fosang et al., 2008). Their pathological relevance was established with the deletion of *ADAMTS-5* in mice, which notably protected against surgically induced OA and inflammatory arthritis (Glasson et al., 2005, Stanton et al., 2005) . Deletion of *ADAMTS-4* in mice did not exhibit the same level of protection against surgically induced OA (Glasson et al., 2004), although there is evidence that *ADAMTS-4* might contribute towards cartilage degradation in human articular chondrocytes (Song et al., 2007, Troeberg and Nagase, 2012).

Under normal circumstances, *MMP-13* and *ADAMTS-5* are both constitutively secreted by chondrocytes, but regulated through their rapid endocytosis by cell surface receptor low-density lipoprotein receptor-related protein 1 (LRP1). Following this they undergo intracellular degradation (Yamamoto et al., 2013). In OA shedding of LRP1 is increased, thereby resulting in a reduction in MMP and ADAMTS-5 uptake and destruction. In recent years, specific sheddases have been identified that have a role in this. Yamamoto

et al. (2017) identified membrane-bound MMPs, ADAM-17 and MMP-14 as LRP1 sheddases using proteinase inhibitor profiling and gene silencing techniques. These sheddases were then selectively inhibited by monoclonal antibodies, which restored endocytosis of MMPs by LRP1 and reduced aggrecan degradation.

Both ADAMTSs and MMPs can be inhibited by tissue inhibitors of metalloproteinases 3 (TIMP3), giving it a significant chondro-protective role (Scilabra et al., 2018). *TIMP3* mRNA levels are largely upregulated and its protein levels are also controlled by LRP1-mediated uptake. The primary cilium also appears to have a role in LRP1-mediated endocytosis (Coveney et al., 2018).

1.1.8 Cartilage injury

Joint injury can affect the soft tissue (e.g. meniscus, cruciate ligaments or bone) and commonly leads to the development of post-traumatic osteoarthritis (Buckwalter and Martin, 2004). While other risk factors such as age and obesity are important in "age-related" osteoarthritis, approximately 12 % of symptomatic osteoarthritis is estimated to be post-traumatic, and in the ankle joint it is the predominant cause of OA (Buckwalter and Martin, 2006). Some studies suggest that up to 80 % of patients with cartilage injury confirmed by imaging develop further degeneration within 5 years (Roos et al., 1995, Sherman et al., 1988).

Injury at the joint resulting in damage to the articular cartilage has implications at the molecular and tissue level. There is currently little evidence from either murine studies

or human injury cohorts that soluble factors release at the time of injury are driving OA development (Vincent, Watt unpublished).

Homeostatic regulation between anabolism and catabolism of the extracellular matrix by the chondrocyte following injury may be managed by factors including growth factors, cartilage loading, matrix matalloproteinases (MMPs), and MMP tissue inhibitors (TIMPs) (Vincent et al., 2004).

The processes that occur after acute joint injury likely involve an immediate cell loss through necrosis, damage to ECM components and bleeding into the joint space. There is a strong inflammatory response that can be detected in the synovial joint (Lohmander et al., 1989, Watt et al., 2016). Beyond the first 3 months symptoms usually resolve, although there may be some residual low-grade inflammation. However, around 50 % of individuals who sustain a destabilising joint injury will subsequently develop OA within 10 to 20 years (Lohmander et al., 2007).

1.1.9 Molecular response to mechanical injury

Mechanical injury has been a particular focus of the Vincent group in view of its importance as an aetiological factor in OA and the exquisite sensitivity of cartilage to injury. Inducing mechanical injury to cartilage enables the study of the molecular response that takes place. By cutting the intact articular cartilage from the joint (explantation) or by removing the femoral head cartilage (avulsion) of young mice,

cellular signalling can be activated. A number of injury-regulated pathways have been explored in detail.

1.1.9.1 Growth factor release: FGF2 and CTGF

Our group has previously shown that FGF2 and latent transforming growth factor beta (TGFβ), along with CTGF, are released upon cartilage injury in mouse (Tang et al., 2018). FGF2, a heparin-binding molecule, is released from the PCM, presumably due to a rapid sodium influx from the sulfated, aggrecan-rich matrix upon mechanical stress (Vincent et al., 2002). This FGF2-mediated injury response in chondrocytes has been reported to delay the development of OA in mice (Chia et al., 2009, Chong et al., 2013). FGFs are a family of 22 proteins that regulate a wide and varied number of processes in the adult human body including tissue repair and regeneration, as well as brain patterning and limb development in the embryo (Ornitz and Itoh, 2015). The involvement of FGFs in human disease processes is complex but can be observed in an array of diseases as a manifestation of FGF signalling dysfunction, which can be a result of loss- or gain-of-function mutations in both the FGF ligand or its receptor (Helsten et al., 2015). Examples include autosomal dominant hypophosphataemic rickets (ligand gain-of-function) and Kallmann syndrome (receptor loss-of-function).

Similarly, Tang et al. (2018) reported that CTGF controlled TGFβ activation of its cell surface receptor in cartilage, by covalent CTGF/TGFβ binding, regulated by TGFβR3. TGFβ, a key growth factor, is implicated in OA progression but also is a vital driver of chondrogenesis (van der Kraan et al., 1990, van der Kraan, 2017, Zhen et al., 2013).

1.1.9.2 Injury signalling pathway

Upon injury, numerous intracellular signalling pathways have been reported to activate. These include TGF-β activated kinase 1 (TAK1) which lies upstream of mitogen-activated protein kinases (MAPK) (c-Jun N-terminal kinase (JNK) and p38 kinase (Ismail et al., 2015, Ismail et al., 2016)), as well as nuclear factor κB (NFκB). However, the injury-dependent upstream mediators of MAPK and NFκB are still unknown in cartilage (Vincent, 2019).

1.1.9.3 Regulation of retinoic acid signalling pathways

Another molecule with a role implicated in OA is retinoic acid (RA). A recent genome-wide association study (GWAS) carried out by Styrkarsdottir et al. (2014) elucidated an association between hypomorphic variants within the *ALDH1A2* gene and severe hand OA. The *ALDH1A2* gene codes for the enzyme responsible for generating RA. Our lab showed (Zhu et al., 2018) that injury to porcine and murine cartilage strongly suppressed RA dependent genes including cytochrome P450 family 19 (*CYP19A1*) (in mouse), cytochrome P450 family 26 (*CYP26A1),* and *CYP26B1* (in porcine) as well as all three retinoic acid receptors (RAR), RAR-alpha (*RARA*), RAR-beta (*RARB*) and RAR-gamma (*RARG*) in porcine.

1.1.9.4 The primary cilium

The cilium does not play a role in acute mechanosensation, however, it has been demonstrated to be an important transducer in other cell types.

Primary cilia are non-motile microtubule-based organelles that generally serve as sensory organelles and are involved in a number of biological processes (Satir et al., 2010). In cartilage, the presence of these cilia on chondrocytes increases with the level of tissue degeneration and disease progression in explanted mature bovine patella articular cartilage (McGlashan et al., 2008).

The primary cilium has been linked to OA pathogenesis, with one study reporting that cilia-mediated hedgehog signalling is abnormally up-regulated during OA pathogenesis, which led to accelerated degradation of the cartilage extracellular matrix (Knight, 2013).

1.1.10 Human mesenchymal stem cells

Mesenchymal stem cells (MSCs) are adult multipotent cells with potential to differentiate into multiple cell types including chondrocytes, osteoblasts, myocytes, and adipocytes (Caplan, 1991). As cells with multipotency and self-renewing abilities, that are additionally easy to isolate, human mesenchymal stem cells (hMSCs) are frequently utilised in a research context. The location source first discovered and most frequently used to this day to harvest hMSCs is the bone marrow, however hMSCs are also found in adipose tissues (Bunnell et al., 2008), in dental tissues (Gronthos et al., 2002), as well as in the umbilical cord (McElreavay et al., 1991) and amniotic fluid (Anker et al., 2003). There is some scope and use for adipose tissue-derived and cord-derived stem cells (Mareschi et al., 2012). Bone marrow-derived stem cells have good potency to differentiate towards osteogenic, chondrogenic and adipogenic cell fates. Importantly, MSCs in the bone marrow are a separate entity from haematopoietic stem cells also

present in the bone marrow, which differentiate into blood cell type lineages. As such, CD34, a marker for hematopoietic stem cells is generally said to not be expressed by MSCs, however there is some controversy surrounding certain expression markers of MSCs (Lin et al., 2012). In 2006, the International Society for Cellular Therapy proposed minimal and essential criteria to meet the definition of a hMSC, in order to standardise and maintain some consistency in the characterisation of cells across studies. These criteria included, in addition to the absence of CD34 expression, the presence of CD73, D90, and CD105, and absence of CD14, CD45 and HLA-DR (Dominici et al., 2006).

In addition to differentiating into multiple cell types, hMSCs have a role in immunomodulation, secreting prostaglandin E2, nitric oxide, IL-6 and other signalling molecules which overall have an inhibitory effect on immune action. A study by Spaggiari et al. (2009) showed that peripheral monocytes in the blood (precursors to dendritic cells, cells involved with T-cell activation) are specifically inhibited from developing the phenotype of immature or mature dendritic cells in the presence of MSCs.

In this study we availed from the potential of stem cells to differentiate into chondrocytes and form cartilage *in vitro*.

1.1.11 Chondrogenesis

Chondrogenesis is the process by which cartilage is formed, and can be grossly summarised as the aggregation and condensation of mesenchymal stem cells within the areas of the embryo mapped to form the skeleton. More specifically, the process of

chondrogenesis can be defined in distinct stages: Presence of MSCs, aggregation and condensation of MSCs, their differentiation into chondrocytes, proliferation of chondrocytes, pre-hypertrophic chondrocyte formation, and hypertrophic chondrocyte differentiation. As mesenchymal stem cells condense, neural cadherin (N-cadherin) (Oberlender and Tuan, 1994) and neural cell adhesion molecule (N-CAM) (Widelitz et al., 1993) are important adhesion molecules that are expressed and facilitate communication between cells and the organisation of the tissue. *SOX9* is a crucial transcription factor that regulates the differentiation of cells towards a chondrocytic cell fate following their aggregation (Bi et al., 1999), with other members of the SOX family, FGF growth factors, TGFβ and others also playing a role (DeLise et al., 2000). A crucial requirement for *in vitro* chondrogenic differentiation of MSCs includes the supplementation of TGFβ. TGFβ has been reported to activate intracellular signalling pathways such as MAPK, p38, extracellular signal-regulated kinases (ERK1), and JNK (Watanabe et al., 2001, Tuli et al., 2002, Nakamura et al., 1999).

Several microRNAs (miRNAs) target and regulate the process at each stage of chondrogenesis by interacting with the aforementioned transcription factors and cytokines. For example, *miR-145* directly targets and regulates *SOX9*, with increased levels of *miR-145* reducing *SOX9* (Martinez-Sanchez et al., 2012). *SOX9* is expressed in all cartilage primordia during embryogenesis, as well as in nerve tissues and urogenital tissue. Mutations in *SOX9* result in profound skeletal defects and dwarfism in mouse (Bi et al., 1999). Martinez-Sanchez et al. (2012) showed that essential genes involved in the production of ECM, *COL2A1* and *ACAN*, but also *miR-140*, were significantly reduced with increased levels of *miR-145*. Essentially, communication between cells and various

receptors, growth factors, transcription factors and enzymes collectively give rise to a spatiotemporally sensitive pattern of signalling that ultimately triggers MSCs to differentiate into chondrocytes, which grow to release ECM components, thereby eventually forming a cartilaginous template for bone to form within.

1.2 MicroRNAs

MicroRNAs (miRNAs), first discovered through the analysis of developmental timing mutants in *Caenorhabditis elegans* (Lee et al., 1993), constitute an extended family of small, highly conserved non-coding RNA molecules (21-23 nucleotides long), which control gene expression at the post-transcriptional level. They are involved in a wide variety of biological processes such as development, metabolism, and cell proliferation (Martinez-Sanchez et al., 2012). Therefore, it is important to understand their underlying mechanism and biological function. In the same context, it is important to identify functional mRNA targets of specific miRNAs, since miRNAs are also involved in many diseases such as cancer (Lujambio and Lowe, 2012) and cardiovascular disorders (Small and Olson, 2011).

1.2.1 miRNA biogenesis

By 2017, 1881 pre-miRNAs in humans had been identified, with half being produced from the DNA of intergenic regions (transcribed independently from a host gene with

their own promotors), and the other half from intragenic regions (mostly introns, very few exons) in the genome (De Rie et al., 2017).

RNA polymerase II transcribed genes are responsible for encoding the vast majority of mammalian miRNAs (Lee et al., 2004, Cai et al., 2004). miRNA biogenesis is structured in canonical and non-canonical pathways (Fig. 1.3). The more dominant, canonical pathway begins in the nucleus, when a precursor hairpin structure (pre-miRNA), consisting of roughly 60-80 nucleotides, is first cleaved from the primary miRNA transcript (pri-miRNA) by the microprocessor complex, consisting of RNAse III enzyme Drosha and RNA binding protein DiGeorge syndrome critical region 8 (DGCR8). While Drosha cleaves the pri-miRNA duplex at the base of its hairpin structure, two DGCR8 proteins recognise a specific N6-methyladenylated GGAC and other motifs at the stem and ensure correct cleavage by binding at them (Lee et al., 2003, Denli et al., 2004, Gregory et al., 2004) . At this step, different miRNA isomiRs can originate from different Drosha cleavage patterns. isomiRs bear small sequence variations compared with the reference sequence (Kim et al., 2017, Neilsen et al., 2012). Cleavage by Drosha leads to the formation of a 2 nucleotide (nt) 3' overhang on the pre-miRNA. An exportin 5 (XPO5)/RanGTP (Okada et al., 2009) complex recognises this overhang and transports the pre-miRNA to the cytoplasm, where it undergoes further processing to form a mature miRNA. The pre-miRNA is processed by Dicer, which is another RNAse III enzyme, as well as TAR RNA-binding protein (TARBP) (in vertebra) (Kim et al., 2016). Dicer recognises the 3' overhang, removes the terminal loop by cleaving the pre-miRNA at a species-specific length, and thereby produces a small RNA duplex consisting of

roughly 20 nucleotides with another 2 nucleotide 3' overhang (Lau et al., 2012). Again, alternative Dicer cleavage can lead to various isomiR formations (Neilsen et al., 2012). The 3' end of the pre-miRNA is responsible for generating the 3p strand of the mature miRNA, and the 5p strand derives from the 5' end. While the guide strand of the mature miRNA duplex is loaded into Argonaute (AGO) and incorporated into a miRNA-induced silencing complex (miRISC) to form mature miRNAs of 21-23 nucleotides in length, the passenger strand is released and degraded (Rand et al., 2005, Matranga et al., 2005). Whether the 3p or 5p strand is loaded into AGO can be dictated by cellular environment or the cell type itself; usually preference is determined by the thermodynamic stability at their 5' ends (the less stable being loaded into AGO) (Suzuki et al., 2015), as well as having an adenine or uracil nucleotide as the first nucleotide at the 5'-terminal (Frank et al., 2010). Subsequently, the fraction of AGO-loaded 3p or 5p can range from equal distribution to a predominantly just one-sided loading.

AGO2 is the main protein in the miRISC complex, and mainly responsible for miRNA loading. AGO proteins consist of four different domains which bear different tasks: 1) amino (N) terminal domain, 2) Piwi-Argonaute-Zwille (PAZ) domain, 3) middle (MID) domain, and 4) P-element wimpy tests (PIWI) domain, as well as two linker domains (L1 and L2) (Fig. 1.4). While the 5' end of the miRNA is recognised and bound by the MID and PIWI domains, the miRNA's 3' nucleotide is held by the PAZ domain (Gebert and MacRae, 2019).

The MID and PIWI domains facilitate the first step of a 2-step target recognition mechanism, by arranging nucleotides 2-6 of the so-called seed (binding sequence). In a

second step, this quick, initial binding is further strengthened in the case of additional complementarity of nucleotides 7-8 (Chandradoss et al., 2015).

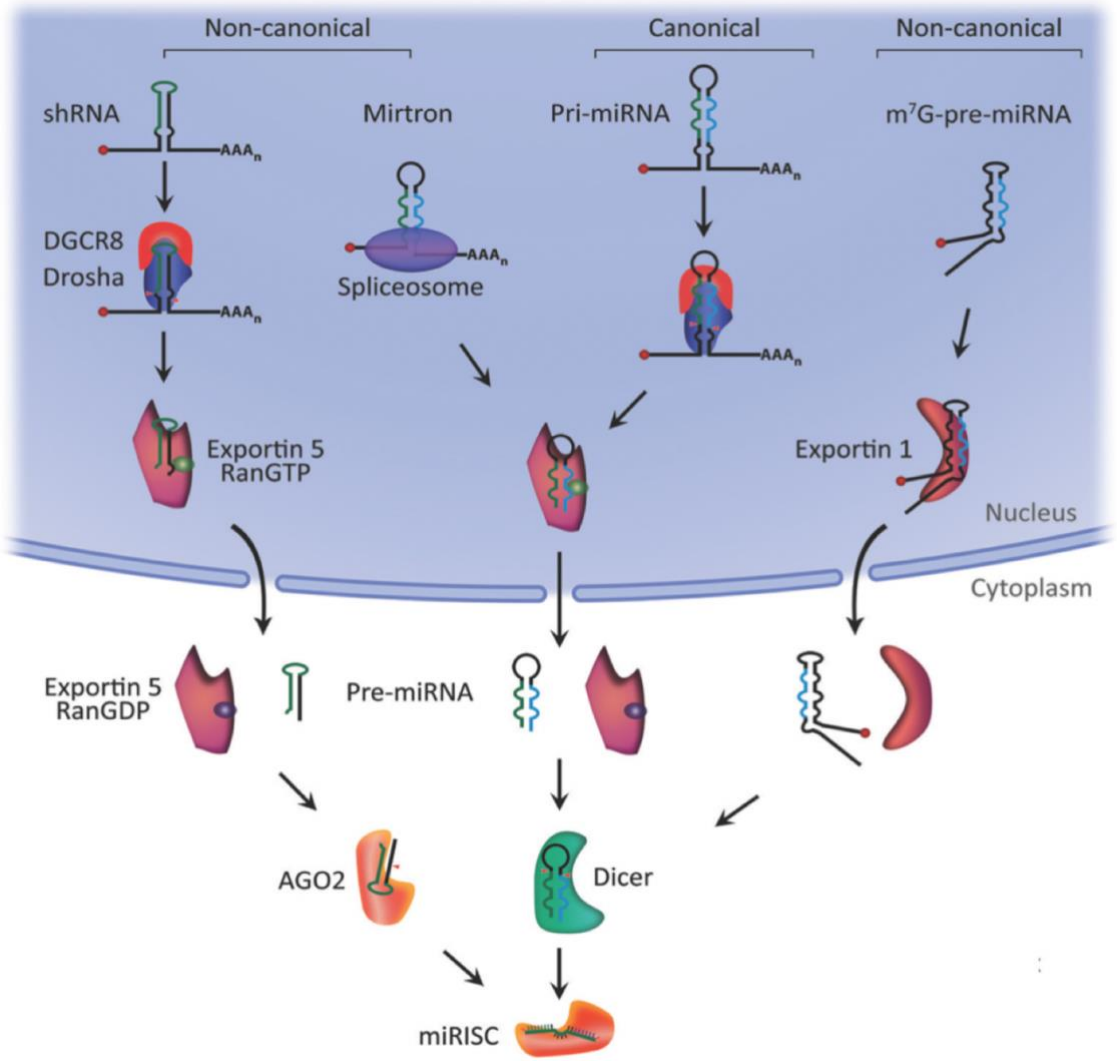


Fig. 1.3: Process of miRNA biogenesis involves numerous processing steps.

miRNA synthesis can occur through different pathways: the canonical (dominant) and non-canonical (non-dominant) pathway. Both the canonical and non-canonical pathways share overlapping features. Adapted from O'Brien et al. (2018).

Non-canonical pathways mainly rely on the same proteins as in the canonical pathway, namely Drosha, Dicer, exportin 5, and AGO2. However, they can be independent of either Drosha/DGCR8 or Dicer, which are replaced by similar substrates. Drosha replacements include mirtrons or 7-methylguanosine (m7G) (Havens et al., 2012). These Drosha-independent mirtrons and m7G capped-pre-miRNAs require Dicer for their cytoplasmic maturation, but are directly imported into the cytoplasm without Drosha cleavage, by either the exportin 5/RanGTP or exportin1 complex, respectively. Dicer replacements include short hairpin RNA transcripts (shRNA) (Havens et al., 2012). These pre-miRNAs are initially cleaved by the microprocessor complex (Drosha/DGCR8), exported to the cytoplasm by exportin 5/RanGTP, and undergo maturation by AGO2-dependent, Dicer independent cleavage.

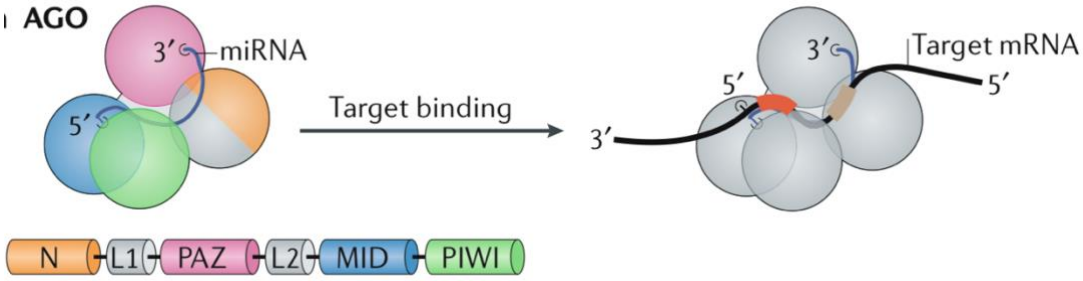


Fig. 1.4: Loading of mature miRNAs into AGO proteins, which consist of four domains

AGO proteins consist of four different domains which allow different tasks: 1) amino (N) terminal domain, 2) Piwi-Argonaute-Zwille (PAZ) domain, 3) middle (MID) domain, and 4) P-element wimpy tests (PIWI) domain, as well as two linker domains (L1 and L2). Modified from Gebert and MacRae (2019).

1.2.2 Mechanisms of miRNA mediated gene silencing

Since their discovery, miRNAs have spawned a dedicated field focused on unravelling the mechanisms and target selection of miRNAs. Nucleotides 2-8 at the 5'-end of miRNAs (seed sequences) are most critical for target selection, and in some cases, even fully sufficient for Watson/Crick base pairing between miRNA and the 3'-end of the target (Mendell and Olson, 2012) . While binding sites are most commonly found in the 3'-UTR of target mRNAs, in recent years this classical understanding has been expanded by new findings. Recent examples have shown that miRNA binding can also occur outside the 3'-UTR, and furthermore may even lack perfect seed sequence pairing (which is compensated for 3' complementarity, or by centred pairing (Martinez-Sanchez et al., 2012)). In other words, stable and specific miRNA binding can occur in the 5'-UTR, coding sequences, and gene promotors (Broughton et al., 2016).

The exact mechanism surrounding miRNA-mediated regulation of gene expression has been subject to controversy, however, there is evidence for consensus on several mechanisms: i) mRNA decay through AGO2 mediated endonucleolytic cleavage, ii) mRNA decay through deadenylation, and iii) inhibition of translation initiation (Fig. 1.5). Firstly, if there is perfect complementarity between a miRNA and a target site, AGO2, a core protein in the miRISC complex, mediates endonucleolytic cleavage of the target mRNA and negatively regulates its expression through mRNA degradation (Wang et al., 2009). At sites of imperfect complementarity (via the 5' seed region with nucleotides 2-8, and potentially additionally pairing at the 3' end), which constitute for the majority of

cases, miRNAs direct RISC to their target mRNA and promote its destabilisation and/or inhibition of its translation, which also results in gene silencing (Pasquinelli, 2012).

Secondly, the most dominant route to mRNA decay, mRNA degradation by deadenylation, begins with the recruitment of the glycine-tryptophan protein (GW182) by AGO. In case of miRNA/mRNA interaction, GW182 in return interacts with polyadenylate-binding proteins 2 and 3 (PAN2 / PAN3), as well as carbon catabolite repressor protein 4 (CCR4-NOT) complexes. These in turn deadenylate the target mRNA, efficiently assisted by poly(A)-binding protein C (PABPC) (Braun et al., 2011, Fabian et al., 2011). Deadenylation is followed by decapping, which is facilitated by the mRNA-decapping enzyme subunits 1 and 2 (DCP1 and DCP2) (Chen et al., 2009). Subsequently, the mRNA is rapidly degraded by 5'–3' exoribonuclease 1 (XRN1) (Braun et al., 2012).

Finally, another mechanism of miRNA mediated gene silencing is the initiation of translation inhibition. This less dominant mechanism involves the recruitment of eukaryotic initiation factors, such as 4 A-I (eIF4A-I) and eIF4A-II (Gebert and MacRae, 2019) .

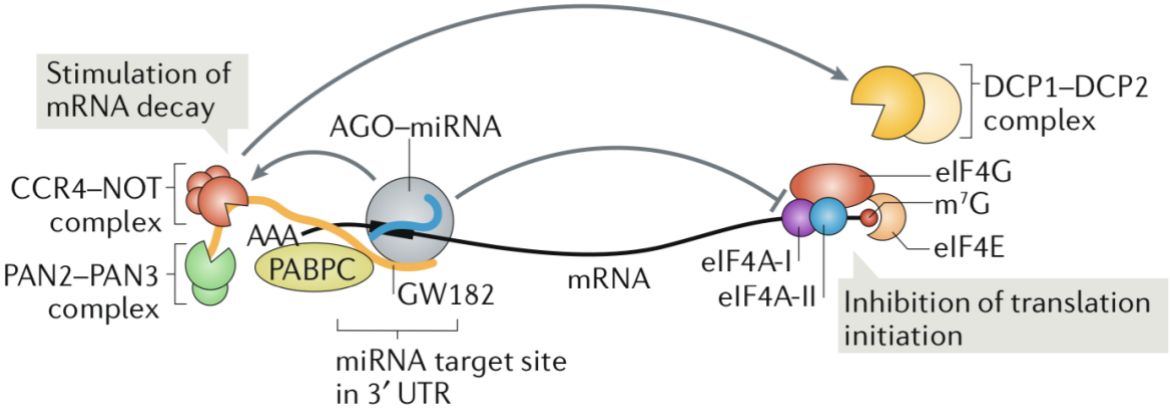


Fig. 1.5: Mechanism of target regulation by miRNAs.

miRNAs can regulate genes through mRNA degradation or translational inhibition. Both mechanisms require the recruitment of numerous factors. Modified from O'Brien et al. (2018).

On the contrary, studies have shown that instead of silencing a gene at post transcriptional level, miRNAs activate gene expression under certain conditions (Vasudevan, 2012). This gene upregulation can occur through two mechanisms: activation of gene expression through direct miRNA binding in either the 3'-UTR or 5'-UTR, or through relief of miRNA-mediated repression. Additionally, miRNAs have been shown to finetune transcription and translation rates, by moving between different subcellular compartments (Makarova et al., 2016). Through this movement, the miRNA can bind to different AGO proteins, which in turn interact with various effector proteins and cause wholly opposing responses.

1.2.3 miRNA turnover and regulation

Gene regulation is controlled in numerous processes by miRNAs, and dysregulation of miRNA activity has been linked with many diseases (Chang and Mendell, 2007). Hence, it is of importance to understand how miRNAs themselves are regulated, and how their expression profiles change over time. In general, miRNAs are known to be stable, however their regulation is controlled by multiple factors. As a result, their turnover rates can vary between several minutes to several days, depending on both biosynthetic and rapid degradation processes (Rüegger and Großhans, 2012). Neglecting these cellular factors, miRNA turnover can be simply determined by miRNA-specificity - their actual sequence and isomiR variants. Guo et al. (2015) reported how a 5' uracil is more stable compared with a 5' guanine or cytosine.

Many miRNAs have been described to display a tissue or stage specific expression pattern, while most miRNAs are regulated by the promotor of their host genes (Endisha et al., 2018). Transcriptions factors stimulate the transcription of these genes in *cis* or *trans* (Endisha et al., 2018).

Several cellular conditions are thought to influence the regulation of miRNA expression. Different stages of the cell cycle can affect miRNA levels. For example, *miR-29b* has been shown to have higher levels in mitotically arrested cells compared with cycling cells (Hwang et al., 2007). Interestingly, this regulation may be reciprocal as many miRNAs target components of the cell cycle (Chen et al., 2010).

Another well-documented regulator of miRNA levels is growth factor stimulation. For example, Avraham et al. (2010) demonstrated strong and rapid downregulation of

several miRNAs in human MCF10 breast and HEK293T cells upon serum starvation and subsequent stimulation with epidermal growth factor (EGF). Predicted targets of the downregulated miRNAs in MCF10 cells included numerous genes that were known to be upregulated in response to EGF stimulation, suggesting that these miRNAs might play a role in EGF cell response.

In the same year, Krol et al. (2010) reported that certain miRNAs were light sensitive and regulated by light in the mouse retina. Upon adaption from light to dark, miRNA levels roughly halved. Other miRNAs, which did not appear to be regulated by light, were characterised by fast turnover rates. In fact, rapid miRNA turnover is described as a distinct feature of miRNAs in human neurons *in vitro* as well as in post mortem human brain tissues (Sethi and Lukiw, 2009).

Specific enzymes that are responsible for miRNA decay have been identified. These miRNA RNases include several exoribonucleases (both 3' to 5' and 5' to 3'), such as small RNA degrading nucleases (SDNs), exoribonucleases 1 or 2 (XRNX1/2), ribosomal RNA-processing protein41 (RRP41), or the human polynucleotide phosphorylase (PNPase) (Rüegger and Großhans, 2012).

1.2.4 miRNAs and their role in cartilage biology

Dicer, an RNase III enzyme, is an essential component in the biogenesis of miRNAs. It cleaves the 70-nucleotide long precursor miRNA into a small RNA duplex of around 20 nucleotides. Bernstein et al. (2003) carried out a universal Dicer knockout in mouse with a PGK-ned expression cassette replacing the first domain of the enzyme. The disruption of Dicer led to a lack of development in Dicer-null mice which were depleted of stem cells, suggesting that miRNAs are essential for normal development.

A few years later, Kobayashi et al. (2008), demonstrated that miRNAs play a vital role in cartilage by specifically deleting the Dicer gene in cartilage, through use of a cartilage specific promotor, *Col2A1* (Col2-Cre). These mice were born with severe skeletal growth defects and died prematurely by the time of weaning. Skeletal growth defects were characterised by lack of chondrocyte proliferation and enhanced hypertrophy.

Certain miRNAs show dysregulated expression in OA (Jones et al., 2009). In fact, several miRNAs along with their predicted target genes were identified to be up- or downregulated in osteoarthritic chondrocytes, compared with healthy chondrocytes (Iliopoulos et al., 2008). Work in my previous lab has identified other important miRNAs and their targets in human chondrocytes. For example, *miR-675* has been shown to control expression of the key cartilage matrix gene *COL2A1* (Dudek et al., 2010), while *miR-145* has been identified as a direct regulator of *SOX9* (Martinez-Sanchez et al., 2012), which, as previously mentioned, is essential for normal cartilage development and function (Bi et al., 1999). Another miRNA, *miR-1247*, binds in the coding sequence of *SOX9* and alters its expressions levels (Martinez-Sanchez and Murphy, 2013b). Lastly,

the lab showed that derepression of *miR-138* contributes to the loss of the human articular chondrocyte phenotype, and identified *HIF-2α* as an important target regulating matrix synthesis (Seidl et al., 2016). Additionally, miRNAs have been identified that control cellular maintenance processes in cartilage such as autophagy (Endisha et al., 2018). The expression of such autophagy genes (autophagy associated gene 5 (*ATG5*) or beclin1 (*BECN1*)), which are directly regulated by miRNAs, has been shown to be reduced in osteoarthritic cartilage tissue. In healthy cartilage, autophagy is a vital process that controls chondrocyte homeostasis. It also regulates cell metabolism after exposure to stress (Endisha et al., 2018).

Similarly, other studies have shown that miRNAs can be regulated by proinflammatory cytokines, which contribute towards OA progression, but also directly regulate proinflammatory cytokine expression themselves (Endisha et al., 2018). In response to Il-1β *in vitro*, *miR-92a-3p* expression is downregulated in OA human cartilage (Mao et al., 2017). *ADAMTS-4* and *ADAMTS-5* are reported direct targets of *miR-92a-3p and miR-140-5p,* and contribute towards the degradation of aggrecan during OA progression (Mao et al., 2017, Miyaki et al., 2009). In contrast, *miR-203* has been shown to bear an anti-inflammatory role in OA. Increased levels of *miR-203* led to decreasing levels of proinflammatory cytokines, while in reverse the inhibition of *miR-203* in mouse chondrocytes increased the production of proinflammatory cytokines (Wang et al., 2018).

As various miRNAs are involved in a wide range of biological processes, naturally many miRNAs have been discovered in recent years that contribute towards cartilage-protective and -destructive processes, which can lead to complex signalling cascades.

These processes include the modulation of inflammatory signalling pathway genes, nitric oxide production (which promotes cartilage degeneration), senescence regulation (which can accelerate age-related cartilage degeneration), cell differentiation, apoptosis (chondrocyte apoptosis can contribute towards cartilage degradation), but also various transcription regulating factors (Endisha et al., 2018).

One miRNA deserves special attention. The cartilage-specific *miR-140* has been subject to broad research. *miR-140* is encoded in the E3 ubiquitin protein ligase producing gene *WWP2*, between exons 16 and 17 on chromosome 16 in humans (murine chromosome 8). *WWP2* possesses three isoforms: The full-length *WWP2* (WWP2-FL, 870aa), an N-terminal isoform (*WWP2-N*, 336aa), likely the result of splicing out intron 9-10, and a C-terminal isoform (*WWP2-C*, 440aa), which presumably is produced from a second promotor between intron 10-11 (Soond and Chantry, 2011). *miR-140* is co-expressed with the host *WWP2-C* gene (Yang et al., 2011). While the *miR-140* gene is not present in invertebrates, it is highly conserved amongst vertebrates, including human, mice and zebrafish.

In one of their earliest publications, Wienholds et al. (2005) carried out microarray analysis of 115 conserved vertebrate miRNAs in zebrafish embryos. Here they reported *miR-140* to be highly tissue specific and exclusively expressed in the cartilage of the jaw, head, and fins of zebrafish. One year later, Tuddenham et al. (2006) reported *miR-140* to be specifically expressed in cartilage of mouse embryos during both long and flat bone development. Through luciferase and western blot analysis, they established that histone deacetylase 4 (*HDAC4)* is regulated by *miR-140,* and for the first time, *miR-140* was proposed to contribute towards the formation and maintenance of cartilage

through this regulation of *HDAC4*. *HDAC4* is seen to control chondrocyte hypertrophy during skeletogenesis (Vega et al., 2004).

Miyaki et al. (2010), generated *miR-140* KO mice to study the role of *miR-140 in vivo*. These mice displayed a mild, osteoarthritic-like phenotype with age, characterised by a short stature, proteoglycan loss, and fibrillation of articular cartilage, particularly due to elevated *ADAMTS-5* expression levels. In reverse, transgenic mice with antigen-induced arthritis, which overexpressed *miR-140*, did not display any OA related phenotype (Miyaki et al., 2010). Similarly, the overexpression of *miR-140-5p* in synovial mesenchymal stem cell exosomes in an OA rat model led to increased chondrocyte proliferation and migration, and delayed OA progression (Tao et al., 2017). These studies suggest *miR-140* is required for normal articular cartilage development, but also normal endochondral bone development (Nakamura et al., 2011). Interestingly, *miR-140* expression is reduced in osteoarthritic tissue, but exhibits higher levels of expression in normal articular cartilage *in vitro* (Miyaki et al., 2009). This study further proposed *miR-140* to regulate *ADAMTS-5 in vitro* upon using a dsRNA mimicking *miR-140*. IL-1β, an inflammatory cytokine, was described to contribute towards suppression of *miR-140* in OA pathogenesis. Lastly, the same study also reported in the increased expression of *miR-140* upon chondrogenic differentiation of human bone marrow MSCs *in vitro*. Karlsen et al. (2013) described both *miR-140-3p* and *miR-140-5p* to be drastically upregulated during chondrogenic differentiation of hMSCs *in vitro*.

The role of *miR-140* in chondrogenesis is principally through its regulation of bone morphogenetic protein (BMP) signalling (Buechli et al., 2012). BMPs (bone morphogenetic proteins) are members of the TGFβ family that bind to serine-threonine

kinase receptors. In chondrogenesis, BMPs 2 and 4 in particular have a substantial role in promoting the differentiation process through Smad signalling pathways (Soullier et al., 1999). BMP2 has been identified as a direct target of *miR-140* (Hwang et al., 2014), with *miR-140*-null mice showing growth defects in bones, dwarfism, and craniofacial abnormalities, and importantly, reduced chondrocyte BMP signalling (Nakamura et al., 2011). A possible second mechanism by which *miR-140* contributes to chondrogenesis is through the regulation of SOX9 and aggrecan protein levels. It has been reported that inhibition of *miR-140* led to reduced SOX9 and aggrecan protein, and conversely, overexpression led to increased protein levels, while mRNA levels remained unchanged (Martinez-Sanchez et al., 2012). The small GTPase, *RALA*, was identified as a *miR-140* target and shown to regulate SOX9 protein levels (Karlsen et al., 2013). Thus, *miR-140* may promote chondrogenesis through indirectly stabilising these crucial cartilage proteins.

The knockdown of *miR-140* impairs chondrocyte proliferation due to increased Sp1 protein levels (despite no changes in mRNA levels) (Yang et al., 2011). *miR-140* regulates chondrocyte proliferation, partially by suppressing *Sp1*, a crucial transcription factor modulating the cell cycle regulator *p15* (Yang et al., 2011). Taken together, these studies suggest that *miR-140* plays a protective role in the development of osteoarthritis.

1.2.5 miRNA targeting

Using classical methods such as microarray screening, candidate miRNAs can be identified, which in turn, can bioinformatically predict mRNA targets that require

experimental validation. By subsequently validating such miRNA targets, proteins and pathways can be identified that may impact OA pathogenesis. However, identifying functional miRNA target sites is a challenge. A single miRNA can and does regulate several mRNA targets, while the converse is also true in that one gene can be targeted by multiple miRNAs (Bartel, 2004).

The complexity and difficulty of mapping direct interaction between a given miRNA and its target has resulted in the generation of a wide variety of experimental approaches (Martinez-Sanchez and Murphy, 2013a). However, these techniques, which include overexpression or inhibition of miRNAs, genome-wide approaches, and immunoprecipitation of the miRISC complex, are subject to limitations (Martinez-Sanchez and Murphy, 2013a, Long and Lahiri, 2012). While some of these limitations are obviously method-specific, a commonly important one is that miRNA targeting is not assessed under endogenous conditions. One way to address this is to modify the miRNA *in vivo,* through gene deletion or by modifying binding to target sites. CRISPR-mediated genetic modification is one way to achieve this.

1.3 CRISPR-Cas9

1.3.1 CRISPR-Cas9 is derived from the bacterial immune system

CRISPR stands for clustered regularly interspaced short palindromic repeats. These DNA sequence repeat clusters are separated by so-called spacer regions, which were identified to belong to viruses and other mobile genetic elements. The host bacterium integrates invader DNA into these spacer regions, keeping a genetic print of prior infection, and making the host immune against future infections by the same invader (Makarova et al., 2006, Barrangou et al., 2007). Adjacent to these CRISPR arrays are CRISPR-associated (cas) genes, which determine different CRISPR systems (Jansen et al., 2002). These cas nucleases cleave foreign DNA upon second infection from the same invader and are guided by the matured, so-called CRISPR RNA (crRNA) (which is transcribed from the CRISPR array) to their target (Maeder et al., 2013, Perez-Pinera et al., 2013). Based upon biochemical activities, comparative genomic analyses, and structures**,** two different classes of CRISPR systems have been proposed: class 1 CRISPR systems consist of type I and III (they utilise one large protein complex consisting of several cas proteins), while class 2 CRISPR systems consist of type II, IV, V, and VI (Makarova et al., 2015, Shmakov et al., 2015).

The most applicable CRISPR system for genome engineering is the class 2 type II CRISPR-Cas9 system from *Streptococcus pyogenes.* This bacterial immune system-derived CRISPR-Cas9 system has been widely adapted for genome engineering, and

implemented to demonstrate single guide RNA (sgRNA)-mediated Cas9 cleavage first *in vitro* (Jinek et al., 2012), then in human and mouse cells (Cong et al., 2013).

1.3.2 Cas9 endonuclease

The Cas9 endonuclease (1386 amino acids) uses a sgRNA to introduce a site-specific double strand break (DSB) in the DNA. This sgRNA includes the mature CRISPR RNA (crRNA) and a trans-activating crRNA (tracrRNA). The crRNA confers target specificity through a seed sequence, which is located in its 3' end of its spacer region (10 – 12 nucleotides). In contrast, the tracrRNA plays a critical role in Cas9 recruitment. This dual-component system, which specifies the DNA target site by Watson-Crick base pairing and binds to the Cas9 nuclease, enables targeted DNA cleavage through Cas9 (Doudna and Charpentier, 2014).

Cas9 is composed of the alpha-helical recognition lobe (REC) and nuclease lobe (NUC), which are mainly connected through an arginine rich bridge helix structure (Fig. 1.6). While the REC domain is structured in three alpha-helical domains (Hel-I, Hel-II, and Hel-III), the NUC lobe contains both nuclease domains (HNH and RuvC), as well as the C-terminal domain (CTD) (Jinek et al., 2014). This CTD includes sites for protospacer-adjacent motif (PAM; 5'-NGG) interaction, but only undergoes substantial structural rearrangement after binding to a guide RNA. This structural change is mainly manifested in the REC lobe, particularly Hel-III, which moves ~65 A° towards the HNH domain upon sgRNA binding (Jiang and Doudna, 2017). Only then does Cas9 turn from an inactive

complex to an active complex which is able to recognise target DNA through its PAM interaction sites at the CTD (Jinek et al., 2014).

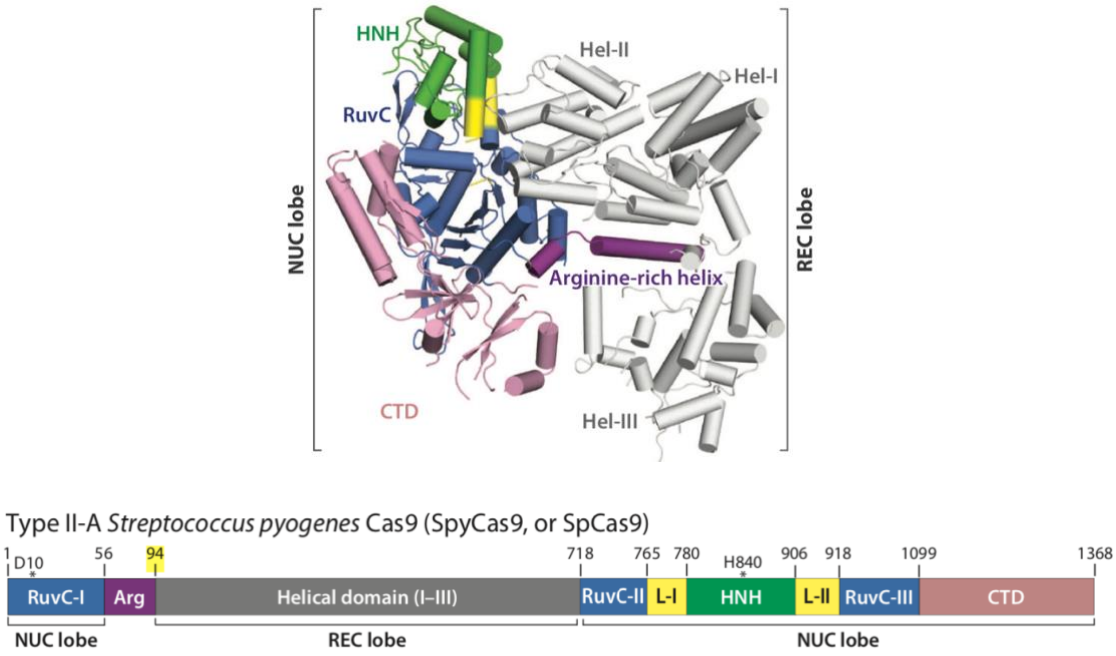


Fig. 1.6: Structure of *Streptococcus pyogenes* Cas9 from type II CRISPR.

The nuclease lobe (NUC) and alpha-helical recognition lobe (REC) are both connected through an arginine rich bridge helix structure. Together, they form the Cas9 endonuclease. The REC domain is organised in three alpha-helical domains (Hel-I, Hel-II, and Hel-III). In contrast, the NUC lobe contains both nuclease domains (HNH and RuvC), as well as the C-terminal domain (CTD). Modified from Jiang and Doudna (2017).

Cas9 (through its Hel-1, arginine-rich bridge helix, and CTD domain) mainly interacts with the sgRNA and its so-called repeat-antirepeat duplex as well as stem loop 1. Stem loops 2 and 3 of the sgRNA contribute towards stability and increasing catalytic efficiency (Jiang and Doudna, 2017).

1.3.3 Targeted DNA cleavage through Cas9

The RNA-guided Cas9 endonuclease recognises a PAM through its CTD domain, which creates contact through Watson Crick base pairing with NGG nucleotides. This includes forming hydrogen bonds with the deoxyribose-phosphate backbone of the PAM containing non-target nucleotides. This quick binding to the PAM enables Cas9 to screen adjacent DNA for potential sequence complementarity, but also destabilises adjacent DNA through structural changes (Anders et al., 2014). A so-called phosphate lock loop, located in the Cas9-CTD, stabilises the target strand immediately upstream of the PAM with a +1 phosphate group rotation. Additionally it facilitates the transition of the target strand DNA from pairing with the non-target strand, to pairing with the guide RNA through readily flipping the first nucleobase of the target DNA sequence upwards to the guide RNA (Anders et al., 2014, Jiang et al., 2016). This process is initiated by local DNA melting (which itself is initiated by the phosphate group rotation) at the target site, and is known as the so-called R-loop formation. The R-loop is a three stranded nucleic acid structure that is comprised of an RNA-DNA hybrid and a displaced, non-template single-stranded DNA strand (Szczelkun et al., 2014, Sternberg et al., 2014). In short, PAM-Cas9 initiated structural changes of the target DNA, and its subsequent stabilisation through phosphate lock loop, facilitates Watson-Crick base pairing between the sgRNA and target DNA strand (Anders et al., 2014, Jiang et al., 2016).

Upon sgRNA and target DNA complementarity, Cas9 cleaves each DNA strand by its RuvC and HNH domains to generate a DSB at a position three base pairs upstream of the 3'-edge protospacer, measuring from the PAM (Jinek et al., 2012) (Fig. 1.7). The RuvC

domain cleaves the DNA strand opposite the complementary strand (non-target strand) using a two-metal-ion catalytic mechanism. In contrast, the HNH domain is responsible for cleaving the DNA strand that is complementary to the guide RNA sequence (target strand) using a one-metal ion mechanism (Jiang and Doudna, 2017, Yang, 2008).

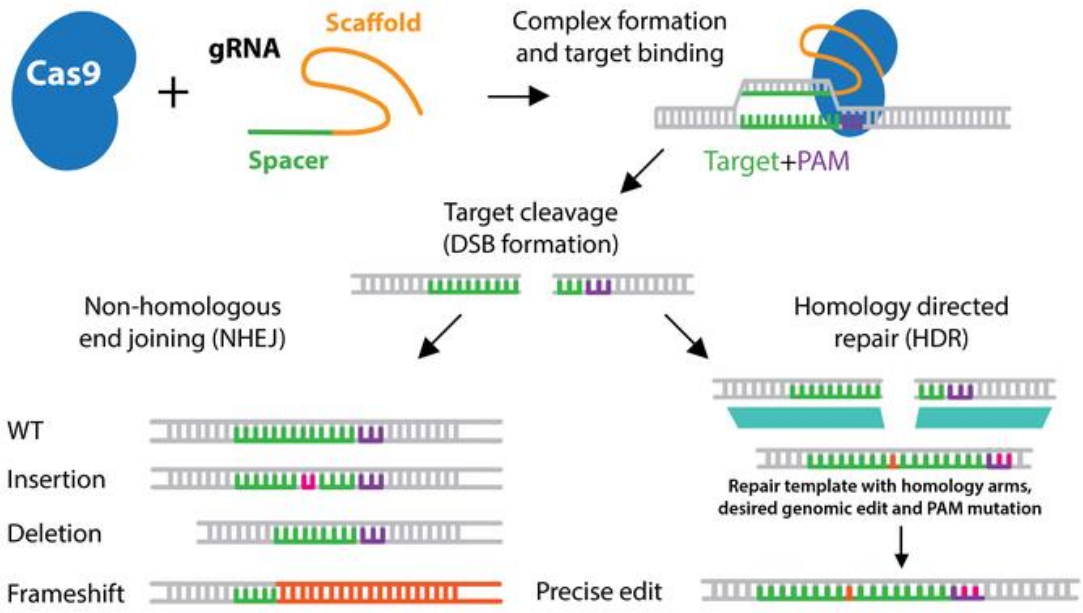


Fig. 1.7: Overview of the CRISPR-Cas9 mechanism.

gRNA-mediated Cas9 cleavage leads to a DNA double-strand break (DSB) 3 bp upstream of the PAM. In response, cells typically employ two mechanisms to repair the break; non-homologous end joining (NHEJ) (which is the primary way to repair the DSB) leads to indel (insertion and deletion) formation at the cleavage site. The same site can be repaired in the presence of a homologous DNA template through homology directed repair (HDR). Adapted from Addgene (https://addgene.org/guides/crispr).

1.3.4 DNA double-strand break repair through NHEJ and HDR

In response to DNA double-strand breaks (DSBs), cells typically utilise two main repair mechanisms: non-homologous end joining (NHEJ) and homology-directed repair (HDR). NHEJ causes random insertions and deletions (indels) at the site surrounding the cleavage, whereas the same site can be repaired in the presence of a homologous DNA template through HDR (Ran et al., 2013b) (Fig. 1.7).

NHEJ is classified into canonical NHEJ (c-NHEJ) and alternative NHEJ (alt-NHEJ or microhomology-mediated end-joining (MMEJ)) pathway. The more dominant, c-NHEJ pathway, is initiated by assembly of the core complex, which comprises of numerous proteins (Fig. 1.8). A main component of this core complex, the Ku heterodimer (Ku80/70) recognises broken, blunt ends of the DNA and holds them together for further processing (Walker et al., 2001). Ku further recruits and forms a complex with the DNA-dependent protein kinase catalytic subunit (DNA-PKcs), which as a result becomes active (Liu et al., 2018). In response, DNA-PK phosphorylates itself and a number of DNA damage response proteins (Neal et al., 2014), thus recruiting Artemis nuclease (Moshous et al., 2001). Artemis possesses 5'-3' single-strand DNA exonuclease and endonuclease (DNA-PKcs dependent) activity (Moshous et al., 2001). Gaps at the DSB ends are filled with nucleotides by Pol X family polymerases Pol4 and others (Liu et al., 2018). Re-joining of the DNA ends occurs through various ligases, such as ligase IV/XRCC4/XLF (Tsai et al., 2007).

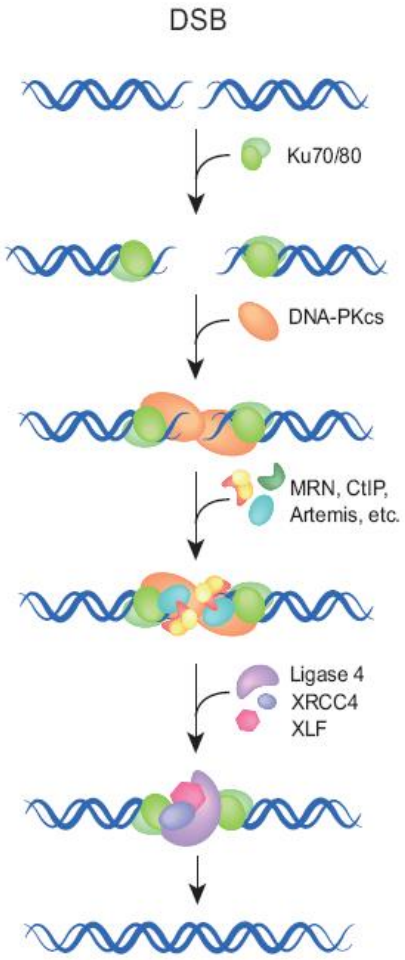


Fig. 1.8: DNA double-strand breaks are predominantly repaired by the c-NHEJ pathway.

In response to DNA double-strand break (DSB) a number of proteins are involved in the repair through the canonical non-homologous end joining (cNHEJ) pathway. The cNHEJ pathway is the most dominant way to repair the break. It recruits proteins that are involved in break recognition, complex formation, nucleotide filling, and ligation. Adapted from Dueva and Iliakis (2013).

Alt-NHEJ has been reported to act as a backup repair pathway, which is mainly involved in chromosomal translocations in mammalian cells. The main proteins contributing towards Alt-NHEJ are PARP1, the MRN complex, and its partner CtiP (Simsek and Jasin, 2010). While alt-NHEJ is predominantly active during S phase and G2 phase of the cell cycle, c-NHEJ operates throughout the cell cycle (Truong et al., 2013).

In contrast, the HDR pathway is only active during late S phase and G2 phase of the cell cycle in the presence of a DNA template with homology sequences (Truong et al., 2013). Upon DSB, a cascade involving several proteins is initiated: Histone variant H2AX is phosphorylated throughout the area surrounding the DSB and sets off a signal cascade, recruiting further proteins (Rogakou et al., 1998). MDC1 acts as a large scaffold and interacts with ATM and NBS1. At the same time, the MRN (MRE11-RAD50-NBS1) complex keeps the broken ends together and forms a nick 15-20 bp from the 5' end of the DSB to resect DNA (Lamarche et al., 2010). After the DNA DSB ends have been coated with a number of these proteins, one end starts searching for homologous DNA sequences by creating a displacement loop (D-loop) structure (Kim and Mirkin, 2018). Homologous template DNA can be incorporated through either the Holliday Junction (HJ) pathway or the synthesis-dependant strand annealing (SDSA) pathway.

D-loop structures usually undergo extension through DNA synthesis; if the second end begins extension as well, a characteristic double Holliday junction is created through ligation. SLX1 and MUS81-EME1 nucleases or BTR complex (BLM helicase-Topoisomerase IIIα-RMI1-RMI2) cleave these double Holliday junctions into either crossover or non-crossover products (Liu et al., 2018).

The SDSA pathway also originates from a D-loop structure and 3'-extension through DNA synthesis (Daley et al., 2014). This translocating D-loop is unstable, collapses, and as a result may anneal with the resected, complementary ssDNA from the other end. Subsequently, both ends undergo replicative extension and ligation. This process creates non-crossover products (Liu et al., 2018).

To potentially treat genetic diseases, it is important to efficiently and precisely correct the mutation in the target locus, rather than cause stochastic disruption of the gene. However, the efficiency of HDR is relatively low. To address this challenge, many studies (for example Komor et al. (2016) and Richardson et al. (2016)) have been undertaken to increase the efficiency of HDR.

1.3.5 Enhancing the CRISPR portfolio

Exploring different CRISPR systems has given rise to new ways that make CRISPR more effective and versatile. Researchers have characterised more than 10 different nucleases which have been repurposed for genome editing (Adli, 2018). The most prominent nucleases alongside Cas9 are Cpf1 (also known as Cas12a) (Class II type V) (Zetsche et al., 2015), and C2C2 (Class II type VI) (Abudayyeh et al., 2016). Cpf1 is smaller in size (1307 amino acids), targets DNA molecules with a single crRNA alone without the need for a coupled tracrRNA, cleaves DNA via a staggered DNA DSB, and targets different PAM sites. On the other hand, C2C2 (928 amino acids) is a programmable RNA targeting nuclease, finding use in degradation of transcripts and in translational inhibition.

Notably, point mutations in both Cas9 domains, HNH (H840A) and RuvC (D10A), result in a catalytically inactive, so-called dead Cas9 (dCas9)(Qi et al., 2013). This dCas9 loses its capability to cleave the DNA, but can still be guided to the target sequence. Here it can efficiently suppress the expression of targeted genes by creating strong DNA

recognition complexes, which can obstruct RNA polymerase/transcription factor binding or translational elongation (Qi et al., 2013). This method is also known as CRISPR interference (CRISPRi). The dCas9 can also be utilised to induce robust expression by fusing strong transcriptional activators to itself (Maeder et al., 2013, Perez-Pinera et al., 2013). A third way to exploit dCas9 is through live cell chromatin imaging by fluorescently labelling the dCas9 (Chen et al., 2013).

Another major milestone in evolving CRISPR technology has been the development of base editing technologies without causing a DNA DSB. Instead, by exploiting a nickase Cas9 (nCas9) that only generates a targeted single-strand DNA break (due to a point mutation at one of its domains), the precise and direct conversion of Cytosine (C) to Thymine (T) (Komor et al., 2016) or Adenine (A) to Guanin (G) (Gaudelli et al., 2017) can be achieved. A nCas9 fused to an APOBEC1 deaminase enzyme and uracil glycosylase inhibitor (UGI) protein successfully converts C into T, while fusion of a transfer RNA adenosine deaminase to a nickase Cas9 results in A-G conversion at the target site (Gaudelli et al., 2017, Komor et al., 2016). C-T conversions have been exploited to amend the genetic code by introducing premature STOP codons in target genes in a controlled and less deleterious way compared with DSB brakes (Kuscu et al., 2017, Billon et al., 2017).

Possibly the most important improvements to CRISPR technology in the last few years is a further advancement of CRISPR-Cas9, known as prime editing (Anzalone et al., 2019). Prime editing enables precise gene editing not only without the need to generate a DNA DBS, but also without the requirement for donor DNA. It directly writes new genetic information at a target site by relying on a three-component system: an engineered

reverse transcriptase, programmed with a prime editing guide RNA (pegRNA), which is fused to a dsCas9 (Anzalone et al., 2019). Importantly, the pegRNA encompasses both target specificity and the desired coding sequence edit. This technology was only published 2 months prior to thesis submission, and its potential and impact are yet to be assessed.

1.3.6 CRISPR-Cas9, a novel approach to investigate miRNA targeting

One problem that classical gain-of-function and loss-of-function experiments face is that a single miRNA can and does regulate several mRNA targets (Bartel, 2004). As a result, therapeutic approaches that manipulate miRNA levels are risky, due to probable unwanted and usually unknown effects on other targets.

Bassett et al. (2014) developed a new technique to validate functional miRNA response elements (MREs) in cultured cells (Fig. 1.9). This technique is based on the RNA-guided Cas9 nuclease from type II CRISPR. Crucially, this technique enables us to assess miRNA targeting in an endogenous context in the 3'-UTR, without perturbing miRNA levels by overexpression or inhibition. This is particularly important as all other targets are left unaffected – except that which is being investigated by gene editing.

In the first step, a putative MRE is targeted by sgRNA and Cas9 mediated genome editing in order to cause a DSB at the MRE binding site. HDR introduces a very small proportion of two roughly 140 nucleotide long single-stranded DNA (ssDNA) oligonucleotide templates, the T3 and T7 promotor sequences, which act as primer binding sites. Either

the T3 or T7 barcode gets integrated in the gDNA. The T3 barcode, which contains the sequence for an intact MRE and the barcode itself, is integrated downstream of the putative MRE site. Conversely, the T7 barcode deletes the endogenous MRE site while introducing a primer-binding site. The final read-out is attained through qPCR, which assesses the transfected barcode in an heterogenous cell population with an endogenous forward primer and universal, barcode-specific T3 or T7 primers.

A functionally active MRE will be characterised by increased levels of mRNA transcript being produced (due to the deletion of an active MRE through the T7 barcode) compared with the T7/T3 ratio from gDNA.

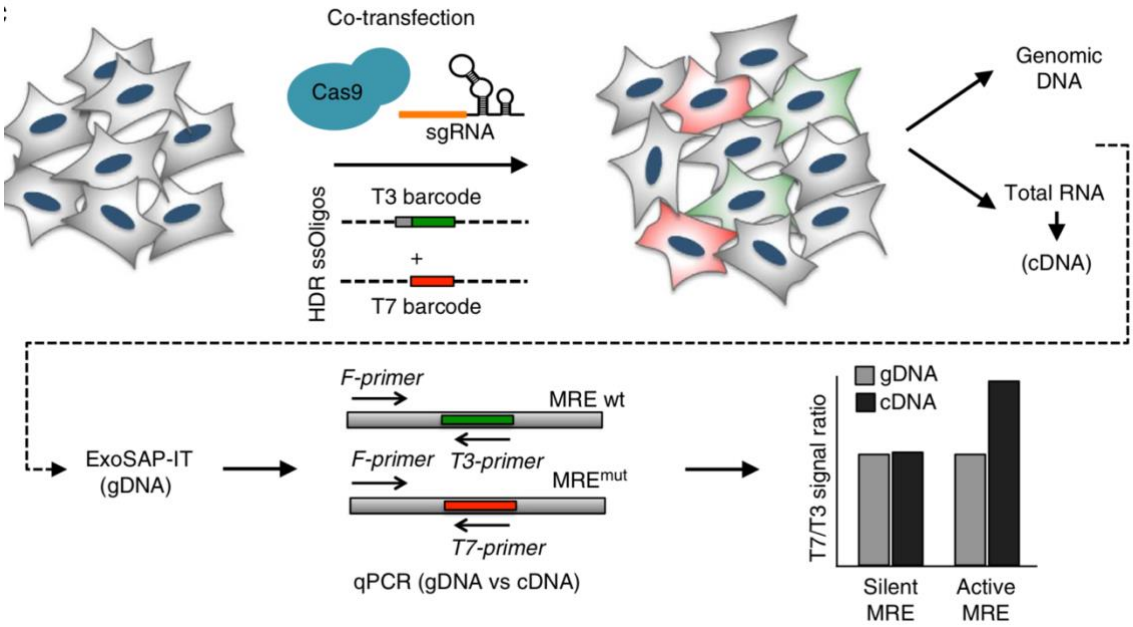


Fig. 1.9: Workflow of MRE analysis through CRISPR-mediated HDR.

The experimental outline includes pX330-mediated transfection of Cas9/sgRNA along with two barcodes, T3 and T7. gDNA and RNA are isolated from transfected cells and gDNA is ExoSAP-IT treated to get rid of redundant ssODNs. The read-out is based on qPCR with an endogenous forward primer and universal, barcode-specific, reverse primers. An increase in cDNA levels compared with gDNA levels suggests an active MRE due to increase in transcript levels, which results from barcode-mediated deletion of the MRE. Modified from Bassett et al. (2014).

1.3.7 CRISPR-Cas 9 off-targeting

Ever since its discovery, the CRISPR technology has been subject to discussion and criticism surrounding its off-target effects. Target similarity as well as enzyme concentration are amongst the two most significant factors that can determine the extent to which a given guide RNA exhibits off-target activity (Ran et al., 2013b). As a result, to minimise the likelihood of off-targeting, an ideal way to design the guide is by comparing off-target sequences in the genome with the guide sequence through computational analysis, and by reducing sequence complementary.

Moreover, off-targeting can be governed by the amounts of transfected Cas9 and sgRNA, and hence reduced by carefully titrating the amount of Cas9 and sgRNA transfected (Ran et al., 2013b). In fact, the delivery method of Cas9 and sgRNA is another determining parameter. Different studies in induced pluripotent stem cells (iPS), human embryonic kidney (HEK) 293 FT cells and K562 cells have revealed how Ribonucleoprotein (RNP) delivery of Cas9 and sgRNA can significantly decrease off-targeting compared with plasmid-mediated delivery approaches of Cas9 and sgRNA into the cell (Kim et al., 2014, Zuris et al., 2015, Liang et al., 2015). The RNP method produces rapid but more transient levels of Cas9, compared with a plasmid, where Cas9/sgRNA are expressed by the cell through transcription and translation following uptake in the host cell. Over the years, many attempts have been made to increase Cas9 specificity while reducing the off-targeting potential. For example, Ran et al. (2013a) demonstrated that by using a mutant Cas9 nickase (nCas9), off-targeting could be significantly reduced. This nCas9 only generates single strand breaks which are repaired with high fidelity and

without indel formation. Simultaneous nicking through a pair of guide RNAs is required to generate a DSB, hence this approach effectively doubles the number of specifically recognised bases at the target site, and therefore increases the genome editing specificity (Ran et al., 2013a).

Further studies to increase Cas9 specificity have involved specific Cas9 point mutations (Kleinstiver et al., 2016), the use of small chemicals such as doxycycline (Dow et al., 2015), optical light (Nihongaki et al., 2015, Hemphill et al., 2015), as well as temporal and spatial regulation of the Cas9/sgRNA complex through ligands (Oakes et al., 2016). Finally, means to modify the sgRNA scaffold to increase targeting specificity have also been undertaken (Fu et al., 2014, Adli, 2018).

Material & Methods

2

Chapter 2. Material & Methods

2.1 Isolation of osteoarthritic human articular chondrocytes

Osteoarthritic human articular chondrocytes were isolated from tissue which was obtained from patients undergoing unicompartmental knee arthroplasty (UKA), or total knee replacement (TKR). These patients were diagnosed with advanced or end stage OA and consented to the Oxford Musculoskeletal Biobank (OMB) following local ethics committee guidelines. Tissue samples and/or data were obtained from the Oxford Musculoskeletal Biobank and were collected with informed donor consent in full compliance with national and institutional ethical requirements, the UK Human Tissue Act, and the Declaration of Helsinki (HTA Licence 12217 and Oxford REC C 09/H0606/11). In a first step, cartilage was briefly washed with PBS and then dissected from the tibial plateau and femoral condyle of the osteoarthritic knee joint. Subsequently the cartilage was cut into 1-2 mm small rectangles, from which human articular chondrocytes were isolated by enzymatic digestion in Dulbecco's modified eagle's medium (DMEM) containing contained collagenase A (2 mg / ml) (Merck, Germany), 1 % penicillin / streptomycin as well as 20 ng / ml Fungizone and incubated between 18 h - 22 h at 37 °C with constant shaking (80 rpm).

Following incubation, the digested material was passed through a 70 µm nylon mesh cell strainer to remove any undigested tissue material from the cell suspension.

To neutralise and wash away the collagenase, the digested and strained suspension was placed in an equal volume of DMEM containing 10 % fetal bovine serum (FBS) and 1 % penicillin / streptomycin, then centrifuged twice at 300 g for 5 minutes with the supernatant being removed and cells resuspended in fresh medium each time. Cells were then quantified and cell viability was assessed by Trypan blue exclusion. Finally, cells were seeded at roughly 25.000 cells / cm^2 in a 10 cm culture treated dishes and incubated at 37 °C, 20 % O_2 and 5 % CO_2.

2.1.1 Donor details

Donors (low gene edit)	**Sex**	**Age**
Donor 1	male	62
Donor 2	male	64
Donor 3	male	57
Donor 4	male	71
Donor 5	female	70
Donor 6	male	54
Donors (high gene edit)		
Donor 1	female	63
Donor 2	female	70
Donor 3	male	78
Donor 4	male	61
Donor 5	female	61

Donor 6	male	69
Donors (unedited)		
Donor 1	Female	72
Donor 2	Female	61
Donor 3	Male	78
Donor 4	Male	78
Donor 5	Male	61
Donor 6	Male	69

2.2 Cell culture

2.2.1 Osteoarthritic human articular chondrocytes

Osteoarthritic human articular chondrocytes (OA hACs) were cultured at 37 °C, 20 % CO_2 and 5 % CO_2 in Dulbecco's modified Eagle's medium (DMEM) containing 4.5 g/l of glucose and L-Glutamine (Lonza, Verviers, Belgium) and were supplemented with 10 % fetal bovine serum (FBS) (Gibco, NY, USA) and 1 % Penicillin and Streptomycin (Gibco, NY, USA).

2.2.2 HEK293T cells

HEK293T cells were cultured at 37 °C, 20 % CO_2 and 5 % CO_2 in Dulbecco's modified Eagle's medium (DMEM) containing 4.5 g/l of glucose and L-Glutamine (Lonza, Verviers,

Belgium) and were supplemented with 10 % FBS (Gibco, NY, USA) and 1 % Penicillin and Streptomycin (Gibco, NY, USA).

2.2.3 Human mesenchymal stem cells

Human mesenchymal stem cells (hMSCs) were purchased from Lonza (donor details: 26 years old female) and cultured at 37 °C, 20 % CO_2 and 5 % CO_2 in Minimum Essential Medium (MEM Alpha) containing 4.5 g/l of glucose and L-glutamine (Lonza, Verviers, Belgium) and were supplemented with 20 % FBS performance plus (Gibco, NY, USA) and 1 % Penicillin and Streptomycin (Gibco, NY, USA). Performance plus FBS was additionally sterile filtered through 0.2 µm filters. For performing the scratch assay, hMSCs (24 years old male) were obtained from Prof. Oppermann's group (Botnar Research Centre, University of Oxford).

2.3 sgRNA design

2.3.1 for plasmid-based (pX330) transfections

Oligonucleotides for targeting C9orf7, including primers (Table 2.9), sgRNA (Table 2.1) and T3 and T7 barcodes (Table 2.10), were obtained from the original publication (Bassett et al., 2014) while *SOX9* primers and sgRNAs were designed to target putative microRNA response elements (MREs), which were predicted to be in the 3'-UTR of *SOX9* and *C9orf7* respectively, close to a PAM sequence (NGG).

Target sequences were subcloned into the pX330 vector (obtained from Addgene (plasmid ID: 48137), Watertown, MA, USA; initially constructed by the Zhang lab, Harvard University) which contains the open reading frame (ORF) for the Cas9 protein and restriction sites to subclone sgRNAs. In order to clone the target sgRNA into the linearized pX330 (digested with BbsI) backbone, the sgRNA was annealed and subsequently ligated with the backbone and transformed into competent E. coli TOP10 cells. 10 colonies were picked and isolated for gDNA using the ZymoPURETM Plasmid Midiprep Kit (Zymo Research, Irvine, CA) following the manufacturer's recommendation. Correct sgRNA insertion was verified through sequencing.

Table 2.1: Designed sgRNA and their target sequences for C9orf7 and SOX9 for plasmid-based transfections.
Seed sequences are bold, single guide RNA (sgRNA) target sequences highlighted blue, and PAM sequences underlined.

Gene	Target sequence 5'-3'	sgRNA 5'-3'
C9orf7 (Bassett et al., 2014)	TCTGCGTCTGAGGCTGGGAGTGGCATCTG**A GGCCGGGA**GTGGCATCTGAGGCCAGGAGT GGCAGGCTGGTG	GTGGCATCTGAGGCCGGGAG
SOX9 MRE1	GGCCAGGCCAACCTTGGCTAAATGGAGCAG CGAAATCAACGAGA**AACTGGA**CTTTTTAAAC CCTCTTCAGAGCAAGCGTGGAGGATGATGG AGAATCGTG	CTCCACGCTTGCTCTGAAGA

2.3.2 Ribonucleoprotein (RNP)-based transfections

For gene editing, the ALT-R® CRISPR-Cas9 system from Integrated DNA Technologies (IDT, Coralville, IA, USA) was applied. The Cas9 protein, tracrRNA and CRISPR RNA (crRNA) were all acquired from the same company. To form a functional sgRNA duplex, 3 µl of tracrRNA (5 nmol) were mixed with 3 µl of target-specific crRNA (2 nmol) in 94 µl IDT nuclease free duplex buffer and quantified. While the universal tracrRNA forms the backbone, the crRNA is custom designed and target-specific. The required amount was then incubated for 5 mins at 95 °C and slowly annealed at room temperature for 10 mins. Subsequently the final single guide RNA (sgRNA) (Table 2.2) was complexed with Cas9 for transfection.

Table 2.2: Designed sgRNA and their target sequences for *miR-140* and *SOX9* for RNP transfections.
miR-140 hairpin region is highlighted in yellow. sgRNA target sequences in blue and PAM sequences underlined. *SOX9* seed sequence is bold.

Gene	Target sequence 5‘-3‘	sgRNA 5‘-3‘
WWP2 miR-140 L1	GCCCTGTGTGTGTCTCTCTCTGTGTCCTGCCA GTGGTTTTACCCTATGGTAGGTTACGTCATG CTGTTCTACCACAGGGTAGAACCACGGACAG GATACCGGGGCACCCTCTGCGTCGACGGACT CCTCGTCTGCCCAGCCACAAACAGC	CTCTCTCTGTGTCCTGCCAG
WWP2 miR-140 L2	GCCCTGTGTGTGTCTCTCTCTGTGTCCTGCCA GTGGTTTTACCCTATGGTAGGTTACGTCATG CTGTTCTACCACAGGGTAGAACCACGGACAG GATACCGGGGCACCCTCTGCGTCGACGGACT CCTCGTCTGCCCAGCCACAAACAGC	CTGCCAGTGGTTTTACCCTA

WWP2 miR-140 L3	GCCCTGTGTGTGTCTCTCTCTGTGTCCTGCCA GTGGTTTTACCCTATGGTAGGTTACGTCATG CTGTTCTACCACAGGGTAGAACCACGGACAG GATACCGGGGCACCCTCTGCGTCGACGGACT CCTCGTCTGCCCAGCCACAAACAGC	ACGTCATGCTGTTCTACCAC
WWP2 miR-140 L4	GCCCTGTGTGTGTCTCTCTCTGTGTCCTGCCA GTGGTTTTACCCTATGGTAGGTTACGTCATG CTGTTCTACCACAGGGTAGAACCACGGACAG GATACCGGGGCACCCTCTGCGTCGACGGACT CCTCGTCTGCCCAGCCACAAACAGC	CGAGGAGTCCGTCGACGCAG
SOX9 MRE1	GGCCAGGCCAACCTTGGCTAAATGGAGCAG CGAAATCAACGAGA**AACTGGA**CTTTTTAAAC CCTCTTCAGAGCAAGCGTGGAGGATGATGG AGAATCGTG	CTCCACGCTTGCTCTGAAGA

2.4 Transfections

2.4.1 HEK293T

The Cas9/sgRNA expression vector (pX330, 1 µg), along with the homology oligonucleotides (0.5 µg each of T3/T7) (Table 2.10), was co-transfected into 40 % confluent HEK293T cells in 12-well plates using 3 µl Lipofectamine 2000 (Thermo Fisher Scientific, Northumberland, UK) following the manufacturer's protocol.

2.4.2 OA human articular chondrocytes

OA hACs were seeded into 12 well plates at a density of roughly 20.000 cells / cm^2 and transfected 48 h later at a confluency between 40 % - 60 %. Prior to transfection cells were cultured for 1.5 h in antibiotic free DMEM with 10 % FBS.

To carry out RNP transfections, single guide RNA (sgRNA) molecules were formed, by annealing crRNA with tracrRNA (see 2.3.2). These single guide RNAs (sgRNAs) were complexed with Cas9 for 5 mins in reduced serum medium (Opti-MEM, Gibco, NY, USA). Cas9 Plus reagent was added before incubation. Cas9 Plus is part of Lipofectamine CRISPRMAX (Invitrogen, CA, USA). In another tube, Lipofectamine CRISPRMAX was added to Opti-MEM, and incubated for 5 mins. The Opti-MEM media containing the sgRNA/Cas9 complex was carefully added to the Opti-MEM containing Lipofectamine CRISPRMAX and incubated for another 10 mins before slowly pipetting the Cas9/sgRNA/Lipofectamine into the cell media. Quantities were used according to Table 2.3. RNP transfections involving HDR-mediated barcode optimisation (Chapter 5) included 0.5 µg of each ssODN (Table 2.10) in tube 1.

Table 2.3: Table detailing the amounts of reagents used for RNP transfection (in pipetting order) for osteoarthritic, human articular chondrocytes (OA hACs) and human mesenchymal stem cells (hMSCs).

Tube 1	**OA hACs**	**hMSCs**
Opti-MEM reduced serum media	50 µl	50 µl
Alt-R S.p. Cas9 Nuclease V3	1 µg (diluted in Opti-MEM)	1.1 µg (diluted in Opti-MEM)

Alt-R sgRNA (tracrRNA + crRNA)	400 ng	450 ng
Cas9 PLUS reagent	2 µl	2.5 µl
	Mix carefully by pipetting up and down. Incubate tube 1 for 5 mins.	Mix carefully by pipetting up and down. Incubate tube 1 for 5 mins.
Tube 2		
Opti-MEM reduced serum media	50 µl	50 µl
CRISPRMAX Lipofectamine	3.5 µl	4 µl
	Mix by vortexing. Incubate tube 2 for 5 mins.	Mix by vortexing. Incubate tube 2 for 5 mins.
	Add tube 1 into tube 2. Incubate for 10 mins.	Add tube 1 into tube 2. Incubate for 10 mins.

The Alt-R S.p. Cas9 Nuclease V3 aliquot requires replacement every ~4 - 5 weeks as its cleavage activity decreases over time.

2.4.3 Human mesenchymal stem cells

Human mesenchymal stem cells (hMSCs) were transfected exactly in the same way as osteoarthritic human articular chondrocytes (OA hACs) but with slightly increased concentrations (Table 2.3). Prior to transfection, hMSCs were cultured for 1.5 h in 20 % FBS containing DMEM medium without antibiotics.

hMSCs were either removed for analysis at 48 h, or medium was exchanged for 20 % / DMEM for 24 h and cells were then re-transfected as before. Cells were incubated for a further 48 h at 37 °C, 5 % CO_2 and subsequently harvested for analysis.

2.5 Genomic DNA (gDNA) and RNA extraction low-throughput

Genomic DNA (gDNA) was extracted from cells by addition of 300 µl lysis buffer (3 M NaCl, 1 M Tris-HCL pH 7.4, Igepal CA-630 (Sigma-Aldrich, Dorset, UK), 0.5 M EDTA, 1x Protein Inhibitor Complex (Roche, Basel, Switzerland)) for one well in a 12 well plate. Cell lysate was spun down at 3,000 g for 5 mins. Genomic DNA was isolated from the cell pellet using the Wizard® SV Genomic DNA Purification System as per the manufacturer's instructions, quantified and assessed for purity using a digital spectrophotometer (Nanodrop 1000, ThermoFisher, MA, USA), and Exo-SAP-IT (Affymetrix, Santa Clara, USA) treated to remove any contaminating donor oligonucleotides for downstream applications as per the manufacturer's instructions.

RNA was isolated from cells using the supernatant with TRIzol® reagent according to the manufacturer's recommendations. RNA was resuspended in 30 µl of RNase free water and quantified using a digital spectrophotometer (Nanodrop 1000, ThermoFisher, MA, USA). 1 µg of RNA in a total volume of 10 µl RNase free water was treated with Turbo™ DNase (Ambion, ThermoFisher, MA, USA) to remove any contaminating DNA from the

sample according to the manufacturer's guidelines. RNA samples were also run on a 1 % agarose gel in order to test for integrity.

2.6 Genomic DNA (gDNA) and RNA extraction high-throughput

For simultaneous purification of DNA and RNA from the same biological sample, the AllPrep DNA/RNA mini kit from Qiagen (Hilden, Germany) was used. Samples were purified according to the manufacturer's instructions. RNA was on-column treated with DNase for 15 minutes. Genomic DNA and RNA samples were eluted in each 20 µl nuclease-free water.

2.7 T7 Endonuclease 1 (T7E1) assay to verify gene editing

Prior to determining the gene editing efficiency with T7E1, the region of interest in gDNA or cDNA from targeted cells was amplified via polymerase chain reaction (PCR) using a high-fidelity DNA Polymerase (Q5® Hot Start High-Fidelity Polymerase; New England Biolabs, Ipswich, MA, USA) and gene-specific primers (Table 2.7 and Table 2.8) following the manufacturer's guidelines. In brief, between 10 – 15 ng gDNA or cDNA template, 0.4 µl forward and reverse primers (10 nM) and 5 µl of 2xQ5® Hot Start High-Fidelity

polymerase were used and the volume made up to 10 µl per reaction with nuclease-free water. Samples were amplified according to Table 2.4 using a C1000 touch thermocycler.

Table 2.4: PCR programme used to amplify genomic regions of interest prior to T7E1 assay

PCR protocol			
Initial denaturing	2 min	98°C	
Denaturing	10 s	98°C	
Annealing	20 s	Primer dependent	40 cycles
Elongation	35 s	72°C	
Final elongation	2 min	72°C	

The T7 Endonuclease 1 (New England Biolabs, Ipswich, USA) recognises and cleaves non-perfectly matched DNA such as heteroduplexes and nicked DNA. In a first step, DNA is denatured at 95 °C for 5 mins, then allowed to reanneal by slowly cooling it down (95 - 85 °C: -2 °C/second; 85 - 25 °C: -0.1 °C/second). This allows heteroduplex formation between wild-type DNA and CRISPR-Cas9-mutated DNA. 2 units of T7E1, which recognizes and cleaves mismatched DNA, was added in a final step to digest heteroduplexes for 45 mins at 37 °C.

2.8 Agarose gel electrophoresis

1.5 % – 2.5 % agarose gels were poured by mixing agarose powder in 1x tris-acetate-EDTA (TAE) buffer. Agarose was completely dissolved through boiling. 10,000x SYBR Safe DNA Gel Stain (Invitrogen, CA, USA) was added to a final concentration of 1x. SYBR Safe DNA Gel Stain binds to the DNA and visualises DNA under ultraviolet (UV) light.

Once solidified, the agarose gel was covered in 1x TAE buffer and PCR samples were loaded with 6x Gel Loading Dye, purple (New England Biolabs, Ipswich, USA), to a final concentration of 1x. Agarose gels were run at 100 – 180 V (depending on size of chamber) for 20 - 30 mins. The resulting cleaved and full-length PCR products were visualized under UV light. The 50 bp DNA ladder (New England Biolabs, Ipswich, USA) acted as molecular weight ladder.

2.9 Verification of successful integration of ssODN barcodes

Upon successful DNA cleavage, maintaining ssODNs (T3, M1, M2), which contain 60 nt homology arms flanking a miRNA response element (MRE), introduced a primer-binding site downstream of the putative MRE in HEK293T cells or OA hACs and maintained the MRE. The deleting ssODNs (T7, D1, D2) contained the same 60 nt homology arms but deleted a miRNA seed sequence while also introducing a primer-binding site. Cells were collected 48 h post transfection for extraction of DNA and RNA.

Genomic DNA (gDNA) and complementary DNA (cDNA) of targeted cells were checked for successful integration of both the maintaining oligonucleotide as well as the deleting oligonucleotides. In HEK293T cells, verification happened with an endogenous forward primer and barcode specific T3 and T7 reverse primers using a SYBR Green JumpStart Taq DNA Polymerase mix (Sigma-Aldrich, Dorset, UK). Additionally, in OA hACs, barcode verification was attempted with New England Biolabs Q5 Hot Start High-Fidelity 2x MM (Ipswich, MA, USA), Bioline SensiFAST™ Real-Time PCR kit (Boston, MA, USA) and Integrated DNA Technologies PrimeTime Gene Expression® Master Mix (Coralville, IA, USA).

In brief, 10 - 15 ng gDNA or cDNA template, 0.4 µl forward and reverse primers (80 - 400 nM), optionally 0.5 µl DMSO and 5 µl of Polymerase mix were used and filled up with nuclease-free water to 10 µl per reaction. Samples were amplified according to Table 2.5 using a C1000 touch thermocycler.

Samples were visualised by loading with gel loading dye (final concentration: 1x) and separation on a 1.5 % – 2.5 % agarose gel electrophoresis in 1x TAE buffer.

Table 2.5: PCR protocol for verifying barcode integrations in gDNA and cDNA of CRISPR-Cas9 targeted cells.

PCR protocol			
Initial denaturing	2 min	98°C	
Denaturing	10 s	98°C	
Annealing	20 s	Primer dependent	40 cycles
Elongation	35 s	72°C	
Final elongation	2 min	72°C	

2.10 Sequencing

The region of interest was PCR amplified with the help of the Q5® Hot Start High-Fidelity Polymerase (New England Biolabs, Ipswich, MA, USA) which generates blunt ends. The PCR product was transformed (with the help of the Zero Blunt TOPO PCR cloning kit, Invitrogen, Ipswich, MA, USA) into competent E. coli as per the manufacturer's instructions and plated on lysogeny broth (LB) plates with kanamycin selection marker. 16 h post plating, 15 clones were picked and inoculated in LB + kanamycin media overnight. Cells were mini-prepped with the help of the QIAprep Spin Miniprep kit (Qiagen, Hilden, Germany) as per the manufacturer's instructions, and sent off for sequencing to Eurofins Genomics in Germany.

MiSeq sequencing and analysis were carried out in collaboration with Dr. Damien Downes, MRC Weatherall Institute of Molecular Medicine, University of Oxford. The approach consisted of two PCR reactions. I carried out the first PCR reaction using modified, locus specific primers (Fw 5' ACACTCTTTCCCTACACGACGCTCTTCCGATCT-Target primer sequence (top strand) Rv: 5' GACTGGAGTTCAGACGTGTGCTCTTCCGATCT-Target primer sequence (bottom strand)) in standard conditions (Table 2.4) using the Q5® Hot Start High-Fidelity Polymerase (New England Biolabs, Ipswich, MA, USA). I then completed a second PCR amplification with Illumina sequencing adapters as per the manufacturer's instructions (NEBNext Multiplex Oligos for Illumina Index Primers, New England Biolabs, Ipswich, MA, USA). After the second PCR a DNA clean-up with AMPure XP beads (Beckman Coulter, Brea, CA, USA) was carried out by Dr. Downes. Sequencing was completed using the MiSeq Reagent Kit V2 (San Diego, CA, USA) on an Illumina

MiSeq Next Generation Sequencer (San Diego, CA, USA). Analysis was carried out by CRISPResso (Pinello et al., 2016) by Dr. Downes.

A more detailed workflow can be found from “Illumina: And introduction to Next-Generation Sequencing Technology”, which is available in the appendix (8.1).

2.11 Reverse Transcription Polymerase Chain Reaction (RT-PCR)

2.11.1 Gene quantifications

Complementary DNA (cDNA) was synthesized from the DNase treated RNA using a high capacity reverse transcription kit (Applied Biosystems, Massachusetts, USA) following the manufacturer’s guidelines (in a total of a 20 µl mixture, 100 ng – 500 ng RNA were added to 13.2 µl RNase free water. A 1x master mix solution contained 2 µl of 10x RT buffer, 0.8 µl 10 mM deoxyribonucleotides (dNTPs), 2 µl of random primer, 1 µl RNase inhibitor (20 U) and 1 µl reverse transcriptase (50 U) per reaction) using a C1000 touch thermocycler (BioRad, Hertfordshire, UK). A programme of 25 °C for 10 mins, 37 °C for 120 mins and finally 85 °C for 5 mins was used for reverse transcribing RNA into cDNA. cDNA was diluted in RNAse free water according to their initial concentration (100 ng 1:5, 500 ng 1:25).

2.11.2 microRNA quantification

Stem-loop structure primers with specific sequences were used for microRNA RT-PCR (Table 2.6). Prior to reverse transcription, the mature miRNA of choice is extended through miRNA-specific primers on each end through 5' ligation and 3' poly-A tailing of an adaptor sequence. These universal sequences on both end of the extended miRNA are then recognised by universal RT primers. Complementary DNA (cDNA) was synthesized from the DNase-treated RNA using a high capacity reverse transcription kit (Applied Biosystems, Massachusetts, USA). 15 ng RNA were added to a total of a 15 µl 1x master solution containing 1.5 µl of 10x RT buffer, 0.15 µl 10 mM deoxyribonucleotides (dNTPs), 0.19 µl RNase inhibitor (20 U), 1 µl reverse transcriptase (50 U), 1.5 µl 5x miRNA-specific RT primer and 1.5 µl 5x miRNA-specific housekeeper primer (either RNU24 or 44), and 9.16 µl RNase free water per reaction using a C1000 touch thermocycler (BioRad, Hertfordshire, UK). A programme of 16 °C for 30 mins, 42 °C for 30 min and finally 85 °C for 5 min was used for reverse transcribing RNA into cDNA.

2.12 Real-Time quantitative Polymerase Chain Reaction (RT-qPCR)

2.12.1 Barcode verification in targeted HEK293T cells

A total of 10 µl of reaction mixture containing 2 µl cDNA template, 0.4 µl of each forward and reverse primer (Table 2.9), 0.1 µl of ROX stain, 5 µl of SYBR Green JumpStart Taq DNA Polymerase (2x mix containing buffer and dNTPs) and 2.1 µl of nuclease-free water was added to each well of a 384 well plate. Each reaction was run in a triplet. The plate was sealed and centrifuged for 1 min at 1,500 g. Real Time-qPCR was carried out on a ViiA™ 7 Real-Time PCR System (Applied Biosystems). The final signal ratio between T7 and T3 primers on DNA and RNA level was obtained by calculating the difference in Ct (ΔCt method) between T7 and T3 primers during the exponential amplification phase and subsequently transforming the resulting values 2-ΔCt. The ratio of T7 signal to T3 signal in the gDNA is used to determine the relative integration efficiency of the T7 and T3 oligonucleotides. An increase towards T7 integration (which results in the deletion of the MRE) in this ratio in the cDNA compared to the gDNA suggests that the MRE is active, as its deletion results in a decrease in degradation (Bassett et al., 2014).

2.12.2 Gene TaqMan assays in osteoarthritic human articular chondrocytes and human mesenchymal stem cells

A total of 10 µl of reaction mixture containing 1 µl cDNA template, 0.5 µl of 20x TaqMan® Gene Expression Assay (Table 2.6), 5 µl 2x TaqMan universal PCR master mix (Applied Biosystems, Foster City, CA, USA) and 3.5 µl of nuclease-free water was added to each well of a 384 well plate. Each reaction was run in triplicate. The plate was sealed and centrifuged for 1 min at 1500 g. Real Time-qPCR was carried out on a ViiA™ 7 Real-Time PCR System (Applied Biosystems) under the following conditions; 1X stage 1: 95 °C for 10 mins; 40X stage 2: 95 °C for 15 s; 60 °C for 1 min. Relative gene expression was calculated using the ΔΔCt method, and depending on the cell type, 18S, RPLP0 or β2M were used as the control gene. For expression analysis, it was ensured that these control genes were stably expressed and did not show any changes under the experimental conditions or disease state. Generally, 18S was used in HEK293T cells, RPLP0 in OA hACs, and β2M in hMSCs.

2.12.3 microRNA TaqMan assays in osteoarthritic human articular chondrocytes and human mesenchymal stem cells

A total of 10 µl of reaction mixture containing 0.7 µl cDNA template, 0.5 µl of 20x TaqMan® Small RNA Assay (Table 2.6), 5 µl 2x TaqMan universal PCR master mix (Applied Biosystems, Foster City, CA, USA) and 3.8 µl of nuclease-free water was added to each well of a 384 well plate. Each reaction was run in a triplet. The plate was sealed and centrifuged for 1 min at 1500 g. Real Time-qPCR was carried out on a ViiA™ 7 Real-Time PCR System (Applied Biosystems) under the following conditions; 1X stage 1: 95 °C for 10 mins; 40X stage 2: 95 °C for 15 s; 60 °C for 1 min. Relative gene expression was calculated using the ΔΔCt method with RNU24 or RNU44 as the control gene.

2.13 Western blot analysis

Targeted HEK293T cells were collected 72 h post transfection by addition of 300 µl lysis buffer (3 M NaCl, 1 M Tris-HCL pH 7.4, Igepal CA-630 (Sigma-Aldrich, Dorset, UK), 0.5 M EDTA, 1x Protein Inhibitor Complex (Roche, Basel, Switzerland)) for one well in a 12 well plate. OA hACs were collected in the same way at 3h, 6h, 9h and 18h post transfection. The lysate was spun down at 3,000 g for 5 mins and the pellet (nuclear fraction) again resuspended in lysis buffer. A small fraction of the resuspended nuclear fraction (6 µl) was used for western blot analysis by adding 2 µl of a 4x SDS protein gel loading buffer and boiling the samples for 10 minutes before loading. Samples were resolved over a 8

% sodium dodecyl sulfate (SDS)–polyacrylamide (PAGE) gel and the resolved proteins then transferred onto a PVDF membrane (Millipore, Bedford, USA). The membrane was blocked in a blocking solution (5 % dried milk and 1 % Tween-20 in 1x PBS-T) for 1 h, followed by overnight incubation at 4 °C with a primary antibody (Cas9 (Active Motif, Belgium), mouse monoclonal, 1:2,000 or α-Tubulin (Sigma Aldrich, Dorset, UK), mouse monoclonal, 1:5,000, diluted in blocking solution).

After incubation with the primary antibody, the membrane was washed three times for 10 minutes each time in 1x PBS-T, followed by 60 minutes incubation in the peroxidase conjugated secondary antibody (mouse, 1:10,000) and subsequently washed again three times for 10 minutes in 1x PBS-T.

Finally, protein was detected via enhanced chemiluminescence (ECL Select™, Amersham Biosciences, Buckinghamshire, UK).

2.14 Hip evulsion in mouse

The procedure was carried out by a senior post-doc in the lab, Dr. Linyi Zhu as described by Stanton et al. (2011). In brief, mice (5 – 6 weeks old) were culled by CO2. Subsequently the acetabulofemoral joint was exposed in a laminar flow hood. Through applying pressure from behind the femur, the hip joint was dislocated and the femoral cap was avulsed using forceps. Hip cartilage was snap frozen (0 h) in liquid nitrogen or cultured in serum free medium for 4 h, and then frozen.

Two hip joints from each mouse were pooled together and stored at -80°C. 8 mice were used in total. Hip cartilage was ground to a powder using Cryo-Cup Grinder (Biospec,

Bartlesville, USA), and RNA was extracted using the RNeasy Mini kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions.

2.15 Chondrogenic disc assay in human mesenchymal stem cells

Human MSCs (wildtype), *miR-140* KD hMSCs and NT sgRNA hMSCs, all passage matched at P10, were cultured in 20 % FBS and 1 % P/S containing DMEM. Cells were trypsinised, quantified and resuspended in supplemented medium (0 h cells were resuspended in standard medium without supplements). 5 x 10^5 hMSCs of each cell type were added to the 6.5 mm insert of a 12 well transwell plate with 0.4 µm pores (Costar/Corning, Wiesbaden, Germany). Subsequently, transwell plates were centrifuged at 200 g for 5 mins to help sediment cells on the membrane. The bottom chamber of each well was filled with 600 µl of supplemented DMEM to incubate at 37 °C and 5 % CO2 overnight for cell discs to form. For 21 days the medium was replaced every 48 h by carefully replacing the top and bottom chambers. DMEM was supplemented with 1x ITS (Corning, Wiesbaden, Germany), 10 ng/µl TGFβ3 (R&D systems, Minneapolis, USA), 100 µg/ml sodium pyruvate, 40 µg/ml L-proline 25 µg/ml ascorbate-2-phosphate and 1x L-glutamine 100 nM dexamethasone (Sigma, Dorset, UK). Discs were either stored in TRIzol at -80 °C for RNA isolation or in 70 % EtOH to be processed by the histology department.

2.15.1 RNA isolation from chondrogenic discs

Chondrogenic discs were transferred into 1.5 ml Eppendorf tubes with a wide bore pipette and homogenised with disposable pestle containing 50 µl of molecular grinding resin (G-Biosciences, MO, USA). Molecular grinding resin was resuspended in TRIzol according to the manufacturer's manual. Subsequently the corresponding TRIzol from the original Eppendorf tubes was added to the homogenised disc/resin sample. After pelleting the resin at 12,000 RPM and 4 °C for 5 mins, the supernatant was transferred into a new 1.5 ml Eppendorf tube which, in order to disrupt nucleosome links, were placed in an orbital mixer for 5 min at room temperature. 150 µl 1-Bromo-3-chloropropane (Sigma-Aldrich, Dorset, England) was added and samples vortexed and left for 5 mins at room temperature before they were phase separated by centrifugation at 13,000 RPM for 10 mins at 4 °C. Further steps were carried out as per the manufacturer's instructions (RNeasy Mini Kit, Qiagen, Hilden, Germany). Samples were eluted in 12 µl RNase-free water and subsequently quantified and assessed for purity at 230 nm, 260 nm and 280 nm using a digital spectrophotometer (Nanodrop One, ThermoFisher, MA, USA).

2.16 Histology

All chondrogenic pellet disks were processed by the histology team at the Kennedy Institute of Rheumatology, Oxford. Following fixation in 10 % neutral buffered formalin,

chondrogenic discs at 10 days and 21 days were embedded in paraffin and sectioned at 5 µm intervals. Sections were stained with safranin-O, (stains proteoglycans orange/red), Alcian blue (stains collagen green), and Sirius red (stains collagen fibres red).

2.16.1 Safranin-O staining

Protocol in brief: Tissue processing was carried out in the Tissue Tek VIP (Sakura, Alphen aan den Rijn, Netherlands) automated machine set to the following run: dewaxing (2 x 5 mins xylene, 3 x 1 min 100 % ethanol), 20 s tap water, 30 s harris haematoxilyn, 1 min wash (all washes in water with the tap running), 20 s acid alcohol (70% ethanol plus 10 ml hydrochloric acid per litre), 1 min wash, 1 min bluing water (600 ml water plus 1.6 ml ammonia), 1 min wash, 2 mins safranin-o, 1 min wash water, dehydrate to xylene, 3 x 1 min ethanol, 3 x 1 min Xylene. Cover slips were mounted with Di-n-butylphthalate (DPX) mounting media.

2.16.2 Alcian Blue staining

Protocol (Atom Scientific, Manchester, UK) in brief: Tissue processing was carried out in the Tissue Tek VIP (Sakura, Alphen aan den Rijn, Netherlands) automated machine set to the following run: 2x Xylene 5 mins, 2x 100 % Ethanol 1 min, transferred sections to water, stained in Alcian Blue (8GX 1 % in 0.1 M HCL) for 30 mins, washed well in distilled water, stained nuclei lightly with Haemalum Mayer for 1 min, differentiated quickly in

running tap water, dehydrated in the staining machine as follows: 3x 100 % Ethanol for 1 min, 2x Xylene 1 min, cover slips were mounted with DPX mounting media.

2.16.3 Sirius Red

Staining in brief: Tissue processing was carried out in the Tissue Tek VIP (Sakura, Alphen aan den Rijn, Netherlands) automated machine set to the following run: deparaffinise the slides into water, stained in Picrosirius red for 1 h, washed in two changes of acidified water (2 mins each wash), physically most of the water was removed from the slides by vigorous shaking, dehydrated into xylene, using the staining machine, and cover slips were mounted with DPX mounting media.

All slides were imaged using an Olympus BX61 microscope.

2.17 Scratch assay in human mesenchymal stem cells

Human MSCs (wildtype), *miR-140* KD hMSCs and NT sgRNA hMSCs, all passage matched at P4, were expanded in 20 % FBS and 1 % P/S containing DMEM until ~70-80 % confluency. The medium was carefully replaced with serum-free medium. A 10 µl pipette tip was used to create a scratch in the middle of each well of a 12 well plate and placed in an incubator at 37 °C, 5 % CO2 on a JuLi stage microscope. Time-lapsed images

with images taken every 20 mins for 48 h of the scratch at two different points from each well were taken. To quantify the gap closure area Image J was applied for each image by drawing around it and by applying the formula 100 - ((X h area / 0 h area) x 100).

2.18 Statistical analysis

Prism 8 software (GraphPad Software Inc., USA) was used to perform statistical analysis. Levels of significance were reported as follow: when p-value ≤ 0.05: significant (*), when p-value ≤ 0.01: very significant (**), when p-value < 0.001: highly significant (***), when p-value ≤ 0.0001: extremely significant (****). Analysis were marked with non-significant (n.s.), when p > 0.05.

Using Prism, samples were tested for Gaussian normality. Unpaired, two-tailed t-tests were performed to compare two sets of measurements. For three, or more groups, one-way ANOVA was applied instead. A two-way ANOVA was applied, when the response was determined by two factors.

Following one-way or two-way ANOVA, multiple comparison analyses were carried out using Turkey and Dunnett tests. The Turkey test was performed to compare every mean to every other mean. The Dunnett test was performed to compare every mean to a control mean (normalised to 1).

2.19 TaqMan probes

TaqMan probes were used to quantify miRNA and gene expressions in OA hACs and hMSCs.

Table 2.6: TaqMan probes used in this study

Primer	Assay ID
hsa-miR-140-3p	002234
mmu-miR-140	001187
hsa-miR-145-5p	002278
hsa-miR-92a-3p	000431
RNU24	001001
RNU44	001094
18S	Hs99999901_s1
B2M	Hs06637353_s1
COL2A1	Hs00264051_m1
COL10A1	Hs00166657_m1
ACAN	Hs00153936_m1
COMP	Hs00164359_m1
CDKN2A (p16)	Hs00923894_m1

2.20 Primers

Primers were designed with the help of the PrimerQuest Tool from Integrated DNA Technologies (IDT) and checked for uniqueness and low primer secondary structure index.

2.20.1 *miR-140* (chapters 3 and 4)

Table 2.7: Primer used for *miR-140* in gDNA and cDNA of targeted OA hACs and hMSCs cells. ROI = region of interest; MM = mismatch; OT = off target. Fw = Forward. Rv = reverse.

Name	Sequence 5'-3'	Details
miR-140 Fw	CCGCTTGGTGTTGGGTTAACTTG	Fw primer to amplify ROI for subsequent T7E1
miR-140 Rv	AGTGTGGGAACGGGGATGG	Rv primer to amplify ROI for subsequent T7E1
WWP2E28Fw	CACAGAGGAGAACAAGGAAGAG	qPCR fw primer to check expression
WWP2E28Rv	CTTTCTCTGCAACCCACTAGAT	qPCR rv primer to check expression
miR140_L1_OT1_Fw	ACAGCTACAACCTGGGACTTTCA	Fw primer for predicted 2-MM OT1 against sgRNA L1
miR140_L1_OT1_Rv	AACGGCATGGCATTAGCGTTTAG	Rv primer for predicted 2-MM OT1 against sgRNA L1
miR140_L1_OT2_Fw	CCCAGACCAATCCTGAGCAGAAA	Fw primer for predicted 2-MM OT2 against sgRNA L1

miR140_L1_OT2_Rv	TCCAGAAATCCCTGACCTACCTACA	Rv primer for predicted 2-MM OT2 against sgRNA L1
miR140_L1_OT3_Fw	CAGGCCACACTCTTCTCGATCTT	Fw primer for predicted 2-MM OT3 against sgRNA L1
miR140_L1_OT3_Rv	ATTGGCAGCCAGGGAGTAGTT	Rv primer for Predicted 2-MM OT3 against sgRNA L1
miR140_L1_OT4_Fw	GTCAAGCCTCAACAATCCCTCTCT	Fw primer for predicted 2-MM OT4 against sgRNA L1
miR140_L1_OT4_Rv	AGACTCCAAAGCATCCGCTCTT	Rv primer for predicted 2-MM OT4 against sgRNA L1
miR140_L1_OT5_Fw	ACGCCACAGAGAAGGAGTGAAG	Fw primer for predicted 2-MM OT5 against sgRNA L1
miR140_L1_OT5_Rv	ACGTAGACCTGGGAGAGCTAGAA	Rv primer for Predicted 2-MM OT5 against sgRNA L1
miR140_L1_OT6_Fw	GGTGATTGGATCTTGCGGGTAGT	Fw primer for predicted 2-MM OT6 against sgRNA L1
miR140_L1_OT6_Rv	ATGGTTCCCAGGTCCTTGAGAAAG	Rv primer for predicted 2-MM OT6 against sgRNA L1
miR140_L1_OT7c_Fw	AGCCTCTCGAGTAGCTGGGATTA	Fw primer for predicted 2-MM OT7 against sgRNA L1
miR140_ L1_OT7c_Rv	ACACTGGCTTGGGCAACATAGTA	Rv primer for Predicted 2-MM OT7 against sgRNA L1

miR140_L1_OT8c_Fw	AGTCCACTCAGGGTTGGAAATGG	Fw primer for predicted 2-MM OT8 against sgRNA L1
miR140_L1_OT8c_Rv	GGATGGTGGCACACACCTATGA	Rv primer for predicted 2-MM OT8 against sgRNA L1
miR140_L1_OT9e_Fw	TGTGCACCAGCTCAGGAAAGGT	Fw primer for predicted 2-MM OT9 against sgRNA L1
miR140_L1_OT9e_Rv	ACCCACCTGCTACCCAGCTATCTAAT	Rv primer for predicted 2-MM OT9 against sgRNA L1
miR140_L2_OT1_Fw	TCGGAGGAGTCAGGAAGGAAGAA	Fw primer for predicted 2-MM OT1 against sgRNA L2
miR140_L2_OT1_Rv	GCCATGATTCCTTGGAGCAGAGA	Rv primer for predicted 2-MM OT1 against sgRNA L2
miR140_L2_OT2_Fw	TTTGCCACACGGCCCTAAATC	Fw primer for predicted 2-MM OT2 against sgRNA L2
miR140_L2_OT2_Rv	GCTACTTGGGAGGCTGATCTCTT	Rv primer for predicted 2-MM OT2 against sgRNA L2
miR140_L3_OT1c_Fw	ATGGTGGCGGGCACCTGTAATC	Fw primer for predicted 2-MM OT1 against sgRNA L3
miR140_L3_OT1c_Rv	GAGCATCAGGCCAAACAAGGGATGAA	Rv primer for predicted 2-MM OT1 against sgRNA L3
miR140_L3_OT2_Fw	ACCAGCACACTCTTTCTCCTCTTG	Fw primer for predicted 2-MM OT2 against sgRNA L3

miR140_L3_OT2_Rv	GGTCTGGTTGGACCTCGTTGTATT	Rv primer for predicted 2-MM OT2 against sgRNA L3
miR140_L3_OT3_Fw	AACCACAGGGTCAGAGCACATAAG	Fw primer for predicted 2-MM OT3 against sgRNA L3
miR140_L3_OT3_Rv	CCTGGCAGTGTATGAAGAGGCATTA	Rv primer for predicted 2-MM OT3 against sgRNA L3

2.20.2 *SOX9* (chapter 5)

Table 2.8: Primers used for *SOX9* in gDNA and cDNA of targeted HEK293T cells and OA hACs.
ROI = region of interest, T_A = Primer annealing temperature. Fw = Forward. Rv = reverse.

Name	Code	Sequence 5'-3'	Details
SOX9 out HDR-MRE1F	C29	GCTAAAGGCAACTCGTACCC	Fw primer to amplify ROI
SOX9 out HDR-MRE1R	C30	ATCCCCTCAAAATGGTAATGAATC	Rv primer to amplify ROI
Sox9-en-MRE1-f2	C15	TAAAGGCAACTCGTACCCAA	Endogenous fw primer for validating T3/T7 integration
T3 Rv-Primer	N1	TCCCTTTAGTGAGGGTTAATT	Universal T3 barcode rv Primer
T7rev(+6)Sox9M1	N17	CCTATAGTGAGTCGTATTACTCGTT	Elongated T7 rv primer for SOX9 MRE1 to detect T7 barcode
SOX9-MRE1-endgF	N18	AGCTAAAGGCAACTCGTACCCA	Endogenous fw primer for validating barcode integration
Rv_BC_A	N19	GATTACGCGCGCGCGATATG	Reverse primer to validate D1 integration

Rv_BC_B	N20	ATATCGCGCGTTACTGCCGT	Reverse primer to validate D2 integration
Rv_BC_C	N21	CGTCGACGATTATTCGCGCG	Reverse primer to validate M1 integration
Rv_BC_D	N22	ACGCGATTCGCGCGTATTC	Reverse primer to validate M2 integration
Sox9-mre1-ENDG68-FW	N23	GGGCGAAGATGGCCGAGATGATC	Endogenous fw primer for validating barcode integration, higher T_A
D1-UNIV68-RV	N24	TCGATTACGCGCGCGCGATATG	Reverse primer to validate D1 integration, higher T_A
D2-UNIV67-RV	N25	ATATCGCGCGTTACTGCCGTCG	Reverse primer to validate D2 integration, higher T_A
M1-UNIV66-RV	N26	TCGTCGACGATTATTCGCGCGA	Reverse primer to validate M1 integration, higher T_A
M2-UNIV67-RV	N27	TACGCGATTCGCGCGTATTCGG	Reverse primer to validate M2 integration, higher T_A
D1-MRE1-RV	N33	GCGCGCGCGATATGCTCGTT (6+14)	Semi-endogenous reverse primer to validate D1 insertion
D2-MRE1-RV	N34	CGCGTTACTGCCGTCGCTCG (6+14)	Semi-endogenous reverse primer to validate D2 insertion
M1-MRE1-RV	N35	TCGACGATTATTCGCGCGATCCAGT (6+19)	Semi-endogenous reverse primer

			to validate M1 insertion
M2-MRE1-RV	N36	CGATTCGCGCGTATTCGGTCCAGT (6+18)	Semi-endogenous reverse primer to validate M2 insertion

2.20.3 *C9ORF7* (chapter 5)

Table 2.9: Primers used for *C9orf7* in gDNA and cDNA of targeted HEK293T cells.
Fw = Forward. Rv = reverse

Name	Code	Sequence 5'-3'	Details
C9ORF7-Fw	N3	GGGCGGGTGTGGAAGATA	Fw primer to amplify ROI
C9ORF7-Rv	N4	CACACACACAGCCCTTTGAA	Rv primer to amplify ROI
C9orf7 Fw	N7	GCTTCTGGAGCGCAGGTACTG	Endogenous fw primer for validating T3/T7 integration
T3 Rv-Primer	N1	TCCCTTTAGTGAGGGTTAATT	Universal T3 barcode rv Primer
C9orf7_2_Fw	N10	TGGTCAGGCCTGGCTTAG	Fw Primer to amplify ROI
C9orf7_2_Rv	N11	GCACATGTCCACCAGGCTAT	Rv Primer to amplify ROI
C9orf7_3_Fw	N12	GGAGGATGCAGAGAGCTGGT	Fw Primer to amplify ROI
C9orf7_3_Rv	N13	CAAGTCCCCAGGGCAGTA	Rv Primer to amplify ROI
T7rev(+6)C9orf7	N15	CCTATAGTGAGTCGTATTACAGATG	Elongated T7 rv primer for C9orf7 to detect T7 barcode

Table 2.10: Designed single-stranded oligo DNA nucleotides (ssODNs) with ~ 60 nt homology arms either side of the inserted sequence for the *C9orf7* and *SOX9* target sites.

Target	ssODN 5‘-3‘
C9orf7 ssODN T3	AACTGTTTCCCAGGAACACCTCTCGGGCCCATCTGCGTCTGAGGCTGGGAGTG GCATCTGTAATACGACTCACTATAGGGGTGGCATCTGAGGCCAGGAGTGGCA GGCTGGTGGGCTGGGCGTGGGGTTTTCTGGGCCCT
C9orf7 ssODN T7	AACTGTTTCCCAGGAACACCTCTCGGGCCCATCTGCGTCTGAGGCTGGGAGTG GCATCTGAGGCCGGGAAATTAACCCTCACTAAAGGGAGTGGCATCTGAGGCC AGGAGTGGCAGGCTGGTGGGCTGGGCGTGGGGTTTTCTGGGCCCT
SOX9 ssODN T3	ACTTGTGGCCAATCAGTGGCCAGGCCAACCTTGGCTAAATGGAGCAGCGAAA TCAACGAGAAACTGGAAATTAACCCTCACTAAAGGGACTTTTTAAATACTCTTC AGAGCAAGCGTGGAGGATGATGGAGAATCGTGTGATCA
SOX9 ssODN T7	ACTTGTGGCCAATCAGTGGCCAGGCCAACCTTGGCTAAATGGAGCAGCGAAA TCAACGAGTAATACGACTCACTATAGGGCTTTTTAAATACTCTTCAGAGCAAGC GTGGAGGATGATGGAGAATCGTGTGATCAGTGTG
SOX9 ssODN M1	ACTTGTGGCCAATCAGTGGCCAGGCCAACCTTGGCTAAATGGAGCAGCGAAA TCAACGAGAAACTGGATCGCGCGAATAATCGTCGACGACTTTTTAAACCCTCTT CAGAGCAAGCGTGGAGGATGATGGAGAATCGTGTGATCAGTGTG
SOX9 ssODN M2	ACTTGTGGCCAATCAGTGGCCAGGCCAACCTTGGCTAAATGGAGCAGCGAAA TCAACGAGAAACTGGACCGAATACGCGCGAATCGCGTACTTTTTAAACCCTCTT CAGAGCAAGCGTGGAGGATGATGGAGAATCGTGTGATCAGTGTG
SOX9 ssODN D1	ACTTGTGGCCAATCAGTGGCCAGGCCAACCTTGGCTAAATGGAGCAGCGAAA TCAACGAGCATATCGCGCGCGCGTAATCGATTTTTAAACCCTCTTCAGAGCAA GCGTGGAGGATGATGGAGAATCGTGTGATCAGTGTG
SOX9 ssODN D2	ACTTGTGGCCAATCAGTGGCCAGGCCAACCTTGGCTAAATGGAGCAGCGAAA TCAACGAGCGACGGCAGTAACGCGCGATATTTTTTAAACCCTCTTCAGAGCAA GCGTGGAGGATGATGGAGAATCGTGTGATCAGTGTG

Results

3

Targeting *miR-140* with CRISPR-Cas9 in OA human articular chondrocytes: discovery of novel *miR-140* targets

Chapter 3. Targeting *miR-140* with CRISPR-Cas9 in osteoarthritic human articular chondrocytes: discovery of novel *miR-140* targets

3.1 Introduction

Classical approaches to study miRNA and mRNA interactions include miRNA overexpression and inhibition. These approaches, mediated through expression vectors/synthetical mimics or inhibitors respectively, have furthered insight into miRNA targeting. However, these approaches are also subject to some limitations. The overexpression of a given miRNA in a non-endogenous context will simultaneously perturb up to thousands of target genes and raise questions surrounding its physiological relevance.

Inhibition studies of miRNAs bear limitations as well. In many cases miRNA inhibitors do not downregulate miRNA levels at all but rather inhibit their function. Lower levels in a qPCR after an inhibitor experiment are due to the complementary sequence of the

inhibitor; the actual miRNA escapes qPCR detection, thus artificially suggesting lower levels. In both cases, direct targets cannot be differentiated from indirect targets.

In this study we utilise the CRISPR-Cas9 technology. The introduction of CRISPR has allowed alteration of the genome of various species virtually almost anywhere, and hence has emerged as a powerful tool for editing the genome in a simplistic and time-efficient way. By using customised sgRNAs, which target the biogenesis processing sites of a given miRNA, these miRNAs are targeted in an endogenous context. Potential off-targeting can be controlled by carefully designing miRNA targeting sgRNAs, and targeting can be carried out with high specificity and long-term stability.

As proof of concept, we chose *miR-140*, which is highly and exclusively expressed in chondrocytes. *miR-140* expression is reduced in osteoarthritic tissue, but exhibits high levels of expression in normal articular cartilage (Miyaki et al., 2009). Another study from the same group showed that *miR-140* knockout mice develop an osteoarthritic-like phenotype with age (Miyaki et al., 2010), characterised by fibrillation and proteoglycan loss. *miR-140* directly targets *ADAMTS-5 in vitro*, a major cartilage degrading proteinase (Miyaki et al., 2009).

miR-140 is a well-studied miRNA in mice, however, only a few studies have been conducted at the cellular level.

As a result, the aims of this chapter included (i) to successfully knockdown *miR-140* in primary, human articular chondrocytes, and (ii) to determine which potential genes were regulated by *miR-140* in chondrocytes.

Osteoarthritic human articular chondrocytes were edited by CRISPR-Cas9 according to the workflow in Fig. 3.1.

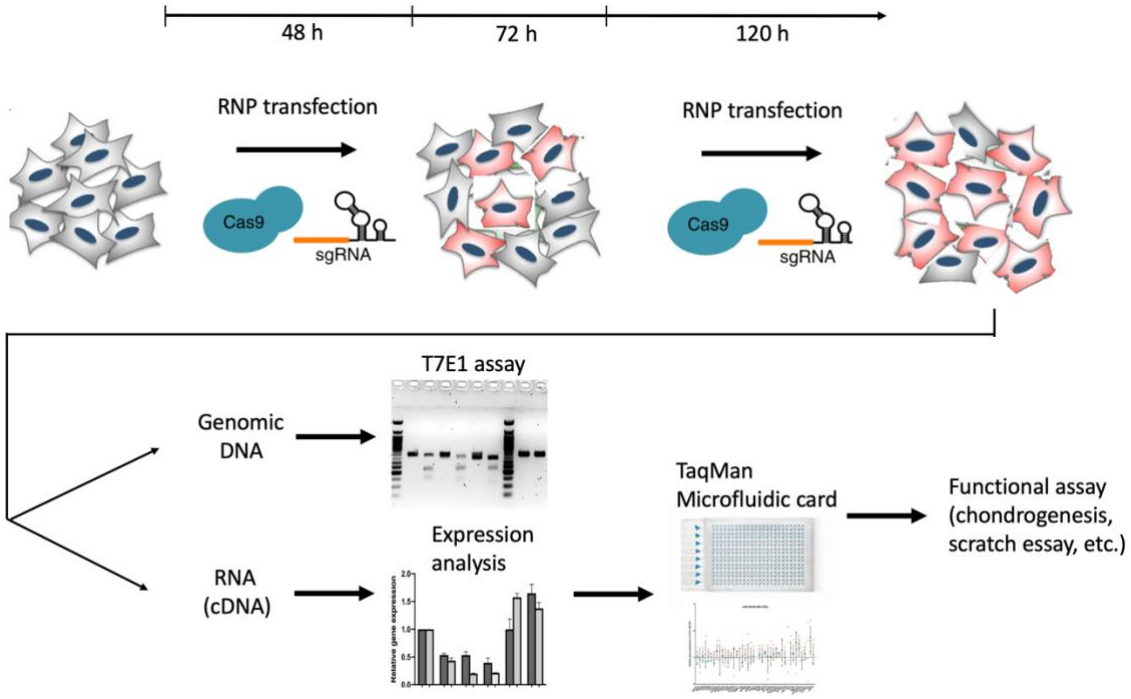


Fig. 3.1: Main workflow describing RNP double transfection of osteoarthritic, human articular chondrocytes to knockdown *miR-140* and their functional readout.

Osteoarthritic, human articular chondrocytes (OA hACs) were RNP transfected to knockdown *miR-140*. 48 h post transfection the media was exchanged, and cells were RNP re-transfected as before. Again, 48 h post second transfection, cells were isolated for gDNA and RNA (cDNA). While the gDNA was used to assess gene editing efficiency, the RNA was reverse transcribed into cDNA using a miRNA-specifc stem loop primer, and quantified for *miR-140* expression. cDNA samples of different donors (reverse transcribed the standard way (without miRNA-specific stem loop primer)) were used to quantify expression levels of candidate genes by TaqMan Microfluidic card. Adapted from (Bassett et al., 2014).

3.2 Results

3.2.1 Ribonucleoprotein (RNP) transfection of osteoarthritic human articular chondrocytes does not cause cell toxicity, and may reduce potential off-targets

In order to assess the impact of two different transfection methods on isolated, human osteoarthritic chondrocytes (OA hACs), cells were isolated from patients undergoing total knee replacement, and cultured in 10 % FBS containing DMEM. 5 x 10^4 cells at passage 1 were plated in 12 - well plates, and after 48 h transfected either with pX330 or using a Ribonucleoprotein (RNP) complex. The pX330 plasmid is an ~8.5 kb plasmid with the coding sequence for Cas9 and a targeting single guide RNA (sgRNA; subcloned into the plasmid). The RNP complex consists of the Cas9 protein and targeting sgRNA. This sgRNA consists of a custom-designed, target specific CRISPR RNA (crRNA), and is annealed to a universal, trans-activating crRNA (trRNA). Both the pX330 plasmid and RNP complexes targeted the same gene, and were transfected into the cell using Lipofectamine RNAiMAX and Lipofectamine CRISPRMAX, respectively. RNP transfected and untreated control OA hACs exhibited continuous growth and increased confluency at 24 h and 48 h post transfection compared with 0 h (Fig. 3.2A). However, plasmid transfected OA hACs detached from the surface and were floating in the cell media 24 h post transfection. Cell media including non-adherent cells was removed, centrifuged and quantified for absolute number of blue (trypan-positive) cells. Cell media of RNP

transfected and untreated cells was also collected. The number of non-adherent cells from plasmid transfected wells (n = 3) was around 3.4 x 10^4 cells and significantly higher compared with non-adherent cells from RNP transfected (n = 3, ~1200 cells), and untreated wells (n = 3, ~950 cells) (Fig. 3.2B).

At 24 h post transfection, all wells were replaced with new 10 % FBS containing DMEM. RNP transfection was established to be the favourable strategy for transfecting primary OA chondrocytes, considering its reduced associated cell toxicity. To check for presence of Cas9 in OA hACs, RNP transfected OA hACs with Cas9/sgRNA complexes were collected 3 h, 6 h, 9 h and 18 h post transfection, resolved by 8 % SDS-PAGE, and immunoblotted for Cas9. α-Tubulin was used as loading control (Fig. 3.2C). Cas9 was detected between 3 h and 9 h post transfection, and levels were reduced by 18 h. Unlike a plasmid, where Cas9/sgRNA are expressed by the cell following uptake into the host cell, the RNP method produces rapid, but more transient levels of Cas9. This has been shown to reduce off-targets otherwise observed for plasmid mediated CRISPR-Cas9 approaches in induced pluripotent stem cells (iPS), human embryonic kidney (HEK) 293 FT cells and K562 cells (Kim et al., 2014, Zuris et al., 2015, Liang et al., 2015).

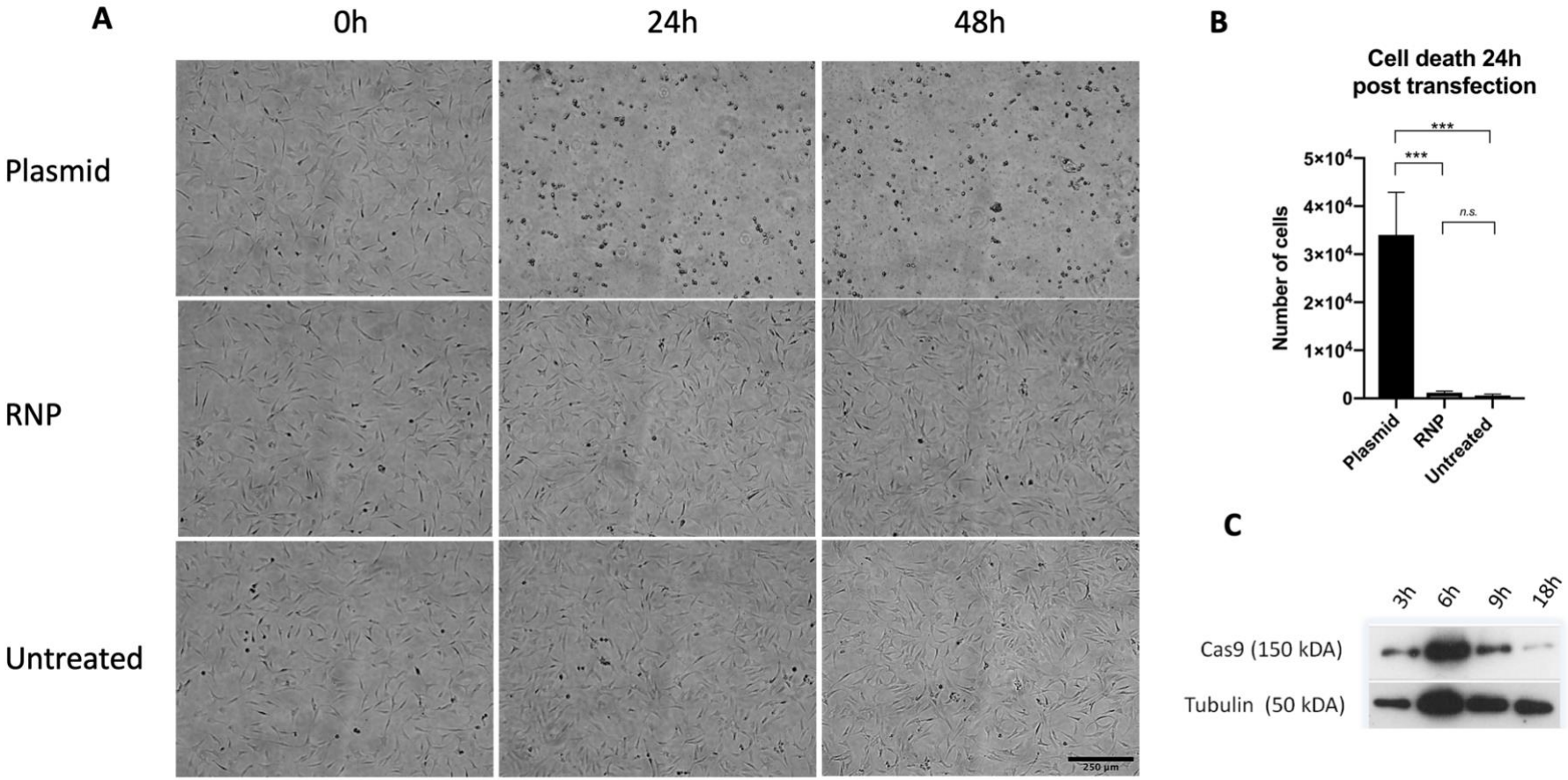

A
0h
24h
48h
Plasmid
RNP
Untreated
250 µm
B
Cell death 24h
post transfection
***
***
n.s.
5×10^4
4×10^4
3×10^4
2×10^4
1×10^4
0
Number of cells
Plasmid
RNP
Untreated
C
3h
6h
9h
18h
Cas9 (150 kDA)
Tubulin (50 kDA)

Fig. 3.2: Ribonucleoprotein (RNP) transfection of osteoarthritic human articular chondrocytes is not toxic

OA hACs were isolated from patients undergoing total knee replacement (TKR) and plated at $5x10^4$ cells/well in 10 % FBS containing DMEM at 37 °C and 5 % CO2. Using Lipofectamine, cells were transfected by two different methods: (i) using pX330, a plasmid which encodes for Cas9 and a subcloned single guide RNA (sgRNA) or (ii) using a Ribonucleoprotein (RNP) complex; Cas9 with sgRNA. Scale bar = 250 µm. **A**. Cell morphology at 0 h, 24 h and 48 h post transfection by microscopy. **B**. Non-adherent cells were stained with trypan blue and dead cells quantified 24 h post transfection. Absolute numbers were shown. The statistical significance of comparisons between treatment groups were analysed with one-way ANOVA, with post hoc multiple comparisons corrected using Turkey's test. *n* = 3. **C**. RNP transfected hACs were collected 3 h, 6 h, 9 h and 18 h post transfection, resolved by 8 % SDS-PAGE, and immunoblotted for Cas9. α-Tubulin was used as loading control. Immunoblotting was visualised using ECL detection reagent on X-ray film.

3.2.2 Both *miR-140-3p* and *miR-140-5p* are expressed in osteoarthritic human articular chondrocytes and their encoding sequence is cleaved by custom designed sgRNAs

miR-140 plays an important role in cartilage biology and was therefore a good candidate microRNA to target as proof of concept. I first checked that both arms of *miR-140*, namely *miR-140-3p* and *miR-140-5p*, were expressed in OA hACs. We resorted to OA hACs due to being constantly supplied with human OA tissue samples (from which OA hACs were isolated) and a lack of constant availability of healthy human tissue. For each donor (n = 18) and each microRNA arm, 15 ng RNA were reverse transcribed and assessed. Samples were normalised to *RNU24* and are displayed as ΔCT values (Fig. 3.3). On average, *miR-140-5p* demonstrated a slightly higher but non-significant (p = 0.0729) expression compared with *miR-140-3p*. Raw CT values for *miR-140-5p* and *miR-140-3p* varied between 28 and 30.5. Each donor was colour coded, so they would be marked in subsequent knockdown experiments.

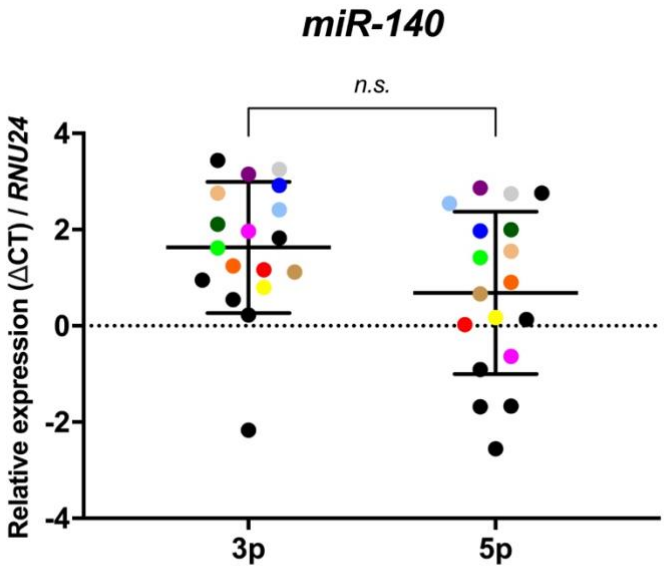


Fig. 3.3: Expression levels of *miR-140-3p* and *miR-140-5p* in osteoarthritic, human articular chondrocytes (OA hACs) are moderate.

RNA was purified from passage 1 (P1) OA hACs, the same quantity of RNA (15 ng) always reverse transcribed, and gene expression of *miR-140-3p* and *miR-140-5p* assessed by qPCR and microRNA TaqMan probes. Samples were normalised to *RNU24* and are displayed as ΔCT values. Data are shown as mean ± SD. The statistical significance of comparisons was analysed with Student's t-test (two-tailed). Samples used for subsequent experiments were colour coded. Low gene editing donors: Donor 1: m, 62. Donor 2: m, 64. Donor 3: m, 57. Donor 4: m, 71. Donor 5: f, 70. Donor 6: m, 54. High gene editing donors: Donor 1: f, 63. Donor 2: f, 70. Donor 3:m, 78. Donor 4: m, 61. Donor 5: f, 61. Donor 6: m, 69. Donors that were not edited: Donor 1: f, 72. Donor 2: f, 61. Donor 3: m, 78. Donor 4: m, 78. Donor 5: m, 61. Donor 6: m, 69. *n* = 18.

To assess Cas9 cleavage activity, four different single guide RNAs (sgRNAs), with bioinformatically low predicted off-target scores[1], were designed to target the hairpin structure of *miR-140* (Fig. 3.4A), which is encoded within an intron of *WWP2*. Each sgRNA consisted of a universal trans-activating CRISPR RNA (tracrRNA) and a target specific CRISPR RNA (crRNA).

miR-140 targeting sgRNAs were complexed with Cas9 as RNPs, and introduced with Lipofectamine CRISPRMax into the cell. A 439 bp long PCR product was amplified with the help of *miR-140* flanking primers. Subsequently, non-homologous end joining (NHEJ) induced gene editing efficiencies were assessed 48 h post transfection by the T7 Endonuclease 1 (T7E1) assay. The T7E1 assay recognises mismatched regions in double-stranded DNA (dsDNA), and cleaves the DNA at this site to generate smaller PCR products (in this case between roughly 150 bp and 250 bp). All four sgRNAs displayed varying gene editing efficiencies, with sgRNA locus (L)3 displaying the highest gene editing efficiencies, followed by sgRNA L2 and sgRNA L1. In contrast, sgRNA L4 displayed very low editing efficiencies. A non-targeting (NT) sgRNA was used as control (Fig. 3.4B).

[1] https://www.atum.bio/eCommerce/cas9/input

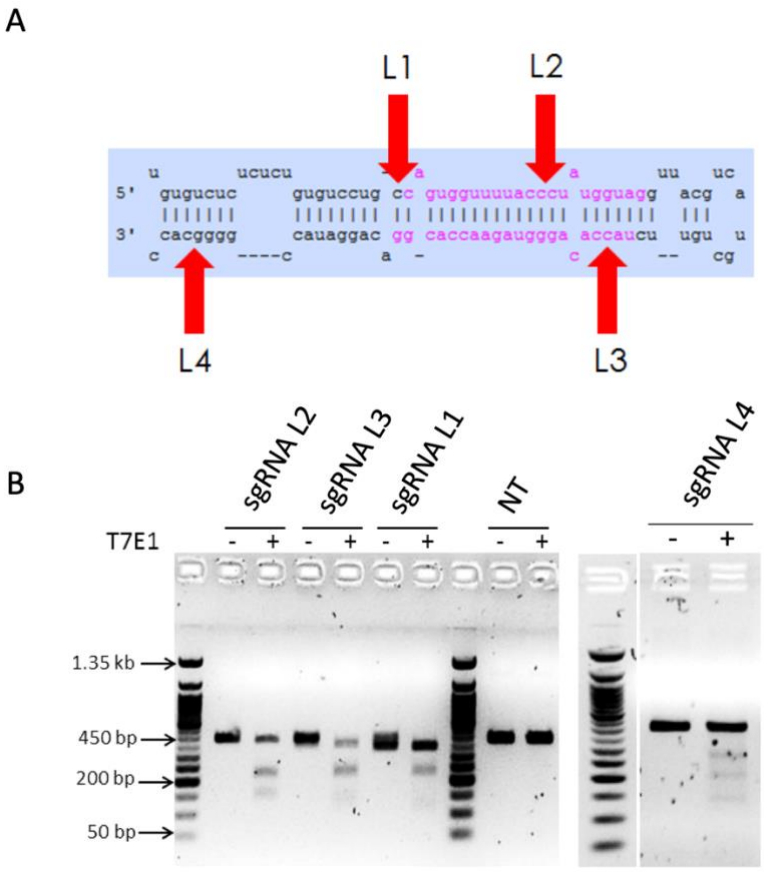


Fig. 3.4: *miR-140* targeting sgRNAs reveal different degrees of DNA cleavage and gene editing efficiencies in OA hACs.

Four different sgRNAs were designed with the help of ATUM software[1], to target the *miR-140* hairpin structure. **A** Red arrows display predicted cutting sites of designed sgRNAs within the miR-140 hairpin structure **B** Primers flanking the *miR-140* hairpin structure were designed to amplify a 439 bp long PCR product. This PCR product was assessed by the T7 Endonuclease 1 (T7E1) for DNA cleavage, and cleaves the DNA at this site to generate products between roughly 150 bp and 250 bp. A non-targeting (NT) sgRNA was used as control. Representative agarose gel shown. *n* = 2.

3.2.3 T7E1 confirms DNA cleavage at the *miR-140* site in six different osteoarthritic human articular chondrocyte donors

Based on the individual gene editing efficiencies of *miR-140* targeting sgRNAs, *miR-140* was targeted in 6 different OA hACs donors using sgRNA L3, or the combination of sgRNA L3 and L2, or the combination of sgRNA L3, L2 and L1. If more than one sgRNA was used, the final required sgRNA concentration was reduced accordingly to ensure the final concentration of total sgRNA was constant. A NT sgRNA was used as control along with only Lipofectamine (LF) and Cas9 treated cells (no sgRNA), as well as completely untreated cells. OA hACs at P1 were seeded in 12 well plates, cultured for 48 h in 10 % FBS containing DMEM, and were RNP transfected at a confluency of between 40 - 60 %. 48 h post-transfection, cells were isolated for gDNA and RNA. For each donor, each experimental condition was run either in duplicate or triplicate. DNA cleavage was assessed by the T7E1 assay, after PCR amplifying a 439 bp long product with primers flanking the targeted miR-140 hairpin region. Gene editing was confirmed for all targeting sgRNAs. Applying the combination of sgRNA L3+L2 or the combination of sgRNA L3+L2+L1 resulted in the strongest gene editing efficiency according to the T7E1 assay across most donors, while gene editing efficiency across biological replicates were largely consistent (Fig. 3.5). Cases of non-consistency could be both the result of biological variation, and/or pipetting. Control cells did not exhibit DNA cleavage at the *miR-140* region.

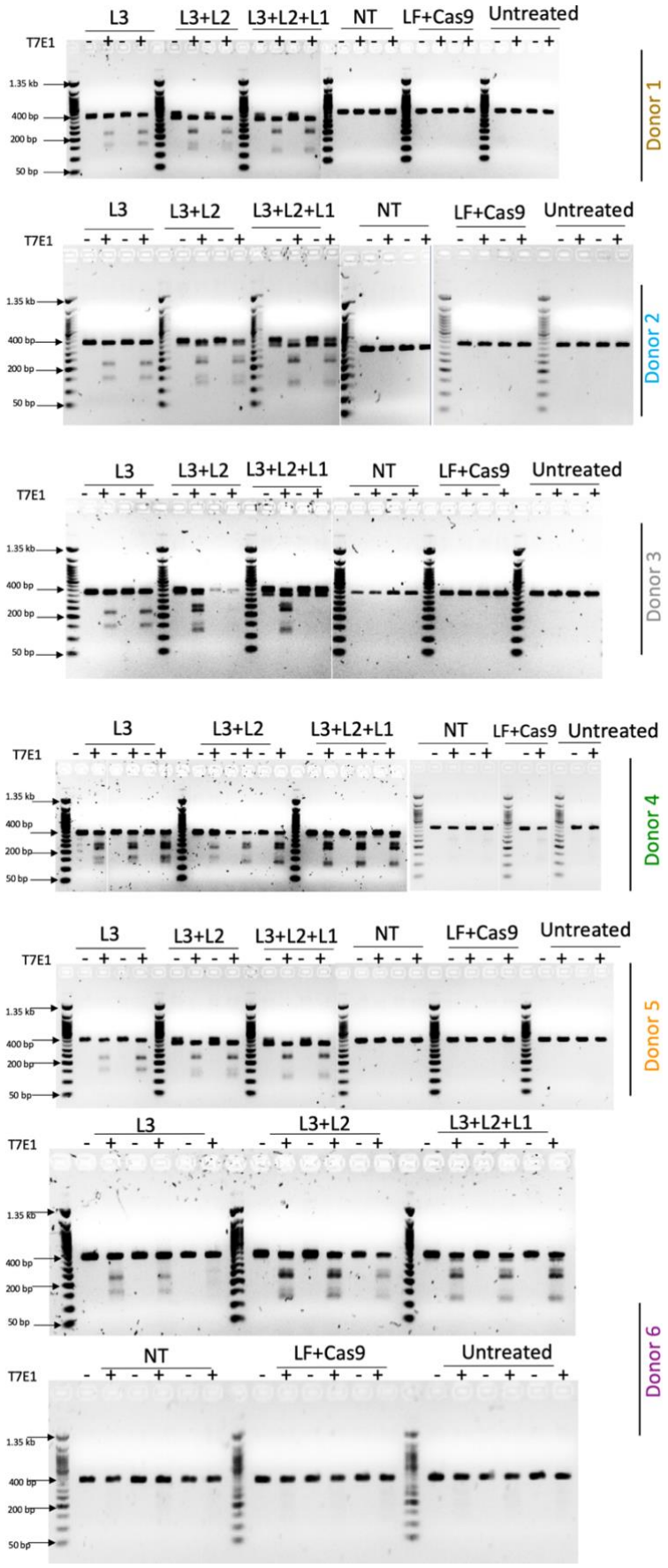

L3
L3+L2
L3+L2+L1
NT
LF+Cas9
Untreated
T7E1
1.35 kb
400 bp
200 bp
50 bp
Donor 1
Donor 2
Donor 3
Donor 4
Donor 5
Donor 6

Fig. 3.5: OA hACs from six different donors display DNA cleavage and miR-140 gene editing after RNP transfection.

After RNP transfecting *miR-140* with CRISPR-Cas9, gDNA was isolated from OA hACs from six different donors 48 h post transfection, and gene editing efficiencies were assessed by the T7E1. For each donor, each experimental condition was run either in duplicate or triplicate. *miR-140* was targeted using a single sgRNA (L3), the combination of two different sgRNAs (L3+L2), or the combination of 3 different sgRNAs (L3+L2+L1). If more than one sgRNA was used, the individual amounts of sgRNA were reduced accordingly to ensure consistency between total final concentrations of sgRNA used. Controls included a non-targeting (NT) sgRNA, only Lipofectamine (LF) and Cas9 treated cells (no sgRNA), as well untreated samples. Donors are colour coded for identification. *n* = 6.

3.2.4 Targeting *miR-140* with CRISPR-Cas9 RNP transfection leads to its downregulation in osteoarthritic human articular chondrocytes

The next aim was to determine how T7E1 gene editing efficiencies of *miR-140* translated into its expression levels. For each miR-140 arm, 15 ng of isolated RNA from *miR-140* targeted OA hACs donors (Fig. 3.5) were reverse transcribed into cDNA, and their expression levels assessed. Samples were normalised to *RNU24* and expressed relative to their normalised NT sgRNA controls. The combination of sgRNA L3+L2 led to the strongest and a statistically significant reduction of *miR-140-3p* levels in OA hACs, by 60 % (*n* = 6), followed by sgRNA L3 alone, which led to a statistically significant reduction of *miR-140-3p* levels by 53 % (*n* = 6) (Fig. 3.6). The triple sgRNA combination L3+L2+L1 led to a statistically significant reduction of *miR-140-3p* levels by 45 % (*n* = 6). Single guide RNA L3 led to the strongest and a statistically significant reduction of *miR-140-5p* levels in OA hACs (by 58 %, *n* = 6), followed by the combination of sgRNA L3+L2 (50 %, statistically significant, *n* = 6) and the triple combination of sgRNA L3+L2+L1 (by 48 %, statistically significant, *n* = 6). The *miR-140* expression levels for both arms of the NT sgRNA control were slightly, but non-significantly decreased, compared with their Lipofectamine+Cas9 (no sgRNA) or untreated cells controls.

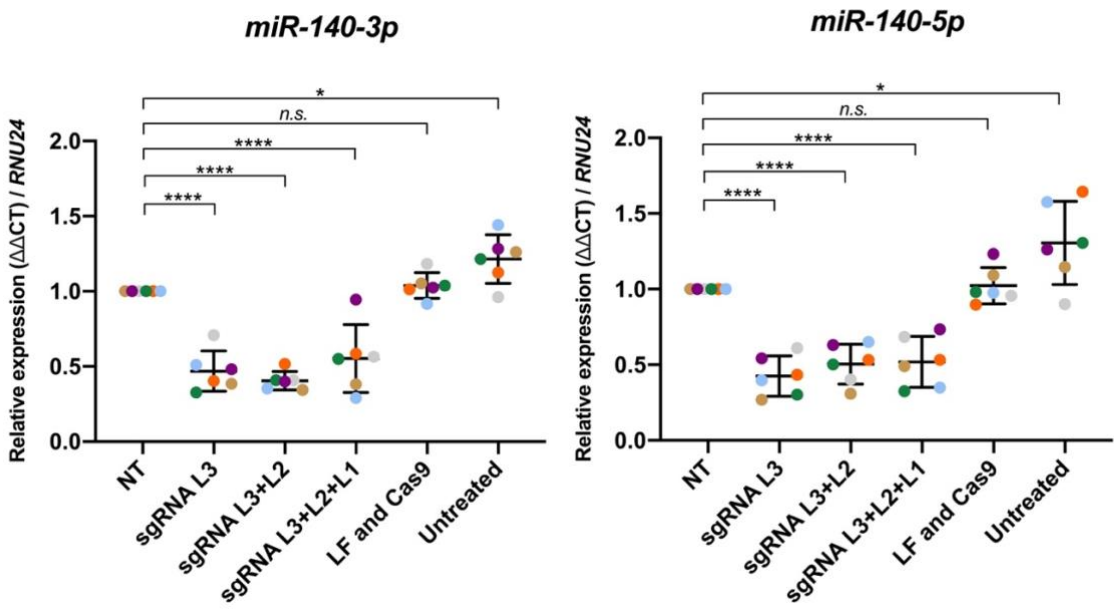


Fig. 3.6: OA hACs from six different donors show downregulation of *miR-140-3p* and *miR-140-5p* after *miR-140*-targeting.

48 h after *miR-140* targeting, RNA was isolated from OA hACs from six different donors, and reverse transcribed to quantify *miR-140-3p* and *miR-140-5p* expression by qPCR.
Expression levels for *miR-140-3p* and *miR-140-5p* were determined after cells were transfected with one sgRNA (L3), the combination of two different sgRNAs (L3+L2), or the combination of three different sgRNAs (L3+L2+L1). If more than one sgRNA was used, the individual sgRNA amounts were reduced accordingly to ensure consistency between total final concentrations of sgRNA used. Controls included a non-targeting (NT) sgRNA, only Lipofectamine (LF) and Cas9 treated cells (no sgRNA), as well as untreated samples. Data are shown as mean ± SD. The statistical significance of comparison between treatment groups were analysed with one-way ANOVA, with post hoc multiple comparisons corrected using Dunnett test. Donor 1: m, 62. Donor 2: m, 64. Donor 3: m, 57. Donor 4: m, 71. Donor 5: f, 70. Donor 6: m, 54. *n* = 6.

3.2.5 CRISPR-Cas9 RNA double RNP transfection increases gene editing efficiency and *miR-140-3p/5p* knockdown in osteoarthritic human articular chondrocytes

OA hACs are not amenable to single cell expansion and they change their phenotype beyond P3. In order to improve gene editing efficiencies for stronger knockdown of expression levels within a cell population, I performed double transfection of primary OA hACs. The rationale being that if one transfection edited approximately 50 % of alleles, a second transfection might edit a further ~50 %, leading to an overall editing efficiency of 75 % of all alleles.

In order to test out this hypothesis, OA hACs from one donor were plated in 12 - well plates and RNP transfected targeting *miR-140* with the same protocol as previously described. Cells were either removed for analysis at 48 h, or medium was exchanged for 10 % / DMEM for 24 h and then re-transfected as before. Cells were incubated for a further 48 h, and subsequently isolated for gDNA and RNA. Genomic DNA (gDNA) was used to PCR amplify a 439 bp long product, flanking the miR-140 hairpin structure. The T7E1 assay was used to compare gene editing efficiencies between single and double transfected cells for each targeting condition. OA hACs that were double RNP transfected exhibited stronger DNA cleavage compared with OA hACs that were only singly RNP transfected (Fig. 3.7A).

RNA samples of each of the transfected OA hACs was reverse transcribed into cDNA, and assessed for *miR-140-3p/5p* expression. Samples were normalised to *RNU24* and

expressed relative to their normalised NT sgRNA controls. *miR-140-3p* expression levels were reduced by 50 % by sgRNA L3, and 60 % by sgRNA L3+L2 and sgRNA L3+L2+L1 for single RNP transfected OA hACs. Double RNP transfected OA hACs showed stronger reduction of *miR-140-3p* levels: sgRNA L3 targeted cells were reduced by 70 %, and sgRNA L3+2 as well as sgRNA L3+2+1 were reduced by approximately 80 % (Fig. 3.7B). Similar ratios also applied to single and double RNP transfected *miR-140-5p* expression levels, in which single RNP transfected cells reduced expression levels by between 50 - 60 %, while double RNP transfected cells reduced expression levels by between 60 - 80 %, with sgRNA L3+L2 and sgRNA L3+L2+L1 causing better reductions in mRNA expression levels compared with sgRNA L3 (Fig. 3.7B).

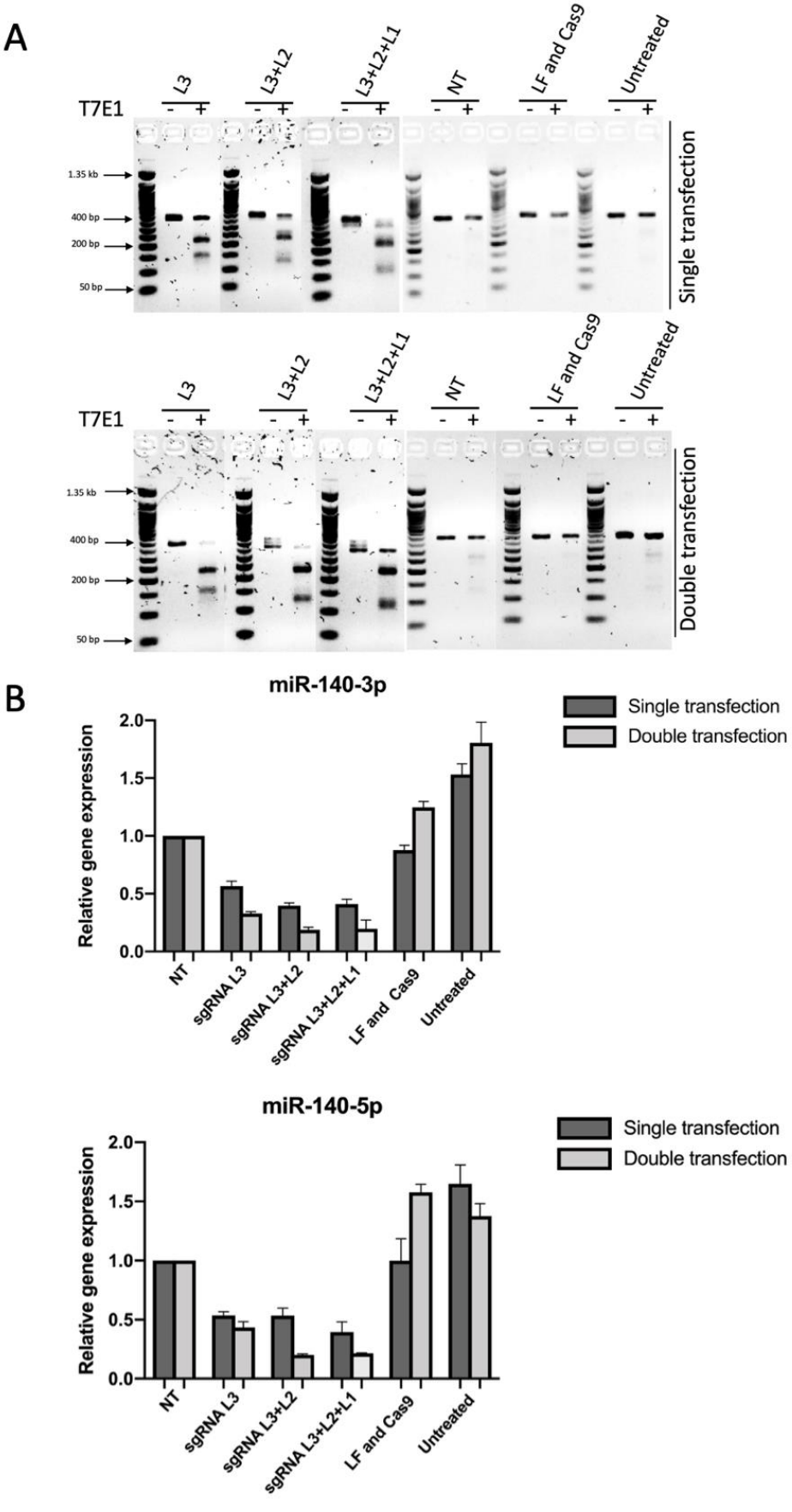

A
T7E1
L3
L3+L2
L3+L2+L1
NT
LF and Cas9
Untreated
- +
1.35 kb
400 bp
200 bp
50 bp
Single transfection
Double transfection
B
miR-140-3p
Relative gene expression
2.0
1.5
1.0
0.5
0.0
NT
sgRNA L3
sgRNA L3+L2
sgRNA L3+L2+L1
LF and Cas9
Untreated
Single transfection
Double transfection
miR-140-5p

Fig. 3.7: CRISPR-Cas9 RNP double transfection results in higher gene editing efficiencies and *miR-140-3p* and *miR-140-5p* knockdowns in OA hACs.

OA hACs were transfected with purified Cas9 in a complex with 1-3 different sgRNAs to target *miR-140*. Cells were either removed for analysis at 48 h, or medium was exchanged for 10 % FBS / DMEM for 24 h and then re-transfected as before. Cells were incubated for a further 48 h at 37 °C, 5 % CO_2. **A** T7E1 assay to determine editing efficiencies. Cells were targeted with one sgRNA (L3), the combination of two different sgRNAs (L3+L2), or the combination of three different sgRNAs (L3+L2+L1). If more than one sgRNA was used, the individual sgRNA amounts were reduced accordingly to ensure consistency between total final concentrations of sgRNA used. Controls included a non-targeting (NT) sgRNA, only Lipofectamine (LF) and Cas9 treated cells (no sgRNA), as well untreated samples. **B** RNA was isolated from the same samples, reverse transcribed, and assessed for *miR-140-3p* and *miR-140-5p* expression by qPCR. Samples were normalised to *RNU24* and expressed relative to their normalised NT sgRNA control. Other controls included only LF and Cas9 treated cells (no sgRNA), as well untreated samples. *n* = 1. Error bars are determined by technical replicates (*n* = 3).

3.2.6 T7E1 confirms strong DNA cleavage at the *miR-140* site in six different osteoarthritic human articular chondrocyte donors after RNP double transfection

As RNP double transfection leads to stronger gene editing efficiencies and therefore better reduction of *miR-140-3p/5p* expression levels, OA hACs from six different donors were plated in 12-well plates. After 48 h, cells were RNP transfected with three different combinations of *miR-140* targeting sgRNAs. After a further 48 h, medium was exchanged for 10 % FBS / DMEM for 24 h and cells subsequently re-transfected as before. Cells were incubated for a further 48 h, and harvested for extraction of gDNA and RNA. A 439 bp long PCR product was amplified with primers flanking the targeted miR-140 region. Gene editing efficiencies were assessed by the T7E1 assay. For each donor, each experimental condition was run in duplicate. Strong DNA cleavage was confirmed for all targeting sgRNAs (either alone or in combination), while gene editing efficiencies within each biological replicate were consistent. Control cells did not exhibit DNA cleavage at the *miR-140* region (Fig. 3.8). Donors are colour coded for identification. *n* = 6.

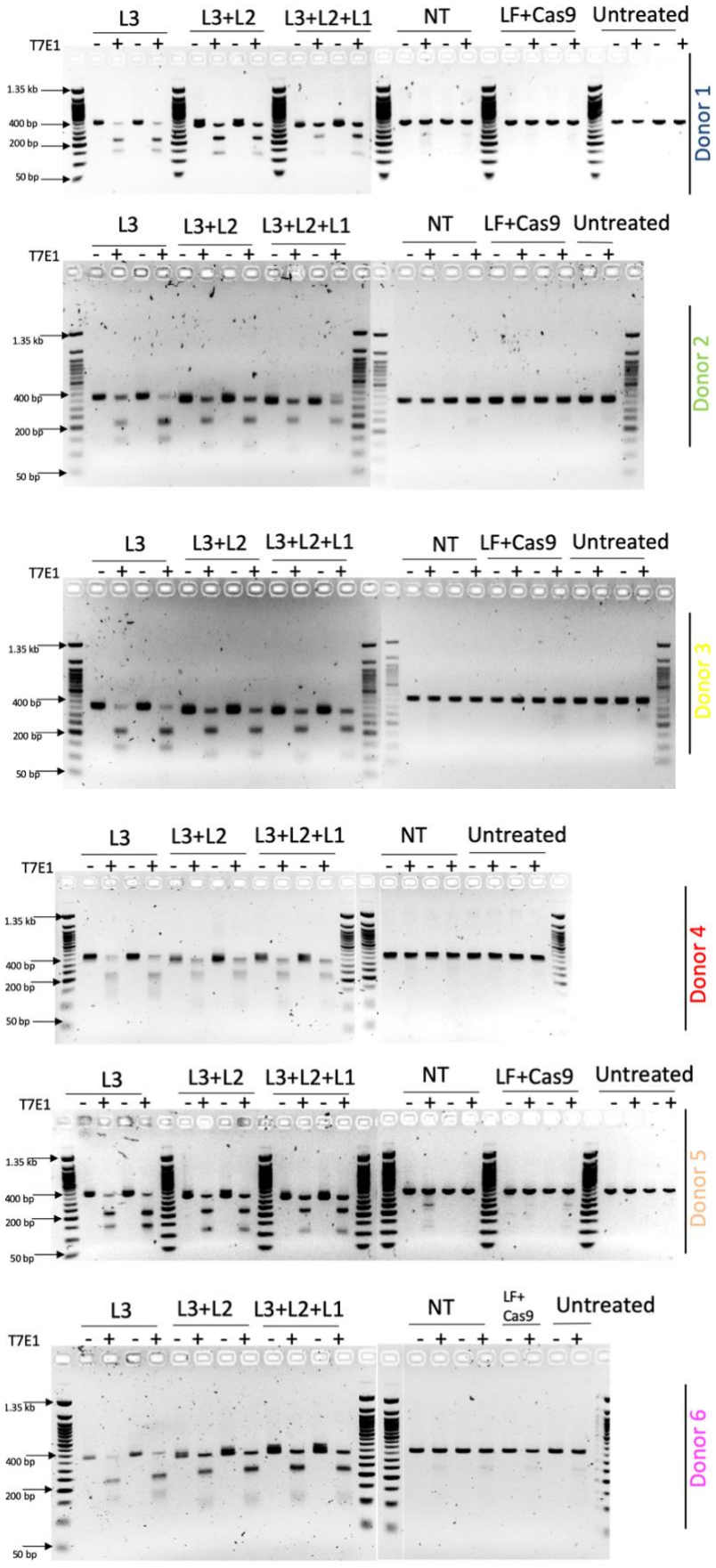

L3
L3+L2
L3+L2+L1
NT
LF+Cas9
Untreated
T7E1
1.35 kb
400 bp
200 bp
50 bp
Donor 1
Donor 2
Donor 3
Donor 4
Donor 5
Donor 6

Fig. 3.8: RNP double transfection targeting *miR-140* in OA hACs from six different donors.

After RNP double transfection, gDNA was isolated from OA hACs from six different donors. 48 h post transfection, cell medium was exchanged for 10 % FBS / DMEM for 24 h and cells re-transfected as before. Cells were incubated for a further 48 h at 37 °C, 5 % CO_2.
Gene editing efficiencies were assessed by the T7E1 assay. For each donor, each experimental condition was run in duplicate. *MicroRNA-140* was targeted using a single sgRNA (L3), the combination of two different sgRNAs (L3+L2), or the combination of 3 different sgRNAs (L3+L2+L1). If more than one sgRNA was used, the individual sgRNA amounts were reduced accordingly to ensure consistency between total final concentrations of sgRNA used. Controls included a non-targeting (NT) sgRNA, only Lipofectamine and Cas9 treated cells (no sgRNA), as well as untreated cells. Donors are colour coded for identification. *n* = 6.

3.2.7 Double RNP transfection leads to improved downregulation of *miR-140* in osteoarthritic human articular chondrocytes

RNA from the same six double transfected OA hACs donors was reverse transcribed to assess how increased T7E1 gene editing efficiency translates into *miR-140-3p* and *miR-140-5p* expression levels. Samples were normalised to *RNU24* and expressed relative to their normalised NT control. The combination of sgRNA L3+L2 led to the strongest and a statistically significant reduction of *miR-140-3p* (by 88 %, *n* = 6) and *miR140-5p* levels (by 85 %, *n* = 6) in OA hACs. The combination of sgRNA L3+L2+L1 led to a statistically significant reduction of *miR-140-3p* by 75 % (*n* = 6), and of *miR-140-5p* by 76 % (*n* = 6). Lastly, the use of sgRNA L3 alone led to a statistically significant reduction of *miR-140-3p* by 64 % (*n* = 6) and *miR-140-5p* by 73 % (*n* = 6) (Fig. 3.9).

The *miR-140* expression levels for both arms of the NT sgRNA control were slightly but non-significantly decreased compared with their Lipofectamine+Cas9 (no sgRNA) and untreated cells controls.

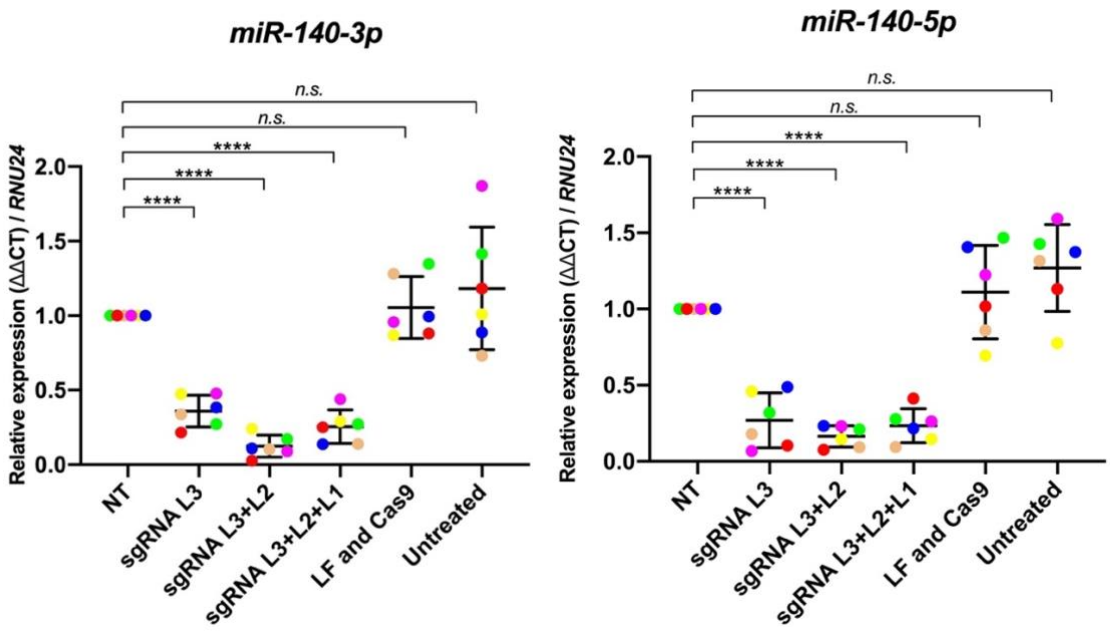


Fig. 3.9: Double RNP transfection results in efficient downregulation of miR-140-3p and *miR-140-5p* in six different OA hACs donors.

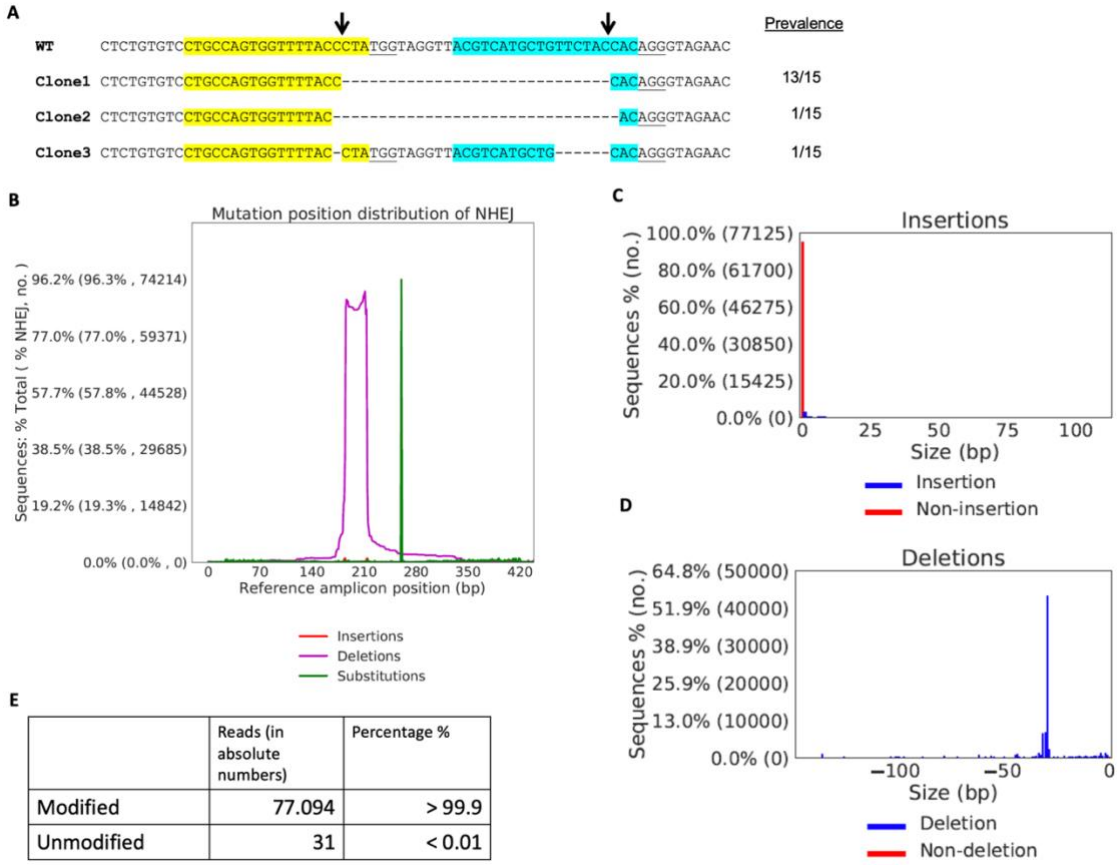
48 h after the first transfection, the cell medium was exchanged for 10 % FBS / DMEM for 24 h, and cells were then re-transfected as before and incubated for a further 48 h.
After RNP double transfection, RNA was isolated from OA hACs from six different donors and reverse transcribed to quantify *miR-140-3p* and *miR-140-5p* expression by qPCR.
MicroRNA-140 was targeted using a single sgRNA (L3), the combination of two different sgRNAs (L3+L2), or the combination of 3 different sgRNAs (L3+L2+L1). If more than one sgRNA was used, the individual sgRNA amounts were reduced accordingly to ensure consistency between total final concentrations of sgRNA used. Samples were normalised to *RNU24* and expressed relative to their normalised NT sgRNA control. Additional controls included only Lipofectamine and Cas9 (no sgRNA) treated cells. The statistical significance of comparison between treatment groups were analysed with one-way ANOVA, with post hoc multiple comparisons corrected using Dunnett test. Donor 1: f, 63. Donor 2: f, 70. Donor 3:m, 78. Donor 4: m, 61. Donor 5: f, 61. Donor 6: m, 69. n = 6.

3.2.8 MiSeq analysis confirms precision of designed *miR-140* targeting sgRNAs in osteoarthritic human articular chondrocytes

To determine how double CRISPR-Cas9 RNP transfection affected the genomic structure of the targeted *miR-140* region, the region of interest was PCR amplified, transformed with a TOPO cloning vector into competent *E. coli,* and plated on LB plates with kanamycin selection marker. 16 h post plating, 15 clones were picked and inoculated in LB media. The following day their DNA was isolated using mini prep and sent for sequencing. Sequencing results revealed that 13 out of 15 clones showed a deletion of 29 bp between the cleavage sites of sgRNA L2 and sgRNA L3 (Fig. 3.10A black arrows). One clone exhibited a deletion of 31 bp, one base pair upstream of the cutting site of sgRNA L2 (highlighted in yellow) and downstream of sgRNA L3 (highlighted in blue). A third clone showed a deletion of 1 bp upstream of the cleavage site of sgRNA L2 and 6 bp downstream of the cleavage site of sgRNA L3.

I amplified a 439 bp long PCR product covering the double RNP targeted *miR-140* region for Illumina MiSeq analysis in a two-step process (Fig. 3.10B). The MiSeq procedure and analysis were carried out by Dr. Damien Downes (Weatherall Institute for Molecular Medicine (WIMM), University of Oxford) in collaboration. At amplicon position 267, a small nucleotide polymorphism (SNP) was detected suggesting a heterogenous cell population. Between amplicon position 191 (the cleavage site of sgRNA L2) and amplicon position 220 (the cleavage site of sgRNA L3), MiSeq analysis confirmed a

deletion of NHEJ induced base pairs for over 90 % of reads (77,125 total reads). Deletions of base pairs, in smaller numbers, were also detected upstream of amplicon position 191 and downstream of amplicon position 220. A very small number of insertions was detected at these positions. In fact, insertions between 1-9 bp occurred but in very small numbers (1 bp insertion: 1,200 reads; 2-9 bp insertion < 100 reads). A significant majority (> 97% of reads) did not exhibit any NHEJ induced bp insertion (Fig. 3.10C). The majority of reads amongst deleted reads (> 90 %) exhibited a deletion of 29 bp or larger, with a peak at -29 bp (55 %) (Fig. 3.10D). Individual reads (< 500) showed bp deletions between -1 bp and -28 bp. The absolute number of modified reads amounted to 77,094 (> 99.9 %), compared with only 31 (< 0.01 %) unmodified reads, suggesting a very high gene editing efficiency through RNP double transfection in OA hACs (Fig. 3.10E).



E

	Reads (in absolute numbers)	Percentage %
Modified	77.094	> 99.9
Unmodified	31	< 0.01

Fig. 3.10: Sequencing confirms the deletion of nucleotides between the cutting sites of sgRNA L3 and L2 at the hairpin structure of *miR-140*.

Genomic DNA of OA hACs, after being targeted by sgRNA L3 and sgRNA L2 in combination to knockdown *miR-140*, was sequenced. **A** 13 out of 15 examined clones show a deletion of 29 bp between the cutting site of sgRNA L2 and sgRNA L1. Two remaining clones show slightly different patterns. Binding site of sgRNA L2 is highlighted in yellow, binding site of sgRNA L3 is highlighted in blue. Respective PAM sequences are underlined, black arrows indicate Cas9 cutting site. **B** Sequence coverage as determined by Illumina MiSeq. Number of sequences are plotted against the unmodified sequence in reference to the amplicon position. Number of NHEJ-induced **(C)** insertions and **(D)** deletions with indication of their size of the insertion / deletion. **E** Number of modified and non-modified reads indicating high efficiency of editing. Figures 9B/C/D were generated by Dr. Damien Downes in collaboration.

After confirming the specificity of the *miR-140* targeting sgRNAs, the next step was to determine whether the designed *miR-140* targeting sgRNAs exhibited any potential off-targeting in OA hACs.

Table 3.1: Bioinformatically predicted 2 mismatch off-targets for *miR-140* targeting sgRNA L1, sgRNA L2, and sgRNA L3, and their loci in the human genome

sgRNA 5'-3'	Sequence	Locus
L1 CTCTCTCTGTGTCCTGCCAG	TTCT-TCTGTGTCCTGCCAG	chr5:+122022578
L1	CTCTCTCTGTGCCCTGCCAA	chr11:-1745319
L1	CTCTTCTCTGTGTTCTGCCAG	chr2:+155357068
L1	CTCTATCT-TGTCCTGCCAG	chr2:+37077756
L1	CTCTCCCTGT-TCCTGCCAG	chr16:+848710
L1	CTCTCTCTTTCTCCTGCCAG	chr7:-113827073
L1	CTCTCTCTGTTGTCCTGCCTG	chr9:-93381813
L1	CTCTCTCTGTGTCCCTCCAG	chr6:18325347
L1	CTCTCTCTGTGTGCTGCCTG	chr10:43213966
L2 CTGCCAGTGGTTTTACCCTA	CTGCCAGTG-TCTTACCCTA	chr3:-72372389
L2	CTGCCAGTGGTTTTACT-TA	chr10:-48497712
L3 ACGTCATGCTGTTCTACCAC	ACATCATGCT-TTCTACCAC	chr21:-30356681
L3	ACGTCTTGCTGTTCCTACCAC	chr11:-26232278
L3	AGGTCAGGCTGTTCTACCAC	chr9:+125984599

Aligning the results of two different off-target algorithm softwares[2], no off-targets with just one sequence mismatch towards sgRNA L3 and L2 were predicted. For two sequence mismatches, two potential off-targets for sgRNA L2 and three potential off-targets for sgRNA L3 were predicted (Table 3.1). Specific primers were designed for each predicted off-target site to be PCR amplified for three different donors (D1, D2 and D3), which had been double RNP transfected with sgRNA L3+L2. T7E1 assay revealed no visible DNA cleavage for any of the predicted sgRNA L3 or L2 two sequence mismatch off-targets, suggesting that sgRNA L3 and L2 were specific to their *miR-140* targets. The likelihood of off-targeting effects of three sequence mismatches was regarded as low and was not experimentally validated. Predicted sgRNA L1 off-targeting for two sequence mismatches were not assessed in OA hACs, as sgRNA L1 samples were not used in subsequent experiments in OA hACs. On-targeting positive controls are shown as a comparison (Fig. 3.11).

[2] IDT - https://www.idtdna.com/site/order/designtool/index/CRISPR_SEQUENCE

SANGER - https://www.sanger.ac.uk/htgt/wge/find_off_targets_by_seq

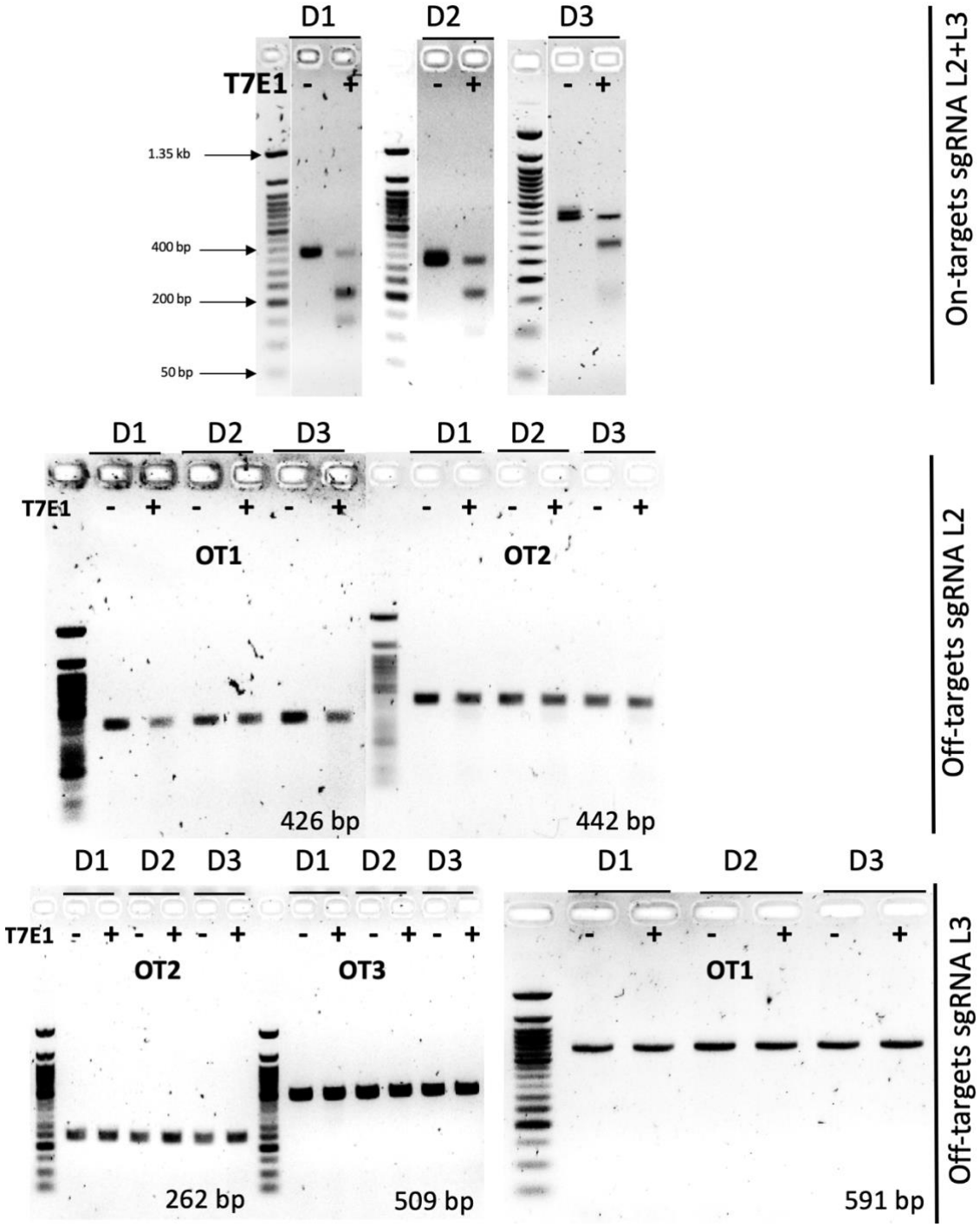


Fig. 3.11: Bioinformatically predicted 2 mismatch off-targets for *miR-140* targeting sgRNA L2 and sgRNA L3 show no off-targeting.

Specific primers were designed for all predicted 2 mismatch off-targets for sgRNA L2 and sgRNA L3. Regions of interest were PCR amplified and their off-targeting assessed by the T7E1 assay for 3 different donors (D1, D2 and D3), which had been transfected with sgRNA L2 and sgRNA L3. On-target controls were performed to assess the gene editing efficiency by amplifying the *miR-140* hairpin structure and subsequent T7E1 gene editing assay.

3.2.9 *miR-140* is expressed in the intronic region of *WWP2*, which does not change expression levels after *miR-140* targeting

WWP2 is the host of *miR-140*. It is located in chromosome 16 of the human genome and encodes for a member of the Nedd4 family of E3 ligases. E3 ligases are involved in protein ubiquitination. The precursor *miR-140* is located between exons 16 and 17 in intron 16 of the *WWP2* gene. To check for *WWP2* expression levels in miR-140 targeted OA hACs, primers were designed, amplifying a 97 bp product in exon 28. *WWP2* was normalised to *RPLP0* and expressed relative to its normalised NT sgRNA control (n = 4). While *miR-140* expression levels significantly decreased (Fig. 3.9), *WWP2* expression levels stayed constant after being targeted by *miR-140* targeting sgRNA L3, or sgRNA L3+L2, or sgRNA L3+L2+L1 (Fig. 3.12).

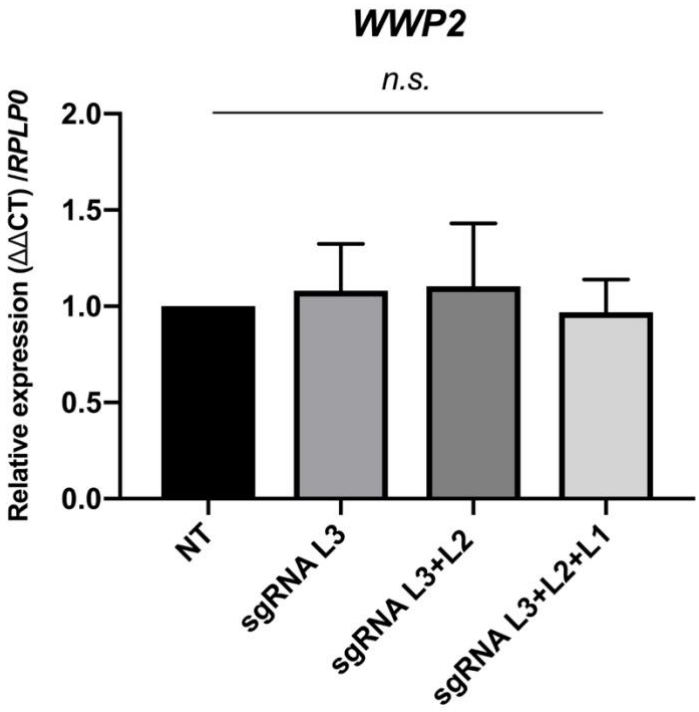


Fig. 3.12: *miR-140* host *WWP2* shows no difference in expression levels.

WWP2, host of *miR-140*, was assessed for expression after being RNP transfected with miR-140 targeting sgRNA L3, sgRNA L3 + L2 or sgRNA L3 + L2 + L1. *WWP2* was normalised to RPLP0 and expressed relative to its normalised NT (non-targeting sgRNA) control. Data are shown as mean ± SD. The statistical significance of comparisons between treatment groups were analysed with one-way ANOVA, with post hoc multiple comparisons corrected using Dunnett test. *n* = 4.

3.2.10 Injury to mouse cartilage reveals downregulation of miR-*140-3p* and *miR-140-5p*

An important focus in our lab is the molecular response of articular cartilage to injury. Injury is the most important aetiological factor in OA development, and mechanical injury *in vitro* is sufficient to activate molecular signalling and drive a number of pathways involved in matrix turnover and repair. To investigate whether *miR-140-3p* and *miR-140-5p* were affected by injury, mouse hips were injured by avulsion of the femoral head from the femur of 6 weeks old mice. This procedure was performed by a senior post-doc in the lab, Dr. Linyi Zhu. The mouse model was used as it is accessible and robust, and due to difficulties performing these types of experiments in human tissue.

Upon injury, both *miR-140-3p* and *miR-140-5p* were downregulated at 4 h compared with 0 h, albeit statistically significant only for *miR-140-3p (p* = 0.0249*; miR-140-5p p* = 0.0541) (Fig. 3.13), suggesting that *miR-140* plays a role in mouse injury response.

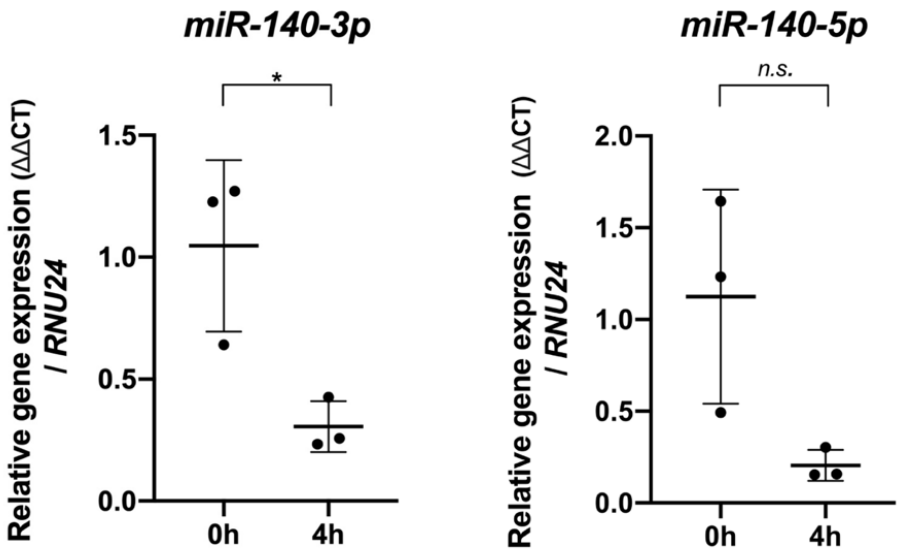


Fig. 3.13: *miR-140-3p* and *miR-140-5p* are downregulated upon injury to mouse cartilage.

Hips were explanted from WT mice aged 6 weeks, and snap frozen in liquid nitrogen (0 h) or cultured in serum-free media for 4 h at 37 °C, 5 % CO_2. RNA was reverse transcribed, and *miR-140-3p* and *miR-140-5p* expression were quantified by qPCR, normalised against β2M, and expressed relative to "0h". Data are shown as mean ± SD. The statistical significance of comparisons were analysed with Student's t-test (two-tailed). *n* = 3.

3.2.11 *miR-140* knockdown in five donors of osteoarthritic human articular chondrocytes reveals statistically significant upregulation of three genes

In line with our observations of *miR-140* being regulated during injury, a TaqMan Microfluidic gene expression card was designed to include genes that were either identified *miR-140* targets, or those that had previously been shown by our group to be strongly regulated by injury (Table 3.2). These included genes involved in cartilage repair pathways, retinoic acid metabolism, chondrogenesis, cilia biology, and cartilage catabolism. A total of 46 genes were investigated in OA hACs (n = 5) for their expression after *miR-140* KD. All genes were normalised to *RPLP0* and expressed relative to their respective NT controls, and displayed according to their statistical significance starting from left (Fig. 3.14). *SEPT2* (Septin 2, 2-fold), *BMP2* (Bone Morphogenetic Protein 2, 5.5-fold) and *RARG* (Retinoic Acid Receptor Gamma, 2.5-fold) were upregulated upon *miR-140* KD and displayed statistical significance after correcting for multiple comparisons (Table 3.3). Both *SEPT2* and *BMP2* were previously predicted targets of *miR-140*, hence validating the robustness of our system.

Table 3.2: Chosen gene targets for TaqMan Microfluidic gene expression card and *miR-140* KD osteoarthritic, human articular chondrocyte donors. Note: Some genes have more than one function, hence they were allocated towards two functional groups.

Function	Genes	References
Direct *miR-140* targets	SEPT2, BMP2, FGF2, VEGFA, FZD6, HDAC4, DNMT1, RALA, NRIP1, PDGFRA, IGFBP5	SEPT2: (Yu et al., 2016) BMP2: (Hwang et al., 2014) FGF2: (Ornitz and Itoh, 2015) VEGFA: (Lu et al., 2017) FZD6: (Barter et al., 2015) HDAC4: (Vega et al., 2004) DNMT1: (Takata et al., 2013) RALA: (Karlsen et al., 2013) NRIP1: (Asano et al., 2017) PDGFRA: (Lan et al., 2015) IGFBP5: (Tardif et al., 2009)
Cilia biology	SEPT2, IFT88, TTBK2	(Yu et al., 2016) (Coveney et al., 2018) (Wann, unpublished)
Retinoic acid	RARG, CYP26B1, CRABP2, RARA, RARB, ALDH1A2, CYP19A1, CYP26A1, AGT	(Zhu et al., 2018) (Sparks et al., 2011)

Anabolic genes	BMP2, AGRN, FGF2, BMP6, COL1A1, SOX9, COL10A1, COL2A1	(Hwang et al., 2014) (Eldridge et al., 2016) (Ornitz and Itoh, 2015) (Bi et al., 1999) (Sophia Fox et al., 2009)
Pain	NGF	(von Loga et al., 2019)
Injury response	FGF2, FGFR1, TGFA, TGFBR2, TGFBR3, TGFBR1, CTGF, FGFR3	(Tang et al., 2018) (Vincent et al., 2002)
Cartilage homeostasis	TNFAIP6, AGRN, TIMP3, COL1A1, INHBA, SOX9, COL10A1, LRP1, MMP13, ADAMTS5, COL2A1	(Liew et al., 2008) (Eldridge et al., 2016) (Kevorkian et al., 2004) (Sophia Fox et al., 2009) (Wei et al., 2010) (Bi et al., 1999) (Yamamoto et al., 2013) (Neuhold et al., 2001) (Glasson et al., 2005)

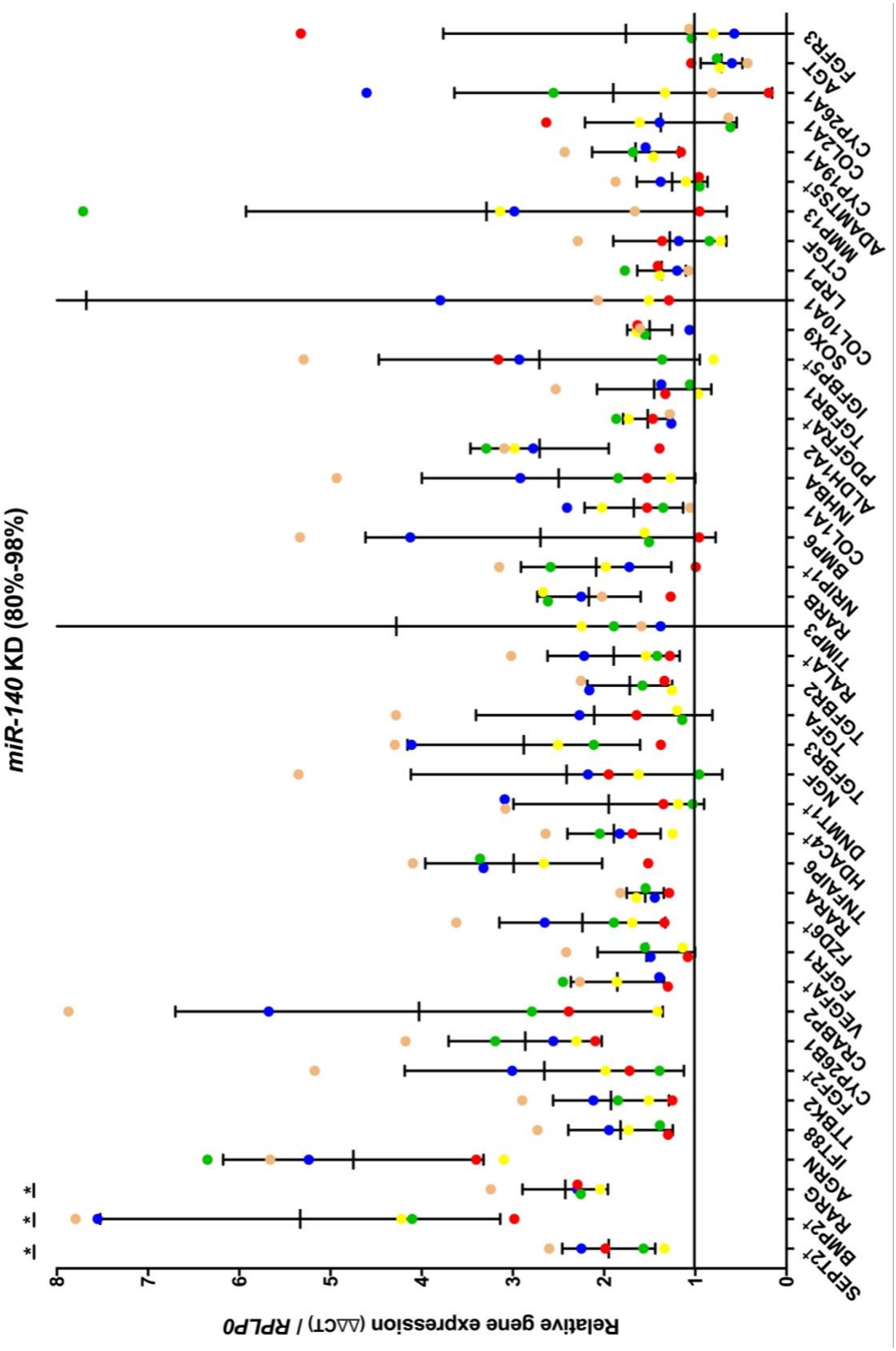

miR-140 KD (80%-98%)
Relative gene expression (ΔΔCT) / RPLP0
0
1
2
3
4
5
6
7
8
* * *
SEPT2†
BMP2†
RARG
AGRN
IFT88
TBK2
FGF2†
CYP26B1
CRABP2
VEGFA†
FGFR1
FZD6†
RARA
TNFAIP6
HDAC4†
DNMT1†
NGF
TGFBR3
TGFA
TGFBR2
RALA†
TIMP3
RARB
NRIP1†
BMP6
COL14A1
INHBA
ALDH1A2
PDGFRA†
TGFBR1
IGFBP5†
SOX9
COL10A1
LRP1
CTGF
MMP13
ADAMTS5†
CYP19A1
COL2A1
CYP26A1
AGT
FGFR3

Fig. 3.14: Genes involved in anabolic pathway repair, retinoic acid pathway and cilia biology are upregulated upon *miR-140* knockdown.

RNA was obtained from donors with 80 % - 96 % knockdown of *miR-140-3p* and *miR-140-5p* in OA hACs. Gene expression of 46 genes was analysed via qPCR on pre-designed microfluidic cards. All genes were normalised to *RPLP0* and expressed relative to their respective NT (non-targeting sgRNA) controls. Genes are displayed according to their statistical significance from left to right. Gene names marked with "†" are targets of *miR-140* previously experimentally validated by other groups. Each colour represents the same donor. Statistical significance determined using the Bonferroni-Dunn method with corrections for multiple comparison. $n = 5$.

3.2.12 Modest *miR-140* knockdown does not alter gene regulation in five donors of osteoarthritic human articular chondrocytes

Another step of validation included re-assessing the expression levels of the same genes by using RNP single transfected OA hACs from OA donors (*n* = 5) with only 40 – 70 % *miR-140* knockdown. All genes were normalised to *RPLP0* and expressed relative to their respective NT controls, and were displayed in the same order as previously. Not a single gene was statistically upregulated upon partial *miR-140* knockdown (Fig. 3.15). Those genes, including those previously statistically upregulated by *miR-140* KD (*SEPT2*, *BMP2,* and *RARG*) were not regulated, suggesting that high *miR-140* knockdown is required to regulate susceptible chondrocyte genes.

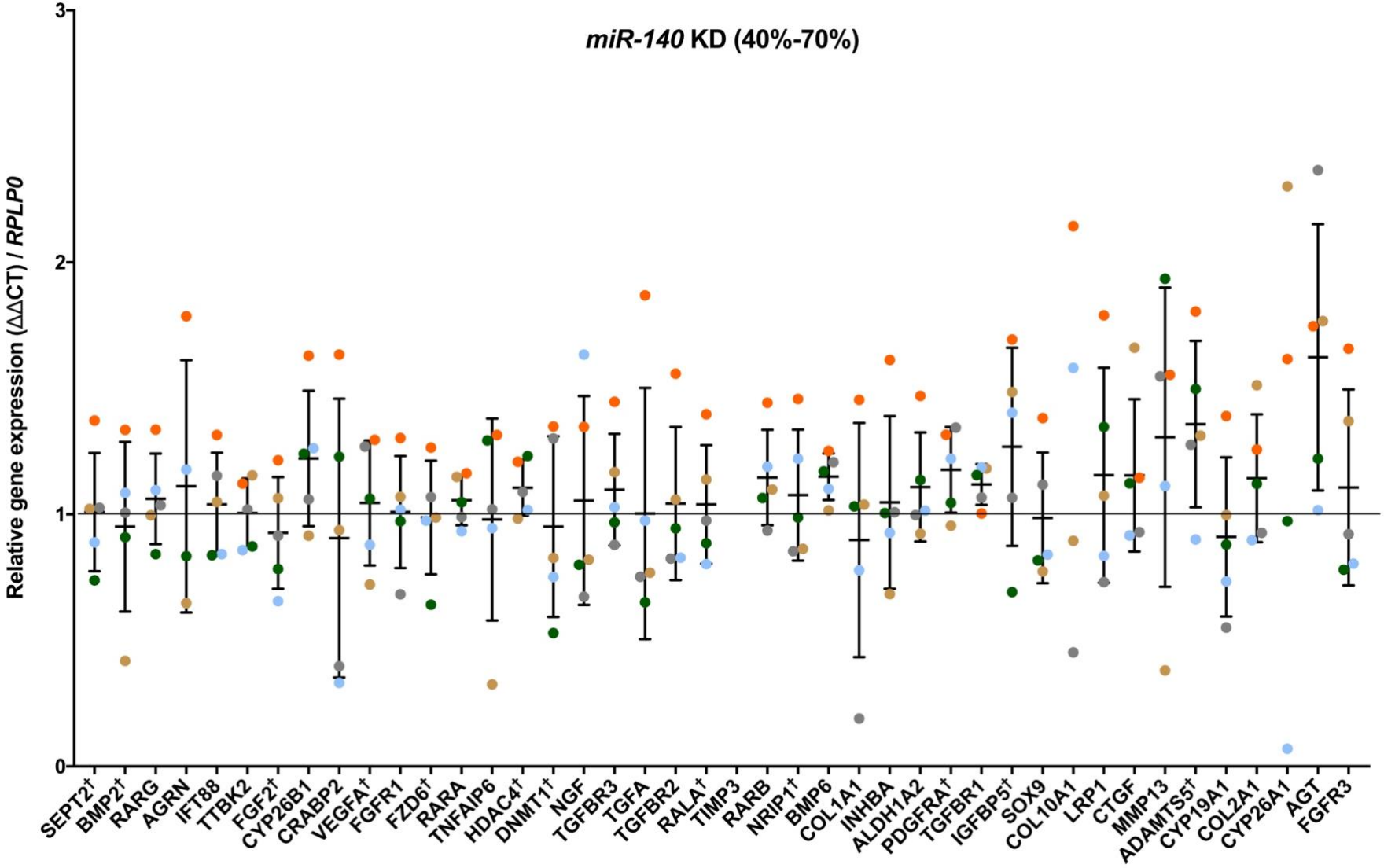

miR-140 KD (40%-70%)
Relative gene expression (ΔΔCT) / RPLP0
3
2
1
0
SEPT2†
BMP2†
RARG
AGRN
IFT88
TTBK2
FGF2†
CYP26B1
CRABP2
VEGFA†
FGFR1
FZD6†
RARA
TNFAIP6
HDAC4†
DNMT1†
NGF
TGFBR3
TGFA
TGFBR2
RALA†
TIMP3
RARB
NRIP1†
BMP6
COL1A1
INHBA
ALDH1A2
PDGFRA†
TGFBR1
IGFBP5†
SOX9
COL10A1
LRP1
CTGF
MMP13
ADAMTS5†
CYP19A1
COL2A1
CYP26A1
AGT
FGFR3

Fig. 3.15: Modest (inefficient) downregulation of *miR-140* does not alter gene regulation.

RNA was obtained from donors with 40 % - 70 % knockdown of *miR-140-3p* and *miR-140-5p* in OA hACs. Gene expression of 46 genes was analysed via qPCR on pre-designed microfluidic cards. All genes were normalised to *RPLP0* and expressed relative to their respective NT (non-targeting sgRNA) controls. Genes are displayed according to their statistical significance from left to right. Genes marked with "†" are targets of *miR-140* previously experimentally validated by other groups. Each colour represents the same donor. Statistical significance determined using the Bonferroni-Dunn method with corrections for multiple comparisons. $n = 5$.

Table 3.3: List of genes that show significant (*) upregulation and tendency of upregulation upon deletion of *miR-140* in osteoarthritic, human articular chondrocytes.
Student's t-test (two-tailed) (P value) and multiple comparison testing (P value adjusted) were performed. Compared to 7-day-old miR-140 KO mice determined by RNA-sequencing. Presence of miR-140 seed sequences displayed.

	***miR-140* KD in osteoarthritic, human articular chondrocytes**		**Costal RNA from 7-day-old-*miR-140* KO mice**		**Presence of *miR-140* seed sequence**
GENE	**P value**	**P value adjusted**	**P value**	**P value adjusted**	
SEPT2 *	0.0006	0.0268	1.03E-11	7.81E-10	140-5p
BMP2 *	0.0008	0.0349	0.9729	0.9880	140-5p
RARG *	0.0009	0.0405	0.29772	0.5394	-
AGRN	0.0019	0.0830	4.75E-06	9.15E-05	-
IFT88	0.0023	0.1005	0.0002	0.0024	-
TTBK2	0.0029	0.1244	0.7826	0.8992	-
FGF2	0.0038	0.1632	0.0033	0.0219	140-5p
CYP26B1	0.0089	0.3750	1.71E-19	4.49E-17	-

Further validation of these targets was examined by studying RNA sequencing (RNA-seq) analysis of costal cartilage RNA from 7-day old miR-140 KO mice (in collaboration with Prof. David Young, Newcastle University) (Table 3.3).

SEPT2, our top hit, showed statistical significance in both systems, while *AGRN*, *IFT88*, *FGF2* and *CYP26B1* (all genes showing tendency towards upregulation in our study) were statistically significantly regulated in the RNA-seq analysis of costal RNA from 7-day old *miR-140* KO mice. Of the top 8 candidate targets from our study (including some where the adjusted *p* value did not reach statistical significance, $p < 0.05$), 5 genes were strongly regulated in the *miR-140* KO mouse.

Interestingly, only *SEPT* and *FGF2* contain a *miR-140* seed sequence in their 3'-UTRs, suggesting that all other targets are indirectly regulated by *miR-140*.

3.3 Discussion

miRNAs are small, non-coding RNAs that regulate gene expression at post-transcriptional/translational levels. In this chapter I applied a state-of-the-art technology to study miRNA loss-of-function, by CRISPR-Cas9 mediated knock down of *miR-140* in OA human articular chondrocytes *in vitro*.

3.3.1 Efficient CRISPR-Cas9 mediated gene editing in osteoarthritic human articular chondrocytes

RNP transfection containing Cas9 and a sgRNA was established as an effective way to transfect primary cells. RNP transfection of OA human articular chondrocytes yielded high gene editing efficiencies which were confirmed by T7E1 assays and MiSeq analysis. Additionally, I used double transfection to increase gene editing efficiencies. This was particularly important in the cell type we used, as OA human articular chondrocytes transition towards a "fibroblast-like" phenotype in monolayer with progressive passaging (Brodkin et al., 2004), and are not amenable to single cell expansion. High gene editing efficiencies are required in such populations to effectively knock down a gene. Similarly, I demonstrated that, by carefully designing, selecting, and combining various sgRNAs, the gene editing impact on a given gene can be optimised.

There were some discrepancies in gene editing efficiencies for a given donor between the T7E1 gene editing assay, the qPCR read-outs, as well as the MiSeq analysis. In general, the editing effectiveness of the T7E1 read-out were less prominent compared

with the expression levels of miR-140-3p/5p, which were determined by qPCR for each donor. At the same time, MiSeq analysis demonstrated high gene editing efficiencies (> 99.9 %) for the donors that were sequenced.

The T7E1 assay has a number of limitations that others have documented previously (Sentmanat et al., 2018). It is important to understand the mechanism by which the T7E1 assay works. The T7E1 assay is initiated by the denaturation of PCR-amplified DNA from CRISPR-Cas9 targeted cells. During the slow re-annealing process, CRISPR-Cas9 mutated DNA forms mismatches with wildtype DNA. The T7E1 recognises these mismatches and cleaves the DNA at these sites to generate bands of different lengths. Importantly, during the process of re-annealing, CRISPR-Cas9 mutated DNA can form mismatches with CRISPR-Cas9 mutated DNA, hence underestimating the actual gene editing read-out.

At the same time, Ilumina MiSeq sequencing employs paired-end sequencing and alignment: both ends of the DNA fragment are sequenced and aligned to generate full sequencing across the amplicon. Each read corresponds to a single PCR amplicon, although each read may not be a unique mutation event due to duplication through amplification. The sequencing patterns were consistent with the anticipated CRISPR-Cas9 affect. However, it is likely that gene editing efficiencies are overestimated by MiSeq. One reason for this can be the presence of a common polymorphic variant at amplicon position 267, resulting in homology mismatch between the endogenous sequence and the reference sequence (Fig. 3.10). This single nucleotide polymorphism (SNP) was present in each of the four donors, so this may account for some of the apparent NHEJ scores. Based on the allele table, ~90 % of the deletions were at the

predicted Cas9 target sites for the two sgRNAs for all donors. This is probably a slight underestimation as there are also cleavages occurring up- and down-stream of these sites. The qPCR readouts suggest the actual efficiency may be between 90 - 98 %.

3.3.2 CRISPR-Cas9 mediated off-targeting in osteoarthritic human articular chondrocytes

Off-targeting takes place when Cas9 creates DNA double strand breaks at places other than the anticipated locus. Off-targets are usually harmless unless they occur in the coding or regulatory regions, through which they could result in cellular changes. The impact of Cas9 off-targeting can be controlled by using sgRNAs that possess less sequence similarity to other sites in the genome. Putative off-targets were identified by searching the genome for sequences with up to 2 mismatches. sgRNA selection had already excluded single sequence mismatches. Off-targeting of the putative mismatches was not evident for either sgRNAs. As the T7E1 assay is not highly sensitive, this result does not exclude there being low numbers of cells where off-targeting has occurred. Additionally, the ability to detect off-targeting is likely to be affected by target area accessibility because the chromatin binding area is structurally less open in the assay chondrocytes (Singh et al., 2015).

3.3.3 Identification of *miR-140* target genes in osteoarthritic human articular chondrocytes

I observed that, upon cartilage injury, both *miR-140-3p* and *miR-140-5p* were downregulated, suggesting that *miR-140* may contribute to the molecular injury response. In line with our observation of *miR-140* being regulated during injury, we decided to investigate *miR-140* KD responsiveness of chondrocyte genes that are known to be strongly regulated by injury, in addition to genes that had previously been identified *miR-140* targets by other groups. Several of these genes were upregulated in OA human articular chondrocytes upon *miR-140* KD. These included genes involved in the retinoic acid pathway (*RARG, CYP26B1*, and *CRABP2*), primary cilia biology (*SEPT2, IFT88,* and *TTBK2*), and *FGF2*, a growth factor released upon cartilage injury in mouse (Vincent et al., 2002). I am next going to briefly elaborate on the role of these genes within their respective pathways.

3.3.3.1 Retinoic acid pathway

Our lab has previously reported that injury to murine and porcine cartilage strongly suppresses retinoic acid (RA) dependent genes, including *CYP19A1* (in murine), *CYP26A1* and *CYP26B1* (in porcine), as well as all three retinoic acid receptors (*RAR*), RAR-alpha (*RARA*), RAR-beta (*RARB*) and RAR-gamma (*RARG*) (in murine and porcine) (Zhu et al., 2018). The cytochrome P450 enzymes of the CYP26 family (*CYP26A1* and *CYP26B1* possess 40 % sequence similarity with overlapping expression profiles) metabolize RA to reduce cellular levels (Topletz et al., 2012). They are believed to be induced by RA

binding to RAR (Lampen et al., 2001). RARs directly interact with cellular retinoic acid binding protein 2 (CRABP2), which carries retinoic acid from the cytoplasm to the nucleus after binding to it (Takazawa et al., 2019).

CYP19A, another member of the cytochrome P450 superfamily, and a gene which did not appear to be regulated in our experiment, encodes for Aromatase, an estrogen synthase (Bulun et al., 1994).

An anti-inflammatory role for RA is suggested in the literature, including a paper by Nozaki et al. (2006). This group showed reduced macrophage infiltration and 34 % less histopathological evidence of joint damage after RA treatment in mice with collagen-induced arthritis. None of the *miR-140* regulated RA genes contained a *miR-140-3p* or *miR-140-5p* seed sequence, suggesting that these genes were indirectly regulated by *miR-140*.

3.3.3.2 The primary cilium

Three of the genes I identified as being regulated by *miR-140* are associated with the primary cilium, which has previously been linked with OA through its established role in mediating hedgehog signalling. It may also act as a mechano-effector molecule in cartilage, as has been reported in other tissues such as the kidney.

Disruption of Intraflagellar transport protein 88 (IFT88), which is a key ciliary trafficking protein, increases aggrecanase activity in chondrocytes *in vitro,* accompanied with changes to the distribution of the endocytosis receptor LRP1 over the cell membrane. This suggests that IFT88 controls protease activity by optimising LRP1 mediated

endocytosis, and indeed, LRP1 is co-localised with IFT88 in WT cells (Coveney et al., 2018).

Another gene I identified to be altered by *miR-140* knockdown was *SEPT2*. *SEPT2* encodes for Septin 2, which is a filamentous GTPase that directly binds to myosin II, a molecular motor driving muscle contraction (Joo et al., 2007). Inhibition of the SEPT2-myosin II interaction in dividing cells leads to loss of stress fibres (Joo et al., 2007). Additionally, Ghossoub et al. (2013) reported that SEPT2 forms equal complexes with SEPT7 and SEPT9, and these complexes are concentrated along the axoneme (central cilium strand) in retinal pigmented epithelial (RPE) cells. Deletion of this complex resulted in the inhibition of ciliogenesis in retinal pigmented epithelial cells as well as in kidney cells. The role of SEPT2 in cartilage is unclear. It possesses a seed region in its 3'-UTR for *miR-140-5p*. Luciferase assays confirmed that *SEPT2* levels are directly regulated by *miR-140-5p* in primary tumor specimens and cell lines (Yu et al., 2016).

Another critical regulator of ciliogenesis is Tau Tubulin Kinase 2 (TTBK2), a serine/threonine kinase (Bowie and Goetz, 2019). *TTBK2* is reported to regulate chondrocyte protease activity through modifying inflammatory signalling (Wann, unpublished). Neither *TTBK2* nor *IFT88* possess *miR-140-3p/5p* seed sequences, suggesting they are indirectly regulated by *miR-140*.

3.3.3.3 ADAMTS-5

Contrary to previous reports, *ADAMTS-5* was not regulated in *miR-140* KD hACs in my hands, even though the 3'-UTR of *ADAMTS-5* contains a putative seed sequence for *miR-140-3p*. Using a *miR-140* mimic for overexpression, Miyaki et al. (2009) reported that

miR-140 directly regulates *ADAMTS-5 in vitro*. In their study, the authors used both healthy and OA human articular chondrocytes although it is unclear as to which one they used in their overexpression experiment. Their *in vitro* experiment, which was carried out in the presence of IL-1β, does not sufficiently support the notion of *ADAMTS-5* being regulated by *miR-140*. For example, it does not discuss the extent to which *miR-140* was upregulated or the possibility of off-target effects. Importantly, it neglects the increased levels of *ADAMTS-5* upon transfection with a non-targeting inhibitor control, which was then compared with a *miR-140* mimic to demonstrate a reduction in *ADAMTS-5* levels. Our work has also failed to show *miR-140-ADAMTS-5* regulation *in vitro* using *miR-140* mimics and/or antagomirs (Seidl / Yamamoto, unpublished).

In a later study, Miyaki et al. (2010) showed that miR-*140* KO mice developed an OA-like phenotype with age, which was characterised by a short stature, proteoglycan loss, fibrillation of articular cartilage, and elevated *ADAMTS-5* expression levels.

3.3.3.4 Anabolic pathway genes

FGF2 is a heparin-bound molecule, that is released from the PCM upon mechanical stress (Vincent et al., 2002). The mechanism for this involves a rapid sodium flux from the sulfated, aggrecan-rich matrix (Keppie, unpublished). This FGF2-mediated injury response in chondrocytes delays the development of OA in mice (Chia et al., 2009, Chong et al., 2013). I have shown the *FGF2* gene to be upregulated upon *miR-140* loss, which in principle contradicts the reported *miR-140* chondroprotective role. However, when studied in articular chondrocytes *in vitro*, FGF2 has been reported to exert both anabolic and catabolic effects, presumably by acting through different receptors: FGFR1

is associated with catabolic responses, while FGFR3 enhances anabolic responses (Ornitz and Itoh, 2015). *FGF2* possesses a *miR-140-5p* seed in its 3'-UTR, and so is predicted to be a direct target of *miR-140*. I will elaborate further on *FGF2* and its contradictory roles in the next discussion chapter.

Similar to *FGF2*, our group has reported that *CTGF* controls activation of latent *TGFβ* in the presence of *TGFβR3,* and their release upon cartilage injury (Tang et al., 2018). TGFβ, a key growth factor, is implicated in OA progression but is also a vital chondrogenic driver (Zhen et al., 2013, van der Kraan, 2017). In my experiments, neither *CTGF* nor *TGFβ* were regulated upon CRISPR-Cas9-mediated *miR-140* KD in OA human articular chondrocytes.

BMP2 (directly) and *AGRN* (indirectly) were also regulated by *miR-140* in my study. As they both have chondrogenic actions (Kramer et al., 2000, Schmitt et al., 2003, Majumdar et al., 2001, Eldridge et al., 2016), they could contribute towards enhancing chondrogenesis / repair. This is explored further in the next chapter.

Results

4

Targeting *miR-140* with CRISPR-Cas9 in human mesenchymal stem cells.

Chapter 4. Targeting *miR-140* with CRISPR-Cas9 in human mesenchymal stem cells

4.1 Introduction

In the previous chapter we established that *miR-140* indirectly regulates three anabolic pathway genes in osteoarthritic human articular chondrocytes (OA hACs): *BMP2*, *FGF2* and *AGRN*. The contributions of *BMP2* (Kramer et al., 2000, Schmitt et al., 2003, Majumdar et al., 2001) and *FGF2* (Lee et al., 2013, Ito et al., 2008) to the formation of cartilage during the process of chondrogenesis have previously been described. Additionally, *FGF2* has been demonstrated to enhance cell velocity in human mesenchymal stem cells (hMSCs) and thus contributes to cartilage repair (Muhammad, Khan et al., in preparation). *AGRN* encodes for Agrin, a proteoglycan component of cartilage that has a purported role in repair (Eldridge et al., 2016).

Therefore, to test roles of these *miR-140* regulated, pro-chondrogenic genes, I next assessed the effect of *miR-140* knockdown on chondrogenesis using human mesenchymal stem cells (hMSCs). The role of *miR-140* and its upregulation during chondrogenesis has also been reported previously by Miyaki et al. (2009) and Karlsen et al. (2013).

The aim of this chapter was (i) to therefore confirm that *miR-140-3p/5p* were regulated during chondrogenesis *in vitro*, (ii) to establish knockdown of *miR-140* in hMSCs and (iii) to comprehend whether *miR-140* KD in hMSCs affects the process of chondrogenesis and activity of hMSCs.

4.2 Results

4.2.1 *miR-140-3p* and *miR-140-5p* are both expressed in human mesenchymal stem cells and are regulated during chondrogenic differentiation

Before targeting *miR-140*, it was important to check if both arms of *miR-140*, namely *miR-140-3p* and *miR-140-5p*, were expressed in hMSCs, and how those expression levels compared with OA hACs.

For each miRNA arm, 15 ng RNA were reverse transcribed and assessed using miRNA specific TaqMan probes. Samples were normalised to *RNU24* and are displayed as ΔCT values (Fig 4.1). On average, *miR-140-3p* (*n* = 16) demonstrated a slightly higher expression compared with *miR-140-5p* (*n* = 7), while both arms were significantly less expressed compared with OA hACs. The difference in n numbers was due to technical failure in obtaining a readout for *miR-140-5p* in some donors. All readouts obtained for *miR-140-5p* exhibited a donor-matched readout for miR-140-3p.

Raw CT values for both microRNA arms varied between 32.5 and 34, which is very low.

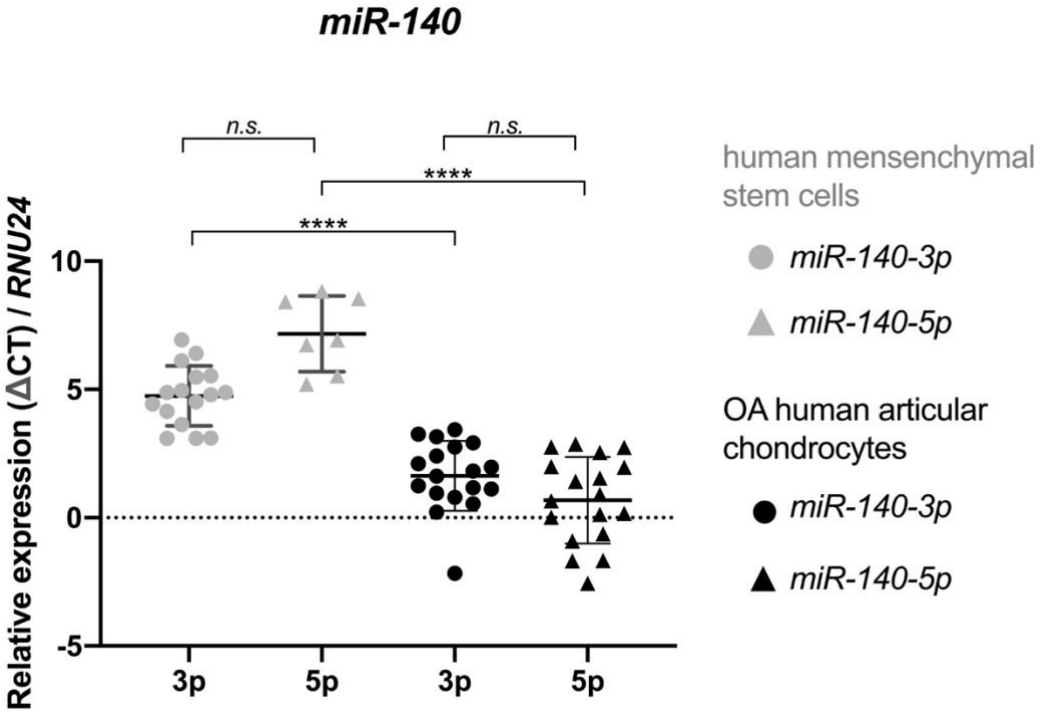


Fig 4.1: Expression of *miR-140-5p* and *miR-140-3p* in human mesenchymal stem cells (hMSCs) compared with osteoarthritic, human articular chondrocytes.

RNA was purified from hMSCs and reverse transcribed, and gene expression of *miR-140-3p* and *miR-140-5p* assessed by qPCR and compared with isolated OA hACs. Samples were normalised to *RNU24* and are displayed as ΔCT values. hMSCs *miR-140-3p*: *n* = 16. *miR-140-5p*: *n* = 7. OA hACs miR-*140-3p* and *miR-140-5p*: *n* = 18.

To investigate whether *miR-140-3p/5p* were regulated during chondrogenesis, RNA samples were obtained from a senior postdoc in the lab, Dr. Hayat Muhammad, who already had a 21 day trans-well chondrogenic disc assay performed, during which wildtype hMSCs at P4 were cultured in chondrogenic medium containg 1x ITS, 100 µg/ml sodium pyruvate, 40 µg/m L-proline, 1X L-glutamine 100 nM dexamethasone, 25 µg/ml ascorbate-2-phosphate, and 10 ng/ml TGFβ3. A negative control was used with chondrogenic medium that did not contain 10 ng/ml TGFβ3. The medium was replaced every 48 h.

RNA for time points 0 h and 21 d (+ TGFβ3 and - TGFβ3) were reverse transcribed, and *miR-140-3p* and *miR-140-5p* expression levels were quantified by qPCR, normalised against *RNU44* (as opposed to *RNU24,* which appeared to be highly regulated during chondrogenesis and hence was unsuitable as a normalising gene) and expressed relative to the normalised "0 h" time point. Both *miR-140-3p* (11-fold) and *miR-140-5p* (13-fold) were strongly upregulated during chondrogenic differentiation at 21 d compared with 0 h (Fig. 4.2).

In contrast, hMSCs that were not treated with TGFβ3 (control), showed a significant decrease in *miR-140-3p* and *miR-140-5p* expression levels at 21 d compared with 0 h. These results indicated that *miR-140-3p/5p* were strongly regulated during chondrogenic differentiation and confirmed observations made by Miyaki et al. (2009) and Karlsen et al. (2013).

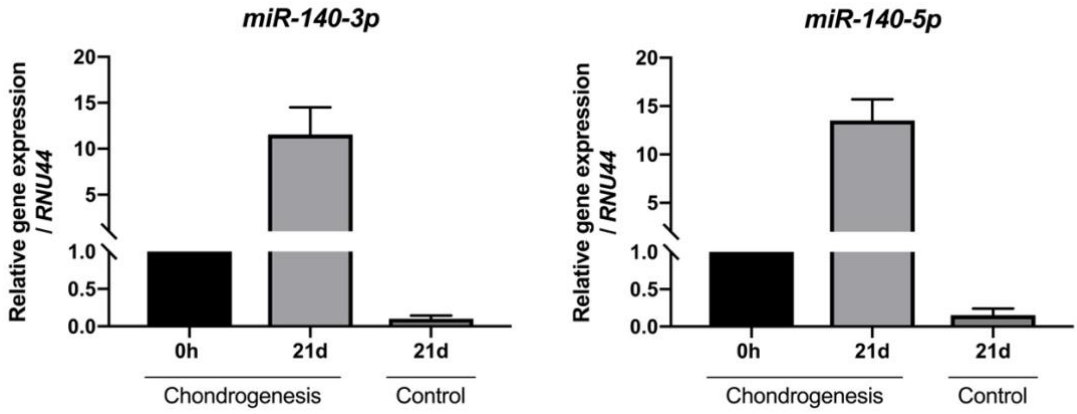


Fig. 4.2: *miR-140-3p* and *miR-140-5p* are upregulated during chondrogenic differentiation.

RNA was obtained from hMSCs undergoing chondrogenesis in trans-well chondrogenic disc assays for 21 days. Cells were cultured with chondrogenic medium. RNA was reverse transcribed, and *miR-140-3p* and *miR-140-5p* expression was quantified by qPCR, normalised against *RNU44,* and expressed relative to the normalised "0 h" time point. Control cells were treated with chondrogenic medium without TGFβ3. Error bars were determined by technical replicates. *n* = 3.

4.2.2 RNP transfection leads to gene editing and *miR-140* knockdown in human mesenchymal stem cells

To assess gene editing efficiencies, hMSCs were RNP single and double transfected with the same *miR-140* targeting sgRNAs and protocol as OA hACs (chapter 3). To assess DNA cleavage, *miR-140* was targeted using sgRNA L3, the combination of sgRNA L3+ L2, or the combination of sgRNA L3+L2+L1. If more than one sgRNA was used, individual sgRNA amounts were reduced accordingly to ensure consistency between total final concentrations of sgRNA used. A NT sgRNA was used as control, along with only Lipofectamine (LF) and Cas9 treated cells (no sgRNA), and completely untreated cells. Human MSCs at P4 were seeded in 12 well plates, cultured for 48 h in 20 % FBS containing DMEM media, and RNP transfected at a confluency between 40 - 60 %. Cells were either removed for analysis at 48 h, or medium was exchanged for 20 % FBS / DMEM for 24 h and cells subsequently re-transfected as before. Cells were incubated for a further 48 h, and then isolated for gDNA and RNA. DNA cleavage was assessed by the T7E1 assay after PCR amplifying a 439 bp long product with primers flanking the targeted *miR-140* hairpin region. Gene editing was confirmed for all targeting sgRNAs. Human MSCs that were double RNP transfected exhibited stronger gene editing efficiencies compared with hMSCs that were only singly RNP transfected (Fig. 4.3A/B). RNA of single and double RNP transfected hMSCs were reverse transcribed into cDNA and assessed for *miR-140-3p* expression levels. Samples were normalised to *RNU24* and expressed relative to their normalised NT sgRNA controls.

miR-140-3p expression levels were reduced by 60 % with the combination of sgRNA L3+L2+L1, by 50 % with the combination of sgRNA L3+L2, and by 40 % with sgRNA L3 for single RNP transfected hMSCs (Fig. 4.3C). Double RNP transfected hMSCs showed marginally stronger reduction of *miR-140-3p* levels: all three targeting conditions were reduced by approximately a further 10 %. Other control samples were higher compared with the NT sgRNA control. The RNP double transfected Lipofectamine and Cas9 were 1.5-fold higher, while the untreated samples, both for single and double RNP transfection, were 2-fold higher. *miR-140-5p* expression levels could not be assessed.

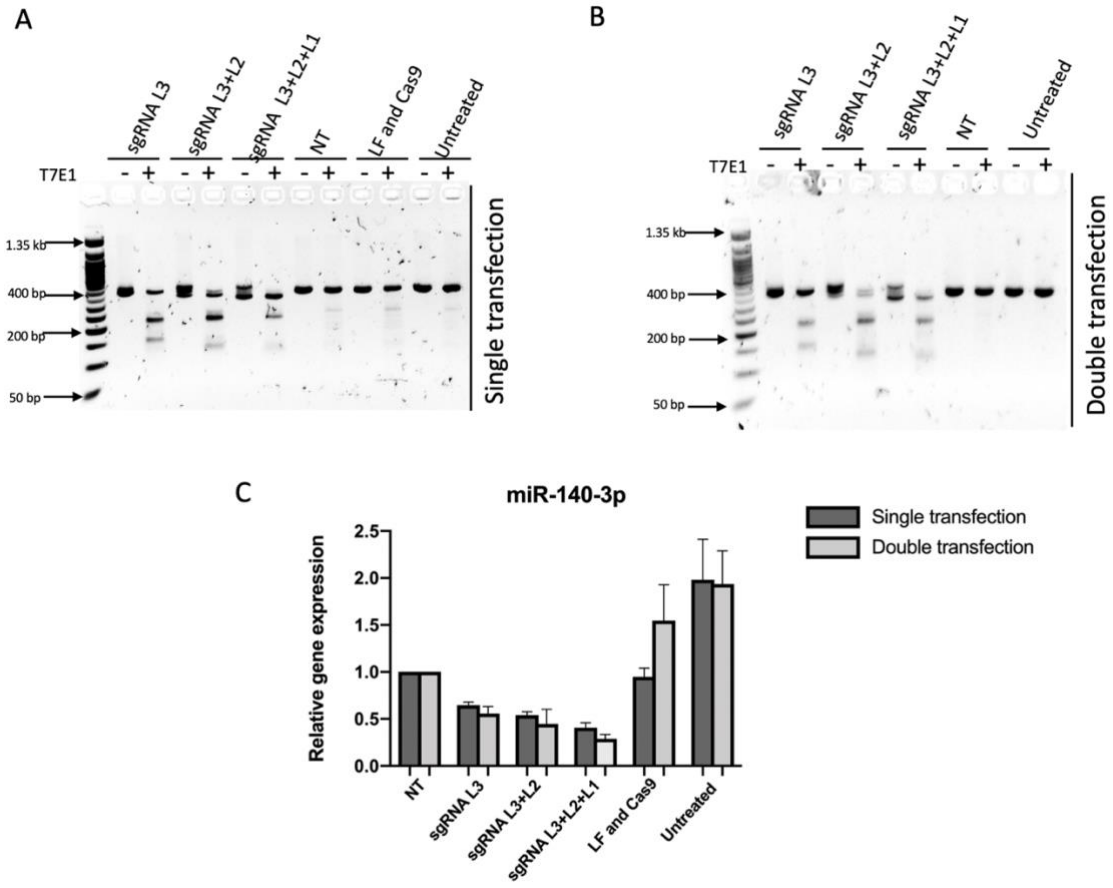


Fig. 4.3: Gene editing efficiencies of hMSCs after single and double RNP transfection.

Human MSCs were transfected according to the previous OA hACs protocol. Cells were either removed for analysis at 48 h, or medium was exchanged for 20 % FBS/ DMEM for 24 h and cells re-transfected as before. Cells were incubated for a further 48 h at 37 °C, 5 % CO_2. **A** gDNA was obtained from single and double RNP transfected hMSCs targeting miR-140 by one single sgRNA (L3), two different sgRNAs (L3+L2), or a combination of three different sgRNAs (L3+L2+L1). A 439 bp long PCR product surrounding the miR-140 hairpin structure was amplified. DNA cleavage and gene editing were assessed by the T7 endonuclease 1 (T7E1) assay, which recognises non-homologous end joining (NHEJ), and cleaves the DNA at this site to generate products between roughly 150 bp and 250 bp. If more than one sgRNA was used, individual sgRNA amounts were reduced accordingly to ensure consistency between final total concentrations of sgRNA used. Controls included a non-targeting (NT) sgRNA, only Lipofectamine (LF) and Cas9 treated cells (no sgRNA), as well untreated samples. **B** RNA from single and double transfected samples was isolated, reverse transcribed and assessed for *miR-140-3p* expression by qPCR. Samples were normalised to *RNU24* and expressed relative to their normalised NT (non-targeting) sgRNA control. Other controls include only Lipofectamine (LF) and Cas9 treated cells (no sgRNA), as well as untreated samples. *n* = 1. Error bars determined by technical replicates (*n* = 3).

In order to increase the extent of *miR-140* knockdown in hMSCs, cells at P5 were plated in 12 well plates, and *miR-140* was targeted through RNP double transfection with an optimised protocol (Table 2.3) after a number of different iterations of Cas9/sgRNA and Lipofectamine doses.

Using the new protocol (but with cells cultured and RNP double transfected in the same timeline as described above), efficient gene editing was confirmed for all targeting sgRNAs, while controls did not exhibit any DNA cleavage at the PCR amplified site (Fig. 4.4A).

RNA of the same samples of double RNP transfected hMSCs was reverse transcribed into cDNA and assessed for *miR-140-3p/5p* expression levels. Samples were normalised to *RNU24* and expressed relative to their normalised NT sgRNA controls. *miR-140-3p* and *miR-140-5p* expression levels were reduced by 94 % and 98 % respectively by the combined efforts of *miR-140* targeting sgRNAs L3+L2+L1 (Fig. 4.4B). sgRNAs L3+L2 reduced *miR-140-3p* expression by 90 %, and sgRNA L3 by 75 %. The combination of sgRNA L3+L2 and sgRNA L3 reduced *miR-140-5p* expression by 80 %.

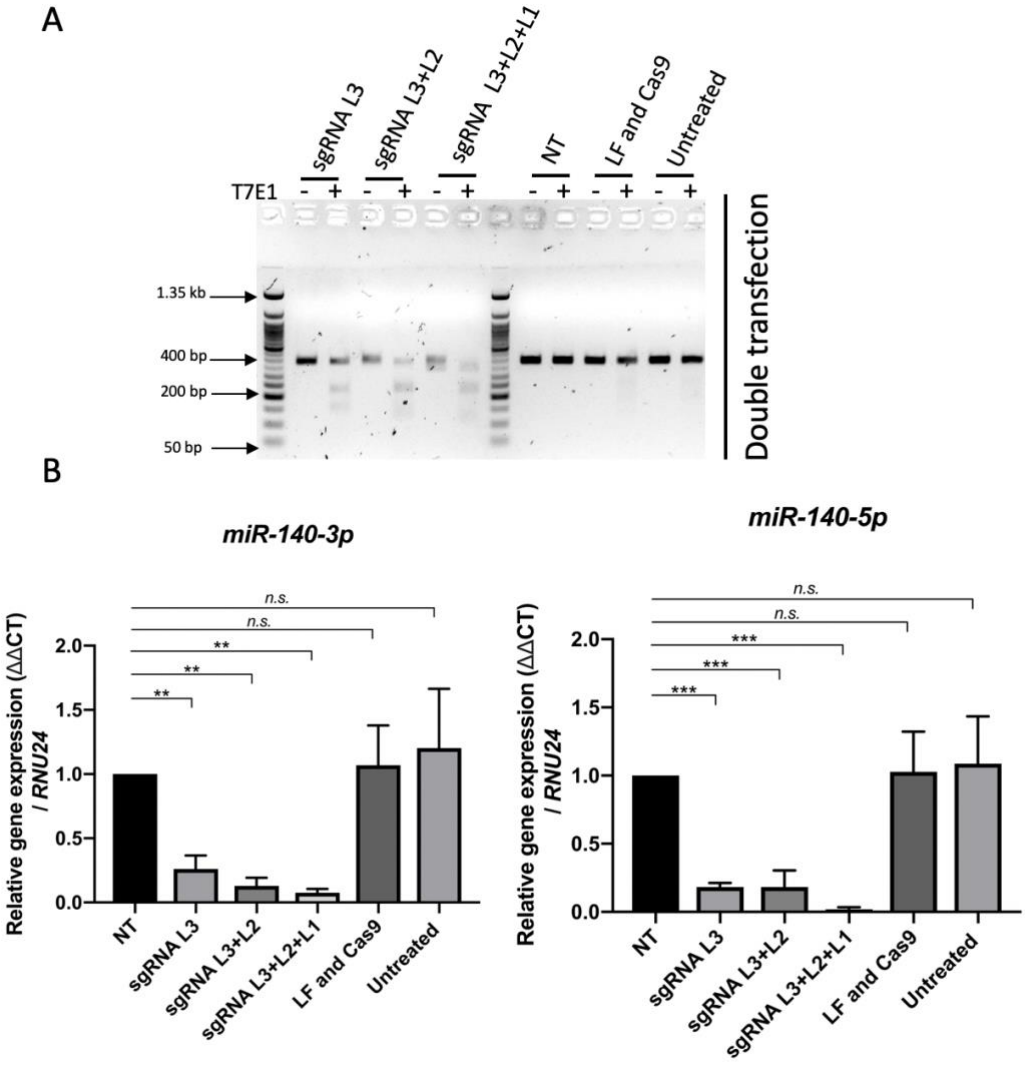


Fig. 4.4: Increased gene editing efficiencies of hMSCs after double RNP transfection.

Amended RNP transfection protocol for targeting *miR-140* was applied for double transfection of hMSCs. **A** gDNA was obtained from double RNP transfected hMSCs 120 h post transfection after targeting *miR-140* by one single sgRNA (L3), two different sgRNAs (L3+L2), or a combination of three different sgRNAs (L3+L2+L1). A 439 bp long PCR product surrounding the *miR-140* hairpin structure was amplified. Gene editing was assessed by the T7 endonuclease 1 (T7E1) assay. If more than one sgRNA was used, individual sgRNA amounts were reduced accordingly to ensure consistency between final total concentrations of sgRNA used. Controls included a NT sgRNA, only Lipofectamine (LF) and Cas9 treated cells (no sgRNA), as well untreated samples. **B** RNA from double RNP transfected samples was isolated, reverse transcribed, and assessed for *miR-140-3p* and *miR-140-5p* expression by qPCR. Samples were normalised to *RNU24* and expressed relative to their normalised NT sgRNA control. Other controls included only Lipofectamine (LF) and Cas9 treated cells (no sgRNA), and untreated samples. Data are shown as mean ± SD. The statistical significance of comparisons between treatment groups were analysed with one-way ANOVA, with post hoc multiple comparisons corrected using Dunnett test. n = 3.

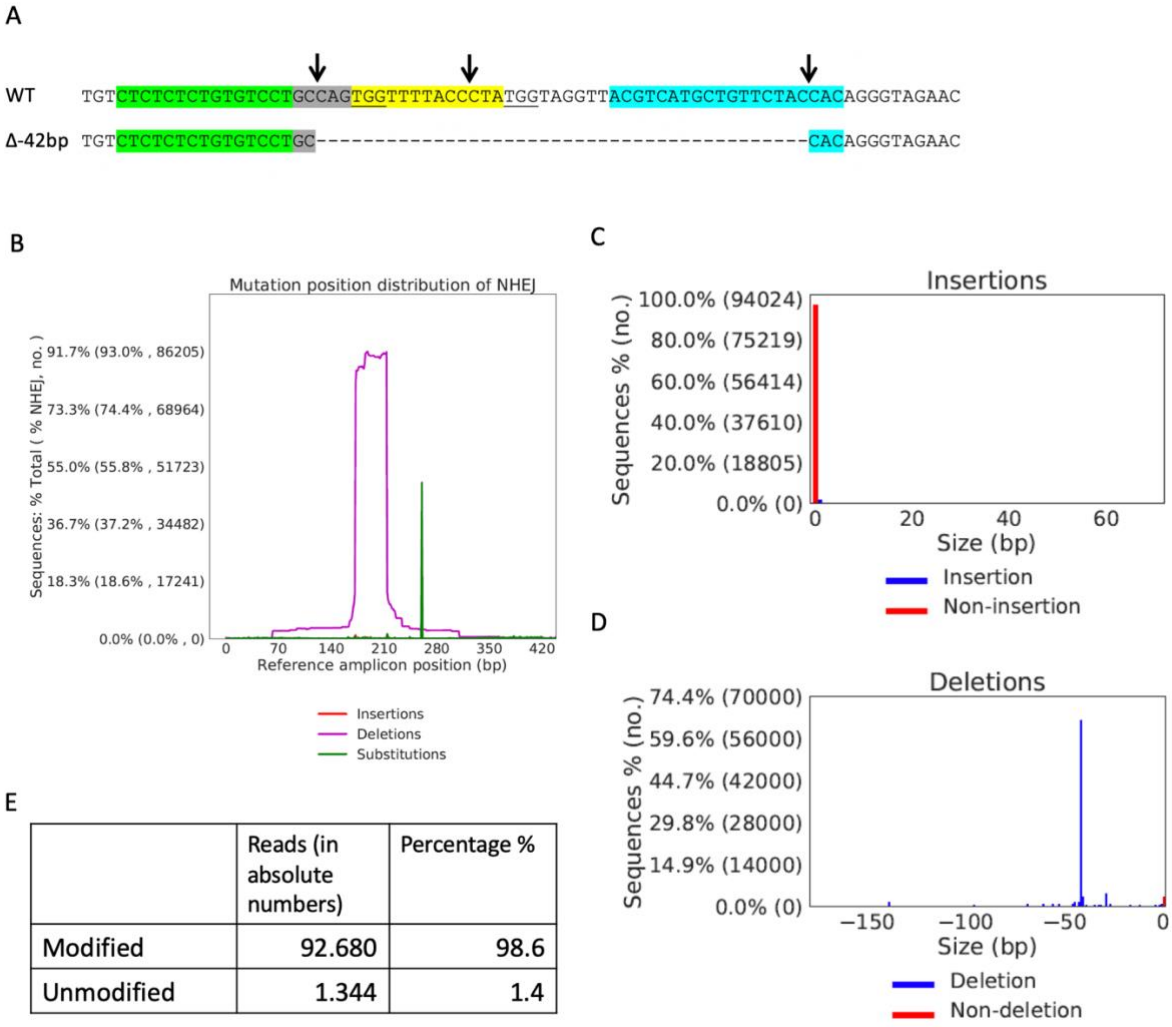
4.2.3 MiSeq and off-targeting analysis confirm precision of designed *miR-140* targeting sgRNAs in human mesenchymal stem cells

To examine how double CRISPR-Cas9 RNP transfection affected the genomic structure of the targeted *miR-140* region in hMSCs, the region of interest was PCR amplified, primer indexed, and analysed via Illumina MiSeq. Analysis of 10 randomly selected reads all showed a deletion of 42 bp between the cleavage sites of sgRNA L1 (highlighted in green), sgRNA L2 (highlighted in yellow), and sgRNA L3 (highlighted in blue) (Fig. 4.5A, black arrows indicate cleavage sites). The area highlighted in grey is an overlap between sgRNA L1 and sgRNA L2.

At amplicon position 267, a small nucleotide polymorphism (SNP) was detected, suggesting a heterogenous cell population. Between amplicon position 178, the cutting site of sgRNA L1, and amplicon position 220, the cutting site of sgRNA L3, MiSeq analysis confirmed a deletion of NHEJ induced base pairs for over 90 % of reads (94,024 total reads) (Fig. 4.5B). Deletions of base pairs, in smaller numbers, were also detected upstream of amplicon position 178 and downstream of amplicon position 220. A very small number of insertions was detected at amplicon position 178 (cleavage site of sgRNA L1) and 191 (cleavage site of sgRNA L2) (Fig. 4.5C).

In fact, NHEJ-induced insertions sized 1 bp occurred but in small numbers (1 bp insertion: ~800 reads). The vast majority (> 98% of reads) did not exhibit any NHEJ

induced bp insertion. The majority of reads amongst those deleted (> 90 %) exhibited a deletion of 42 bp or larger, with a peak at -42 bp (65 %) (Fig. 4.5D). Individual reads (< 500) showed bp deletions between -1 bp and -41 bp, with a small peak at -29 bp (~2,500 reads). The absolute number of modified reads amounted to 92,680 (98.6 %), compared with only 1,344 (1.4 %) unmodified reads, suggesting a very high gene editing efficiency through RNP double transfection in hMSCs (Fig. 4.5E).



E

	Reads (in absolute numbers)	Percentage %
Modified	92.680	98.6
Unmodified	1.344	1.4

Fig. 4.5: Sequencing confirms the deletion of nucleotides between the cutting sites of sgRNA L3, sgRNA L2 and sgRNA L1 at the hairpin structure of *miR-140*.

gDNA of hMSCs, after being targeted by sgRNA L3+L2+L1 in combination to knockdown *miR-140*, was sequenced. **A** Illumina MiSeq confirms a deletion of 42 bp between the cutting sites of sgRNA L3, sgRNA L2 and sgRNA L1. Binding site of sgRNA L1 is highlighted in green, sgRNA L2 in yellow, and binding site of sgRNA L3 in blue. Respective PAM sequences are underlined. Black arrows indicate Cas9 cutting site. The area in grey is the overlap region between sgRNA 1 and sgRNA 2 binding sites. **B** Sequence coverage as determined by Illumina MiSeq. Number of sequences are plotted against the unmodified sequence in reference to the amplicon position. Number of NHEJ-induced **(C)** insertions and **(D)** deletions are plotted with indication of insertion / deletion size. **E** Number of modified and non-modified reads indicating high efficiency of editing. Figures 5B/C/D were created by Dr. Damien Downes in collaboration.

After confirming the specificity of the *miR-140 targeting* sgRNAs, the next step involved determining whether they exhibited any potential off-targeting in hMSCs.

Analysis was carried out similarly to off-target analysis in OA hACs, however, predicted two mismatch off-targets analysis were additionally carried out for sgRNA L1, as the combined usage of sgRNA L3+L2+L1 led to the most efficient *miR-140* KD in hMSCs.

T7E1 assay revealed no DNA cleavage for any of the predicted sgRNA L1, L2 or L3 two sequence mismatch off-targets (Fig. 4.6), suggesting that all three sgRNAs were specific to their *miR-140* targets. The off-targeting likelihood for predicted three sequence mismatches for both targeting sgRNAs, though not experimentally validated, is reduced since two sequence mismatches did not lead to any visible DNA cleavage. Predicted off-target 1 (OT1) for sgRNA L3 could not be PCR amplified with at least three different primer pairs (data not shown), hence no T7E1 assay could be performed. On-targeting controls were performed by assessing DNA cleavage at the *miR-140* site.

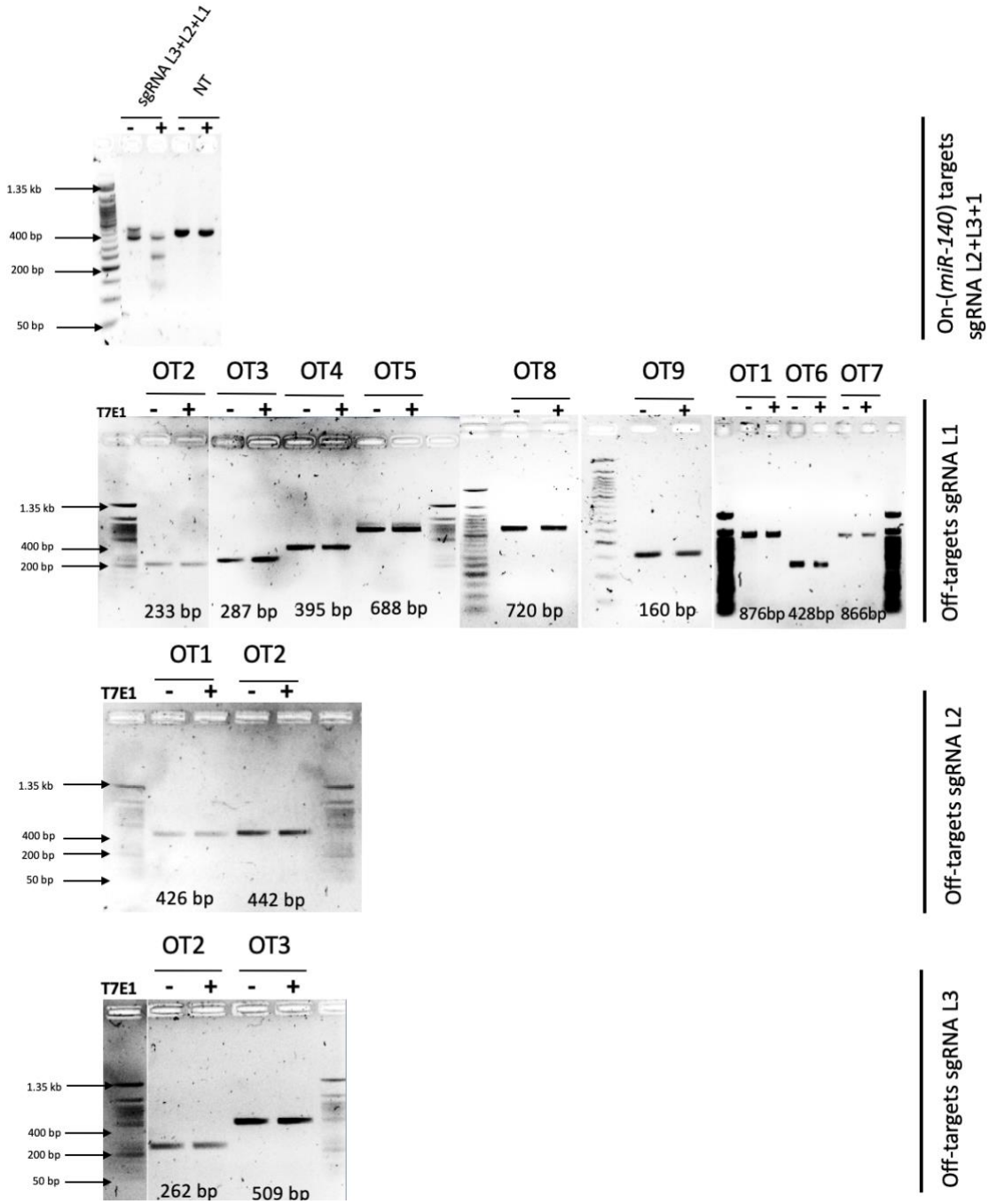


Fig. 4.6: *miR-140* sgRNAs L1, L2 and L3 do not generate bioinformatically predicted 2 mismatch off-targets.

Specific primers were designed for all predicted off-targets with 2 mismatches for sgRNA L1, sgRNA L2 and sgRNA L3. Regions of interest were PCR amplified and their off-targeting assessed by the T7E1 assay, after samples had been transfected with sgRNA L1+L2+L3 in combination. On-target controls were performed to assess the gene editing efficiency by amplifying the miR-140 hairpin structure, and subsequently performing the T7E1 assay for DNA cleavage.

4.2.4 Functional readout of *miR-140* KD in human mesenchymal stem cells via chondrogenic differentiation *in vitro*

As *miR-140* regulated three putative anabolic pathway genes (*BMP2*, *FGF2* and *AGRN*, see 3.2.11) and is itself strongly regulated during chondrogenesis, a 21 day trans-well chondrogenic disc assay was carried out. Under the supervision of Dr. Hayat Muhammad, I cultured *miR-140* KD hMSCs at P10 (modified as in fig. 4; *miR-140-3* levels reduced by 94 % and *miR-140-5p* levels reduced by 98 %) in chondrogenic medium (1x ITS, 100 µg/ml sodium pyruvate, 40 µg/m L-proline, 1X L-glutamine 100 nM dexamethasone, 25 µg/ml ascorbate-2-phosphate and 10 ng/ml TGFβ3). The medium was replaced every 48 h and the chondrogenic potential of the discs was assessed by qPCR analysis and by histology at 10 d and 21 d. NT sgRNA transfected and non-transfected hMSCs (wildtype), both at P10, acted as controls and also underwent chondrogenic differentiation in the same way.

RNA for time points 0 h, 10 d, and 21 d was reverse transcribed and analysed for expression levels for chondrogenic marker genes by qPCR, normalised against *β2M,* and expressed relative to the normalised "0 h" time point. *miR-140-3p* and *miR-140-5p* expressions were quantified by qPCR, normalised against *RNU44* and expressed relative to the normalised "0 h" time point.

All chondrogenic markers (Fig. 4.7 (A) aggrecan *(ACAN)*, (B) cartilage oligomeric matrix protein *(COMP),* (C) collagen type II *(COL2),* and (D) collagen type 10 *(Col10)*) were

significantly upregulated in non-transfected (wildtype) control hMSCs at 10 d and 21 d compared with 0 h.

In contrast, the same increase in chondrogenic expression was not seen in transfected NT sgRNA control hMSCs. This suggested that transfection per se was suppressing chondrogenic cell differentiation *in vitro*. This was also observed in the (*miR-140*) targeted and transfected cells.

Both *miR-140-3p* and *miR-140-5p* were upregulated at 21 d for untreated hMSCs, but not in transfected hMSCs (Fig. 4.7E) Likewise, an increase in *miR-140* levels in NT sgRNA cells was expected, but not evident through our experiment, supporting the notion that RNP transfected hMSCs (at least NT control) did not undergo chondrogenic cell differentiation, while non-transfected (wildtype) hMSCs did. Independent of the differentiation state, an increase in *miR-140* levels in *miR-140* KD cells would not have been expected due to its deletion.

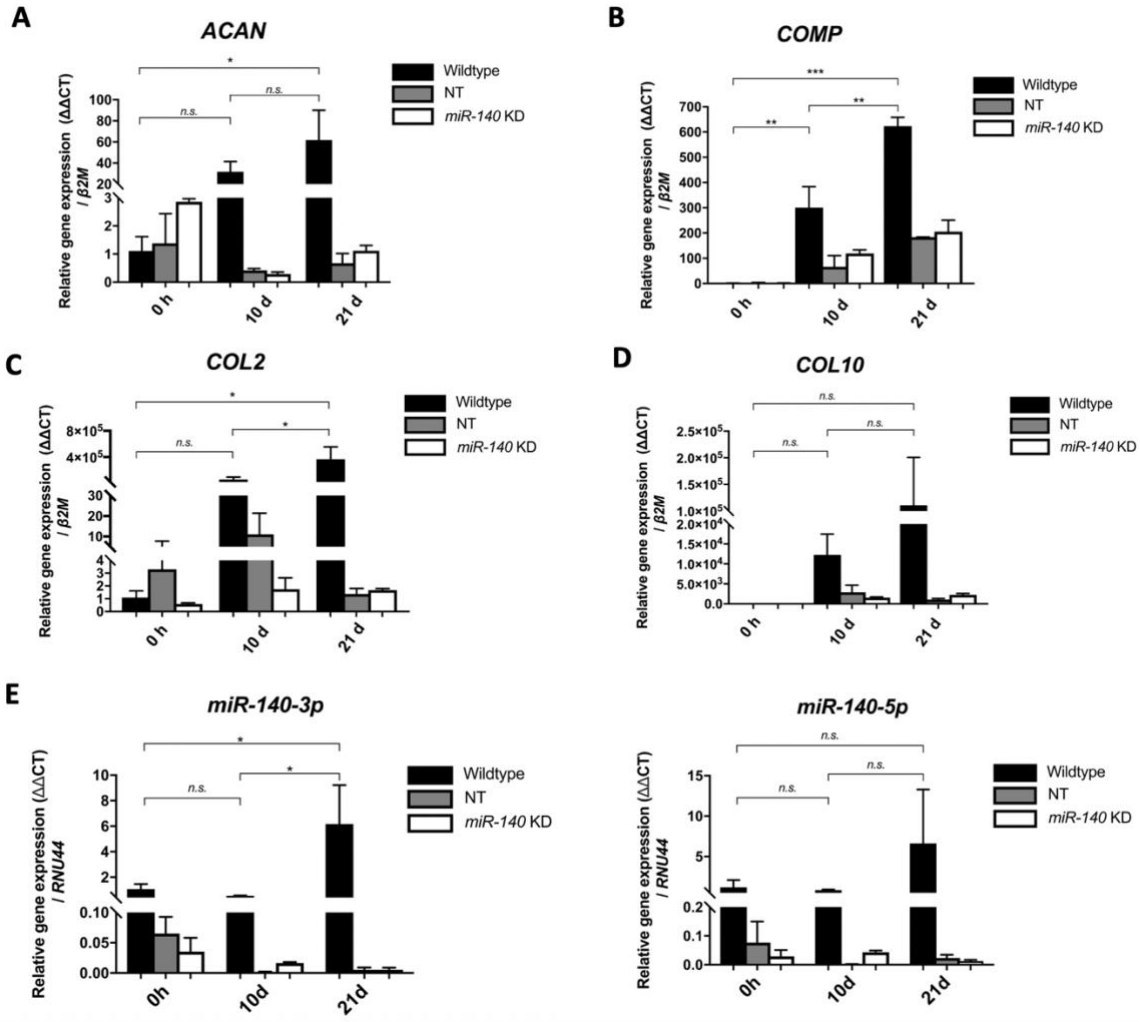


Fig. 4.7: Wildtype hMSCs undergo differentiation into chondrocytes, while transfected hMSCs cells do not.

Untreated (wildtype) human mesenchymal stem cells (hMSCs) at P10, passage matched NT (non-targeting) sgRNA control hMSCs, and passage matched *miR-140* KD hMSCs were cultured in trans-well chondrogenic disc assays for 21 days with chondrogenic medium. RNA was isolated from in TRIzol lysed discs 10 days and 21 days post chondrogenic initiation, reverse transcribed, and used to quantify the expression of chondrogenic marker genes via qPCR: **A** aggrecan *(ACAN)*, **(B)** cartilage oligomeric matrix protein *(COMP)*, **(C)** collagen type II *(COL2)*, and **(D)** collagen type 10 *(Col10)*. Samples were normalised to *β2M* and expressed relative to their "0 h" control. *n* = 3. **E** Expression of *miR-140-3p* and *miR-140-5p*. RNA samples were reverse transcribed and normalised to *RNU44* and expressed relative to their "0 h" time point. Data are shown as mean ± SD. The statistical significance of comparisons between treatment groups of only untreated samples were analysed with one-way ANOVA, with post hoc multiple comparisons corrected using Turkey's test. NT and *miR-140* KD samples were not analysed statistically. *n* = 3.

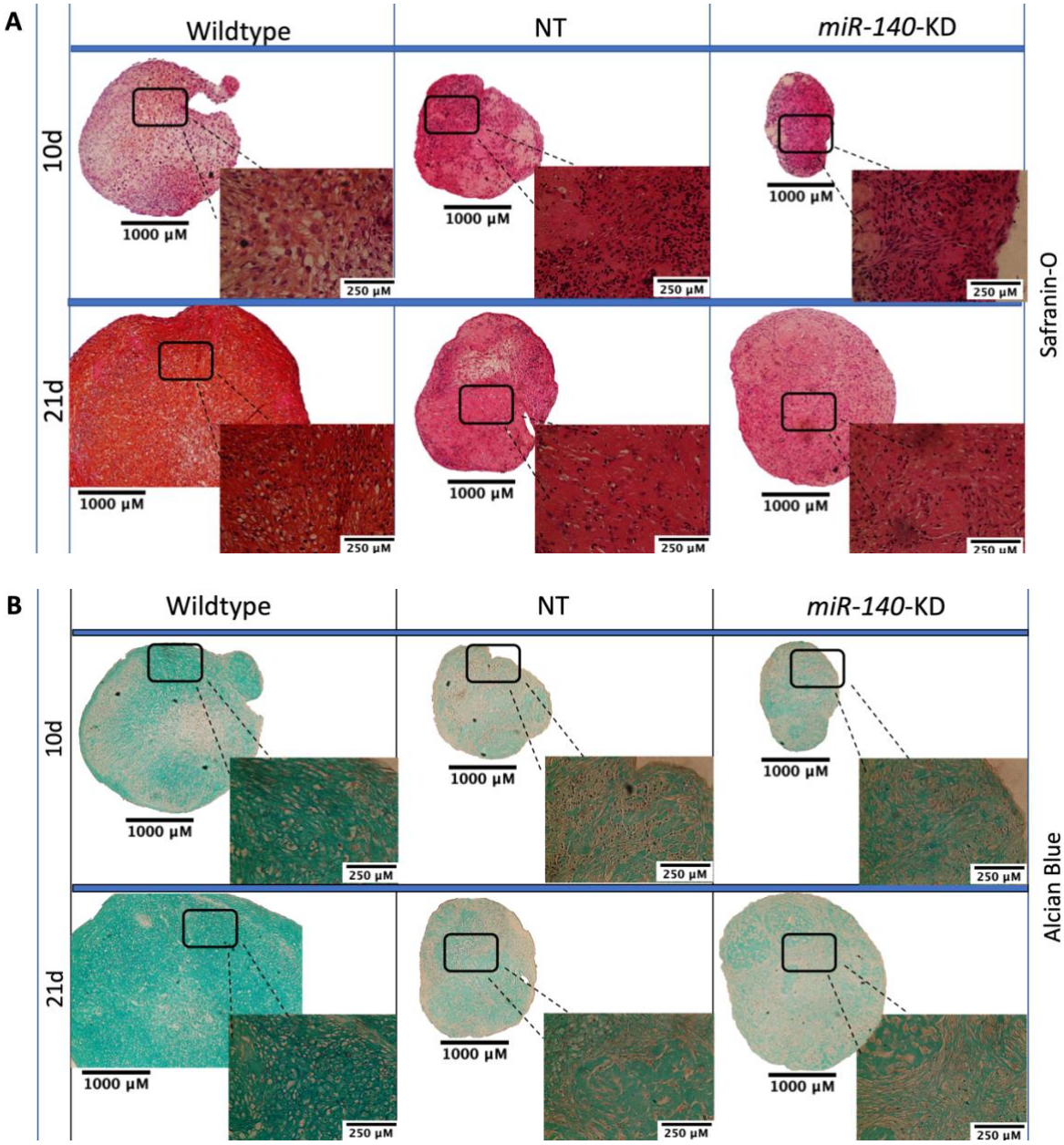
Chondrogenic trans-well discs from the same assay were histologically prepared and assessed for their chondrogenic nature using various staining methods.

Proteoglycan deposition was assessed by Safranin-O staining (orange/red) for all three cell conditions at 10 d and 21 d (Fig. 4.8A). Similarly, collagen deposition was assessed by Alcian blue staining (stains collagen green) (Fig. 4.8B), while Sirius red staining was used as an indicator for collagen fibres (stains collagen fibres red) (Fig. 4.8C). Sirius red stained samples were also assessed under polarised light (Fig. 4.8D). Polarised microscopy differentiates between thin (green) and thick (yellow-orange-red) fibres.

Wildtype hMSCs that were treated for 21 d with chondrogenic medium exhibited a morphological resemblance to chondrocytes, and appeared highly cartilaginous in nature, with strong proteoglycan (Fig. 4.8A) and collagen (Fig. 4.8B) stainings at 21 d, and less substantial stainings at 10 d. Wildtype hMSCs in Fig. 4.8C also revealed some fibrous staining, presumably accounted for by type I collagen fibrils, whilst polarised microscopy also exhibited thick collagen fibres for wildtype hMSCs at 21 d, which were less intense at 10 d.

In contrast, the discs of transfected hMSCs showed no proteoglycan staining, slightly more fibrous (Sirius red) staining, and were much smaller in size; representing cells that had not undergone chondrogenesis.

Of note was the difference in size between chondrogenic discs from untreated and transfected hMSCs. Chondrogenic discs of untreated control cells were much larger in size compared with transfected cells.


A
Wildtype
NT
miR-140-KD
10d
21d
1000 µM
250 µM
Safranin-O
B
Wildtype
NT
miR-140-KD
10d
21d
1000 µM
250 µM
Alcian Blue

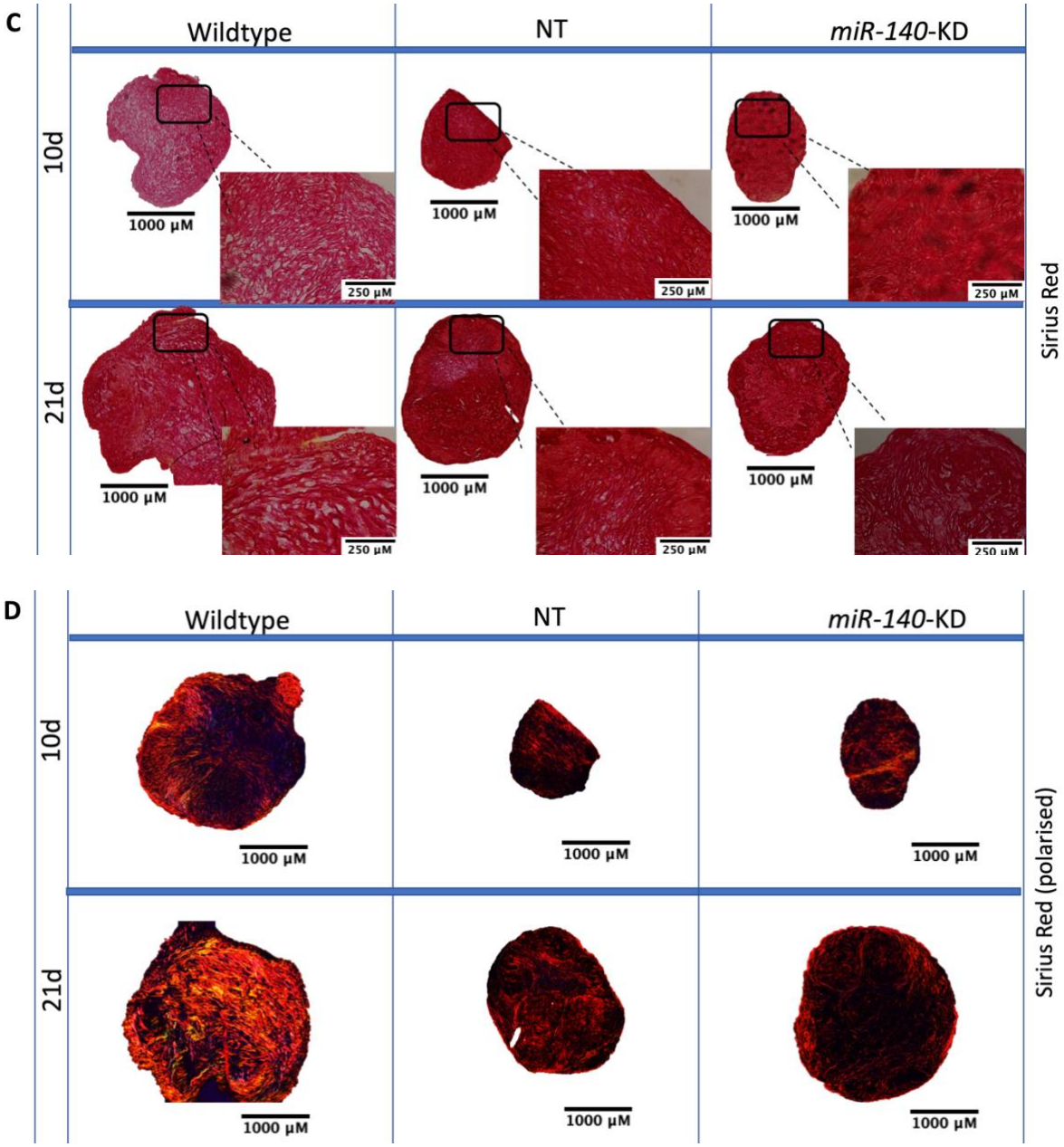


Fig. 4.8: Chondrogenic differentiation of disc assays reveals differences between untreated and RNP transfected human MSCs.

Wildtype hMSCs at P10, passage matched NT sgRNA control hMSCs, and passage matched *miR-140* KD hMSCs were cultured in trans-well chondrogenic disc assays for 21 days with chondrogenic medium. Chondrogenic discs at 10 days and 21 days were embedded in paraffin and sectioned at 5 µm intervals. Sections were stained with **(A)** Safranin-O, (stains proteoglycans orange/red), **(B)** Alcian blue (stains collagen green), and **(C)** Sirius red (stains collagen fibres red). Samples were assessed under light microscopy using an Olympus B56 microscope. **(D)** Using the same microscope, Sirius red samples were also assessed under polarised light. Polarised microscopy differentiates between thin (green) and thick (yellow-orange-red) fibres. Representative images of n = 3. Scale bar of discs = 1000 µm, enlarged discs pictures = 250 µm.

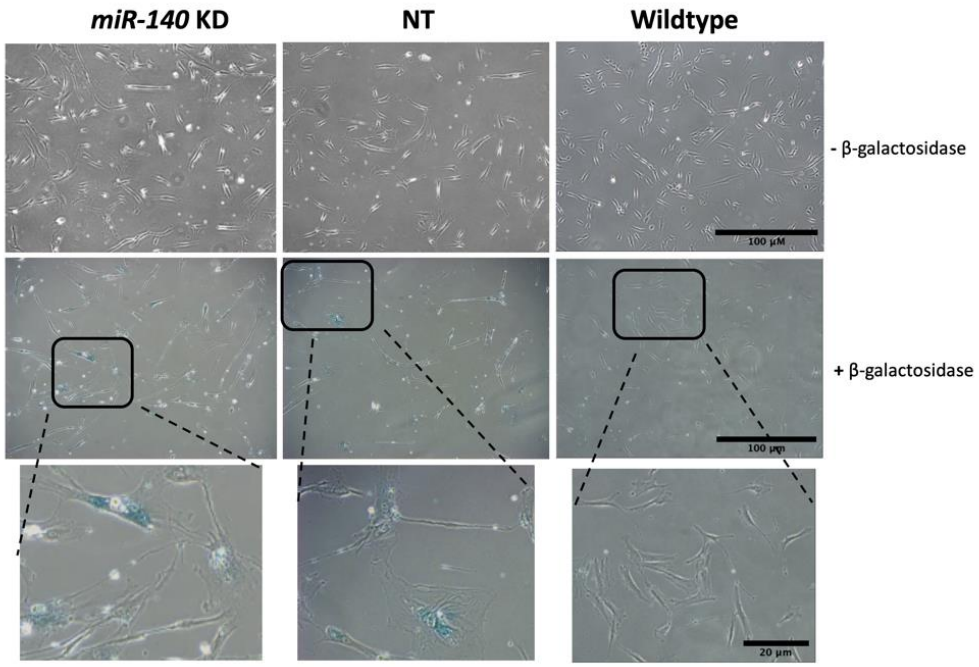
In order to investigate why RNP double transfected (NT sgRNA and *miR-140* targeted) hMSCs at P10 did not undergo chondrogenic differentiation, transfected and non-transfected (wildtype) cells from the same donor were plated at P10 in 12-well plates, and assessed for senescence using the β-galactosidase assay, which is a reliable biomarker for cell senescence (irreversible growth arrest). Endogenous lysosomal β-galactosidase specifically accumulates in senescent cells, where it catalyses the hydrolysis of X-gal into galactose, and produces a blue colour.

Applying the β-galactosidase assay revealed that a large number of RNP double transfected hMSCs were senescent compared with cells that had not been transfected (Fig. 4.9A). Blind quantification (*n* = 4) of the number of blue cells against non-blue cells revealed that at least 2/3 of cells both for *miR-140* KD and NT sgRNA were β-galactosidase positive, and hence senescent. Non-transfected cells showed very low levels of blue staining (Fig. 4.9B).

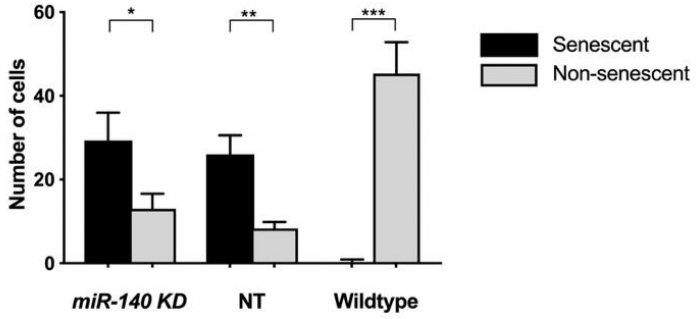
RNA from cells of the chondrogenic assay were analysed for a senescence marker, *p16*. Samples were quantified using qPCR, normalised to *β2M,* and expressed relative to their "0 h" time point. At 10 d, *p16* levels showed a 2-3-fold increase for non-transfected hMSCs while at 21 d the increase was 3-4-fold compared with 0 h. Double RNP transfected hMSCs exhibited a 3-fold increase of *p16* levels after 21 d, compared with 0h and 10 d (Fig. 4.9C).

These results did not necessarily align with the β-galactosidase assay. However, cells undergoing chondrogenic differentiation display hypertrophy, which means that cells enlarge as the differentiation proceeds. As a result, many cells eventually fall into cell cycle arrest, which may explain the rise of *p16* in non-transfected hMSCs.

A


miR-140 KD
NT
Wildtype
- β-galactosidase
+ β-galactosidase
100 µM
100 µm
20 µm

B


Number of cells
60
40
20
0
*
**
***
miR-140 KD
NT
Wildtype
Senescent
Non-senescent

C

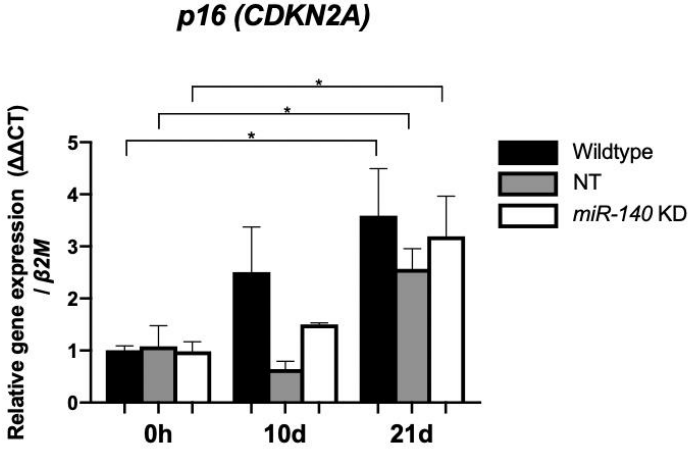

p16 (CDKN2A)
Relative gene expression (ΔΔCT) / β2M
5
4
3
2
1
0
*
*
*
0h
10d
21d
Wildtype
NT
miR-140 KD

Fig. 4.9: Passage 11 RNP transfected human MSCs go into senescence.

A Wildtype (non-transfected) passage 11 human hMSCs, NT (non-targeting) sgRNA transfected hMSCs, and *miR-140* KD human hMSCs from cells of the chondrogenic assay were fixed in 1 x fixative solution for 15 mins, and stained with β-galactosidase overnight. Scale bar = 100 μm, enlarged pictures (bottom row) = 20 μm **B** Blind quantification of blue cells against non-blue cells for each cell condition at 0 h. *n* = 4. **C** Chondrogenic discs were lysed in TRIzol and extracted RNA was reverse transcribed. Expression levels of *p16* senescence marker were quantified using qPCR. Samples were normalised to *β2M* and expressed relative to their 0 h time points. Data are shown as mean ± SD. The statistical significance of comparisons between treatment groups were analysed with two-way ANOVA, with post hoc multiple comparisons corrected using Turkey's test. Unless marked with a "*", comparison between treatment groups was statistically non-significant.

4.2.5 RNP double transfected human mesenchymal stem cells at low passage number are unlikely to be senescent

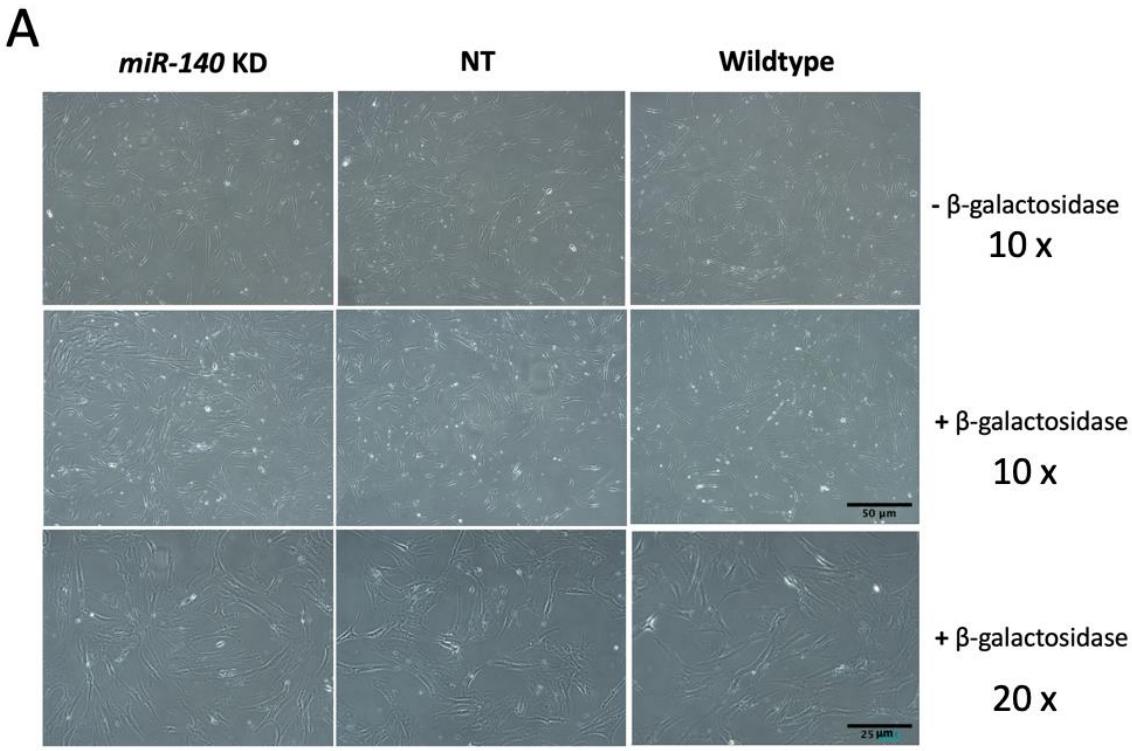
In order to investigate whether RNP single or double transfection of hMSCs at earlier passages led to senescence, and in order to carry out further experiments, hMSCs at P5 from a different donor were seeded into 6 well plates and subsequently single or double RNP transfected with *miR-140* targeting sgRNAs or a NT sgRNA (Fig. 4.10). The same, optimised protocol for hMSCs was applied as previously (Fig. 4.4). Following this, the β-galactosidase assay was applied on both single (Fig. 4.10A) and double (Fig. 4.10B) RNP transfected hMSCs (*miR-140* KD and NT Ctrl). Wildtype (non-transfected), passage matched hMSCs acted as control.

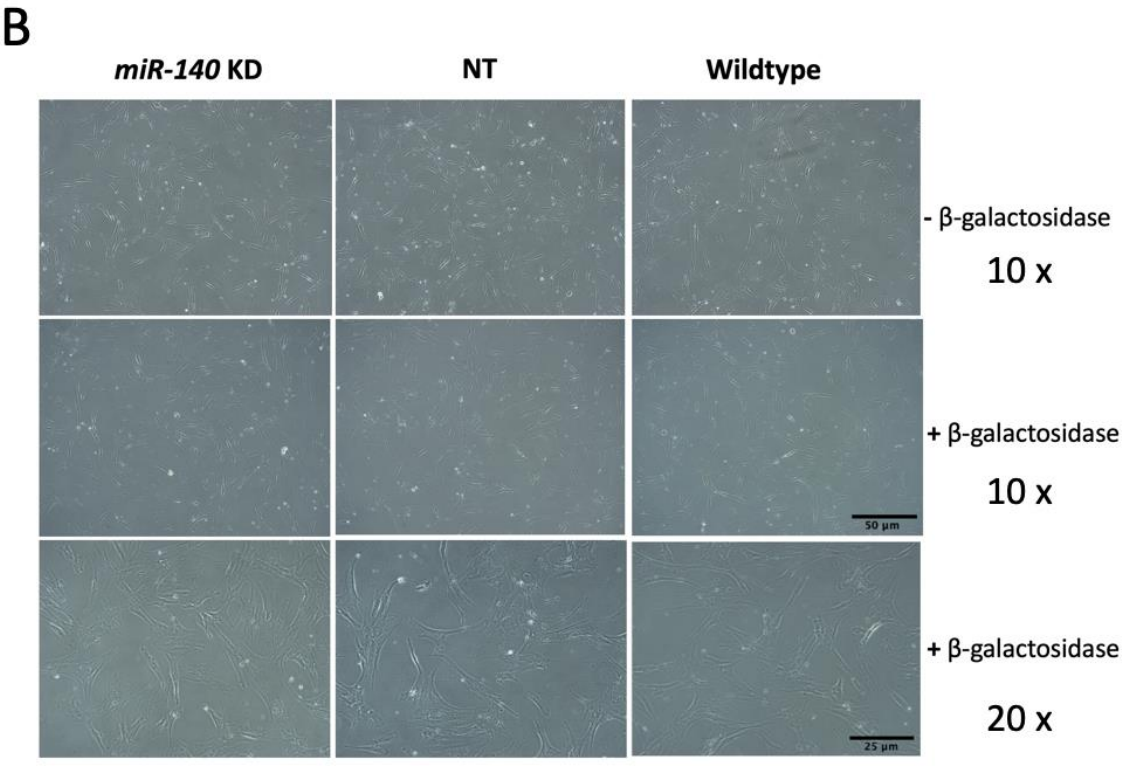
No blue cells were visibly detected for all three cell conditions (regardless of whether single or double RNP transfected); *miR-140* KD, NT sgRNA control, and wildtype (non-transfected) cells. Only very few individual cells (< 5 in total) amongst all three cell types were stained blue during quantification analysis in examined pictures. Blind quantification of blue cells against non-blue cells for each three cell conditions (both for single and double RNP transfection), was displayed as the average absolute numbers from $n = 4$ (Fig. 4.10C).

RNA was isolated from all three cell conditions for RNP double transfected hMSCs at timepoint 0 h. RNA was reverse transcribed and, using qPCR, quantified for expression levels of p16, a senescence marker gene. Samples were normalised to *β2M* and expressed relative to their untreated, non-transfected hMSC control. No significant

differences in expression levels were found for the senescence marker gene *p16* between non-transfected (wildtype) and RNP double transfected cells (NT sgRNA and *miR-140* KD) (Fig. 4.10D).

Both β-galactosidase assay as well as expression analysis suggested that double RNP transfected hMSCs at this passage were unlikely to display a senescent phenotype.


A
miR-140 KD
NT
Wildtype
- β-galactosidase
10 x
+ β-galactosidase
10 x
50 µm
+ β-galactosidase
20 x
25 µm


B
miR-140 KD
NT
Wildtype
- β-galactosidase
10 x
+ β-galactosidase
10 x
50 µm
+ β-galactosidase
20 x
25 µm

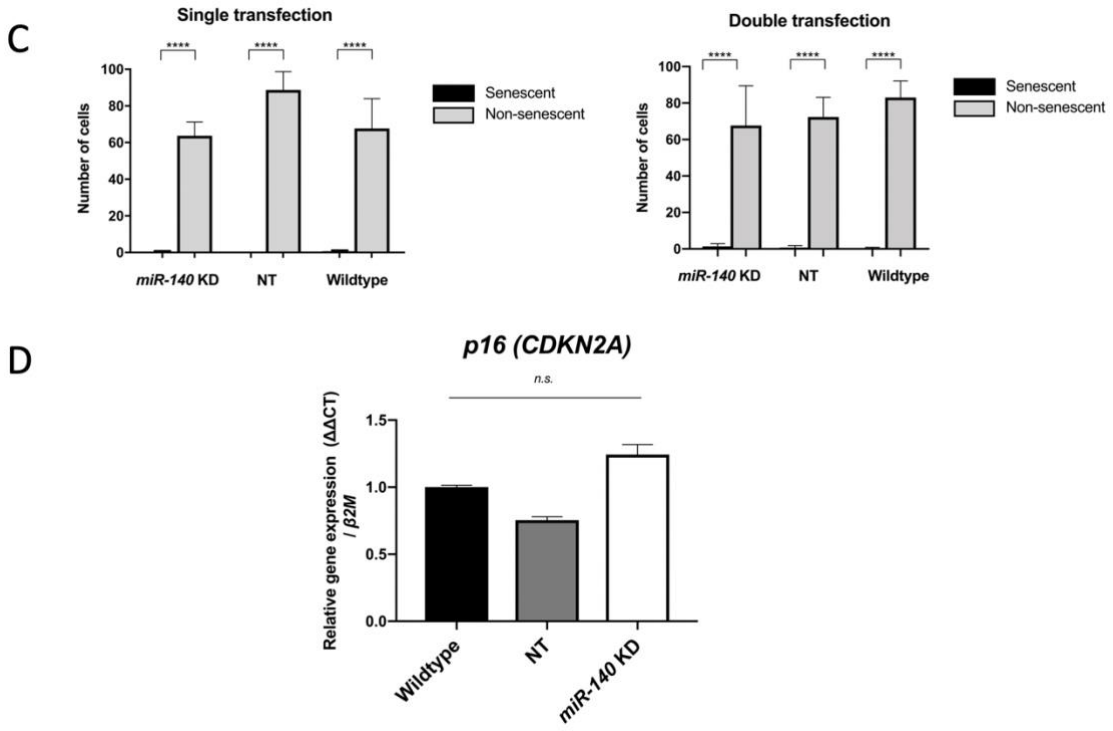


Fig. 4.10: Early passage 5 single and double RNP transfected human MSCs are not senescent.

P5 hMSCs, which were targeted by *miR-140* sgRNAs L1+L2+L3, were collected after **(A)** single RNP transfection and **(B)** double RNP transfection, plated at 50 % confluency, and 48 h later fixed in 1 x fixative solution for 15 mins, stained with β-galactosidase overnight (scale bar first and second rows (10x) = 50 µm, third row (20x) = 25 µm) and **(C)** blind quantified for blue cells (senescent) against non-blue (not senescent) cells. Absolute numbers shown. The statistical significance of comparisons between treatment groups were analysed with one-way ANOVA, with post hoc multiple comparisons corrected using Turkey's test. *n* = 4. **D** Unfixed RNP double transfected cells were lysed in TRIzol, and RNA was reverse transcribed and quantified for expression levels of *p16* (senescence marker) using qPCR. Samples were normalised to *β2M* and expressed relative to their untreated, non-transfected hMSC control. Data are shown as mean ± SD. The statistical significance of comparisons between treatment groups were analysed with one-way ANOVA, with post hoc multiple comparisons corrected using Turkey's test. *n* = 3.

After demonstrating that single and double RNP transfected hMSCs at P5 did not display senescence, they were assessed for DNA cleavage by the T7E1 assay at the *miR-140* site. *miR-140* was targeted using the combination of sgRNA L3+L2+L1.

T7E1 confirmed efficient DNA cleavage for double RNP transfected hMSCs, while the NT sgRNA control as well as non-transfected (wildtype) cells did not exhibit any DNA cleavage (Fig. 4.11A). Genomic DNA samples of single RNP transfected hMSCs could not be amplified.

In order to translate these gene editing efficiencies into mRNA expression levels of *miR-140-3p/5p*, RNA from the same samples was reverse transcribed. Samples were normalised to *RNU24* and expressed relative to their normalised NT sgRNA control. Single RNP transfected hMSCs reduced *miR-140-3p* expression by 50 % and *miR-140-5p* expression by 55 %, while double RNP transfected hMSCs reduced *miR-140-3p* expression by 90 % and *miR-140-5p* expression by 80 % (Fig. 4.11B).

We next intended to repeat the chondrogenesis assay with these healthy *miR-140* KD hMSCs, but were unable to do so due to time constraints, especially in light of our outstanding scratch assay.

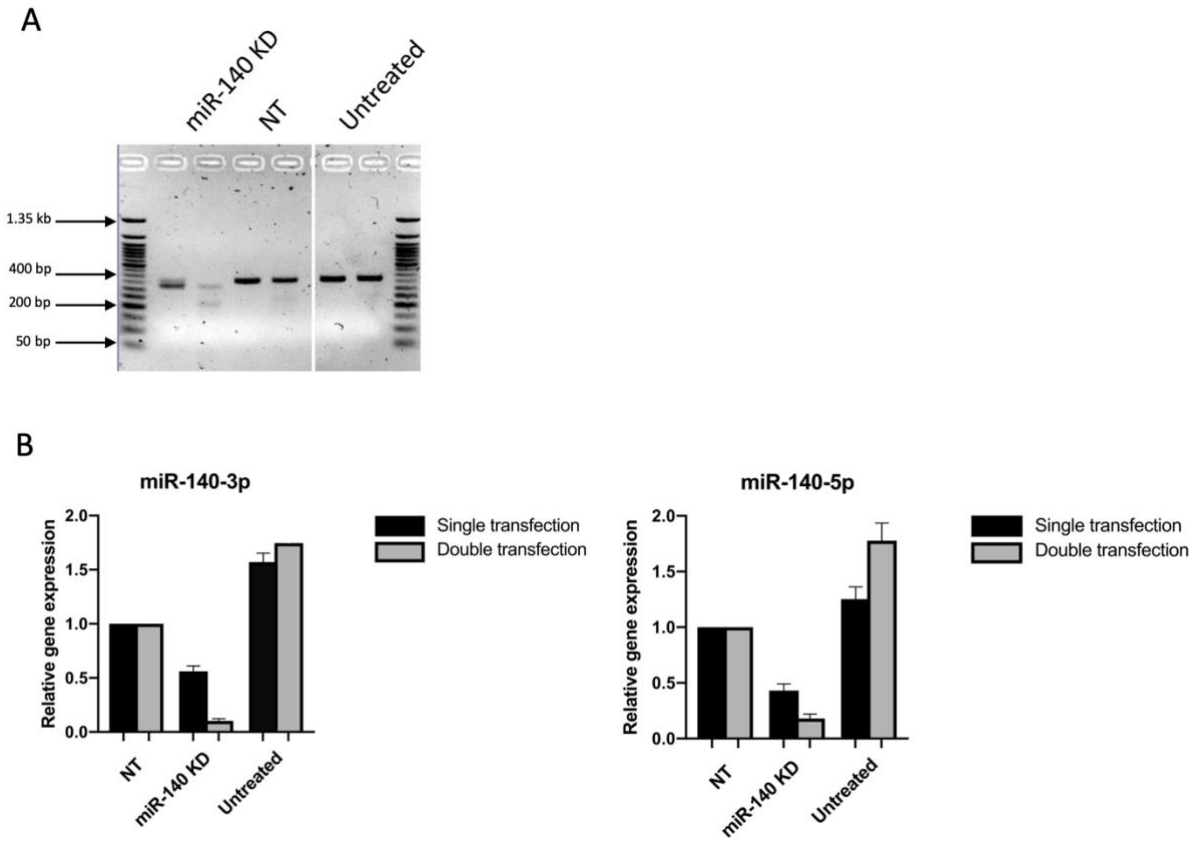


Fig. 4.11: Passage 5 hMSCs reveal efficient gene editing and *miR-140* downregulation upon double RNP transfection.

Cells were either removed for analysis at 48 h, or medium was exchanged for 10 % / DMEM for 24 h and cells subsequently re-transfected as before. Cells were incubated for a further 48 h at 37 °C, 5 % CO_2. **A** Genomic DNA was obtained from double RNP transfected hMSCs, which were targeted by the combination of three different *miR-140* targeting sgRNAs (L3+L2+L1). A 439 bp long PCR product surrounding the *miR-140* hairpin structure was amplified. Gene editing was assessed by the T7E1 assay. A NT sgRNA as well as untreated cells acted as controls. **B** RNA from single and double transfected samples was isolated, reverse transcribed, and assessed for *miR-140-3p/5p* expression by qPCR. Samples were normalised to RNU24 and expressed relative to their normalised NT sgRNA control. Untreated cells acted as additional control. *n* = 1. Error bars determined by technical replicates. *n* = 3.

4.2.6 *miR-140* KD human mesenchymal stem cells delay gap closure after scratch assay

FGF2 has previously been shown to accelerate cell activity and repair in murine MSCs (Muhammad, Khan, unpublished) and we demonstrated *FGF2* to be regulated by *miR-140* in human chondrocytes (Fig. 3.14).

To test whether *miR-140* knockdown impairs hMSC velocity, I used my modified, RNP double transfected, *miR-140* KD hMSCs, NT sgRNA control hMSCs, as well as wildtype hMSCs (all passage matched at P6), and grew them in monolayers until they reached a confluency of roughly 70 %. A 10 µl pipette tip was used to create a scratch in the middle of each well. Gap closure was captured by an automated JuLi stage microscope which was programmed to take pictures in 20-minute intervals for a total of 36 h. Gap closure was calculated by Dr. Hayat Muhammad using Image J software and plotted as the percentage of gap filled against the time.

At 22 h, the WT control and NT control completely closed the gap, whereas *miR-140* KD hMSCs showed a gap closure of only approximately 89 % (Fig. 4.12A), which, however, did not reach statistical significance. Representative images at 12 h intervals between 0h and 36 h are displayed in Fig. 4.12B, with blue, dotted lines representing the gap.

These results therefore indicate that *miR-140 KD* hMSCs delay gap closure after scratch assay compared with their controls.

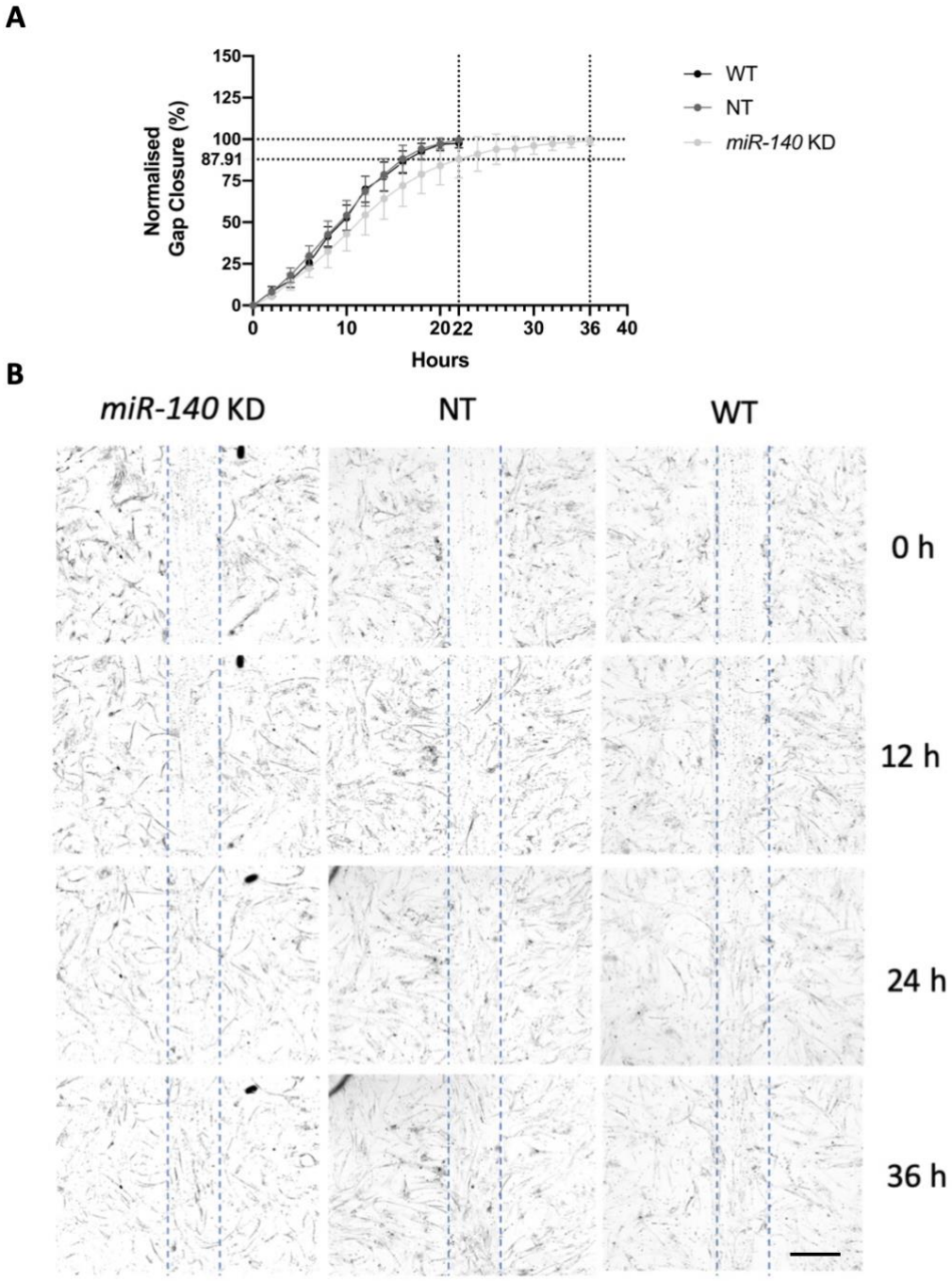


Fig. 4.12: *miR-140* KD hMSCs reveal delay in gap closure.

Passage 6 hMSCs (*miR-140* KD, NT, WT) were cultured in monolayers. A 10 µl pipette tip was used to create a scratch in the middle of each well. **A** Closure of the gap was assessed by using Image J software and plotted as a percentage of the gap filled. Graph shows mean +/- S.D. The statistical significance of comparisons between treatment groups were analysed with two-way ANOVA, with post hoc multiple comparisons corrected using Turkey's test. Comparison between treatment groups was statistically non-significant. *n* = 4, technical replicates. **B** Cells were placed in a Juli stage microscope within an incubator (37 °C, 5 % CO_2), which was programmed to take brightfield images of the scratch at 2 defined points in each well (2 wells for each cell type), every 20 minutes for a total of 36 h. Representative pictures are shown. Scale bar = 200 µm.

4.3 Discussion

In a first step I confirmed previous observations from the literature that *miR-140-3p* and *miR-140-5p* were regulated during chondrogenesis. In resting hMSCs, the expression levels of both miRNA arms were significantly lower than in OA hACs, and these levels increased upon initiation of chondrogenic differentiation. It is not clear whether the levels of *WWP2*, host gene of *miR-140*, were also expressed at lower levels compared with OA hACs.

4.3.1 Modifying human mesenchymal stem cells by CRISPR-Cas9

In order to genetically modify human MSCs via CRISPR-Cas9, I initially adopted the same RNP double-transfection protocol I used for transfecting human chondrocytes. By using the same *miR-140*-targeting sgRNAs, *miR-140* levels were reduced by approximately 60 %. After optimising the protocol for hMSCs by a number of different iterations of Cas9/sgRNA dose and Lipofectamine, *miR-140* levels were reduced by more than 90 % in hMSCs. Notably, it was the combination of 3 different sgRNAs (L3, L2 and L1) that led to the best reduction in hMSCs, whereas the combination of two different sgRNAs (L3 and L2) led to the best reduction in OA hACs.

Additionally, due to the switch in cell type and the use of an additional sgRNA (L1) to facillitate the most efficient gene editing, I investigated potential off-targeting for all three sgRNAs that had 2 sequence mismatches to putative loci in the genome. These putative targets in the genome were identical to those screened in hACs, with the addition of 9 additional sites for sgRNA L1. Consistent with the findings in hACs, no off-

targeting was detected in hMSCs. However, limitations in detection sensitivity as well as structural accessibility of the chromatin binding areas remain.

4.3.2 Chondrogenesis with *miR-140* KD human mesenchymal stem cells

hMSCs were RNP double transfected at P5 (Fig. 4.4), and during subsequent passaging there was no visible difference in phenotype or growth rates. In line with our observation of *miR-140* regulating three anabolic pathway genes (*BMP2*, *AGRN*, and *FGF2*) in OA hACs and additionally being regulated itself during chondrogenic differentiation, a 21-days chondrogenic disc assay was carried out with *miR-140* KD hMSCs. Passage matched NT sgRNA hMSCs and wildtype hMSCs were used as controls (P10). We found that wildtype hMSCs went through chondrogenic differentiation and were highly cartilaginous in nature, whereas the *miR-140* KD hMSCs as well as NT sgRNA hMSCs did not undergo chondrogenic differentiation. This suggested that RNP double transfected hMSCs at passage 10 did not retain the capability to undergo chondrogenic differentiation, as opposed to non-transfected, wildtype hMSCs at passage 10. In fact, transfected cells were proven to be in senescence, a state in which the growth of normally dividing cells is irreversibly arrested. Presumably, stress through transfection, combined with several passages over time, resulted in hMSCs experiencing accelerated senescence compared with non-transfected cells. Indeed, stress has been proven to be a major cause for accelerated senescence *in vitro* (de Magalhães and Passos, 2018).

Another important consideration is that increased passage is associated with loss of stemness (Tsutsumi et al., 2001).

I also attempted to clonally expand individual cells to obtain a pure population with 100 % *miR-140* knockdown for subsequent chondrogenic differentiation. This pure cell population would also prevent potential selective proliferation advantages of the few, non-edited cells over edited cells. Unfortunately, these individual cells did not grow in 96-well plates and the experiment could not be repeated in the given time.

Apart from *miR-140*, there are numerous other miRNAs involved during the process of chondrogenesis. These miRNAs could be subjected to CRISPR-Cas9 mediated knockdown or overexpression. The miRNA profile differs at each stage of chondrogenesis, with each miRNA holding a unique role specific for its corresponding stage. At the mesenchymal stem cell stage, condensation is principally governed by the production of proteoglycans and a switch in synthesis from collagen type I to collagen type II. For example, *miR-488* is upregulated before condensation of cells and down-regulated after condensation. Song et al. (2011) found that by blocking *miR-488* at this stage, integrin and adhesion molecule levels were decreased and cell migration and motility inhibited. Following condensation, cells differentiate towards chondrocyte cell fate and proliferate, secreting ECM components, resulting in each cell being surrounded by matrix. *miR-221* has been identified as an important down regulator in this process (Kim et al., 2010).

4.3.3 Scratch assay in *miR-140* KD human mesenchymal stem cells

I modified new hMSCs by RNP double transfection at a younger passage number (P5) and confirmed them to be senescence free. These new, *miR-140* KD hMSCs were used in an *in vitro* repair-type assay along with passage matched NT sgRNA hMSCs and wildtype hMSCs, both of which were determined to be senescence-free.

My work showed that *miR-140* positively regulates *FGF2,* and our lab has previously shown that FGF2 activates mesenchymal stem cells and enhances scratch assay closure *in vitro* (Muhammad, Khan et al., in preparation). I observed that *miR-140 KD* hMSCs showed delayed gap closure compared with both controls.

There are a number of limitations to this part of the study, which, due to time constraints, could only be carried out once. For example, while it is established that *FGF2* is expressed in hMSCs, the extent to which *FGF2* and its receptors are regulated by *miR-140* in hMSCs is unknown. There is also uncertainty about FGF2 regulation at the protein level. Transcript levels are not always a reliable indicator to predict protein levels (Liu et al., 2016), for which Western Blot analysis could have provided a better understanding. This is a consideration which applies to all our regulated genes upon *miR-140* KD in human chondrocytes.

In articular cartilage, FGF2 in particular is abundant and found within the pericellular matrix of chondrocytes. The role of FGF2 in this region is debated and controversial, with both anabolic and catabolic roles observed. Work by Im et al. (2007) and Yan et al. (2011) and showed that catabolic action of FGF2, namely its induction of the important

cartilage collagenase *MMP13* and *ADAMTS5*, and suppression of matrix production, are mediated through FGFR1. Other catabolic effects observed are the inhibition of anabolic factors insulin-like growth factor (IGF-1) and osteogenic protein 1 (OP-1) (Loeser et al., 2005), and a dose-dependent loss of proteoglycan in response to FGF2 (Li et al., 2012). All these studies were carried out *in vitro*.

In contrast, previous work focusing on anabolic action carried out by our lab group showed FGF2 release following cartilage injury, and accelerated osteoarthritis development amongst *Fgf2* KO mice (Chia et al., 2009). Furthermore, in more recent work, *Fgf2* KO mice failed to repair focal cartilage defects compared with WT controls (Muhammad, Khan et al., in preparation). Work by Valverde-Franco et al. (2006) showed that an accelerated development of OA can also be mimicked through the deletion of receptor *Fgfr3,* while in the *Fgfr1* conditional knockout mice are protected both from spontaneously or surgically acquiring osteoarthritis (Weng et al., 2012). Furthermore, FGF18, a member of the FGF ligand family that is more selective than FGF2 and activates only FGFR3, is shown to stimulate production of proteoglycan (Davidson et al., 2005). Thus, a role is implied specifically for *Fgfr3* in the chondroprotective action of FGF2, and it might be that the driving protective or destructive influence of FGF2 on the articular cartilage tissue is determined by the relative presence of these receptor subtypes.

miRNAs can regulate numerous genes at a time. For example, the knockdown of *miR-140* negatively impairs chondrocyte proliferation due to increased Sp1 protein levels despite no changes in mRNA levels. Sp1 is a crucial transcription factor which modulates the cell cycle regulator *p15* (Yang et al., 2011). *miR-140* regulates chondrocyte proliferation, partially by suppressing *Sp1*. This study suggests that *miR-140* plays a

protective role in the development of osteoarthritis by enhancing chondrocyte proliferation.

Taken together, *miR-140* only exerts a mild effect on *FGF2* in human chondrocytes. It would have been of interest to determine mRNA and protein levels both for *FGF2* and *Sp1* in our *miR-140* knockdown hMSCs.

In light of the potential of miRNAs to regulate thousands of genes (and conversely, one gene can be regulated by multiple miRNAs), the importance of differentiating between direct and indirect targets is paramount in understanding how certain regulatory mechanisms work. In the next chapter I endeavour to adapt a CRISPR-Cas9 based technology that is capable of differentiating between direct and indirect miRNA/mRNA targeting.

Results

5

Can we manipulate miRNA response elements (MRE) by CRISPR-mediated-homology-directed repair (HDR) in human cells?

Chapter 5. Can we manipulate miRNA response elements (MRE) by CRISPR-mediated-homology-directed repair (HDR) in cells?

5.1 Introduction

After successfully knocking down *miR-140* in human articular chondrocytes as well as in human mesenchymal stem cells, a more ambitious approach was tested to study mRNA and miRNA interaction.

Classical miRNA knockout and overexpression studies come with two challenges: (i) physiological relevance (uncertainty about whether the results translate to the endogenous context) and (ii) the inability to distinguish between direct and indirect targets. Bassett et al. (2014) demonstrated a method for targeting the miRNA seed region, also now referred to as the miRNA response element (MRE), to overcome these shortcomings and assess direct mRNA/miRNA targeting in an endogenous context. This novel CRISPR-mediated method targets specific genes and relies on homology-directed repair (HDR) to incorporate specific barcodes. First, a putative MRE is targeted by sgRNA

and Cas9-mediated genome editing to disrupt the MRE binding site. In the absence of added single-stranded oligo DNA nucleotides (ssODNs), this will repair by the process of non-homologous end joining (NHEJ). In the presence of long single-stranded DNA (ssDNA) oligonucleotide templates (roughly 140 nucleotides long) encoding either the T3 or T7 promotor sequences (serving as barcodes), a proportion of these will be incorporated into gDNA, although the % of incorporation will likely be very low. The T3 ssODN contains an intact MRE and the T3 barcode. The T7 ssODN only contains the T7 barcode without the MRE. In other words, it represents the null variant. This transfection of barcodes leads to a heterogenous cell population which can be assessed and quantified with endogenous forward and universal T3 and T7 specific primers (Fig. 5.1A). The T7/T3 ratio in gDNA estimates the relative integration efficiency of the T7 and T3 oligonucleotides. An increase in this ratio for mRNA (cDNA) is expected in the case of a functionally active MRE (Fig. 5.1B), since deletion of an active MRE through the T7 barcode will prevent miRNA-mediated inhibition of the target.

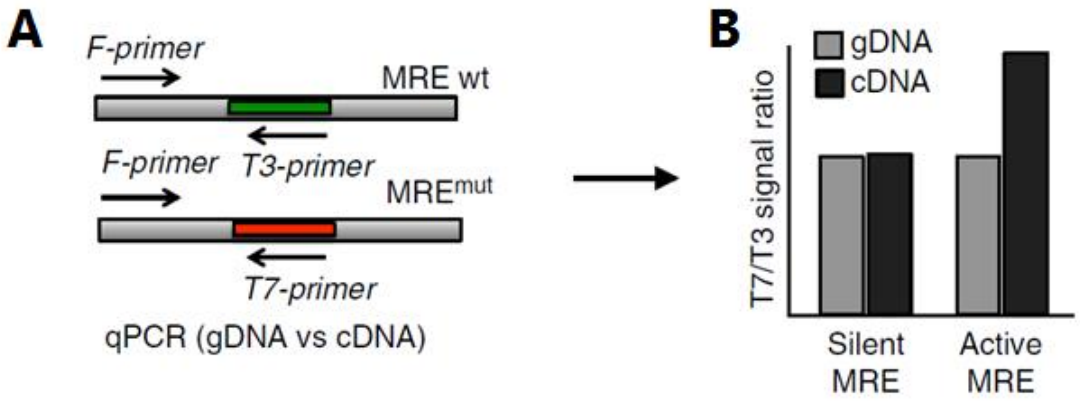


Fig. 5.1: Schematic overview of the analysis step of CRISPR mediated HDR method to investigate mRNA/miRNA interaction.

A Validation of correct insertion and qPCR analysis of the T3 and T7 barcode in gDNA and cDNA using a gene-specific forward primer and T3- or T7-specific reverse primers. **B** T7/T3 signal ratio on genomic DNA level is used to estimate relative integration efficiency of T3 and T7 barcodes. An increase in this ratio on transcript level (cDNA) indicates an active miRNA site (from Bassett *et al.*, 2014).

One key advantage of this method is that it is independent of the efficiency of transfection, as one only cells, that have either T3 or T7 barcodes expressed, will be analysed. The read-out is based on these edited cells only. Additionally, the method is independent of artificial changes in levels of miRNA (i.e. inhibition or overexpression). Lastly, the method demonstrates direct effects of a given miRNA on a specific gene – it targets the binding site, rather than the miRNA itself.

The aim of this chapter was (i) to verify that I could reproduce results published by Bassett et al. (2014) (in my co-supervisor's group) in HEK293T cells using the CRISPR-mediated HDR barcode method, and (ii) demonstrate that the same methodology can be applied to a key chondrocyte gene firstly in HEK293T cells, then in human articular chondrocytes.

5.2 Results

5.2.1 *SOX9* but not *COMP* is expressed in HEK293T cells

First a key chondrocyte gene was selected that is known to be under miRNA control (Martinez-Sanchez et al., 2012). qPCR was performed in resting HEK293T cells for *SOX9*, a key regulator of chondrocytes, and *COMP*, a component of articular cartilage (used as control), and compared with their respective expression levels in OA hACs. RNA from HEK293T cells (*n* = 3) and OA hACs (*n* = 3) was reverse transcribed and quantified for expression levels of *SOX9*.

SOX9 expression levels in HEK293T were detected, albeit at a lower level compared with OA hACs (average CT levels at 28) (Fig. 5.2). In comparison, *COMP*, as expected, was barely detectable in HEK293T cells, and highly expressed in OA hACs. So further analyses were conducted targeting *SOX9*.

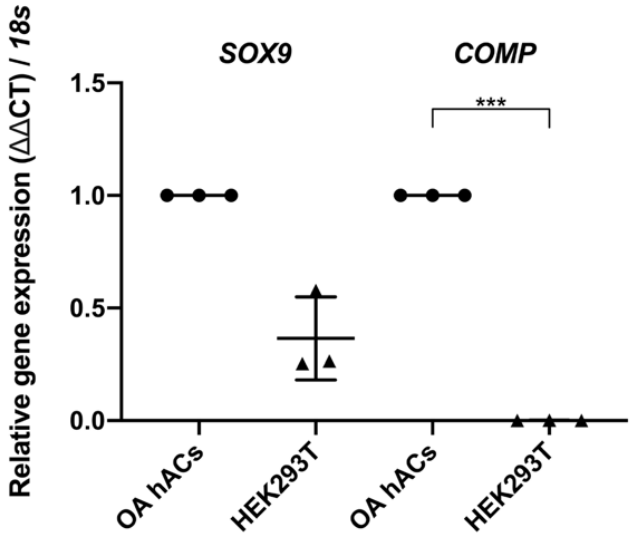


Fig. 5.2: In HEK293T cells *SOX9* expression levels are lower but detectable, compared with osteoarthritic, human articular chondrocytes.

Untreated cells were lysed in TRIzol, and RNA was reverse transcribed and quantified for expression levels of *SOX9* and *COMP* using qPCR. *S*amples were normalised to *18S* and expressed relative to their respective, normalised expression in OA hACs. Data are shown as mean ± SD. The statistical significance of comparisons was analysed with Student's t-test (two-tailed) for each gene between both cell types. *n* = 3.

5.2.2 *C9orf7* and *SOX9* sites possess putative MREs and were chosen to be investigated in parallel

To demonstrate functionality of their CRISPR mediated HDR method, Bassett *et al.,* 2014 chose (amongst others) *C9orf7* as a target gene, and demonstrated its direct regulation by *miR-92a*. *miR-92a* binds to a non-canonical binding site (MRE) at the 3'UTR of *C9orf7* (Fig. 5.3A). As proof of concept the following experiments were performed looking at *C9orf7* but also *SOX9* in parallel. *SOX9* possesses a classical seed region in its 3'-UTR for *miR-145* (Fig. 5.3B) (Martinez-Sanchez et al., 2012). To carry out experiments, ssODNs were designed for the *SOX9* site (Fig. 5.3C) or were derived from Bassett et al. (2014), for the *C9orf7* site (Fig. 5.3D)

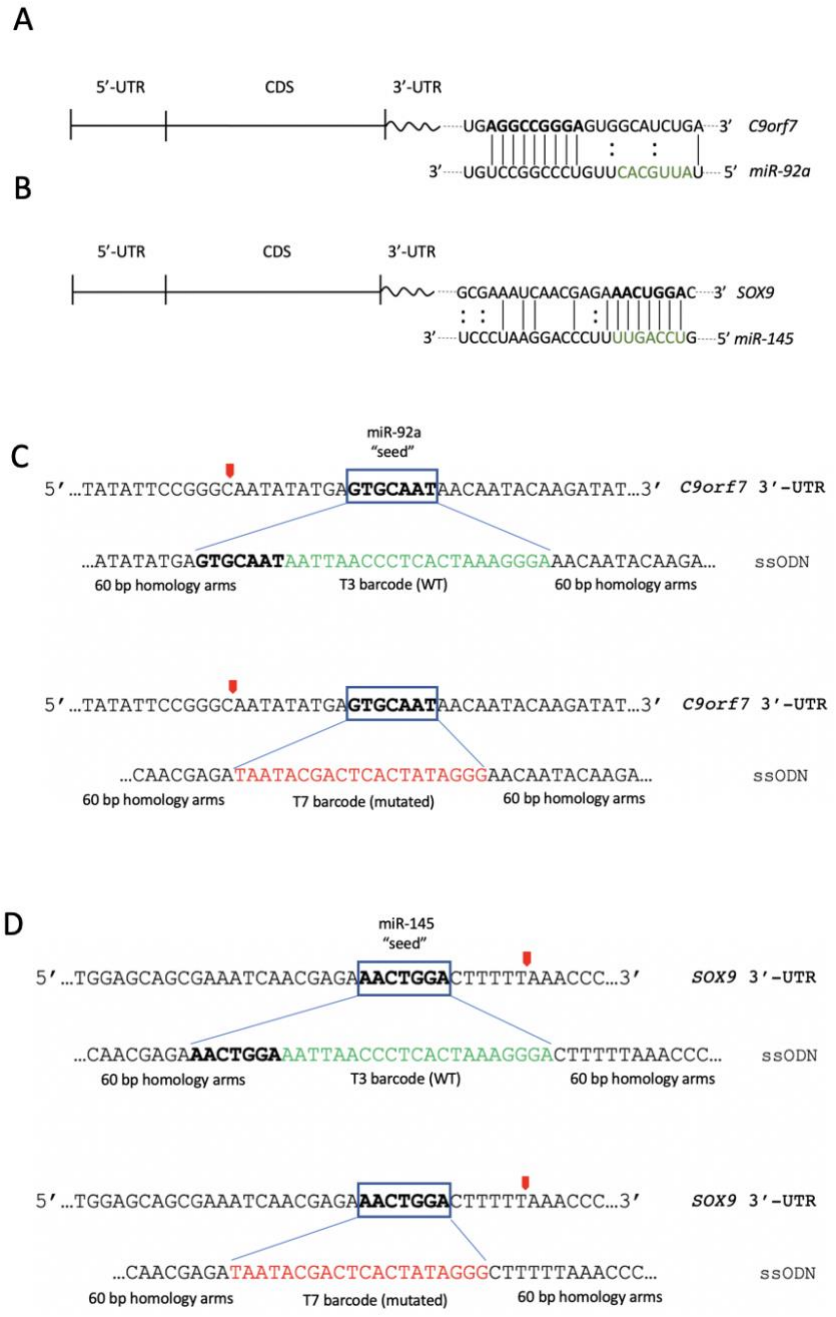


Fig. 5.3: A Schematic view of genomic regions of target genes *C9orf7* and *SOX9*, and their binding miRNAs, *miR-92a* and *miR-145*, respectively.

A *C9orf7* contains a non-canonical binding motif and **(B)** *SOX9* contains a classical seed region (bold). The seed regions of *miR-92a* and *miR-145* are displayed in green. T3 and T7 ssODNs contain 60 bp homology arms, flanking the **(C)** *miR-92a* seed (bold and in blue box), namely MRE (miRNA response element) and **(D)** *miR-145* seed. The T3 barcode (green) replaces and maintains the MRE while introducing a T3 primer binding site. The T7 barcode (red) deletes the MRE and introduces a T7 primer binding site. Red arrows indicate Cas9 cleavage site. *C9orf7* ssODNs were designed as per Bassett et al. (2014).

5.2.3 T7E1 determines efficient gene editing for the *SOX9* site after pX330 plasmid transfection in HEK293T cells

The T7 Endonuclease 1 (T7E1) (not to be mistaken for the T7 barcode) determines the efficiency of gene editing by recognising and cleaving non-perfectly matched DNA, such as heteroduplexes and nicked DNA. A MRE for *miR-145* in the 3'- UTR of the *SOX9* gene was targeted with a sgRNA and Cas9 expressing plasmid (pX330; as per original publication (Bassett et al., 2014)), and co-transfected along with the two ssODN barcodes (T3 and T7). Digesting gDNA from targeted cells with T7E1 shows two bands between 100 – 200 bp, representing indel (insertions and deletions) formation through non-homologous end joining (NHEJ) after Cas9 cleavage at 48 h for the *SOX9* locus (Fig. 4A). An increase in gene editing efficiency for the locus can be observed 72 h post transfection with the sgRNA and Cas9 expressing pX330 plasmid. The T7E1 assay does not indicate whether T3 or T7 have been incorporated, merely that targeting has occurred.

Similar observations were made at the RNA (cDNA) level; however, the bands were less intense (Fig. 5.4B). Lipofectamine only treated cells showed no indel formation at any time point (Fig. 5.4A/B).

In parallel to *SOX9*, a seed sequence for *miR-92a* in the 3'-UTR of the *C9orf7* gene was targeted with pX330, and co-transfected along with the T3 and T7 ssODNs. The T7E1 assay was performed for the *miR-92a* targeting locus in the 3'-UTR of *C9orf7,* but could

not be interpreted. The *C9orf7* PCR product ran as a duplet with apparent cleavage products for control as well as targeted samples (data not shown).

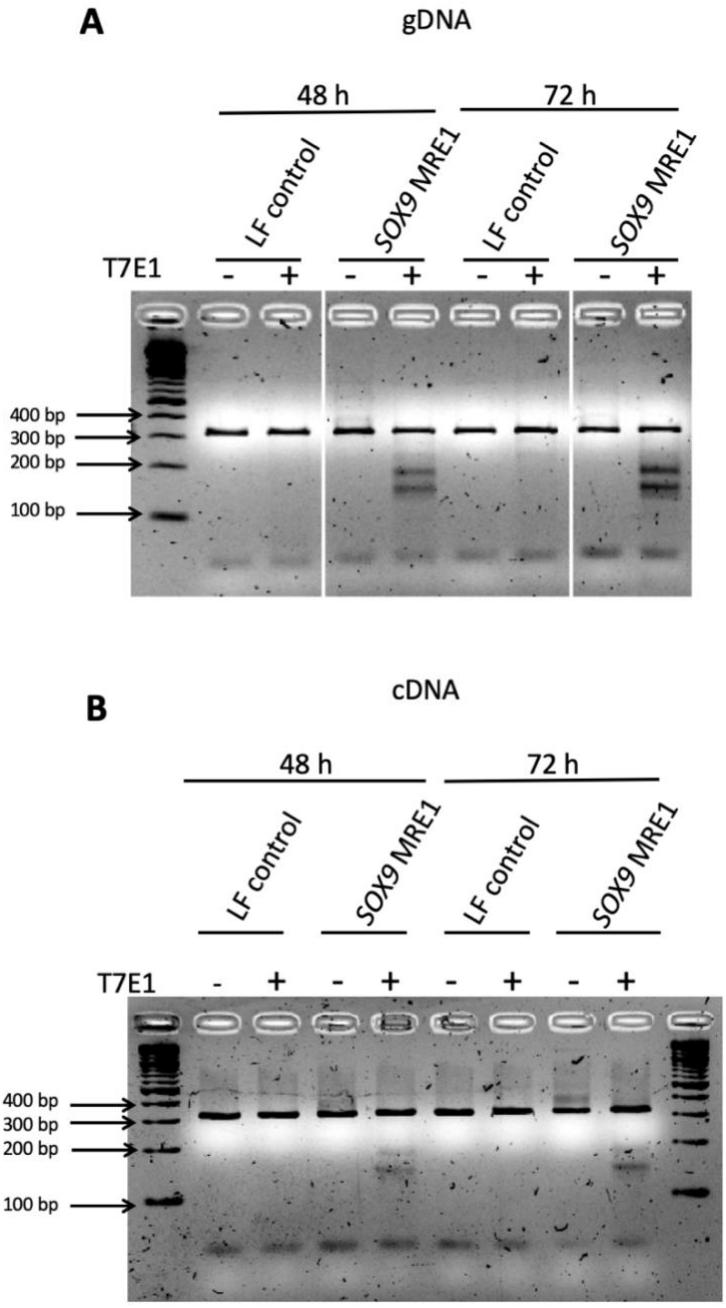


Fig. 5.4: Gene editing efficiencies of HEK293T cells after targeting *SOX9* with plasmid transfection.

T7 Endonuclease 1 (T7E1) assay validation of gene editing in **(A)** gDNA and **(B)** cDNA from targeted HEK293T cells. T7E1 assay run to determine DNA cleavage and indel formation for a *SOX9* target site after 48 h and 72 h. LF - Lipofectamine.

5.2.4 PCR verifies correct barcode insertions for the *C9orf7* and *SOX9* target sites

After successfully validating gene editing and indel formation through NHEJ in *SOX9* targeted cells, the correct insertions of the T7 and T3 ssODN barcodes into gDNA and cDNA by HDR were assessed next.

T3 ssODN successfully integrated in targeted HEK293T cells for the *C9orf7* and *SOX9* target sites via HDR (lane 8 and 12). Correct T3 insertion was confirmed by a band at 121 bp for *C9orf7* and 150 bp for *SOX9* in gDNA (Fig. 5.5A) and cDNA (Fig. 5.5B). Non-template control (lane 1/5/9), no primers control (lane 2/6/10), and endogenous forward primer only control (lane 3/7/11) showed no detectable signal for T3 at its expected sizes (Fig. 5A/B). Furthermore, no T3 integration was observed in the Lipofectamine only control. These results confirm correct CRISPR-Cas9 mediated T3 insertion for both target sites, *SOX9* and *C9orf7*.

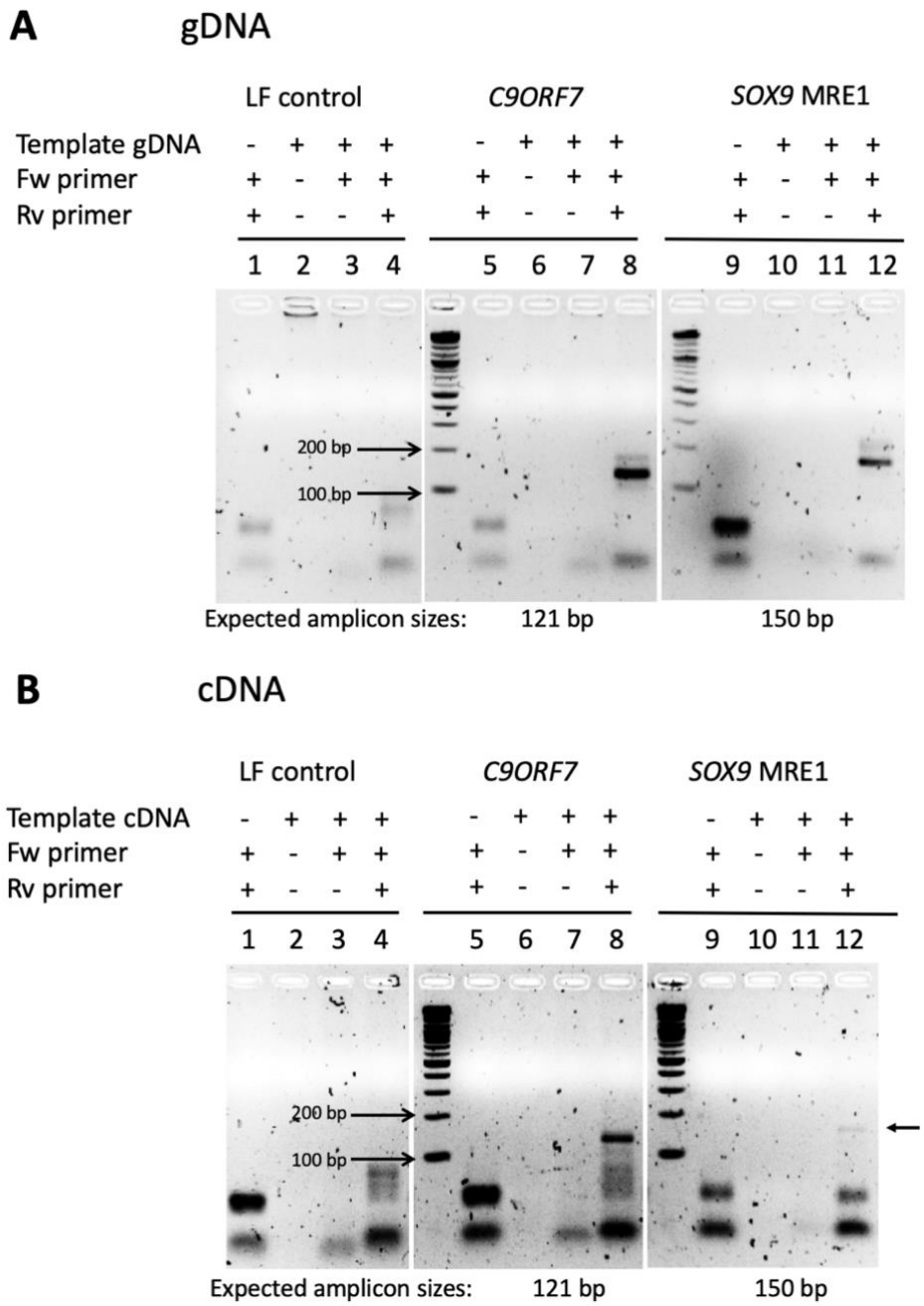


Fig. 5.5: Correct T3 barcode insertion verified by PCR.

PCR amplification with an endogenous forward primer and barcode specific reverse primer to verify correct insertion of the T3 barcode for the *C9orf7* (lane 4, 121 bp) and *SOX9* (lane 12, 150 bp) target in **(A)** gDNA and **(B)** cDNA. LF – Lipofectamine.

Similar observations were made for T7 integration in targeted HEK293T cells for both *C9orf7* and *SOX9* in gDNA (Fig. 5.6A) and cDNA (Fig. 5.6B). Correct T7 insertion was confirmed by a band at 114 bp for *C9orf7* and at 140 bp for *SOX9,* while all controls were negative.

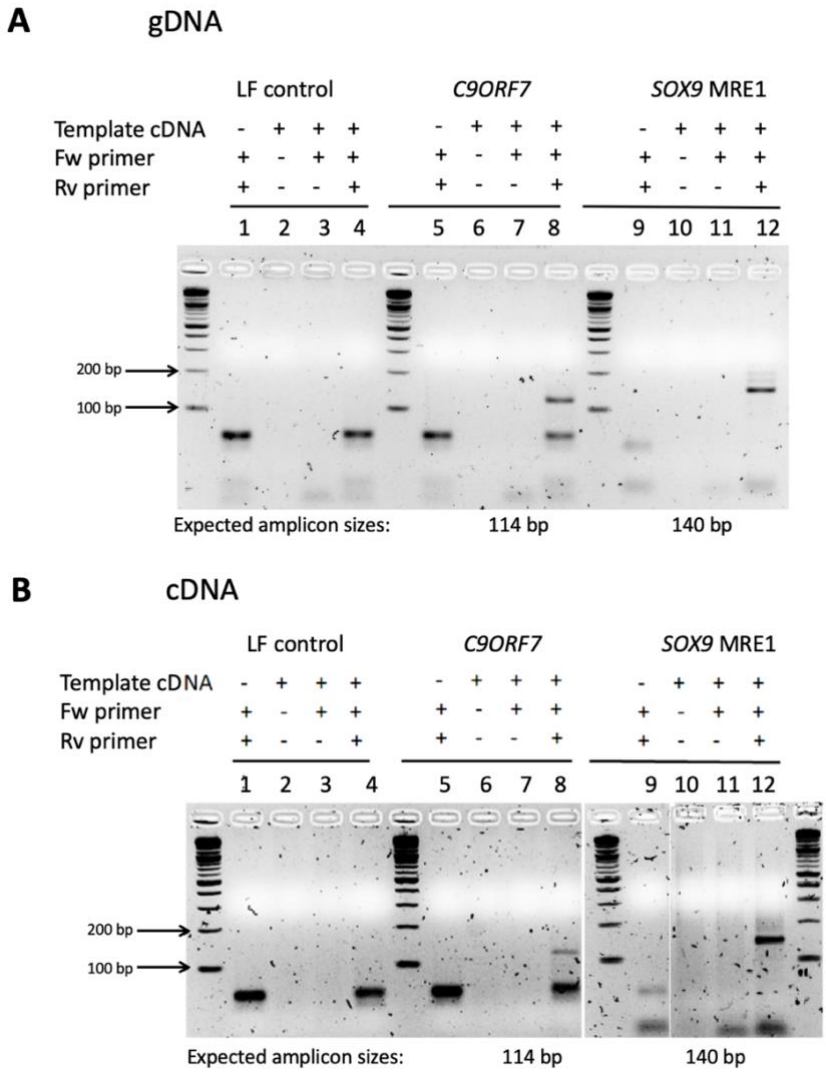


Fig. 5.6: Correct T7 barcode insertion verified by PCR.

PCR amplification with an endogenous forward primer and barcode specific reverse primer to verify correct insertion of the T7 barcode for the *C9orf7* (lane 4, 114 bp) and *SOX9* (lane 12, 140 bp) target in (A) gDNA and (B) cDNA. LF – Lipofectamine.

5.2.5 Functional readouts of *C9orf7* and *SOX9* via qPCR

gDNA and cDNA were subjected to real-time quantitative PCR analysis for T7 and T3 barcode expression levels. The T7/T3 ratio from gDNA estimates the relative integration efficiency. An increase in this ratio for the cDNA suggests an active MRE, as its deletion results in increasing amounts of transcript due to a decrease in miRNA-dependent degradation.

In the case of *C9orf7*, deletion of the putative MRE increases *C9orf7* mRNA from cDNA (Fig. 5.7A), confirming the original observation made by Bassett *et al.*, 2014.

Deletion of the putative MRE of *SOX9* did not change cDNA derived transcript abundance compared with that obtained from gDNA (Fig. 5.7B). These results were not consistent with miRNA-dependent regulation of *SOX9*.

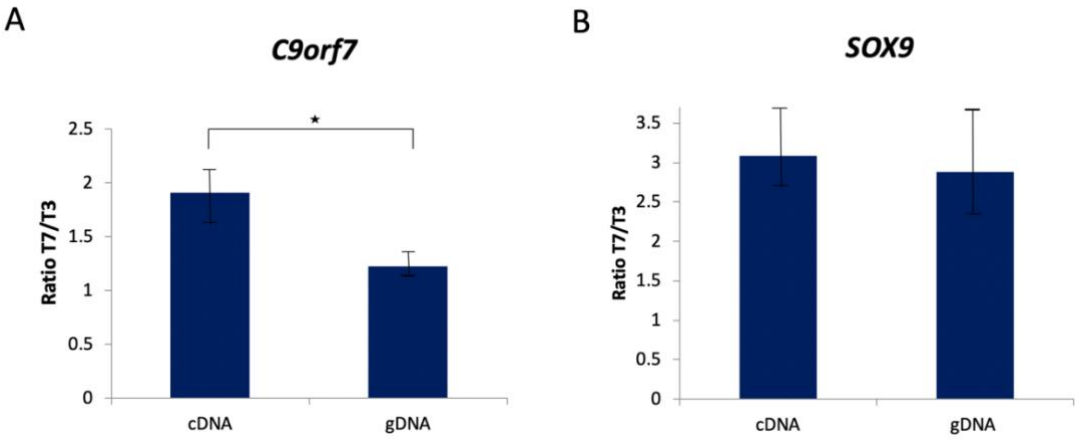


Fig. 5.7: Functional qPCR readouts for *miR-92a-C9orf7* and *miR-145-SOX9* targets.

PCR data showing analysis of MRE functionality in **(A)** *C9orf7* and **(B)** *SOX9*. The relative expression of T7 / T3 is displayed on the y axis, and results from gDNA and cDNA are shown. Results were obtained using the ΔCt method. Results are based on HEK293T cell cultures from three independent transfections ($n = 3$, $p < 0.05$).

Previously, Martinez-Sanchez et al. (2012) demonstrated through Luciferase assay that *SOX9* was regulated by *miR-145*. One possible explanation for why I was unable to demonstrate miRNA-dependent regulation of *SOX9* could be due to low expression levels of *miR-145* in HEK293T cells.

RNA from HEK293T cells (*n* = 3) and OA hACs (*n* = 3) was quantified for expression levels of *miR-145* using qPCR. *miR-145* was expressed in OA hACs, but barely expressed in HEK293T cells, while *miR-92a* (targeting *C9orf7*) was significantly more highly expressed in HEK293T cells compared with OA hACs (Fig. 5.8).

These results support the notion that low/non-existent levels of *miR-145* in HEK293T cells might explain the failure to demonstrate regulation at the MRE of *SOX9* in HEK293T cells.

In the next section I describe my attempts to repeat this methodology in human articular chondrocytes.

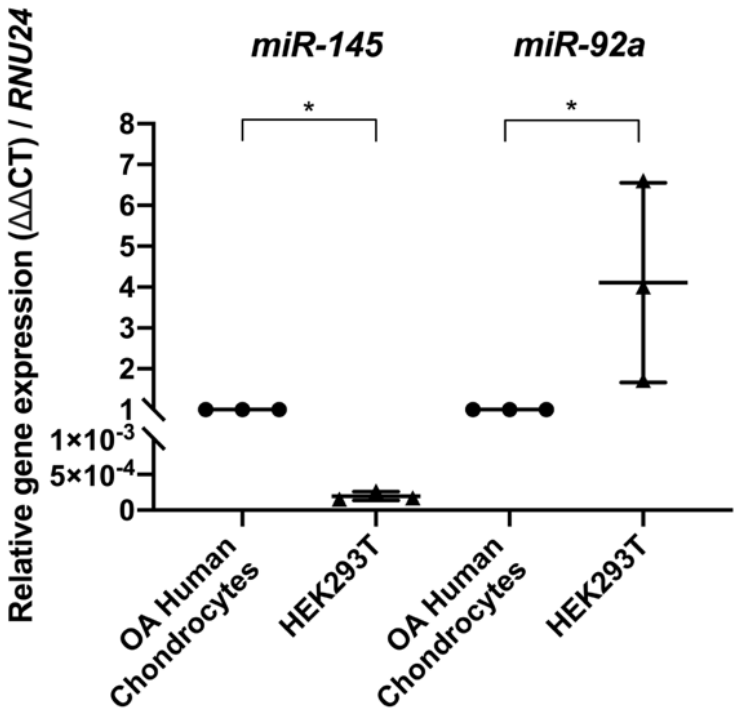


Fig. 5.8: *miR-145* is barely expressed in HEK293T cells.

Untreated cells were lysed in TRIzol, and RNA was reverse transcribed and quantified for expression levels of *miR-145* (and *miR-92a* control) using qPCR. *miR-145 and miR-92a* samples were normalised to *RNU24* and expressed relative to their respective, normalised expression in OA HACs. Data are shown as mean ± SD. The statistical significance of comparisons was analysed with Student's t-test (two-tailed). *n* = 3.

5.2.6 Targeting *miR-145* MRE of *SOX9* in primary osteoarthritic human articular chondrocytes

After successfully implementing and repeating the MRE activity assessment by the CRISPR HDR methodology in HEK293T cells, the next aim was to implement the same technique in primary OA hACs. In view of my findings in chapter 3, the method of transfection was changed to RNP transfection, as transfecting OA hACs by a Cas9 and sgRNA encoding plasmid (pX330) was found to be toxic (Fig. 3.2).

Before starting this, I took the opportunity to refine the barcode-containing ssODN sequences as I had encountered non-specific amplified products during the workflow in HEK293T cells (data not shown).

To counter this, four 22 nucleotide (nt) long barcodes were designed, each one consisting of two 11 nt sequences that only occur in the human genome between one and three times (Fig. 5.9A). Additionally, these barcodes were designed to provide higher primer annealing temperatures T_A, to further increase specificity. The use of nullomers, short DNA sequences, that do not occur in the human genome, was avoided due to potential interference with the human immune system, especially in a therapeutical application.

The designed barcodes were named Maintenance 1 (M1), Maintenance 2 (M2), Deletion 1 (D1) and Deletion 2 (D2), and were flanked with 60 bp homology arms for the *SOX9* MRE target site (Fig. 5.9B).

A

Annotation	Sequences used as barcodes
M1	CATATCGCGCGCGTAATCGA
D1	TCGCGCGAATAATCGTCGACGA
M2	CGACGGCAGTAACGCGCGATAT
D2	CCGAATACGCGCGAATCGCGTA

B

Fig. 5.9: Primer-binding barcodes and their schematic integration at the predicted *miR-145* binding site (MRE).

A Four newly designed barcodes (M1, D1, M2, D2) consist of the combination of 8 different 11 bp long sequences that only occur in the human genome between one and three times. **B** All four ssODNs contain 60 bp homology arms, flanking the miR-145 seed (bold and in blue box), namely MRE (miRNA response element). The M1 (green) and M2 (blue) barcodes replace and maintain the *SOX9* MRE, while introducing a M1 or M2 primer binding site. The D1 (red) and D2 (orange) barcodes delete the *SOX9* MRE and introduce a D1 or D2 primer binding site. Red arrows indicate Cas9 cleavage site.

The *SOX9* MRE1 was targeted with the same sgRNA and Cas9 – this time as an RNP complex (as opposed to a pX330 plasmid), and co-transfected alongside the newly designed ssODNs. *SOX9* was targeted and co-transfected with either M1 + D1, or M2 + D2, or the old barcodes, T3 + T7. All targeted samples were PCR amplified to generate 330 bp long fragments surrounding *SOX9* MRE1. Digested PCR products from targeted cells in the T7E1 assay showed two bands between 100 – 200 bp, representing indel formation through NHEJ (Fig. 5.10).

The respective NT controls did not exhibit DNA cleavage to the same extent.

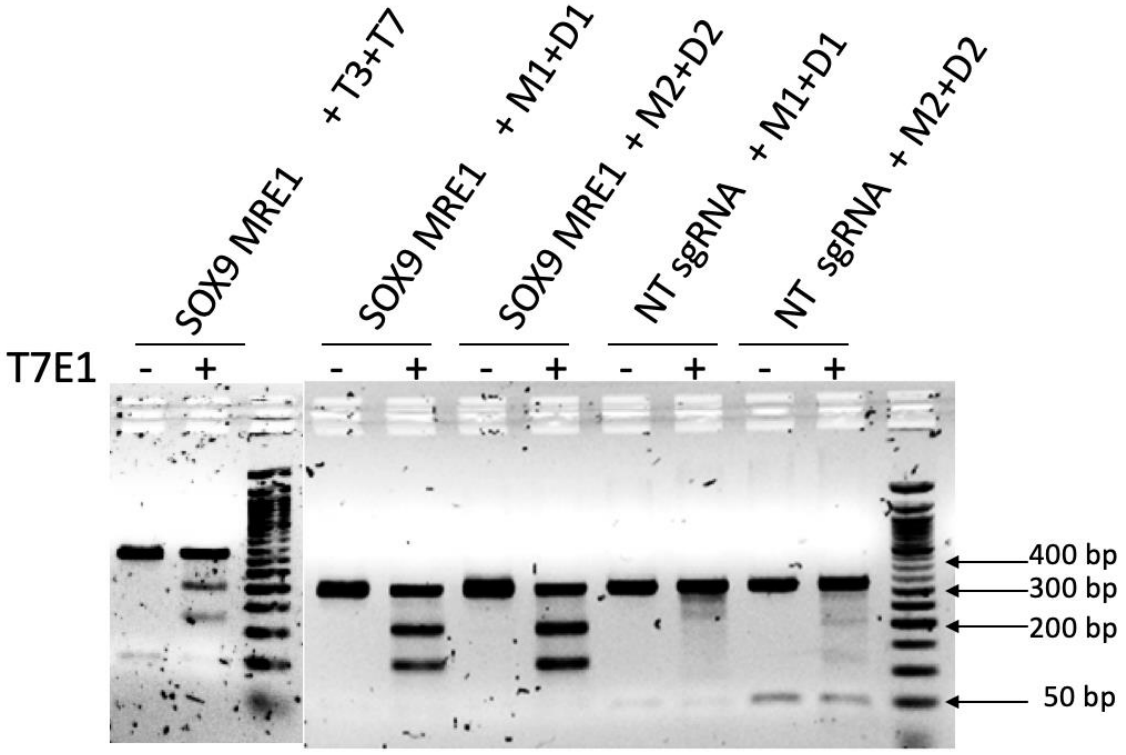


Fig. 5.10: T7E1 validation of DNA cleavage in gDNA from *SOX9* MRE1 targeted OA hACs.

SOX9 MRE1 was RNP transfected with three different barcode combinations. 330 bp-long products were PCR amplified. Subsequently, the T7E1 assay was applied to assess DNA cleavage and indel formation for a SOX9 target site transfected with three different barcode combinations (T3+T7 / M1+D1 / M2+D2) after 48 h. NT – Non-targeting.

5.2.7 Identification of optimal ssODNs (M1 + D1) for insertion into *SOX9* in osteoarthritic human articular chondrocytes

The next aim was to determine whether the co-transfected ssODNs were successfully integrated in the genome via HDR. For each barcode, endogenous forward primers were used in combination with barcode specific reverse primers, in order to verify correct barcode insertion in the gDNA of *SOX9* targeted OA hACs.

Three out of six barcodes gave strong bands at their predicted sizes (D1 = 146 bp, M1 = 155 bp, D2 = 146 bp) (Fig. 5.11). All three also displayed additional bands of unknown identity. Bands were present for barcode M2 but this was only weak at the predicted size (155 bp, black arrow on Fig. 5.11) and a stronger, but smaller band was also apparent.

For T3 + T7 no products were visible, suggesting a lack of barcode amplification. Control samples ("-", no template) did not show barcode insertion.

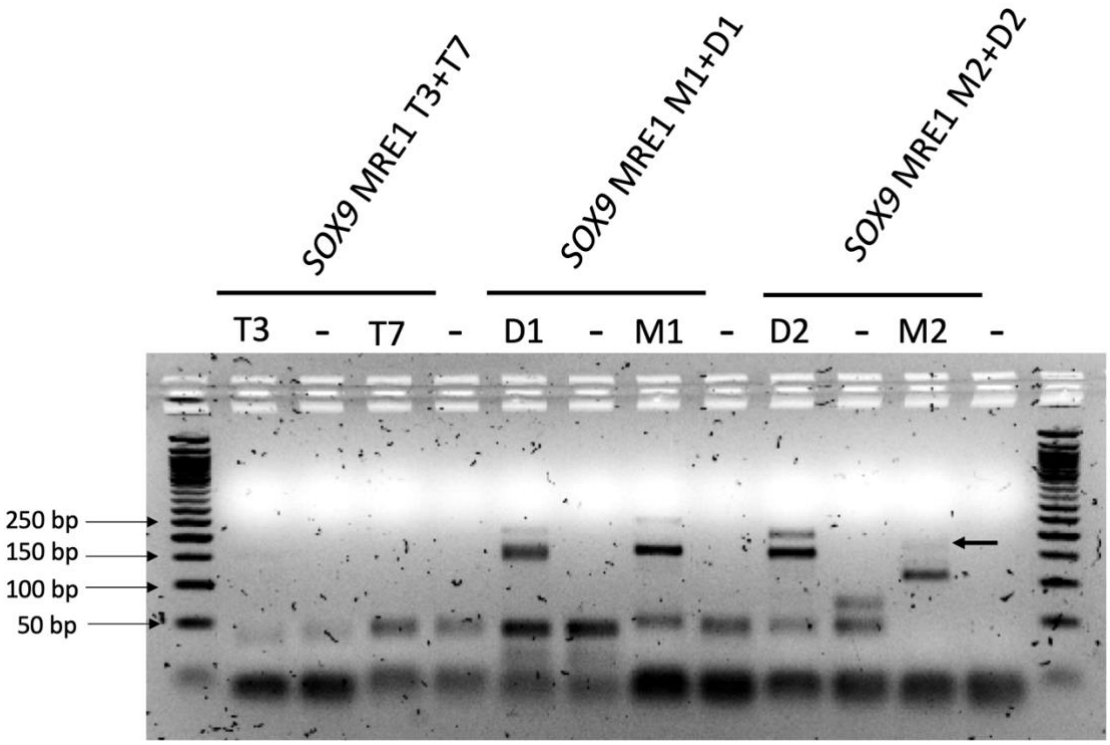


Fig. 5.11: PCR verifies barcode integration for three barcodes (out of six) in gDNA from RNP transfected OA hACs.

SOX9 MRE1 was RNP transfected with three different barcode combinations, and subsequent barcode integration was assessed by Sigma-Aldrich JumpStart™ Taq DNA Polymerase.
PCR amplifications were performed for each transfection to assess barcode insertions. Predicted PCR products: T3 = 150 bp, T7 = 140 bp, D1 = 146 bp, M1 = 155 bp, D2 = 146 bp, M2 = 155 bp. "-" control without gDNA template. Arrow indicates correct M2 barcode size.

5.2.8 Testing M1 / D1 barcodes with non-targeting sgRNA controls

From the previous experiment, barcodes M1 and D1 appeared to be the most robust combination and hence were investigated further.

M1 and D1 were tested for correct insertion into OA hACs from isolated gDNA and cDNA. Each sample was run in duplicate. A satisfactory result would have included a single band at the correct size in targeted but not non-targeted cells. However, (i) bands were variably present in the NT sgRNA controls (Fig. 5.12). (ii) All of the bands present (targeted + non targeted) were duplets or triplets. (iii) There was significant variation between duplicate samples. (iv) All PCR reactions were accompanied by primer dimer formation of around 50 bp.

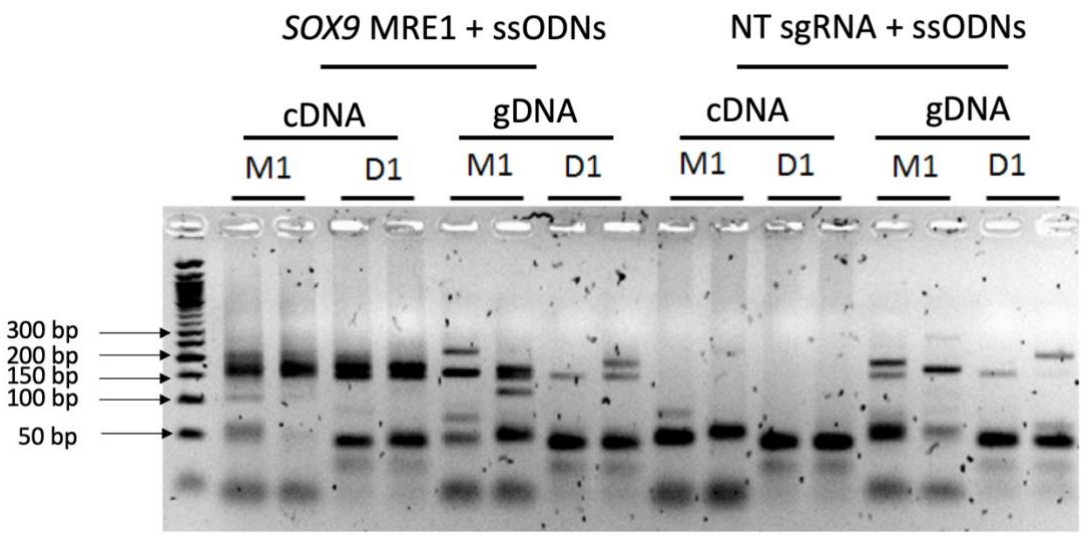


Fig. 5.12: PCR verifies correct barcode insertions for RNP transfected *SOX9* MRE1 in OA hACS, but also amplifies non-specific products.

SOX9 MRE1 was RNP co-transfected with barcodes M1 (155 bp) and D1 (146 bp), and assessed for their integration in duplicate in gDNA and cDNA. A NT sgRNA with co-transfected barcodes acted as control. Barcode integration was assessed by Sigma-Aldrich JumpStart™ *Taq* DNA Polymerase.

5.2.9 Refining the methodology to improve reproducibility.

5.2.9.1 Does changing polymerases affect barcode amplification?

Assessment of M1 and D1 barcode integrations (using Sigma-Aldrich JumpStart™ *Taq* DNA Polymerase) was carried out exactly as described in the previous experiment except this time in single PCR reactions. Three additional polymerases were used for PCR barcode amplification in RNP transfected OA hACs. All three polymerases were used to run PCRs with the same primers. The primer annealing temperatures (T_A) were adapted according to the manufacturer's protocols.

Both the SensiFAST (Fig. 5.13B) and PrimeTime (Fig. 5.13C) polymerases did not amplify M1 and D1 at their predicted sizes. Moreover, they displayed what was observed in previous experiments: duplet formations, non-specific bands, primer dimer formations, and bands in the NT control.

The Q5 Hot Start polymerase however (Fig. 5.13A) amplified both barcodes at their predicted sizes on cDNA and gDNA level, though correct barcode insertions were accompanied by non-specific bands, a faint signal for D1 on cDNA level in the NT sgRNA control, and primer dimer formation throughout. The Q5 Hot Start polymerase appeared to be the most appropriate choice going forward.

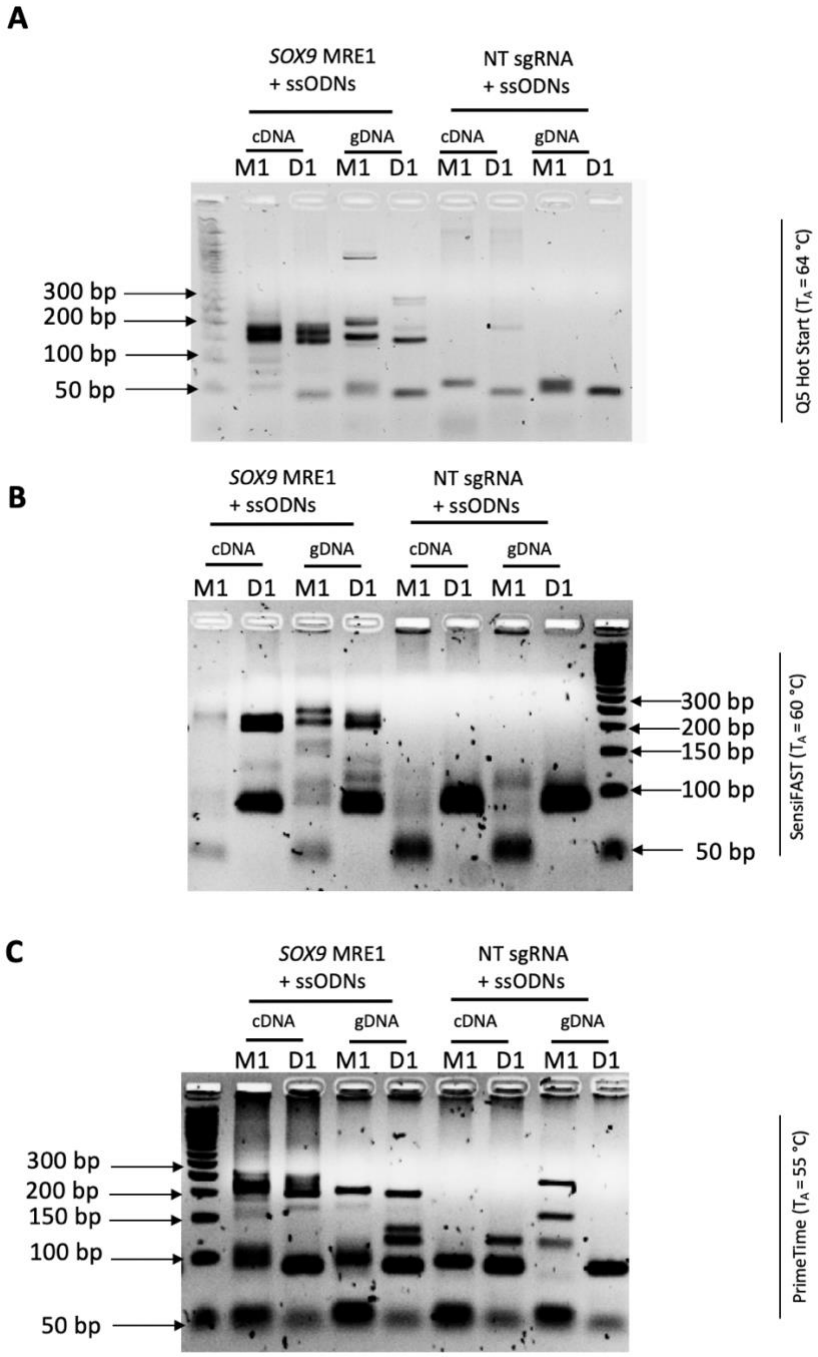


Fig. 5.13: Variable results of PCR barcode insertions using three different polymerases for RNP transfected *SOX9* MRE1 in OA hACs.

SOX9 MRE1 was RNP co-transfected with barcodes M1 (expected amplicon size: 155 bp) and D1 (146 bp), and assessed for their integration in in gDNA and cDNA using **(A)** New England BioLabs Q5 Hot Start High-Fidelity 2x MM, **(B)** Bioline SensiFAST™ Real-Time PCR kit, and **(C)** Integrated DNA Technologies PrimeTime Gene Expression® Master Mix. A NT sgRNA with co-transfected barcodes acted as control.

5.2.9.2 Does increasing primer annealing temperature and primer specificity give a solid PCR barcode readout?

Even when using the Q5 Hot Start High-Fidelity 2x MM polymerase to verify barcode insertion, the obtained results were still unsatisfactory (Fig. 5.13A). Correct barcode insertions were accompanied by non-specific bands, duplet formation, and primer dimerisation. Additionally, a faint barcode signal was detected for the NT sgRNA control. As a consequence, two new primer sets were designed (Fig. 5.14A) to assess barcode insertion for the *SOX9* MRE: i) Primers that had an increased T_A, were 22 nucleotide (nt) long (now covering the entire barcode), and a modified forward primer (23 nt), designed to match the new reverse primer (72 °C rather than 69 °C). ii) The same, newly designed endogenous forward primer was used in conjunction with a new reverse primer, that extended downstream of the barcode into the endogenous *SOX9* sequence (Fig. 5.14A). Specifically, these bound 6 nt on the endogenous SOX9 sequence and 14 nt (for deletion barcodes) or 19 nt (for maintenance barcodes) on the barcode sequence. This held the potential advantage of enhancing target specificity, at the cost of losing the convenience of having universal barcode primers. Primer T_A for semi-endogenous primers was 72 °C as well.

These new primer pairs were then used to assess barcode integration for 4 (M1, M2, D1, D2) barcodes in RNP transfected and *SOX9* targeted OA hACs, either in the presence or absence of dimethyl sulfoxide (DMSO). DMSO is commonly used in PCR reactions to enhance PCR conditions by eliminating secondary structures of the DNA and resolving high GC content. Primer pairs with the higher annealing temperatures revealed specific

bands for all barcodes at their predicted sizes (expected amplificon sizes: M1 = 279 bp; M2 = 279 bp; D1 = 270 bp), except for D2 (270 bp) (Fig. 5.14B). No bands were present in the NT Control. However, all samples were accompanied with strong primer dimer bands at 50 bp.

Similar observations were made when assessing barcode integration with the newly designed semi-endogenous primers (Fig. 5.14C). Strong bands were amplified for all four barcodes at their predicted sizes (expected amplificon sizes: M1 = 275 bp; M2 = 274 bp; D1 = 264 bp; D2 = 264 bp), both in presence and absence of DMSO. However, bands were also present in the NT sgRNA control. Primer dimerisation bands at 50 bp were still present.

Based on the improved ratio between specific bands and primer dimer bands, it was decided to continue further refinements with the semi-endogenous primers, with particular emphasis on eliminating the ssODN signal in the NT sgRNA control.

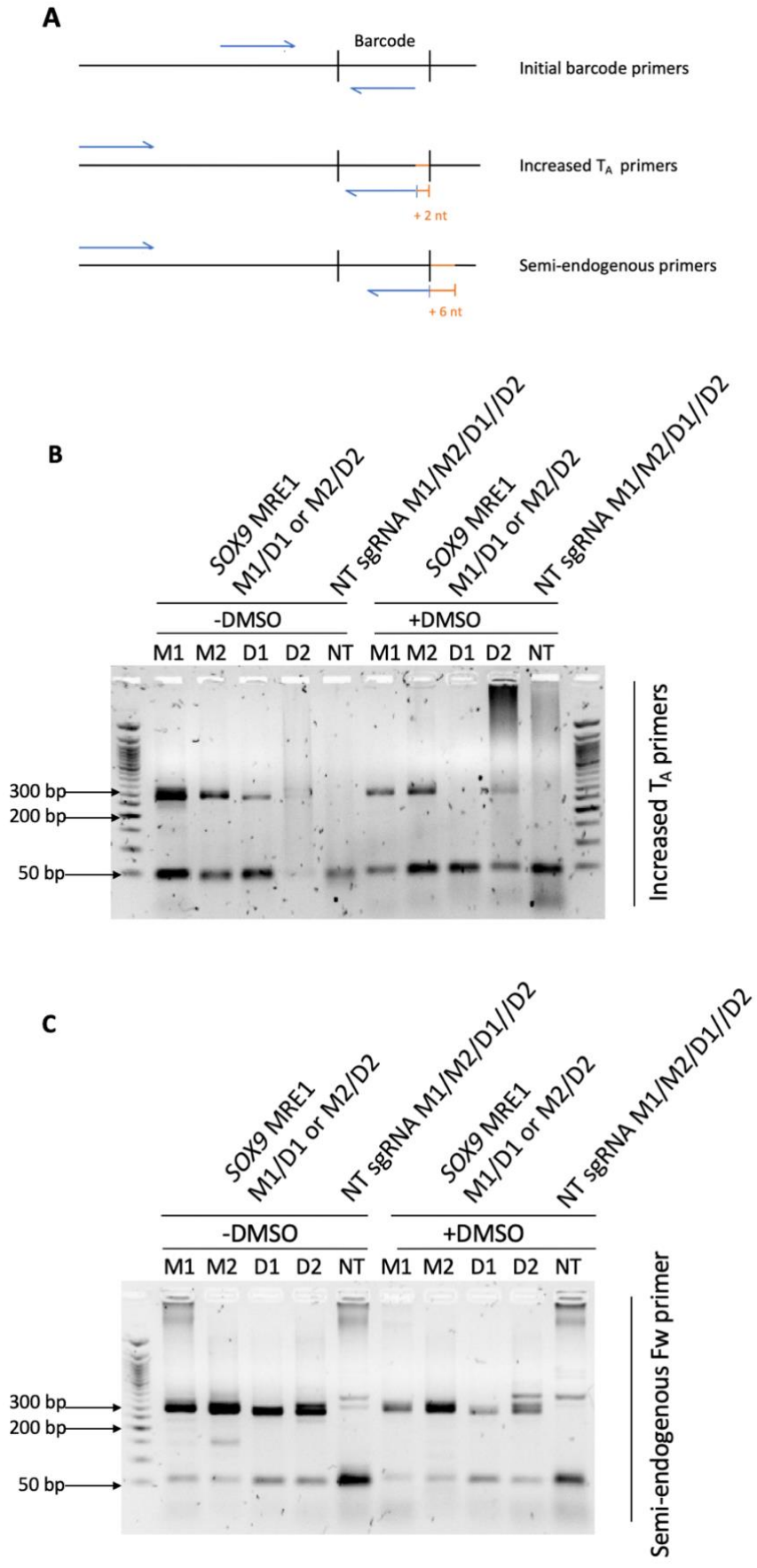

A
Barcode
Initial barcode primers
Increased T_A primers
+ 2 nt
Semi-endogenous primers
+ 6 nt
B
SOX9 MRE1 M1/D1 or M2/D2
NT sgRNA M1/M2/D1//D2
SOX9 MRE1 M1/D1 or M2/D2
NT sgRNA M1/M2/D1//D2
-DMSO
+DMSO
M1 M2 D1 D2 NT M1 M2 D1 D2 NT
300 bp
200 bp
50 bp
Increased T_A primers
C
SOX9 MRE1 M1/D1 or M2/D2
NT sgRNA M1/M2/D1//D2
SOX9 MRE1 M1/D1 or M2/D2
NT sgRNA M1/M2/D1//D2
-DMSO
+DMSO
M1 M2 D1 D2 NT M1 M2 D1 D2 NT
300 bp
200 bp
50 bp
Semi-endogenous Fw primer

Fig. 5.14: Assessment of *SOX9* MRE1 barcode insertion by changing barcode primers in OA hACs gives increased specificity.

SOX9 MRE1 was RNP co-transfected with barcodes M1 and D1 or M2 and D2, and their cDNA was assessed for barcode integration including and excluding DMSO as supplement. **(A)** Schematic view of different primer sets used to verify barcode insertion. In order to verify correct barcode insertion, a newly designed endogenous forward primer was used in conjunction with **(B)** universal reverse primers that exhibit higher primer annealing temperatures (T_A) (expected amplificon sizes: M1 = 279 bp; M2 = 279 bp; D1 = 270 bp; D2 = 270 bp), and additionally **(C)** bind partly on the barcode, partly at the endogenous sequence (expected amplificon sizes: M1 = 275 bp; M2 = 274 bp; D1 = 264 bp; D2 = 264 bp). A non-targeting (NT) sgRNA with all co-transfected barcodes acted as control. New England BioLabs Q5 Hot Start High-Fidelity 2x MM used for barcode integration assays. Please note change in pipetting order (M1-M2-D1-D2).

5.2.9.3 By using specific, semi-endogenous barcode primers, do we get rid of the signal in the non-targeting sgRNA control?

Going forward in RNP transfected and *SOX9* targeted OA hACs, with Q5 Hot Start High-Fidelity 2x MM polymerase and semi-endogenous barcode primers, the next step involved validation of the ssODN bands in transfected control cells. OA hACs were RNP transfected in 4 different conditions and analysed at 48 h: *SOX9* MRE with ssODN barcodes (either M1/D1 or M2/D2) or without barcodes, or NT sgRNA with ssODN barcodes (either M1/D1 or M2/D2) or without barcodes. All conditions were run in the presence or absence of DMSO. Templates were colour coded for easier identification.

Presence of barcodes M1 (275 bp), D1 (264 bp), M2 (274 bp) and D2 (264 bp) were confirmed at their correct sizes (Fig. 5.15), however, bands were also present in the NT controls in some conditions.

To conclude, M1 and D1 in the absence of DMSO provided the most promising read-out. However, once again, all PCR reactions were accompanied by strong primer dimer formations, and strategies were required to remove these as they interfered with qPCR read-outs.

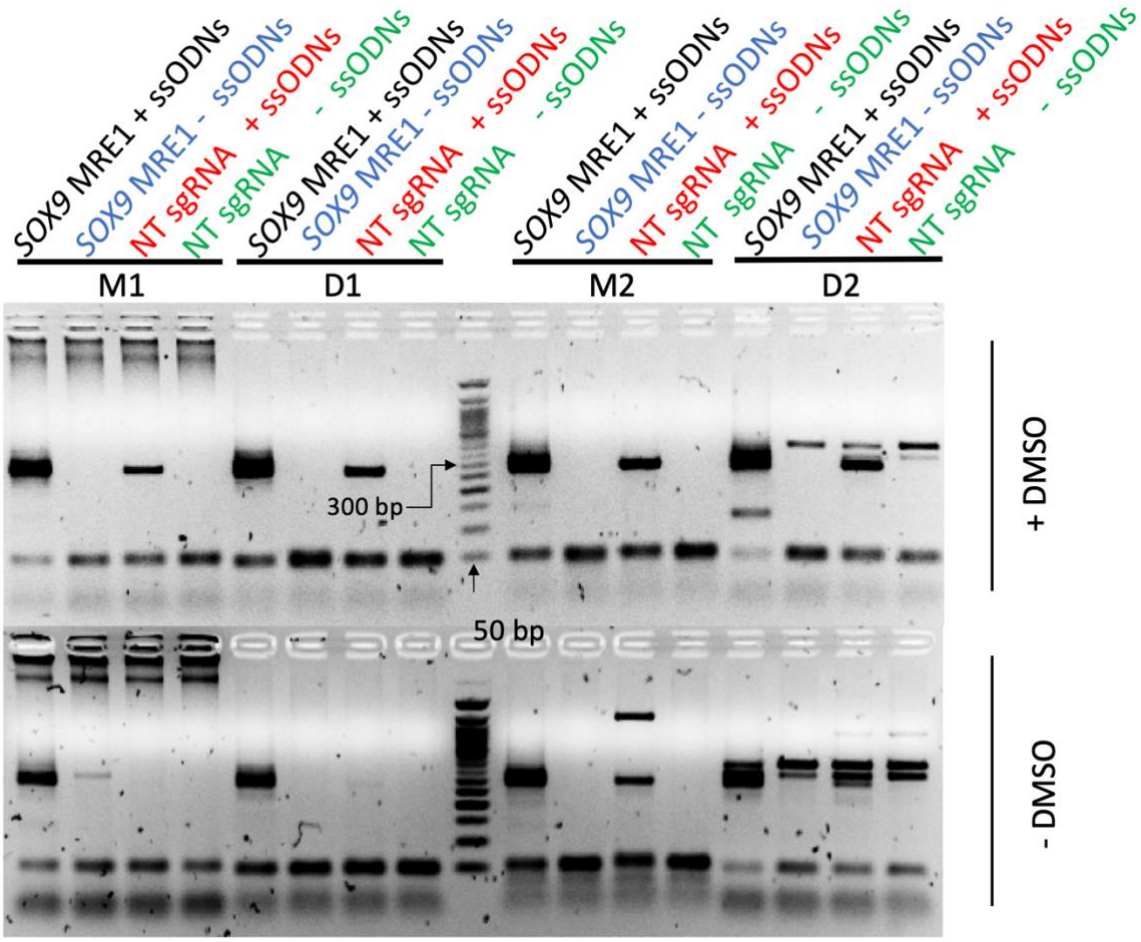


Fig. 5.15: PCR amplifies ssODNs in RNP transfected NT sgRNA controls using semi-endogenous primers in OA hACs.

SOX9 MRE1 and NT sgRNA controls were RNP co-transfected either with i) barcodes M1 (275 bp) and D1 (264 bp), or ii) M2 (274 bp) and D2 (264 bp), or iii) no barcodes at all. cDNA was checked for barcode integration in the presence or absence of DMSO 48 h post transfection. New England BioLabs Q5 Hot Start High-Fidelity 2x MM used for barcode integration assays. Template names colour coded for easy identification.

5.2.9.4 Does using decreased primer concentrations reduce primer dimer formation and barcode signal in non-targeting sgRNA control?

With the objective of removing primer dimer formation, M1 and D1 barcode verification was assessed by PCR with decreasing primer concentrations, using cDNA from RNP transfected, *SOX9* targeted OA hACs. Semi-endogenous primers and Q5 Hot Start High-Fidelity 2x MM polymerase were used.

All barcode insertion PCRs up to this point were carried out at a final concentration of 400 nM for each of the forward and reverse primers. Now, M1 and D1 barcode insertions were assessed with a final forward and reverse primer concentration of each 200 nM (1:2), 80 nM (1:5), and 40 nM (1:10). in addition to the original concentration (400 nM).

PCR results revealed (Fig. 5.16) that by using lower primer concentrations (80 nM and 40 nM final concentration), primer dimer formation was mitigated almost completely (red arrows). M1 (275 bp) and D1 (264 bp) were successfully verified at their correct sizes. Similarly, no band was present for barcode M1 in the NT sgRNA control at lower primer concentrations of 80 nM and 40 nM. Barcode D1 was present albeit with faint bands at those concentrations. These results highlight the variability between experiments.

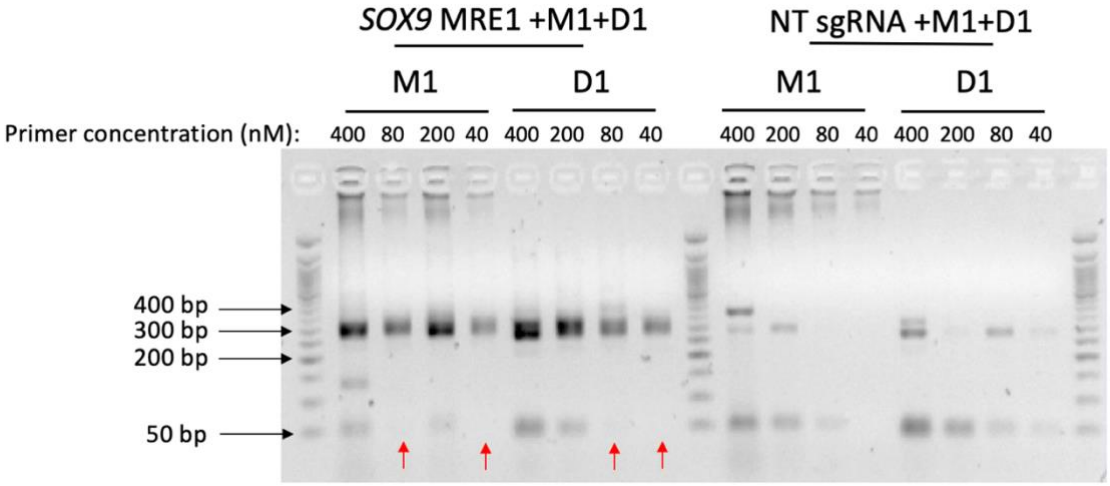


Fig. 5.16: Primer concentration optimisation gets rid of primer duplet formation using semi-endogenous primers for RNP transfected *SOX9* MRE1 in OA hACs.

SOX9 MRE1 was RNP co-transfected with barcodes M1 (275 bp) and D1 (264 bp), and their reverse transcribed cDNA was assessed for barcode integration by PCR with varied concentrations of semi-endogenous primers, and using New England BioLabs Q5 Hot Start High-Fidelity 2x MM. Samples, that were transfected with a non-targeting (NT) sgRNA (including co-transfected barcodes M1 and D1) were used as controls. Red arrows indicate no primer dimer formations. Please note 200 nM and 80 nM primer concentrations for *SOX9* MRE1 M1 were pipetted in the wrong order; labelling is correct.

5.2.9.5 Could changing the annealing temperature improve barcode amplification?

Another challenge which consistently appeared throughout PCR barcode assessments was the presence of duplet formation. Semi-endogenous primers along with the Q5 Hot Start High-Fidelity 2x MM polymerase were used to run a temperature gradient PCR. This assessed barcode M1 and D1 amplification between 62 °C and 72 °C from cDNA of RNP transfected, *SOX9* targeted OA hACs (without DMSO).

72 °C (lane A) and 63 °C (lane G) (red arrows) appeared to be the best conditions for barcode amplification (Fig. 5.17). At 72 °C, M1 (275 bp) was amplified with one dominant band, while D1 showed faint duplet formation. M1 did not show any amplification band at 72 °C for the NT sgRNA control. At most temperatures there was still evidence of some duplet formation and variable signal in the NT control.

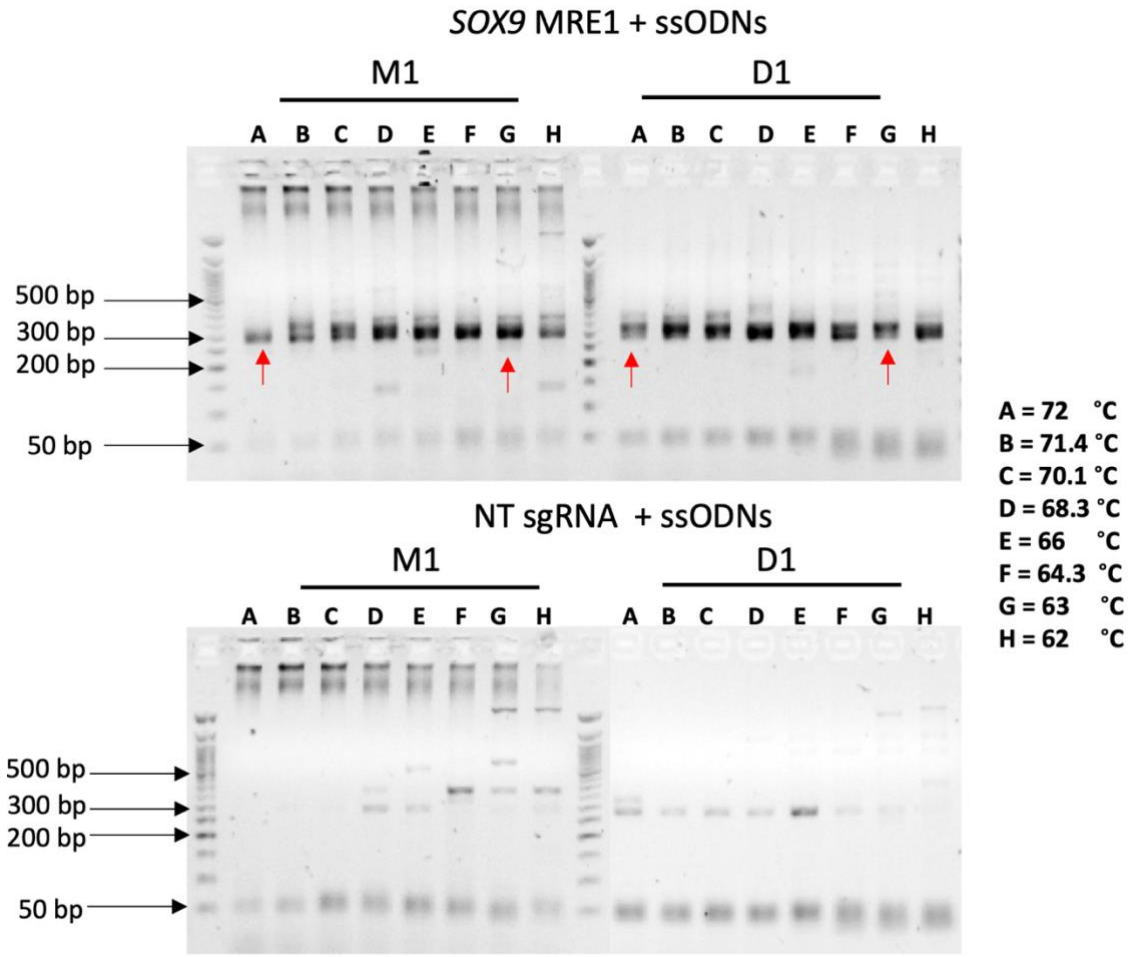


Fig. 5.17: Temperature gradient PCR reveals barcode insertions with different amplification patterns.

SOX9 MRE1 was RNP co-transfected with barcodes M1 (275 bp) and D1 (264 bp), and their cDNA was assessed at 48 h for barcode integration using semi-endogenous primers at 80 nM and New England BioLabs Q5 Hot Start High-Fidelity 2x MM. Red arrows suggest less prominent duplet formation.

5.2.9.6 Does implementing these refined PCR conditions lead to a stable PCR read-out?

Through extensive troubleshooting thus far, certain conditions were established for the optimal assessment of correct barcode insertion from the cDNA of RNP transfected, *SOX9* targeted OA hACs: Q5 Hot Start High-Fidelity 2x MM, along with semi-endogenous primers at 40 nM (or 80 nM) final concentration, and a primer T_A of 72 °C (or 63 °C) appeared to be optimal, albeit not perfect.

Implementing these PCR conditions with 72 °C primer T_A, M1 was characterised by a solid band at 80 nM final primer concentration (275 bp), and a fainter band at 40 nM final primer concentration (Fig. 5.18). Additionally, M1 was not detected in the NT sgRNA control. D1 was characterised by duplet formation (264 bp) at both primer concentrations, and was not detected at the NT sgRNA control.

At 63 °C primer T_A both ssODNs were successfully verified at their expected sizes, but also showed non-specific bands in the NT sgRNA control (Fig. 5.18).

Having now refined a methodology to optimise barcode insertion, the next step was to confirm insertion of the barcode by qPCR using SYBR green. However, this did not lead to any functional read-outs (data not shown). Additionally, using the same conditions and replacing the Q5 Hot Start High-Fidelity 2x MM polymerase + SYBR green dye with a standardised SYBR green qPCR MM at 65 °C or 72 °C primer T_A temperatures, did not provide a functional read-out either (data not shown).

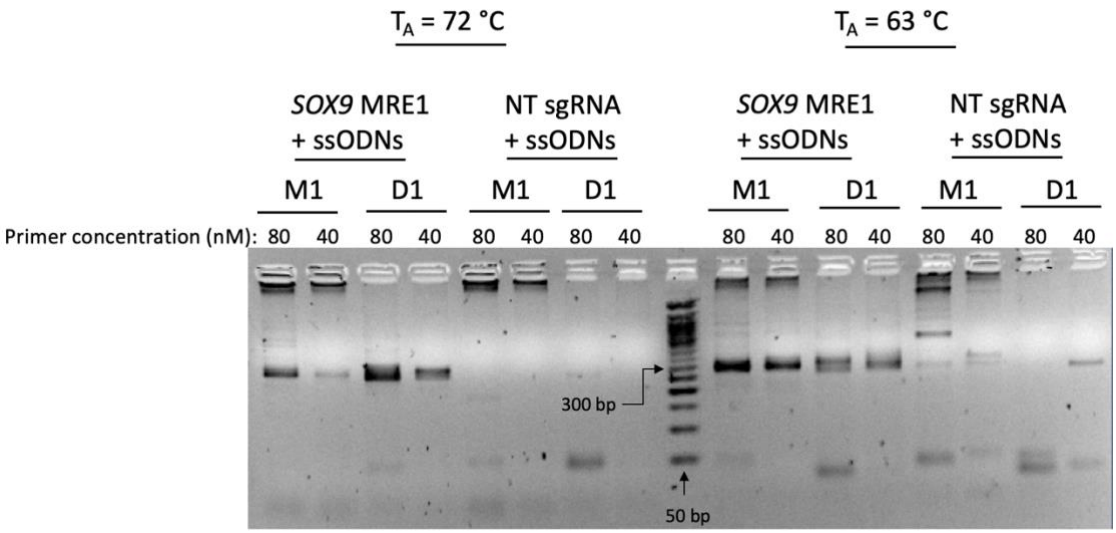


Fig. 5.18: PCR verifies correct barcode insertion for RNP transfected *SOX9* MRE1 in OA hACs using semi-endogenous primers.

SOX9 MRE1 was RNP co-transfected with barcodes M1 and D1, and their cDNA was assessed at 48 h for barcode integration. Troubleshot and prime identified conditions were used to verify M1 (275 bp) and D1 (264 bp) barcode insertion, namely semi-endogenous primers at T_A = 72°C and at 80 nM final concentration, with New England BioLabs Q5 Hot Start High-Fidelity 2x MM polymerase.

5.2.10 Running PCR samples in triplicate reveals the irreproducibility of the results

To investigate why my optimised protocol did not lead to any reliable qPCR read-outs, barcode insertion PCRs were run in triplicate. For each condition, one PCR master mix was prepared. From each master mix, the same amount (10 µl) was carefully pipetted into three different PCR tubes, and these were subsequently amplified next to each other in the same thermocycler.

Both barcodes M1 and D1 showed variable and inconsistent PCR bands when compared within their triplicate reaction at all three different primer concentrations (400 nM / 80 nM / 40 nM) (Fig. 5.19).

As previously observed, barcodes M1 and D1 did not exhibit PCR bands at lower final primer concentrations (40 nM) in the NT sgRNA control, whereas this was seen at higher final primer concentrations (400 nM), though in only one of the tree triplicates. Similarly, both for the targeted and non-targeted samples, primer dimer bands were reduced with decreasing primer concentrations, but were inconsistent within triplicates, especially in the targeted samples.

Collectively, despite laborious protocol refinements, I was unable to achieve a reliable functional read-out using this technique in OA hACs.

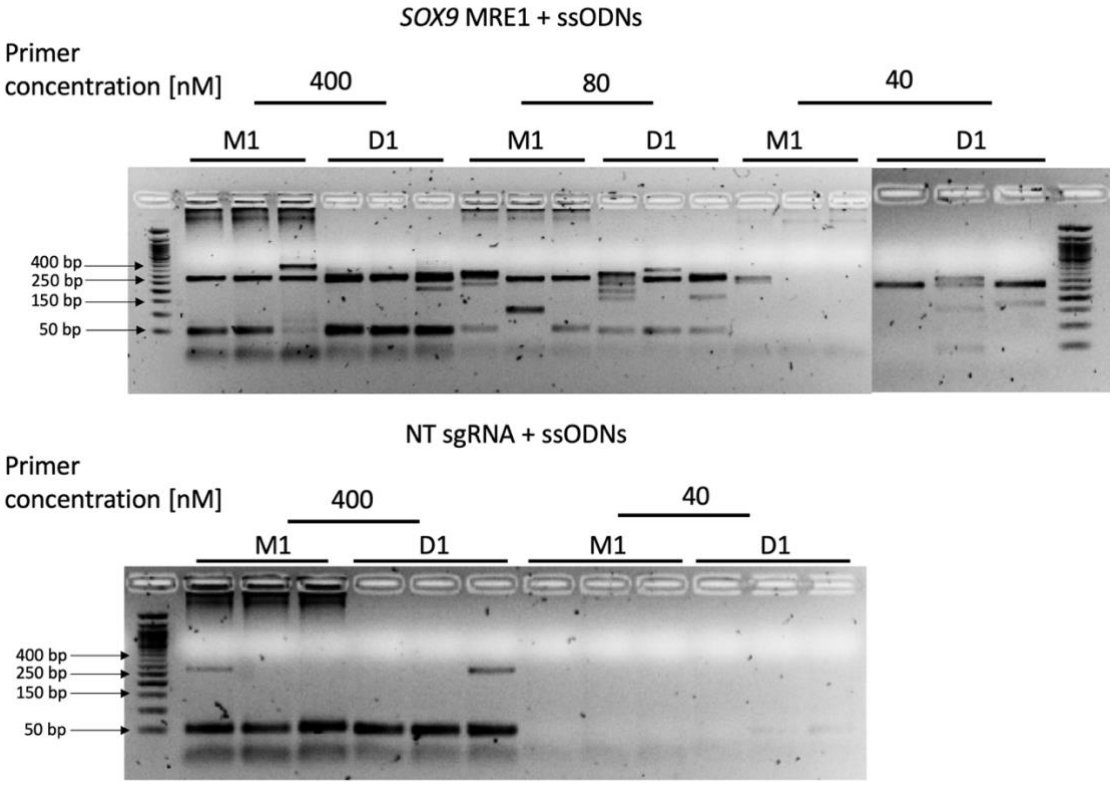


Fig. 5.19: Barcode insertion PCR reveals inconsistent products within a triplicate using troubleshot conditions.

SOX9 MRE1 was RNP co-transfected with barcodes M1 and D1 and their cDNA was assessed at 48 h for barcode integration. PCR amplified M1 (expected amplicon size: 275 bp) and D1 (264 bp) barcodes with the previously identified parameters (semi-endogenous primers, c = 2 µM, T_A = 72 °C, New England BioLabs Q5 Hot Start High-Fidelity 2x MM polymerase) in triplicates.

5.3 Discussion

The aim of this chapter comprised of uncovering the functionality of putative miRNA target sites in human chondrocytes, through CRISPR-Cas9 mediated editing for a key cartilage gene. This approach involves investigation of specific miRNA-mRNA target interaction in an endogenous context, thus allowing evaluation how these interactions contribute to chondrocyte function. Investigating endogenous miRNA target sites is a great challenge and has resulted in the generation of a wide spectrum of experimental approaches (Martinez-Sanchez and Murphy, 2013a). 3'UTR reporter systems such as luciferase assays or miRNA gain- or loss-of function studies are valuable experimental tools; however, these approaches are subject to limitations and are not always sufficient to reliably validate miRNA/mRNA binding.

In order to confirm biologically relevant targets and direct interaction between a given miRNA and its mRNA, it is important to not unduly perturb miRNA levels. It is also evident from the literature that a single miRNA can regulate several mRNA targets (Bartel, 2004). Therefore, instead of manipulating miRNA levels, which leads to probable unwanted (and usually unknown) effects on other targets, Bassett et al., developed a technique which allows putative miRNA binding sites to be directly mutated. This technique was established in human HEK293T cells, but its application in primary cells and specifically human chondrocytes had not yet been tested.

5.3.1 Reproducing analysis of MRE by CRISPR-mediated HDR in HEK293T cells

In order to become familiar with the technique, I first tried to reproduce results from the original publication (Bassett et al., 2014) by subcloning a sgRNA into the pX330 vector, targeting a putative miRNA response element (MRE) in the 3'-UTR of *C9orf7*. I then co-transfected the Cas9 and sgRNA expressing plasmid, along with the proposed T3 and T7 barcodes, into HEK293T cells. *C9orf7*, alternatively known as CACFD1 (Calcium Channel Flower Domain Containing 1), is involved in calcium channel activity and membrane depolarisation (Yao et al., 2009). Interestingly, C9orf7 contains no classical 'seed' sequence within its MRE. However, it contains a non-canonical binding site, which is defined by a sequence that is complementary to the 3'-end of miR-92a (Bassett et al., 2014).

I also targeted a putative MRE in the 3'-UTR of *SOX9* (in HEK293T cells), which is a cartilage master regulator gene. *SOX9* has been shown to be essential in mice for cartilage development (Bi et al., 1999), while mutations in *SOX9* lead to severe skeletal disorders (Foster et al., 1994). My previous lab reported that, using luciferase assays, *miR-145* can directly target *SOX9* in normal human articular chondrocytes through binding a specific site in its 3'-UTR (Martinez-Sanchez et al., 2012). Notably, *SOX9* contains two putative MREs in its 3'-UTR, of which one was shown to be functionally active and the other inactive (Martinez-Sanchez et al., 2012). I targeted and investigated the functionally active site by CRISPR-Cas9.

I was able to successfully replicate data from the original publication on the CRISPR-Cas9 barcode technique for assessing miRNA activity (Bassett et al., 2014), thus confirming that this procedure works in my hands.

5.3.2 Investigating MRE by CRISPR-mediated HDR in osteoarthritic human articular chondrocytes

I next began to apply this approach to target the putative *miR-145* binding site in the 3'-UTR of *SOX9* in OA human articular chondrocytes. Before starting, I intended to optimise the procedure to fully utilise this technique. As a result, I addressed a few weaknesses of the technique that I had identified during my workflow in HEK293T cells. These limitations included the need to re-design the barcodes to increase their T_A, and to further optimise sgRNA and Cas9 delivery especially into primary cells, specifically human chondrocytes. Genome editing using Cas9/sgRNA RNPs has been demonstrated to be effective in comparison to various other Cas9 and sgRNA delivery methods, and provides a better solution for hard-to-transfect cells or cells that die upon transfection (Kim et al., 2014, Zuris et al., 2015, Liang et al., 2015).

In the process of quantifying barcode expression levels by qPCR, I found that although gene editing by RNP transfection resulted in good efficiencies, the successful verification of barcode insertion was accompanied by numerous challenges. These problems can be summarised into three main complications: (i) bands were variably present in the NT

controls, (ii) all of the bands present (targeted + non targeted) were duplets or triplets, and (iii) all PCR reactions were accompanied by primer dimer formation of around 50 bp. After extensive and lengthy refinements, which included designing new primers, temperature gradient-PCRs and primer dilutions, each of these issues, if considered individually, appeared to have been solved. However, these refined PCR conditions could not be translated and implemented onto a qPCR without obtaining unreliable read-outs. To our surprise, when I ran PCR reactions side by side in triplicates, bands were characterised by variability and inconsistency within the triplicate. An important consideration that was neglected was the need for sequencing the amplified PCR products to confirm the authenticity of the bands.

One interpretation of these results could be that due to presumably very low ssODN integration efficiencies via HDR (T3 and T7 integration efficiencies in the Bassett et al. (2014), study were 0.22 % and 0.33 % respectively), primers would bind at different amplification cycles on the template. As a result, bands with different intensities and at different sizes were generated, as the read-out is based on edited cells only. Many studies have been undertaken to address the challenge of low HDR integration efficiency. The most prominent approach includes inhibiting key enzymes for the NHEJ pathway, hence making HDR compensate through repair. In a study using SCR7 (an inhibitor of DNA ligase IV), Maruyama et al. (2015) reported an 19-fold increase for a 21 bp ssODN (with two 100 bp homology arms) in melanoma (MelJuSo) cell lines, and a 13-fold increase for a 800 bp ssODN (with two 80 bp homology arms) into a murine bone marrow-derived dendritic cell line (DC2.4 cells).

Other studies deployed small interfering RNA (siRNA) to promote HDR integration efficiencies by inhibiting expression of Ku proteins, which are key players in the NHEJ pathway. This study improved HDR integration efficiencies up to 3-fold in primary porcine fibroblasts (Li et al., 2018).

Further strategies rely on the timepoint of template delivery. By synchronising cells at the S and G2 phases (by using numerous chemical inhibitors), and timing deliveries of Cas9/sgRNA complexes by RNP nucleofection, HDR integration efficiencies in HEK293T cells were increased 3-6-fold (Lin et al., 2014, Yang et al., 2016).

Additionally, proximity of the region of homology to the actual DSB, as well as rational template designing with sufficient lengths of homology arms, are driving factors that modulate HDR integration efficiencies in any cell type (Liu et al., 2018). Numerous engineering techniques of the Cas9 endonuclease led to 2-3-fold HDR integration efficiency increases in various cell lines (Liu et al., 2018). Of note is the so-called base-editing approach (Komor et al., 2016), which enables the precise mutation of single base pairs, without the need to introduce a DSB or donor DNA.

5.3.3 Advantages and limitations of CRISPR-mediated HDR barcode technology

One of two major positive aspects which emerged from this technique is that the completion of the work flow, from targeting a given site until the final qPCR read out,

only requires a few days. This can potentially can counteract the rapidly changing phenotypes of chondrocytes (Brodkin et al., 2004, Han et al., 2010). Similarly, the final qPCR read-out in our approach is based on edited cells only. As a result, low gene editing efficiencies might be sufficient for a positive read-out, yet presumably extremely low HDR integration efficiencies in our primary cells have led to no reliable read-out.

Likewise, this technique is also subject to another limitation. Apart from its questionable feasibility in primary cells, specifically human chondrocytes, miRNAs silencing genes through translational repression (as opposed to transcriptional silencing) possibly cannot be detected by this technique. This is due to subsequent increase in transcript levels and thereby the lack of distinction between barcode-mediated transcript increase and translational-inhibition mediated transcript increase. However, only a very small number of miRNAs are thought to regulate through translational inhibition, so this may not have had a substantial effect on my experimental outcomes.

Finally, an important consideration in this work (particularly for therapeutic application) is the efficiency of the desired gene editing. While, crucially, the barcode technique allows us to get a read-out exclusively from edited cells, it does not tell us about the efficiency of editing per se. This can, however, be investigated by using barcode-specific primers to amplify the edited genomic region, and by comparing expression levels against any residual, unedited genomic sequence. This will, at least, allow some crude estimation of the editing efficiency of our system in hACs.

In summary, CRISPR-Cas9 enables selective alterations of the genome, and becomes a potent tool in combination with the proposed strategy from Bassett et al. (2014) to

investigate functional miRNA target sites. However, to apply this technique to cartilage-specific genes in chondrocytes, and to ensure reliable interpretation of our read-outs, further optimisation is required. Future work will also have to include strategies which focus on the importance of selecting cells that are edited (or at least transfected) for downstream applications.

Final Discussion

6

Chapter 6. Final discussion

In recent years, CRISPR-Cas9 technology has emerged as the leading method for gene editing, surpassing transcription activator-like effector nucleases (TALEN) and zinc finger nucleases (ZFN). Its simplicity and efficiency have made CRISPR the favourite tool of scientists when it comes to genetic engineering. However, despite its advantages and potential, some unwanted obstructions in the form of off-target effects remain. It is of particular importance to eliminate such effects if CRISPR-Cas9 is to reach its full therapeutic potential in clinical applications. Next-generation sequencing methods, such as exome sequencing (also known as whole exome sequencing) (Choi et al., 2009), or discovery of *in situ* Cas off-targets and verification by sequencing (DISCOVER-Seq) (Wienert et al., 2019), have been developed and enable detection of off-target changes *in vivo*. Although no current methodology can fully guarantee the absence of undesired base editing in any single cell, there has been refinement of the CRISPR technology over the past few years. Such refinements, like modified endonucleases, base-editing or prime-editing, have and are still contributing towards its success by making the technology more and more accurate.

In the present study, I applied the CRISPR-Cas9 technology to investigate miRNA targeting in human chondrocytes with stably reduced expression of *miR-140,* without the need for clonal selection. This is important because human chondrocytes are not amenable to single cell expansion, and as a result they are reliant on efficient editing and knockout strategies.

I developed a protocol involving rapid and efficient CRISPR-Cas9-mediated manipulation of human chondrocytes by RNP double transfection. This reproduceable method should be applicable to any given miRNA, and their knockdown effects in human chondrocytes with stably reduced expression could be studied through agnostic sequencing methods or candidate gene qPCR. In our study we identified several *miR-140* dependent genes, in which *miR-140* knockdown caused upregulation in human chondrocytes. Several targets were validated by studying an RNA sequencing (RNA-seq) analysis of costal cartilage from 7-day old *miR-140* KO mice (in collaboration with Prof. David Young, Newcastle University).

CRISPR-Cas9 mediated gene editing through double transfection results in very high gene editing efficiencies, but single transfection led to inefficient knockdown of 40 – 70 %. Inefficient knockdown of *miR-140* did not lead to any changes in gene expression. This suggests that there is a threshold of deletion which needs to be crossed in order to see changes in gene expression, as opposed to gradual miRNA knockdown being proportional to the degree of gene regulation. This raises an interesting question about whether the knockdown we have seen *in vitro* has any *in vivo* relevance. The levels of *miR-140* are high in articular chondrocytes, albeit somewhat lower in OA chondrocytes (Miyaki et al., 2009). I observed a reduction in *miR-140* upon cartilage injury, leading to suppression of *miR-140* by approximately 70 %. Whether this fold of reduction would cause any *miR-140* dependant effects is unknown. Based on a published report IL-1β also downregulates *miR-140 in vitro,* but only by 40 % (Miyaki et al., 2009).

Therefore, when we envisage our editing techniques to have potential application, it may have greater impact in modifying cell-based therapies, for instance, in autologous

chondrocyte implantation (Harris et al., 2010). hACs are isolated and culture-expanded before re-implantation at the site of cartilage injury. In future therapies, hACs could be edited during the culture-expansion phase of this procedure. If miRNA target site editing results in significant changes in total mRNA (and protein) expression levels for a given target gene, then our approach could have clinical applicability, without the need for selection of edited cells. This will be important to investigate, as the single cell cloning required to select for edited cells would not be feasible for hACs.

6.1 Limitations

There are a number of limitations which should be recognised. it is unknown whether miRNAs from unedited alleles within a CRISPR-Cas9 targeted population exhibit any compensatory effects through increased expression. Similarly, it is also unclear how selective pressure may favour modified cells compared with unmodified cells in monolayer culture over time, and to what extent this affects miRNA expression levels.

Another important consideration is the extent to which a gene is regulated by a single or by multiple miRNAs. It is evident from the literature that a single miRNA tends to exert a somewhat subtle effect on its target (< 2-fold) (Martinez-Sanchez and Murphy, 2013a). An investigation of the simultaneous contribution of multiple miRNAs to the expression of a common target gene might present a solution. Currently, there is evidence in the literature describing synergistic effects between multiple miRNAs at one target in murine glioblastoma models or human cells (Chen et al., 2017, Bhaskaran et al., 2019). In the cartilage context, the analysis of key genes (*SOX9* or *ADAMTS-5,* for

example) could be carried out in this manner, and miRNA activity of predicted (with or without prior experimental evidence) target sites for multiple miRNAs on the gene of interest could be assessed. Herein, one approach could include choosing, for example, up to 5 different miRNAs predicted to target a gene of interest.

Another limitation of this study is the lack of differentiation between direct and indirect targets. To address this, I attempted to adapt a challenging barcode method that employed CRISPR-Cas9 to validate MRE targeting in an endogenous context. Unfortunately, I could not implement this technique in human chondrocytes, despite exhaustive troubleshooting and refinement strategies. Its full potential remains undefined and warrants further investigation.

It is also important to recognise that for these types of experiments one always tends to generalise implications from the cell to tissue, and to an *in vivo* situation. Observations *in vitro* may not necessarily correlate with what actually happens in a tissue or *in vivo*.

An important consideration is which cells were used. In this study I mainly relied on osteoarthritic human articular chondrocytes, which were isolated from patients undergoing knee replacement. Different levels of disease progression are a main caveat of this study, as during disease progression genes expression alters. Within our OA hAC donors a wide spread of the expression of our tested genes was evident. The condition of the diseased tissues, from which chondrocytes were isolated, was not recorded by using a disease scoring system. Healthy chondrocytes could not be used due to lack of regular supply. Two further factors contributing towards the spread of gene expression are the sex and age of the donors. While a different in sex is manifested during OA progression (Srikanth et al., 2005) and pain sensation (Fillingim et al., 2009), sex-

dependent expression in miRNAs has only been recently subject to research. miRNAs may be differently expressed between males and females (Guo et al., 2017).

Finally, cells were cultured and genetically engineered under normoxic conditions, while articular cartilage is aneaural, avascular, and alymphatic. As such, chondrocytes function in a hypoxic environment. Compared to normoxic conditions, matrix gene expression has been demonstrated to be elevated under hypoxia (Murphy and Sambaniss, 2001), while the role of *miR-140* under hypoxic conditions remains unclear.

Taken together, we could use our described RNP double transfection method to simultaneously target multiple miRNAs responsible for targeting a key gene of interest. In order to reduce variability, male doors should be differentiated from female donors. if no healthy donors are available, diseased tissue could be assessed for disease progression using standardised scoring tests. Lastly, chondrocytes could be cultivated and genetically engineered under hypoxic conditions.

References

7

Chapter 7. References

ABUDAYYEH, O. O., GOOTENBERG, J. S., KONERMANN, S., JOUNG, J., SLAYMAKER, I. M., COX, D. B., SHMAKOV, S., MAKAROVA, K. S., SEMENOVA, E. & MINAKHIN, L. 2016. C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science,* 353**,** aaf5573.

ADLI, M. 2018. The CRISPR tool kit for genome editing and beyond. *Nature communications,* 9**,** 1-13.

ANDERS, C., NIEWOEHNER, O., DUERST, A. & JINEK, M. 2014. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature,* 513**,** 569.

ANKER, P. S., SCHERJON, S. A., KLEIJBURG-VAN DER KEUR, C., NOORT, W. A., CLAAS, F. H., WILLEMZE, R., FIBBE, W. E. & KANHAI, H. H. 2003. Amniotic fluid as a novel source of mesenchymal stem cells for therapeutic transplantation. *Blood,* 102**,** 1548-1549.

ANZALONE, A. V., RANDOLPH, P. B., DAVIS, J. R., SOUSA, A. A., KOBLAN, L. W., LEVY, J. M., CHEN, P. J., WILSON, C., NEWBY, G. A. & RAGURAM, A. 2019. Search-and-replace genome editing without double-strand breaks or donor DNA. *Target,* 5.

ASANO, M., UMEZU, T., KATAGIRI, S., KOBAYASHI, C., TAUCHI, T., GOTOH, M., ANDO, K., OKABE, S., OHYASHIKI, J. H. & OHYASHIKI, K. 2017. Up-regulated exosomal miRNA-140-3p in CML patients with musculoskeletal pain associated with discontinuation of tyrosine kinase inhibitors. *International journal of hematology,* 105**,** 419-422.

AVRAHAM, R., SAS-CHEN, A., MANOR, O., STEINFELD, I., SHALGI, R., TARCIC, G., BOSSEL, N., ZEISEL, A., AMIT, I. & ZWANG, Y. 2010. EGF decreases the abundance of microRNAs that restrain oncogenic transcription factors. *Sci. Signal.,* 3**,** ra43-ra43.

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**,** 1709-1712.

BARTEL, D. P. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *cell,* 116**,** 281-297.

BARTER, M. J., TSELEPI, M., GÓMEZ, R., WOODS, S., HUI, W., SMITH, G. R., SHANLEY, D. P., CLARK, I. M. & YOUNG, D. A. 2015. Genome-Wide MicroRNA and Gene Analysis of Mesenchymal Stem Cell Chondrogenesis Identifies an Essential Role and Multiple Targets for miR-140-5p. *Stem Cells,* 33**,** 3266-3280.

BASSETT, A. R., AZZAM, G., WHEATLEY, L., TIBBIT, C., RAJAKUMAR, T., MCGOWAN, S., STANGER, N., EWELS, P. A., TAYLOR, S. & PONTING, C. P. 2014. Understanding functional miRNA–target interactions in vivo by site-specific genome engineering. *Nature communications,* 5**,** 4640.

BERNSTEIN, E., KIM, S. Y., CARMELL, M. A., MURCHISON, E. P., ALCORN, H., LI, M. Z., MILLS, A. A., ELLEDGE, S. J., ANDERSON, K. V. & HANNON, G. J. 2003. Dicer is essential for mouse development. *Nature genetics,* 35**,** 215.

BHASKARAN, V., NOWICKI, M. O., IDRISS, M., JIMENEZ, M. A., LUGLI, G., HAYES, J. L., MAHMOUD, A. B., ZANE, R. E., PASSARO, C. & LIGON, K. L. 2019. The functional synergism of microRNA clustering provides therapeutically relevant epigenetic interference in glioblastoma. *Nature communications,* 10**,** 442.

BHOSALE, A. M. & RICHARDSON, J. B. 2008. Articular cartilage: structure, injuries and review of management. *British medical bulletin,* 87**,** 77-95.

BI, W., DENG, J. M., ZHANG, Z., BEHRINGER, R. R. & DE CROMBRUGGHE, B. 1999. Sox9 is required for cartilage formation. *Nature genetics,* 22**,** 85.

BILLON, P., BRYANT, E. E., JOSEPH, S. A., NAMBIAR, T. S., HAYWARD, S. B., ROTHSTEIN, R. & CICCIA, A. 2017. CRISPR-mediated base editing enables efficient disruption of eukaryotic genes through induction of STOP codons. *Molecular cell,* 67**,** 1068-1079. e4.

BO, N., PENG, W., XINGHONG, P. & MA, R. 2012. Early cartilage degeneration in a rat experimental model of developmental dysplasia of the hip. *Connective tissue research,* 53**,** 513-520.

BOWIE, E. & GOETZ, S. C. 2019. TTBK2 and primary cilia are essential for the connectivity and survival of cerebellar Purkinje neurons. *BioRxiv***,** 689448.

BRAUN, J. E., HUNTZINGER, E., FAUSER, M. & IZAURRALDE, E. 2011. GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets. *Molecular cell,* 44**,** 120-133.

BRAUN, J. E., TRUFFAULT, V., BOLAND, A., HUNTZINGER, E., CHANG, C.-T., HAAS, G., WEICHENRIEDER, O., COLES, M. & IZAURRALDE, E. 2012. A direct interaction between DCP1 and XRN1 couples mRNA decapping to 5′ exonucleolytic degradation. *Nature structural & molecular biology,* 19**,** 1324.

BRIGHTON, C. T. & HEPPENSTALL, R. B. 1971. Oxygen Tension in Zones of the Epiphyseal Plate, the Metaphysis and Diaphysis: AN: in Vitro: AND: in Viro: STUDY IN RATS AND RABBITS. *JBJS,* 53**,** 719-728.

BRODKIN, K., GARCIA, A. & LEVENSTON, M. 2004. Chondrocyte phenotypes on different extracellular matrix monolayers. *Biomaterials,* 25**,** 5929-5938.

BROUGHTON, J. P., LOVCI, M. T., HUANG, J. L., YEO, G. W. & PASQUINELLI, A. E. 2016. Pairing beyond the seed supports microRNA targeting specificity. *Molecular cell,* 64**,** 320-333.

BUCKWALTER, J. & MANKIN, H. 1997. Articular cartilage: part I. *Journal of Bone and joint surgery,* 79**,** 600.

BUCKWALTER, J. A. & MARTIN, J. A. 2004. Sports and osteoarthritis. *Current opinion in rheumatology,* 16**,** 634-639.

BUECHLI, M. E., LAMARRE, J. & KOCH, T. G. 2012. MicroRNA-140 expression during chondrogenic differentiation of equine cord blood-derived mesenchymal stromal cells. *Stem cells and development,* 22**,** 1288-1296.

BULUN, S. E., SIMPSON, E. R. & WORD, R. A. 1994. Expression of the CYP19 gene and its product aromatase cytochrome P450 in human uterine leiomyoma tissues and cells in culture. *The Journal of Clinical Endocrinology & Metabolism,* 78**,** 736-743.

BUNNELL, B. A., FLAAT, M., GAGLIARDI, C., PATEL, B. & RIPOLL, C. 2008. Adipose-derived stem cells: isolation, expansion and differentiation. *Methods,* 45**,** 115-120.

CAI, X., HAGEDORN, C. H. & CULLEN, B. R. 2004. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *Rna,* 10**,** 1957-1966.

CAPLAN, A. I. 1991. Mesenchymal stem cells. *Journal of orthopaedic research,* 9**,** 641-650.

CHANDRADOSS, S. D., SCHIRLE, N. T., SZCZEPANIAK, M., MACRAE, I. J. & JOO, C. 2015. A dynamic search process underlies microRNA targeting. *Cell,* 162**,** 96-107.

CHANG, I.-R. & MARTIN, A. 2018. Anatomy, Cartilage. *StatPearls [Internet].* StatPearls Publishing.

CHANG, T.-C. & MENDELL, J. T. 2007. microRNAs in vertebrate physiology and human disease. *Annu. Rev. Genomics Hum. Genet.,* 8**,** 215-239.

CHEN, B., GILBERT, L. A., CIMINI, B. A., SCHNITZBAUER, J., ZHANG, W., LI, G.-W., PARK, J., BLACKBURN, E. H., WEISSMAN, J. S. & QI, L. S. 2013. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell,* 155**,** 1479-1491.

CHEN, C.-Y. A., ZHENG, D., XIA, Z. & SHYU, A.-B. 2009. Ago–TNRC6 triggers microRNA-mediated decay by promoting two deadenylation steps. *Nature structural & molecular biology,* 16**,** 1160.

CHEN, D., FARWELL, M. A. & ZHANG, B. 2010. MicroRNA as a new player in the cell cycle. *Journal of cellular physiology,* 225**,** 296-301.

CHEN, X., ZHAO, W., YUAN, Y., BAI, Y., SUN, Y., ZHU, W. & DU, Z. 2017. MicroRNAs tend to synergistically control expression of genes encoding extensively-expressed proteins in humans. *PeerJ,* 5**,** e3682.

CHERNOV, A. V. & STRONGIN, A. Y. 2011. Epigenetic regulation of matrix metalloproteinases and their collagen substrates in cancer. *Biomolecular concepts,* 2**,** 135-147.

CHIA, S. L., SAWAJI, Y., BURLEIGH, A., MCLEAN, C., INGLIS, J., SAKLATVALA, J. & VINCENT, T. 2009. Fibroblast growth factor 2 is an intrinsic chondroprotective agent that suppresses ADAMTS-5 and delays cartilage degradation in murine osteoarthritis. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology,* 60**,** 2019-2027.

CHOI, M., SCHOLL, U. I., JI, W., LIU, T., TIKHONOVA, I. R., ZUMBO, P., NAYIR, A., BAKKALOĞLU, A., ÖZEN, S. & SANJAD, S. 2009. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proceedings of the National Academy of Sciences,* 106**,** 19096-19101.

CHONG, K. W., CHANALARIS, A., BURLEIGH, A., JIN, H., WATT, F. E., SAKLATVALA, J. & VINCENT, T. L. 2013. Fibroblast growth factor 2 drives changes in gene expression following injury to murine cartilage in vitro and in vivo. *Arthritis & Rheumatism,* 65**,** 2346-2355.

CONG, L., RAN, F. A., COX, D., LIN, S., BARRETTO, R., HABIB, N., HSU, P. D., WU, X., JIANG, W. & MARRAFFINI, L. A. 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science,* 339**,** 819-823.

COVENEY, C. R., COLLINS, I., MC FIE, M., CHANALARIS, A., YAMAMOTO, K. & WANN, A. K. 2018. Cilia protein IFT88 regulates extracellular protease activity by optimizing LRP-1–mediated endocytosis. *The FASEB Journal,* 32**,** 6771-6782.

DALEY, J. M., GAINES, W. A., KWON, Y. & SUNG, P. 2014. Regulation of DNA pairing in homologous recombination. *Cold Spring Harbor perspectives in biology,* 6**,** a017954.

DAVIDSON, D., BLANC, A., FILION, D., WANG, H., PLUT, P., PFEFFER, G., BUSCHMANN, M. D. & HENDERSON, J. E. 2005. Fibroblast growth factor (FGF) 18 signals through FGF receptor 3 to promote chondrogenesis. *Journal of Biological Chemistry,* 280**,** 20509-20515.

DE MAGALHÃES, J. P. & PASSOS, J. F. 2018. Stress, cell senescence and organismal ageing. *Mechanisms of ageing and development,* 170**,** 2-9.

DE RIE, D., ABUGESSAISA, I., ALAM, T., ARNER, E., ARNER, P., ASHOOR, H., ÅSTRÖM, G., BABINA, M., BERTIN, N. & BURROUGHS, A. M. 2017. An integrated expression atlas of miRNAs and their promoters in human and mouse. *Nature biotechnology,* 35**,** 872.

DELISE, A., FISCHER, L. & TUAN, R. 2000. Cellular interactions and signaling in cartilage development. *Osteoarthritis and cartilage,* 8**,** 309-334.

DENGLER, V. L., GALBRAITH, M. D. & ESPINOSA, J. M. 2014. Transcriptional regulation by hypoxia inducible factors. *Critical reviews in biochemistry and molecular biology,* 49**,** 1-15.

DENLI, A. M., TOPS, B. B., PLASTERK, R. H., KETTING, R. F. & HANNON, G. J. 2004. Processing of primary microRNAs by the Microprocessor complex. *Nature,* 432**,** 231.

DOMINICI, M., LE BLANC, K., MUELLER, I., SLAPER-CORTENBACH, I., MARINI, F., KRAUSE, D., DEANS, R., KEATING, A., PROCKOP, D. & HORWITZ, E. 2006. Minimal criteria for defining multipotent mesenchymal stromal cells. The International Society for Cellular Therapy position statement. *Cytotherapy,* 8**,** 315-317.

DOMM, C., SCHÜNKE, M., CHRISTESEN, K. & KURZ, B. 2002. Redifferentiation of dedifferentiated bovine articular chondrocytes in alginate culture under low oxygen tension. *Osteoarthritis and Cartilage,* 10**,** 13-22.

DOUDNA, J. A. & CHARPENTIER, E. 2014. The new frontier of genome engineering with CRISPR-Cas9. *Science,* 346**,** 1258096.

DOW, L. E., FISHER, J., O'ROURKE, K. P., MULEY, A., KASTENHUBER, E. R., LIVSHITS, G., TSCHAHARGANEH, D. F., SOCCI, N. D. & LOWE, S. W. 2015. Inducible in vivo genome editing with CRISPR-Cas9. *Nature biotechnology,* 33**,** 390.

DUDEK, K. A., LAFONT, J. E., MARTINEZ-SANCHEZ, A. & MURPHY, C. L. 2010. Type II collagen expression is regulated by tissue-specific miR-675 in human articular chondrocytes. *Journal of Biological Chemistry,* 285**,** 24381-24387.

DUEVA, R. & ILIAKIS, G. 2013. Alternative pathways of non-homologous end joining (NHEJ) in genomic instability and cancer. *Translational Cancer Research,* 2**,** 163-177.

ELDRIDGE, S., NALESSO, G., ISMAIL, H., VICENTE-GRECO, K., KABOURIDIS, P., RAMACHANDRAN, M., NIEMEIER, A., HERZ, J., PITZALIS, C. & PERRETTI, M. 2016. Agrin mediates chondrocyte homeostasis and requires both LRP4 and α-dystroglycan to enhance cartilage formation in vitro and in vivo. *Annals of the rheumatic diseases,* 75**,** 1228-1235.

ENDISHA, H., ROCKEL, J., JURISICA, I. & KAPOOR, M. 2018. The complex landscape of microRNAs in articular cartilage: biology, pathology, and therapeutic targets. *JCI insight,* 3.

EYRE, D. The collagens of articular cartilage. Seminars in arthritis and rheumatism, 1991. Elsevier, 2-11.

FABIAN, M. R., CIEPLAK, M. K., FRANK, F., MORITA, M., GREEN, J., SRIKUMAR, T., NAGAR, B., YAMAMOTO, T., RAUGHT, B. & DUCHAINE, T. F. 2011. miRNA-mediated deadenylation is orchestrated by GW182 through two conserved motifs that interact with CCR4–NOT. *Nature structural & molecular biology,* 18**,** 1211.

FILLINGIM, R. B., KING, C. D., RIBEIRO-DASILVA, M. C., RAHIM-WILLIAMS, B. & RILEY III, J. L. 2009. Sex, gender, and pain: a review of recent clinical and experimental findings. *The journal of pain,* 10**,** 447-485.

FOSANG, A. J., ROGERSON, F. M., EAST, C. J. & STANTON, H. 2008. ADAMTS-5: the story so far. *Eur Cell Mater,* 15**,** 11-26.

FOSTER, J. W., DOMINGUEZ-STEGLICH, M. A., GUIOLI, S., KWOK, C., WELLER, P. A., STEVANOVIĆ, M., WEISSENBACH, J., MANSOUR, S., YOUNG, I. D. & GOODFELLOW, P. N. 1994. Campomelic dysplasia and autosomal sex reversal caused by mutations in an SRY-related gene. *Nature,* 372**,** 525.

FRANK, F., SONENBERG, N. & NAGAR, B. 2010. Structural basis for 5′-nucleotide base-specific recognition of guide RNA by human AGO2. *Nature,* 465**,** 818.

FU, Y., SANDER, J. D., REYON, D., CASCIO, V. M. & JOUNG, J. K. 2014. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nature biotechnology,* 32**,** 279.

GAUDELLI, N. M., KOMOR, A. C., REES, H. A., PACKER, M. S., BADRAN, A. H., BRYSON, D. I. & LIU, D. R. 2017. Programmable base editing of A• T to G• C in genomic DNA without DNA cleavage. *Nature,* 551**,** 464.

GEBERT, L. F. & MACRAE, I. J. 2019. Regulation of microRNA function in animals. *Nature reviews Molecular cell biology,* 20**,** 21-37.

GHOSSOUB, R., HU, Q., FAILLER, M., ROUYEZ, M.-C., SPITZBARTH, B., MOSTOWY, S., WOLFRUM, U., SAUNIER, S., COSSART, P. & NELSON, W. J. 2013. Septins 2, 7 and 9 and MAP4 colocalize along the axoneme in the primary cilium and control ciliary length. *J Cell Sci,* 126**,** 2583-2594.

GLASSON, S. S., ASKEW, R., SHEPPARD, B., CARITO, B., BLANCHET, T., MA, H.-L., FLANNERY, C. R., PELUSO, D., KANKI, K. & YANG, Z. 2005. Deletion of active ADAMTS5 prevents cartilage degradation in a murine model of osteoarthritis. *Nature,* 434**,** 644.

GLASSON, S. S., ASKEW, R., SHEPPARD, B., CARITO, B. A., BLANCHET, T., MA, H. L., FLANNERY, C. R., KANKI, K., WANG, E. & PELUSO, D. 2004. Characterization of and osteoarthritis susceptibility in ADAMTS-4–knockout mice. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology,* 50**,** 2547-2558.

GREGORY, R. I., YAN, K.-P., AMUTHAN, G., CHENDRIMADA, T., DORATOTAJ, B., COOCH, N. & SHIEKHATTAR, R. 2004. The Microprocessor complex mediates the genesis of microRNAs. *Nature,* 432**,** 235.

GROFFEN, A. J., RUEGG, M. A., DIJKMAN, H., VAN DE VELDEN, T. J., BUSKENS, C. A., VAN DEN BORN, J., ASSMANN, K. J., MONNENS, L. A., VEERKAMP, J. H. & VAN DEN HEUVEL, L. P. 1998. Agrin is a major heparan sulfate proteoglycan in the human glomerular basement membrane. *Journal of Histochemistry & Cytochemistry,* 46**,** 19-27.

GRONTHOS, S., BRAHIM, J., LI, W., FISHER, L., CHERMAN, N., BOYDE, A., DENBESTEN, P., ROBEY, P. G. & SHI, S. 2002. Stem cell properties of human dental pulp stem cells. *Journal of dental research,* 81**,** 531-535.

GUO, L., ZHANG, Q., MA, X., WANG, J. & LIANG, T. 2017. miRNA and mRNA expression analysis reveals potential sex-biased miRNA expression. *Scientific reports,* 7**,** 39812.

GUO, Y., LIU, J., ELFENBEIN, S. J., MA, Y., ZHONG, M., QIU, C., DING, Y. & LU, J. 2015. Characterization of the mammalian miRNA turnover landscape. *Nucleic acids research,* 43**,** 2326-2341.

HAN, H. S., LEE, S., KIM, J. H., SEONG, S. C. & LEE, M. C. 2010. Changes in chondrogenic phenotype and gene expression profiles associated with the in vitro expansion of human synovium-derived cells. *Journal of Orthopaedic Research,* 28**,** 1283-1291.

HARDINGHAM, T. & FOSANG, A. 1995. The structure of aggrecan and its turnover in cartilage. *The Journal of Rheumatology. Supplement,* 43**,** 86-90.

HARRIS, J. D., SISTON, R. A., PAN, X. & FLANIGAN, D. C. 2010. Autologous chondrocyte implantation: a systematic review. *JBJS,* 92**,** 2220-2233.

HAVENS, M. A., REICH, A. A., DUELLI, D. M. & HASTINGS, M. L. 2012. Biogenesis of mammalian microRNAs by a non-canonical processing pathway. *Nucleic acids research,* 40**,** 4626-4640.

HEIDARI, B. 2011. Knee osteoarthritis prevalence, risk factors, pathogenesis and features: Part I. *Caspian journal of internal medicine,* 2**,** 205.

HELSTEN, T., SCHWAEDERLE, M. & KURZROCK, R. 2015. Fibroblast growth factor receptor signaling in hereditary and neoplastic disease: biologic and clinical implications. *Cancer and Metastasis Reviews,* 34**,** 479-496.

HEMPHILL, J., BORCHARDT, E. K., BROWN, K., ASOKAN, A. & DEITERS, A. 2015. Optical control of CRISPR/Cas9 gene editing. *Journal of the American Chemical Society,* 137**,** 5642-5645.

HILDEBRAND, A., ROMARIS, M., RASMUSSEN, L., HEINEGÅRD, D., TWARDZIK, D., BORDER, W. & RUOSLAHTI, E. 1994. Interaction of the small interstitial proteoglycans biglycan, decorin and fibromodulin with transforming growth factor β. *Biochemical Journal,* 302**,** 527-534.

HURSKAINEN, T. L., HIROHATA, S., SELDIN, M. F. & APTE, S. S. 1999. ADAM-TS5, ADAM-TS6, and ADAM-TS7, Novel Members of a New Family of Zinc Metalloproteases GENERAL FEATURES AND GENOMIC DISTRIBUTION OF THE ADAM-TS FAMILY. *Journal of Biological Chemistry,* 274**,** 25555-25563.

HWANG, H.-W., WENTZEL, E. A. & MENDELL, J. T. 2007. A hexanucleotide element directs microRNA nuclear import. *Science,* 315**,** 97-100.

HWANG, S., PARK, S.-K., LEE, H. Y., KIM, S. W., LEE, J. S., CHOI, E. K., YOU, D., KIM, C.-S. & SUH, N. 2014. miR-140-5p suppresses BMP2-mediated osteogenesis in undifferentiated human mesenchymal stem cells. *FEBS letters,* 588**,** 2957-2963.

ILIOPOULOS, D., MALIZOS, K. N., OIKONOMOU, P. & TSEZOU, A. 2008. Integrative microRNA and proteomic approaches identify novel osteoarthritis genes and their collaborative metabolic and inflammatory networks. *PloS one,* 3**,** e3740.

IM, H.-J., MUDDASANI, P., NATARAJAN, V., SCHMID, T. M., BLOCK, J. A., DAVIS, F., VAN WIJNEN, A. J. & LOESER, R. F. 2007. Basic fibroblast growth factor stimulates matrix metalloproteinase-13 via the molecular cross-talk between the mitogen-activated protein kinases and protein kinase Cδ pathways in human adult articular chondrocytes. *Journal of Biological Chemistry,* 282**,** 11110-11121.

ISMAIL, H. M., MIOTLA-ZAREBSKA, J., TROEBERG, L., TANG, X., STOTT, B., YAMAMOTO, K., NAGASE, H., FOSANG, A. J., VINCENT, T. L. & SAKLATVALA, J. 2016. Brief report: JNK-2 controls aggrecan degradation in murine articular cartilage and the development of experimental osteoarthritis. *Arthritis & Rheumatology,* 68**,** 1165-1171.

ISMAIL, H. M., YAMAMOTO, K., VINCENT, T. L., NAGASE, H., TROEBERG, L. & SAKLATVALA, J. 2015. Interleukin-1 acts via the JNK-2 signaling pathway to induce aggrecan degradation by human chondrocytes. *Arthritis & Rheumatology,* 67**,** 1826-1836.

ITO, T., SAWADA, R., FUJIWARA, Y. & TSUCHIYA, T. 2008. FGF-2 increases osteogenic and chondrogenic differentiation potentials of human mesenchymal stem cells by inactivation of TGF-β signaling. *Cytotechnology,* 56**,** 1-7.

JANSEN, R., EMBDEN, J. D. V., GAASTRA, W. & SCHOULS, L. M. 2002. Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular microbiology,* 43**,** 1565-1575.

JIANG, F. & DOUDNA, J. A. 2017. CRISPR–Cas9 structures and mechanisms. *Annual review of biophysics,* 46**,** 505-529.

JIANG, F., TAYLOR, D. W., CHEN, J. S., KORNFELD, J. E., ZHOU, K., THOMPSON, A. J., NOGALES, E. & DOUDNA, J. A. 2016. Structures of a CRISPR-Cas9 R-loop complex primed for DNA cleavage. *Science,* 351**,** 867-871.

JINEK, M., CHYLINSKI, K., FONFARA, I., HAUER, M., DOUDNA, J. A. & CHARPENTIER, E. 2012. A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. *science,* 337**,** 816-821.

JINEK, M., JIANG, F., TAYLOR, D. W., STERNBERG, S. H., KAYA, E., MA, E., ANDERS, C., HAUER, M., ZHOU, K. & LIN, S. 2014. Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science,* 343**,** 1247997.

JONES, S., WATKINS, G., LE GOOD, N., ROBERTS, S., MURPHY, C., BROCKBANK, S., NEEDHAM, M., READ, S. & NEWHAM, P. 2009. The identification of differentially expressed microRNA in osteoarthritic tissue that modulate the production of TNF-α and MMP13. *Osteoarthritis and cartilage,* 17**,** 464-472.

JOO, E., SURKA, M. C. & TRIMBLE, W. S. 2007. Mammalian SEPT2 is required for scaffolding nonmuscle myosin II and its kinases. *Developmental cell,* 13**,** 677-690.

KARLSEN, T. A., JAKOBSEN, R. B., MIKKELSEN, T. S. & BRINCHMANN, J. E. 2013. microRNA-140 targets RALA and regulates chondrogenic differentiation of

human mesenchymal stem cells by translational enhancement of SOX9 and ACAN. *Stem cells and development,* 23**,** 290-304.

KATSUMI, A., ORR, A. W., TZIMA, E. & SCHWARTZ, M. A. 2004. Integrins in mechanotransduction. *Journal of Biological Chemistry,* 279**,** 12001-12004.

KEVORKIAN, L., YOUNG, D. A., DARRAH, C., DONELL, S. T., SHEPSTONE, L., PORTER, S., BROCKBANK, S. M., EDWARDS, D. R., PARKER, A. E. & CLARK, I. M. 2004. Expression profiling of metalloproteinases and their inhibitors in cartilage. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology,* 50**,** 131-141.

KIANI, C., LIWEN, C., WU, Y. J., ALBERT, J. Y. & BURTON, B. Y. 2002. Structure and function of aggrecan. *Cell research,* 12**,** 19.

KIM, B., JEONG, K. & KIM, V. N. 2017. Genome-wide mapping of DROSHA cleavage sites on primary microRNAs and noncanonical substrates. *Molecular cell,* 66**,** 258-269. e5.

KIM, D., SONG, J. & JIN, E.-J. 2010. MicroRNA-221 regulates chondrogenic differentiation through promoting proteosomal degradation of slug by targeting Mdm2. *Journal of Biological Chemistry,* 285**,** 26900-26907.

KIM, K. P. & MIRKIN, E. V. 2018. So similar yet so different: The two ends of a double strand break. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis,* 809**,** 70-80.

KIM, S., KIM, D., CHO, S. W., KIM, J. & KIM, J.-S. 2014. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome research,* 24**,** 1012-1019.

KIM, Y.-K., KIM, B. & KIM, V. N. 2016. Re-evaluation of the roles of DROSHA, Exportin 5, and DICER in microRNA biogenesis. *Proceedings of the National Academy of Sciences,* 113**,** E1881-E1889.

KLEINSTIVER, B. P., PATTANAYAK, V., PREW, M. S., TSAI, S. Q., NGUYEN, N. T., ZHENG, Z. & JOUNG, J. K. 2016. High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. *Nature,* 529**,** 490.

KNÄUPER, V., LÓPEZ-OTIN, C., SMITH, B., KNIGHT, G. & MURPHY, G. 1996. Biochemical characterization of human collagenase-3. *Journal of Biological Chemistry,* 271**,** 1544-1550.

KNIGHT, M. 2013. The role of primary cilia in cartilage health and disease. *Osteoarthritis and Cartilage,* 21**,** S2.

KOBAYASHI, T., LU, J., COBB, B. S., RODDA, S. J., MCMAHON, A. P., SCHIPANI, E., MERKENSCHLAGER, M. & KRONENBERG, H. M. 2008. Dicer-dependent pathways regulate chondrocyte proliferation and differentiation. *Proceedings of the National Academy of Sciences,* 105**,** 1949-1954.

KOMOR, A. C., KIM, Y. B., PACKER, M. S., ZURIS, J. A. & LIU, D. R. 2016. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature,* 533**,** 420.

KRAMER, J., HEGERT, C., GUAN, K., WOBUS, A. M., MÜLLER, P. K. & ROHWEDEL, J. 2000. Embryonic stem cell-derived chondrogenic differentiation in vitro: activation by BMP-2 and BMP-4. *Mechanisms of development,* 92**,** 193-205.

KROL, J., BUSSKAMP, V., MARKIEWICZ, I., STADLER, M. B., RIBI, S., RICHTER, J., DUEBEL, J., BICKER, S., FEHLING, H. J. & SCHÜBELER, D. 2010. Characterizing light-regulated retinal microRNAs reveals rapid turnover as a common property of neuronal microRNAs. *Cell,* 141**,** 618-631.

KUSCU, C., PARLAK, M., TUFAN, T., YANG, J., SZLACHTA, K., WEI, X., MAMMADOV, R. & ADLI, M. 2017. CRISPR-STOP: gene silencing through base-editing-induced nonsense mutations. *Nature methods,* 14**,** 710.

LAFONT, J. E. 2010. Lack of oxygen in articular cartilage: consequences for chondrocyte biology. *International journal of experimental pathology,* 91**,** 99-106.

LAMARCHE, B. J., ORAZIO, N. I. & WEITZMAN, M. D. 2010. The MRN complex in double-strand break repair and telomere maintenance. *FEBS letters,* 584**,** 3682-3695.

LAMPEN, A., MEYER, S. & NAU, H. 2001. Effects of receptor-selective retinoids on CYP26 gene expression and metabolism of all-trans-retinoic acid in intestinal cells. *Drug metabolism and disposition,* 29**,** 742-747.

LAN, H., CHEN, W., HE, G. & YANG, S. 2015. miR-140-5p inhibits ovarian cancer growth partially by repression of PDGFRA. *Biomedicine & Pharmacotherapy,* 75**,** 117-122.

LAU, P.-W., GUILEY, K. Z., DE, N., POTTER, C. S., CARRAGHER, B. & MACRAE, I. J. 2012. The molecular architecture of human Dicer. *Nature structural & molecular biology,* 19**,** 436.

LEE, R. C., FEINBAUM, R. L. & AMBROS, V. 1993. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *cell,* 75**,** 843-854.

LEE, T.-J., JANG, J., KANG, S., JIN, M., SHIN, H., KIM, D.-W. & KIM, B.-S. 2013. Enhancement of osteogenic and chondrogenic differentiation of human embryonic stem cells by mesodermal lineage induction with BMP-4 and FGF2 treatment. *Biochemical and biophysical research communications,* 430**,** 793-797.

LEE, W., LEDDY, H. A., CHEN, Y., LEE, S. H., ZELENSKI, N. A., MCNULTY, A. L., WU, J., BEICKER, K. N., COLES, J. & ZAUSCHER, S. 2014. Synergy between Piezo1 and Piezo2 channels confers high-strain mechanosensitivity to articular cartilage. *Proceedings of the National Academy of Sciences,* 111**,** E5114-E5122.

LEE, Y., AHN, C., HAN, J., CHOI, H., KIM, J., YIM, J., LEE, J., PROVOST, P., RÅDMARK, O. & KIM, S. 2003. The nuclear RNase III Drosha initiates microRNA processing. *Nature,* 425**,** 415.

LEE, Y., KIM, M., HAN, J., YEOM, K. H., LEE, S., BAEK, S. H. & KIM, V. N. 2004. MicroRNA genes are transcribed by RNA polymerase II. *The EMBO journal,* 23**,** 4051-4060.

LI, G., LIU, D., ZHANG, X., QUAN, R., ZHONG, C., MO, J., HUANG, Y., WANG, H., RUAN, X. & XU, Z. 2018. Suppressing Ku70/Ku80 expression elevates homology-directed repair efficiency in primary fibroblasts. *The international journal of biochemistry & cell biology,* 99**,** 154-160.

LI, H., WANG, D., YUAN, Y. & MIN, J. 2017. New insights on the MMP-13 regulatory network in the pathogenesis of early osteoarthritis. *Arthritis research & therapy,* 19**,** 248.

LI, X., ELLMAN, M. B., KROIN, J. S., CHEN, D., YAN, D., MIKECZ, K., RANJAN, K., XIAO, G., STEIN, G. S. & KIM, S. G. 2012. Species-specific biological effects of FGF-2 in

articular cartilage: Implication for distinct roles within the FGF receptor family. *Journal of cellular biochemistry,* 113**,** 2532-2542.

LIANG, X., POTTER, J., KUMAR, S., ZOU, Y., QUINTANILLA, R., SRIDHARAN, M., CARTE, J., CHEN, W., ROARK, N. & RANGANATHAN, S. 2015. Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. *Journal of biotechnology,* 208**,** 44-53.

LIEW, C.-C., MARSHALL, K. W. & ZHANG, H. 2008. Diagnosis of mild osteoarthritis by determination of TNFAIP6 and TGFBI RNA levels. Google Patents.

LIN, C.-S., NING, H., LIN, G. & LUE, T. F. 2012. Is CD34 truly a negative marker for mesenchymal stromal cells? *Cytotherapy,* 14**,** 1159-1163.

LIN, S., STAAHL, B. T., ALLA, R. K. & DOUDNA, J. A. 2014. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *elife,* 3**,** e04766.

LITTLE, C., BARAI, A., BURKHARDT, D., SMITH, S., FOSANG, A., WERB, Z., SHAH, M. & THOMPSON, E. 2009. Matrix metalloproteinase 13–deficient mice are resistant to osteoarthritic cartilage erosion but not chondrocyte hypertrophy or osteophyte development. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology,* 60**,** 3723-3733.

LIU, M., REHMAN, S., TANG, X., GU, K., FAN, Q., CHEN, D. & MA, W. 2018. Methodologies for improving HDR efficiency. *Frontiers in genetics,* 9.

LIU, Y., BEYER, A. & AEBERSOLD, R. 2016. On the dependency of cellular protein levels on mRNA abundance. *Cell,* 165**,** 535-550.

LOESER, R. F., CHUBINSKAYA, S., PACIONE, C. & IM, H. J. 2005. Basic fibroblast growth factor inhibits the anabolic activity of insulin-like growth factor 1 and osteogenic protein 1 in adult human articular chondrocytes. *Arthritis & Rheumatism,* 52**,** 3910-3917.

LOESER, R. F., GOLDRING, S. R., SCANZELLO, C. R. & GOLDRING, M. B. 2012. Osteoarthritis: a disease of the joint as an organ. *Arthritis & Rheumatism,* 64**,** 1697-1707.

LOHMANDER, L. S., DAHLBERG, L., RYD, L. & HEINEGÅRD, D. 1989. Increased levels of proteoglycan fragments in knee joint fluid after injury. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology,* 32**,** 1434-1442.

LOHMANDER, L. S., ENGLUND, P. M., DAHL, L. L. & ROOS, E. M. 2007. The long-term consequence of anterior cruciate ligament and meniscus injuries: osteoarthritis. *The American journal of sports medicine,* 35**,** 1756-1769.

LONG, J. M. & LAHIRI, D. K. 2012. Advances in microRNA experimental approaches to study physiological regulation of gene products implicated in CNS disorders. *Experimental neurology,* 235**,** 402-418.

LU, Y., QIN, T., LI, J., WANG, L., ZHANG, Q., JIANG, Z. & MAO, J. 2017. MicroRNA-140-5p inhibits invasion and angiogenesis through targeting VEGF-A in breast cancer. *Cancer gene therapy,* 24**,** 386-392.

LUJAMBIO, A. & LOWE, S. W. 2012. The microcosmos of cancer. *Nature,* 482**,** 347.

LUND-OLESEN, K. 1970. Oxygen tension in synovial fluids. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology,* 13**,** 769-776.

MAEDER, M. L., LINDER, S. J., CASCIO, V. M., FU, Y., HO, Q. H. & JOUNG, J. K. 2013. CRISPR RNA–guided activation of endogenous human genes. *Nature methods,* 10**,** 977.

MAJUMDAR, M. K., WANG, E. & MORRIS, E. A. 2001. BMP-2 and BMP-9 promotes chondrogenic differentiation of human multipotential mesenchymal cells and overcomes the inhibitory effect of IL-1. *Journal of cellular physiology,* 189**,** 275-284.

MAKAROVA, J. A., SHKURNIKOV, M. U., WICKLEIN, D., LANGE, T., SAMATOV, T. R., TURCHINOVICH, A. A. & TONEVITSKY, A. G. 2016. Intracellular and extracellular microRNA: an update on localization and biological role. *Progress in histochemistry and cytochemistry,* 51**,** 33-49.

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. *Biology direct,* 1**,** 7.

MAKAROVA, K. S., WOLF, Y. I., ALKHNBASHI, O. S., COSTA, F., SHAH, S. A., SAUNDERS, S. J., BARRANGOU, R., BROUNS, S. J., CHARPENTIER, E. & HAFT, D. H. 2015. An updated evolutionary classification of CRISPR–Cas systems. *Nature Reviews Microbiology,* 13**,** 722.

MAO, G., WU, P., ZHANG, Z., ZHANG, Z., LIAO, W., LI, Y. & KANG, Y. 2017. MicroRNA-92a-3p regulates aggrecanase-1 and aggrecanase-2 expression in chondrogenesis and IL-1β-induced catabolism in human articular chondrocytes. *Cellular Physiology and Biochemistry,* 44**,** 38-52.

MARESCHI, K., RUSTICHELLI, D., CALABRESE, R., GUNETTI, M., SANAVIO, F., CASTIGLIA, S., RISSO, A., FERRERO, I., TARELLA, C. & FAGIOLI, F. 2012. Multipotent mesenchymal stromal stem cell expansion by plating whole bone marrow at a low cellular density: a more advantageous method for clinical use. *Stem cells international,* 2012.

MARTINEZ-SANCHEZ, A., DUDEK, K. A. & MURPHY, C. L. 2012. Regulation of human chondrocyte function through direct inhibition of cartilage master regulator SOX9 by microRNA-145 (miRNA-145). *Journal of Biological Chemistry,* 287**,** 916-924.

MARTINEZ-SANCHEZ, A. & MURPHY, C. L. 2013a. MicroRNA target identification—experimental approaches. *Biology,* 2**,** 189-205.

MARTINEZ-SANCHEZ, A. & MURPHY, C. L. 2013b. miR-1247 functions by targeting cartilage transcription factor SOX9. *Journal of Biological Chemistry,* 288**,** 30802-30814.

MARUYAMA, T., DOUGAN, S. K., TRUTTMANN, M. C., BILATE, A. M., INGRAM, J. R. & PLOEGH, H. L. 2015. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nature biotechnology,* 33**,** 538.

MATRANGA, C., TOMARI, Y., SHIN, C., BARTEL, D. P. & ZAMORE, P. D. 2005. Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. *Cell,* 123**,** 607-620.

MAYNE, R. 1989. Cartilage collagens. What is their function, and are they involved in articular disease? *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology,* 32**,** 241-246.

MCELREAVAY, K. D., IRVINE, A. I., ENNIS, K. T. & MCLEAN, W. I. 1991. Isolation, culture and characterisation of fibroblast-like cells derived from the Wharton's jelly portion of human umbilical cord. Portland Press Limited.

MCGLASHAN, S., CLUETT, E., JENSEN, C. & POOLE, C. 2008. Primary cilia in osteoarthritic chondrocytes: from chondrons to clusters. *Developmental dynamics: an official publication of the American Association of Anatomists,* 237**,** 2013-2020.

MENDELL, J. T. & OLSON, E. N. 2012. MicroRNAs in stress signaling and human disease. *Cell,* 148**,** 1172-1187.

MIYAKI, S., NAKASA, T., OTSUKI, S., GROGAN, S. P., HIGASHIYAMA, R., INOUE, A., KATO, Y., SATO, T., LOTZ, M. K. & ASAHARA, H. 2009. MicroRNA-140 is expressed in differentiated human articular chondrocytes and modulates interleukin-1 responses. *Arthritis & Rheumatism,* 60**,** 2723-2730.

MIYAKI, S., SATO, T., INOUE, A., OTSUKI, S., ITO, Y., YOKOYAMA, S., KATO, Y., TAKEMOTO, F., NAKASA, T. & YAMASHITA, S. 2010. MicroRNA-140 plays dual roles in both cartilage development and homeostasis. *Genes & development,* 24**,** 1173-1185.

MOBASHERI, A., LEWIS, R., MAXWELL, J. E., HILL, C., WOMACK, M. & BARRETT-JOLLEY, R. 2010. Characterization of a stretch-activated potassium channel in chondrocytes. *Journal of cellular physiology,* 223**,** 511-518.

MOSHOUS, D., CALLEBAUT, I., DE CHASSEVAL, R., CORNEO, B., CAVAZZANA-CALVO, M., LE DEIST, F., TEZCAN, I., SANAL, O., BERTRAND, Y. & PHILIPPE, N. 2001. Artemis, a novel DNA double-strand break repair/V (D) J recombination protein, is mutated in human severe combined immune deficiency. *Cell,* 105**,** 177-186.

MURPHY, C. L. & SAMBANISS, A. 2001. Effect of oxygen tension on chondrocyte extracellular matrix accumulation. *Connective tissue research,* 42**,** 87-96.

NAGASE, H., VISSE, R. & MURPHY, G. 2006. Structure and function of matrix metalloproteinases and TIMPs. *Cardiovascular research,* 69**,** 562-573.

NAKAMURA, K., SHIRAI, T., MORISHITA, S., UCHIDA, S., SAEKI-MIURA, K. & MAKISHIMA, F. 1999. p38 mitogen-activated protein kinase functionally contributes to chondrogenesis induced by growth/differentiation factor-5 in ATDC5 cells. *Experimental cell research,* 250**,** 351-363.

NAKAMURA, Y., INLOES, J. B., KATAGIRI, T. & KOBAYASHI, T. 2011. Chondrocyte-specific microRNA-140 regulates endochondral bone development and targets Dnpep to modulate bone morphogenetic protein signaling. *Molecular and cellular biology,* 31**,** 3019-3028.

NEAL, J. A., SUGIMAN-MARANGOS, S., VANDERVERE-CAROZZA, P., WAGNER, M., TURCHI, J., LEES-MILLER, S. P., JUNOP, M. S. & MEEK, K. 2014. Unraveling the complexities of DNA-dependent protein kinase autophosphorylation. *Molecular and cellular biology,* 34**,** 2162-2175.

NEILSEN, C. T., GOODALL, G. J. & BRACKEN, C. P. 2012. IsomiRs–the overlooked repertoire in the dynamic microRNAome. *Trends in Genetics,* 28**,** 544-549.

NEUHOLD, L. A., KILLAR, L., ZHAO, W., SUNG, M.-L. A., WARNER, L., KULIK, J., TURNER, J., WU, W., BILLINGHURST, C. & MEIJERS, T. 2001. Postnatal expression in hyaline

cartilage of constitutively active human collagenase-3 (MMP-13) induces osteoarthritis in mice. *The Journal of clinical investigation,* 107**,** 35-44.

NIHONGAKI, Y., KAWANO, F., NAKAJIMA, T. & SATO, M. 2015. Photoactivatable CRISPR-Cas9 for optogenetic genome editing. *Nature biotechnology,* 33**,** 755.

NOZAKI, Y., YAMAGATA, T., SUGIYAMA, M., IKOMA, S., KINOSHITA, K. & FUNAUCHI, M. 2006. Anti-inflammatory effect of all-trans-retinoic acid in inflammatory arthritis. *Clinical Immunology,* 119**,** 272-279.

O'BRIEN, J., HAYDER, H., ZAYED, Y. & PENG, C. 2018. Overview of microRNA biogenesis, mechanisms of actions, and circulation. *Frontiers in endocrinology,* 9**,** 402.

O'CONOR, C. J., LEDDY, H. A., BENEFIELD, H. C., LIEDTKE, W. B. & GUILAK, F. 2014. TRPV4-mediated mechanotransduction regulates the metabolic response of chondrocytes to dynamic loading. *Proceedings of the National Academy of Sciences,* 111**,** 1316-1321.

OAKES, B. L., NADLER, D. C., FLAMHOLZ, A., FELLMANN, C., STAAHL, B. T., DOUDNA, J. A. & SAVAGE, D. F. 2016. Profiling of engineering hotspots identifies an allosteric CRISPR-Cas9 switch. *Nature biotechnology,* 34**,** 646.

OBERLENDER, S. A. & TUAN, R. S. 1994. Expression and functional involvement of N-cadherin in embryonic limb chondrogenesis. *Development,* 120**,** 177-187.

OKADA, C., YAMASHITA, E., LEE, S. J., SHIBATA, S., KATAHIRA, J., NAKAGAWA, A., YONEDA, Y. & TSUKIHARA, T. 2009. A high-resolution structure of the pre-microRNA nuclear export machinery. *Science,* 326**,** 1275-1279.

ORNITZ, D. M. & ITOH, N. 2015. The fibroblast growth factor signaling pathway. *Wiley Interdisciplinary Reviews: Developmental Biology,* 4**,** 215-266.

PALMOSKI, M., PERRICONE, E. & BRANDT, K. D. 1979. Development and reversal of a proteoglycan aggregation defect in normal canine knee cartilage after immobilization. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology,* 22**,** 508-517.

PASQUINELLI, A. E. 2012. MicroRNAs and their targets: recognition, regulation and an emerging reciprocal relationship. *Nature Reviews Genetics,* 13**,** 271.

PEREZ-PINERA, P., KOCAK, D. D., VOCKLEY, C. M., ADLER, A. F., KABADI, A. M., POLSTEIN, L. R., THAKORE, P. I., GLASS, K. A., OUSTEROUT, D. G. & LEONG, K. W. 2013. RNA-guided gene activation by CRISPR-Cas9–based transcription factors. *Nature methods,* 10**,** 973.

PINELLO, L., CANVER, M. C., HOBAN, M. D., ORKIN, S. H., KOHN, D. B., BAUER, D. E. & YUAN, G.-C. 2016. Analyzing CRISPR genome-editing experiments with CRISPResso. *Nature biotechnology,* 34**,** 695.

POOLE, C. A., FLINT, M. H. & BEAUMONT, B. W. 1987. Chondrons in cartilage: ultrastructural analysis of the pericellular microenvironment in adult human articular cartilages. *Journal of orthopaedic research,* 5**,** 509-522.

QI, L. S., LARSON, M. H., GILBERT, L. A., DOUDNA, J. A., WEISSMAN, J. S., ARKIN, A. P. & LIM, W. A. 2013. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell,* 152**,** 1173-1183.

RAN, F. A., HSU, P. D., LIN, C.-Y., GOOTENBERG, J. S., KONERMANN, S., TREVINO, A. E., SCOTT, D. A., INOUE, A., MATOBA, S. & ZHANG, Y. 2013a. Double nicking by RNA-

guided CRISPR Cas9 for enhanced genome editing specificity. *Cell,* 154**,** 1380-1389.

RAN, F. A., HSU, P. D., WRIGHT, J., AGARWALA, V., SCOTT, D. A. & ZHANG, F. 2013b. Genome engineering using the CRISPR-Cas9 system. *Nature protocols,* 8**,** 2281.

RAND, T. A., PETERSEN, S., DU, F. & WANG, X. 2005. Argonaute2 cleaves the anti-guide strand of siRNA during RISC activation. *Cell,* 123**,** 621-629.

RICHARDSON, C. D., RAY, G. J., DEWITT, M. A., CURIE, G. L. & CORN, J. E. 2016. Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nature biotechnology,* 34**,** 339.

ROGAKOU, E. P., PILCH, D. R., ORR, A. H., IVANOVA, V. S. & BONNER, W. M. 1998. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *Journal of biological chemistry,* 273**,** 5858-5868.

ROOS, H., ADALBERTH, T., DAHLBERG, L. & LOHMANDER, L. S. 1995. Osteoarthritis of the knee after injury to the anterior cruciate ligament or meniscus: the influence of time and age. *Osteoarthritis and Cartilage,* 3**,** 261-267.

RÜEGGER, S. & GROSSHAN, H. 2012. MicroRNA turnover: when, how, and why. *Trends in biochemical sciences,* 37**,** 436-446.

SANCHEZ-ADAMS, J., LEDDY, H. A., MCNULTY, A. L., O'CONOR, C. J. & GUILAK, F. 2014. The mechanobiology of articular cartilage: bearing the burden of osteoarthritis. *Current rheumatology reports,* 16**,** 451.

SAPOLSKY, A. I. & HOWELL, D. S. 1982. Further characterization of a neutral metalloprotease isolated from human articular cartilage. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology,* 25**,** 981-988.

SATIR, P., PEDERSEN, L. B. & CHRISTENSEN, S. T. 2010. The primary cilium at a glance. *J Cell Sci,* 123**,** 499-503.

SCHMITT, B., RINGE, J., HÄUPL, T., NOTTER, M., MANZ, R., BURMESTER, G.-R., SITTINGER, M. & KAPS, C. 2003. BMP2 initiates chondrogenic lineage development of adult human mesenchymal stem cells in high-density culture. *Differentiation,* 71**,** 567-577.

SCILABRA, S. D., PIGONI, M., PRAVATÁ, V., SCHÄTZL, T., MÜLLER, S. A., TROEBERG, L. & LICHTENTHALER, S. F. 2018. Increased TIMP-3 expression alters the cellular secretome through dual inhibition of the metalloprotease ADAM10 and ligand-binding of the LRP-1 receptor. *Scientific reports,* 8**,** 14697.

SEIDL, C. I., MARTINEZ-SANCHEZ, A. & MURPHY, C. L. 2016. Derepression of microRNA-138 contributes to loss of the human articular chondrocyte phenotype. *Arthritis & Rheumatology,* 68**,** 398-409.

SEMENZA, G. L. 2000. HIF-1 and human disease: one highly involved factor. *Genes & development,* 14**,** 1983-1991.

SEMENZA, G. L. & WANG, G. L. 1992. A nuclear factor induced by hypoxia via de novo protein synthesis binds to the human erythropoietin gene enhancer at a site required for transcriptional activation. *Molecular and cellular biology,* 12**,** 5447-5454.

SENTMANAT, M. F., PETERS, S. T., FLORIAN, C. P., CONNELLY, J. P. & PRUETT-MILLER, S. M. 2018. A survey of validation strategies for CRISPR-Cas9 editing. *Scientific reports,* 8**,** 888.

SETHI, P. & LUKIW, W. J. 2009. Micro-RNA abundance and stability in human brain: specific alterations in Alzheimer's disease temporal lobe neocortex. *Neuroscience letters,* 459**,** 100-104.

SHEN, J., LI, J., WANG, B., JIN, H., WANG, M., ZHANG, Y., YANG, Y., IM, H. J., O'KEEFE, R. & CHEN, D. 2013. Deletion of the transforming growth factor β receptor type II gene in articular chondrocytes leads to a progressive osteoarthritis-like phenotype in mice. *Arthritis & Rheumatism,* 65**,** 3107-3119.

SHERMAN, M. F., WARREN, R. F., MARSHALL, J. L. & SAVATSKY, G. J. 1988. A clinical and radiographical analysis of 127 anterior cruciate insufficient knees. *Clinical orthopaedics and related research,* 227**,** 229-237.

SHMAKOV, S., ABUDAYYEH, O. O., MAKAROVA, K. S., WOLF, Y. I., GOOTENBERG, J. S., SEMENOVA, E., MINAKHIN, L., JOUNG, J., KONERMANN, S. & SEVERINOV, K. 2015. Discovery and functional characterization of diverse class 2 CRISPR-Cas systems. *Molecular cell,* 60**,** 385-397.

SIMSEK, D. & JASIN, M. 2010. Alternative end-joining is suppressed by the canonical NHEJ component Xrcc4–ligase IV during chromosomal translocation formation. *Nature structural & molecular biology,* 17**,** 410.

SINGH, R., KUSCU, C., QUINLAN, A., QI, Y. & ADLI, M. 2015. Cas9-chromatin binding information enables more accurate CRISPR off-target prediction. *Nucleic acids research,* 43**,** e118-e118.

SMALL, E. M. & OLSON, E. N. 2011. Pervasive roles of microRNAs in cardiovascular biology. *Nature,* 469**,** 336.

SONG, J., KIM, D. & JIN, E. J. 2011. MicroRNA-488 suppresses cell migration through modulation of the focal adhesion activity during chondrogenic differentiation of chick limb mesenchymal cells. *Cell biology international,* 35**,** 179-185.

SONG, R. H., D. TORTORELLA, M., MALFAIT, A. M., ALSTON, J. T., YANG, Z., ARNER, E. C. & GRIGGS, D. W. 2007. Aggrecan degradation in human articular cartilage explants is mediated by both ADAMTS-4 and ADAMTS-5. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology,* 56**,** 575-585.

SOOND, S. M. & CHANTRY, A. 2011. Selective targeting of activating and inhibitory Smads by distinct WWP2 ubiquitin ligase isoforms differentially modulates TGFβ signalling and EMT. *Oncogene,* 30**,** 2451.

SOPHIA FOX, A. J., BEDI, A. & RODEO, S. A. 2009. The basic science of articular cartilage: structure, composition, and function. *Sports health,* 1**,** 461-468.

SPAGGIARI, G. M., ABDELRAZIK, H., BECCHETTI, F. & MORETTA, L. 2009. MSCs inhibit monocyte-derived DC maturation and function by selectively interfering with the generation of immature DCs: central role of MSC-derived prostaglandin E2. *Blood,* 113**,** 6576-6583.

SPARKS, M. A., CROWLEY, S. D., GURLEY, S. B., MIROTSOU, M. & COFFMAN, T. M. 2011. Classical renin-angiotensin system in kidney physiology. *Comprehensive Physiology,* 4**,** 1201-1228.

SRIKANTH, V. K., FRYER, J. L., ZHAI, G., WINZENBERG, T. M., HOSMER, D. & JONES, G. 2005. A meta-analysis of sex differences prevalence, incidence and severity of osteoarthritis. *Osteoarthritis and cartilage,* 13**,** 769-781.

STANTON, H., GOLUB, S. B., ROGERSON, F. M., LAST, K., LITTLE, C. B. & FOSANG, A. J. 2011. Investigating ADAMTS-mediated aggrecanolysis in mouse cartilage. *Nature protocols,* 6**,** 388.

STANTON, H., ROGERSON, F. M., EAST, C. J., GOLUB, S. B., LAWLOR, K. E., MEEKER, C. T., LITTLE, C. B., LAST, K., FARMER, P. J. & CAMPBELL, I. K. 2005. ADAMTS5 is the major aggrecanase in mouse cartilage in vivo and in vitro. *Nature,* 434**,** 648.

STERNBERG, S. H., REDDING, S., JINEK, M., GREENE, E. C. & DOUDNA, J. A. 2014. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature,* 507**,** 62.

STYRKARSDOTTIR, U., THORLEIFSSON, G., HELGADOTTIR, H. T., BOMER, N., METRUSTRY, S., BIERMA-ZEINSTRA, S., STRIJBOSCH, A. M., EVANGELOU, E., HART, D. & BEEKMAN, M. 2014. Severe osteoarthritis of the hand associates with common variants within the ALDH1A2 gene and with rare variants at 1p31. *Nature genetics,* 46**,** 498.

SUZUKI, H. I., KATSURA, A., YASUDA, T., UENO, T., MANO, H., SUGIMOTO, K. & MIYAZONO, K. 2015. Small-RNA asymmetry is directly driven by mammalian Argonautes. *Nature structural & molecular biology,* 22**,** 512.

SZCZELKUN, M. D., TIKHOMIROVA, M. S., SINKUNAS, T., GASIUNAS, G., KARVELIS, T., PSCHERA, P., SIKSNYS, V. & SEIDEL, R. 2014. Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. *Proceedings of the National Academy of Sciences,* 111**,** 9798-9803.

TAKATA, A., OTSUKA, M., YOSHIKAWA, T., KISHIKAWA, T., HIKIBA, Y., OBI, S., GOTO, T., KANG, Y. J., MAEDA, S. & YOSHIDA, H. 2013. MicroRNA-140 acts as a liver tumor suppressor by controlling NF-κB activity by directly targeting DNA methyltransferase 1 (Dnmt1) expression. *Hepatology,* 57**,** 162-170.

TAKAZAWA, I., LARANGE, A. & CHEROUTRE, H. 2019. Cellular retinoic acid binding protein 2 (CRABP2) serves as a critical regulator in immune regulation and inflammation. Am Assoc Immnol.

TANG, X., MUHAMMAD, H., MCLEAN, C., MIOTLA-ZAREBSKA, J., FLEMING, J., DIDANGELOS, A., ÖNNERFJORD, P., LEASK, A., SAKLATVALA, J. & VINCENT, T. L. 2018. Connective tissue growth factor contributes to joint homeostasis and osteoarthritis severity by controlling the matrix sequestration and activation of latent TGFβ. *Annals of the rheumatic diseases,* 77**,** 1372-1380.

TAO, S.-C., YUAN, T., ZHANG, Y.-L., YIN, W.-J., GUO, S.-C. & ZHANG, C.-Q. 2017. Exosomes derived from miR-140-5p-overexpressing human synovial mesenchymal stem cells enhance cartilage tissue regeneration and prevent osteoarthritis of the knee in a rat model. *Theranostics,* 7**,** 180.

TARDIF, G., HUM, D., PELLETIER, J.-P., DUVAL, N. & MARTEL-PELLETIER, J. 2009. Regulation of the IGFBP-5 and MMP-13 genes by the microRNAs miR-140 and miR-27a in human osteoarthritic chondrocytes. *BMC musculoskeletal disorders,* 10**,** 148.

TOPLETZ, A. R., THATCHER, J. E., ZELTER, A., LUTZ, J. D., TAY, S., NELSON, W. L. & ISOHERRANEN, N. 2012. Comparison of the function and expression of CYP26A1

and CYP26B1, the two retinoic acid hydroxylases. *Biochemical pharmacology,* 83**,** 149-163.

TROEBERG, L. & NAGASE, H. 2012. Proteases involved in cartilage matrix degradation in osteoarthritis. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics,* 1824**,** 133-145.

TRUONG, L. N., LI, Y., SHI, L. Z., HWANG, P. Y.-H., HE, J., WANG, H., RAZAVIAN, N., BERNS, M. W. & WU, X. 2013. Microhomology-mediated End Joining and Homologous Recombination share the initial end resection step to repair DNA double-strand breaks in mammalian cells. *Proceedings of the National Academy of Sciences,* 110**,** 7720-7725.

TSAI, C. J., KIM, S. A. & CHU, G. 2007. Cernunnos/XLF promotes the ligation of mismatched and noncohesive DNA ends. *Proceedings of the National Academy of Sciences,* 104**,** 7851-7856.

TSENG, S., REDDI, A. H. & DI CESARE, P. E. 2009. Cartilage oligomeric matrix protein (COMP): a biomarker of arthritis. *Biomarker insights,* 4**,** BMI. S645.

TSUTSUMI, S., SHIMAZU, A., MIYAZAKI, K., PAN, H., KOIKE, C., YOSHIDA, E., TAKAGISHI, K. & KATO, Y. 2001. Retention of multilineage differentiation potential of mesenchymal cells during proliferation in response to FGF. *Biochemical and biophysical research communications,* 288**,** 413-419.

TUDDENHAM, L., WHEELER, G., NTOUNIA-FOUSARA, S., WATERS, J., HAJIHOSSEINI, M. K., CLARK, I. & DALMAY, T. 2006. The cartilage specific microRNA-140 targets histone deacetylase 4 in mouse cells. *FEBS letters,* 580**,** 4214-4217.

TULI, R., SEGHATOLESLAMI, M., TULI, S., HOWARD, M., DANIELSON, K. & TUAN, R. 2002. p38 MAP Kinase Regulation of AP-2 Binding in TGF-β1-Stimulated Chondrogenesis of Human Trabecular Bone-Derived Cells. *Annals of the New York Academy of Sciences,* 961**,** 172-177.

VALVERDE-FRANCO, G., BINETTE, J., LI, W., WANG, H., CHAI, S., LAFLAMME, F., TRAN-KHANH, N., QUENNEVILLE, E., MEIJERS, T. & POOLE, A. 2006. Defects in articular cartilage metabolism and early arthritis in fibroblast growth factor receptor 3 deficient mice. *Human molecular genetics,* 15**,** 1783-1792.

VAN DER KRAAN, P. M. 2017. The changing role of TGFβ in healthy, ageing and osteoarthritic joints. *Nature Reviews Rheumatology,* 13**,** 155.

VAN DER KRAAN, P. M., VITTERS, E., VAN BEUNINGEN, H. M., VAN DE PUTTE, L. & VAN DEN BERG, W. 1990. Degenerative knee joint lesions in mice after a single intra-articular collagenase injection. A new model of osteoarthritis. *Journal of experimental pathology (Oxford, England),* 71**,** 19.

VANWANSEELE, B., ECKSTEIN, F., KNECHT, H., SPAEPEN, A. & STÜSSI, E. 2003. Longitudinal analysis of cartilage atrophy in the knees of patients with spinal cord injury. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology,* 48**,** 3377-3381.

VASUDEVAN, S. 2012. Posttranscriptional upregulation by microRNAs. *Wiley Interdisciplinary Reviews: RNA,* 3**,** 311-330.

VEGA, R. B., MATSUDA, K., OH, J., BARBOSA, A. C., YANG, X., MEADOWS, E., MCANALLY, J., POMAJZL, C., SHELTON, J. M. & RICHARDSON, J. A. 2004. Histone deacetylase 4 controls chondrocyte hypertrophy during skeletogenesis. *Cell,* 119**,** 555-566.

VINCENT, T., HERMANSSON, M., BOLTON, M., WAIT, R. & SAKLATVALA, J. 2002. Basic FGF mediates an immediate response of articular cartilage to mechanical injury. *Proceedings of the National Academy of Sciences,* 99**,** 8259-8264.

VINCENT, T. L. Mechanoflammation in osteoarthritis pathogenesis. Seminars in arthritis and rheumatism, 2019. Elsevier, S36-S38.

VINCENT, T. L., HERMANSSON, M. A., HANSEN, U. N., AMIS, A. A. & SAKLATVALA, J. 2004. Basic fibroblast growth factor mediates transduction of mechanical signals when articular cartilage is loaded. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology,* 50**,** 526-533.

VINCENT, T. L. & WANN, A. K. 2019. Mechanoadaptation: articular cartilage through thick and thin. *The Journal of physiology,* 597**,** 1271-1281.

VON LOGA, I. S., EL-TURABI, A., JOSTINS, L., MIOTLA-ZAREBSKA, J., MACKAY-ALDERSON, J., ZELTINS, A., PARISI, I., BACHMANN, M. F. & VINCENT, T. L. 2019. Active immunisation targeting nerve growth factor attenuates chronic pain behaviour in murine osteoarthritis. *Annals of the rheumatic diseases,* 78**,** 672-675.

WALKER, J. R., CORPINA, R. A. & GOLDBERG, J. 2001. Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature,* 412**,** 607.

WANG, B., LI, S., QI, H. H., CHOWDHURY, D., SHI, Y. & NOVINA, C. D. 2009. Distinct passenger strand and mRNA cleavage activities of human Argonaute proteins. *Nature structural & molecular biology,* 16**,** 1259.

WANG, H. & ATESHIAN, G. A. 1997. The normal stress effect and equilibrium friction coefficient of articular cartilage under steady frictional shear. *Journal of biomechanics,* 30**,** 771-776.

WANG, Z., CHI, X., LIU, L., WANG, Y., MEI, X., YANG, Y. & JIA, T. 2018. Long noncoding RNA maternally expressed gene 3 knockdown alleviates lipopolysaccharide-induced inflammatory injury by up-regulation of miR-203 in ATDC5 cells. *Biomedicine & Pharmacotherapy,* 100**,** 240-249.

WATANABE, H., DE CAESTECKER, M. P. & YAMADA, Y. 2001. Transcriptional cross-talk between Smad, ERK1/2, and p38 mitogen-activated protein kinase pathways regulates transforming growth factor-β-induced aggrecan gene expression in chondrogenic ATDC5 cells. *Journal of Biological Chemistry,* 276**,** 14466-14473.

WATT, F. E., PATERSON, E., FREIDIN, A., KENNY, M., JUDGE, A., SAKLATVALA, J., WILLIAMS, A. & VINCENT, T. L. 2016. Acute molecular changes in synovial fluid following human knee injury: association with early clinical outcomes. *Arthritis & rheumatology,* 68**,** 2129-2140.

WEI, T., KULKARNI, N., ZENG, Q., HELVERING, L., LIN, X., LAWRENCE, F., HALE, L., CHAMBERS, M., LIN, C. & HARVEY, A. 2010. Analysis of early changes in the articular cartilage transcriptisome in the rat meniscal tear model of osteoarthritis: pathway comparisons with the rat anterior cruciate transection model and with human osteoarthritic cartilage. *Osteoarthritis and cartilage,* 18**,** 992-1000.

WENG, T., YI, L., HUANG, J., LUO, F., WEN, X., DU, X., CHEN, Q., DENG, C., CHEN, D. & CHEN, L. 2012. Genetic inhibition of fibroblast growth factor receptor 1 in knee

cartilage attenuates the degeneration of articular cartilage in adult mice. *Arthritis & Rheumatism,* 64**,** 3982-3992.

WIDELITZ, R. B., JIANG, T. X., MURRAY, B. A. & CHUONG, C. M. 1993. Adhesion molecules in skeletogenesis: II. Neural cell adhesion molecules mediate precartilaginous mesenchymal condensations and enhance chondrogenesis. *Journal of cellular physiology,* 156**,** 399-411.

WIENERT, B., WYMAN, S. K., RICHARDSON, C. D., YEH, C. D., AKCAKAYA, P., PORRITT, M. J., MORLOCK, M., VU, J. T., KAZANE, K. R. & WATRY, H. L. 2019. Unbiased detection of CRISPR off-targets in vivo using DISCOVER-Seq. *Science,* 364**,** 286-289.

WIENHOLDS, E., KLOOSTERMAN, W. P., MISKA, E., ALVAREZ-SAAVEDRA, E., BEREZIKOV, E., DE BRUIJN, E., HORVITZ, H. R., KAUPPINEN, S. & PLASTERK, R. H. 2005. MicroRNA expression in zebrafish embryonic development. *Science,* 309**,** 310-311.

WILKINS, R., BROWNING, J. & ELLORY, J. 2000. Surviving in a matrix: membrane transport in articular chondrocytes. *The Journal of membrane biology,* 177**,** 95-108.

WILUSZ, R. E., DEFRATE, L. E. & GUILAK, F. 2012. A biomechanical role for perlecan in the pericellular matrix of articular cartilage. *Matrix Biology,* 31**,** 320-327.

YAMAGUCHI, Y., MANN, D. M. & RUOSLAHTI, E. 1990. Negative regulation of transforming growth factor-β by the proteoglycan decorin. *Nature,* 346**,** 281.

YAMAMOTO, K., SANTAMARIA, S., BOTKJAER, K. A., DUDHIA, J., TROEBERG, L., ITOH, Y., MURPHY, G. & NAGASE, H. 2017. Inhibition of Shedding of Low-Density Lipoprotein Receptor–Related Protein 1 Reverses Cartilage Matrix Degradation in Osteoarthritis. *Arthritis & Rheumatology,* 69**,** 1246-1256.

YAMAMOTO, K., TROEBERG, L., SCILABRA, S. D., PELOSI, M., MURPHY, C. L., STRICKLAND, D. K. & NAGASE, H. 2013. LRP-1-mediated endocytosis regulates extracellular activity of ADAMTS-5 in articular cartilage. *The FASEB Journal,* 27**,** 511-521.

YAN, D., CHEN, D., COOL, S. M., VAN WIJNEN, A. J., MIKECZ, K., MURPHY, G. & IM, H.-J. 2011. Fibroblast growth factor receptor 1 is principally responsible for fibroblast growth factor 2-induced catabolic activities in human articular chondrocytes. *Arthritis research & therapy,* 13**,** R130.

YANG, D., SCAVUZZO, M. A., CHMIELOWIEC, J., SHARP, R., BAJIC, A. & BOROWIAK, M. 2016. Enrichment of G2/M cell cycle phase in human pluripotent stem cells enhances HDR-mediated gene repair with customizable endonucleases. *Scientific reports,* 6**,** 21264.

YANG, J., QIN, S., YI, C., MA, G., ZHU, H., ZHOU, W., XIONG, Y., ZHU, X., WANG, Y. & HE, L. 2011. MiR-140 is co-expressed with Wwp2-C transcript and activated by Sox9 to target Sp1 in maintaining the chondrocyte proliferation. *FEBS letters,* 585**,** 2992-2997.

YANG, W. 2008. An equivalent metal ion in one-and two-metal-ion catalysis. *Nature structural & molecular biology,* 15**,** 1228.

YAO, C.-K., LIN, Y. Q., LY, C. V., OHYAMA, T., HAUETER, C. M., MOISEENKOVA-BELL, V. Y., WENSEL, T. G. & BELLEN, H. J. 2009. A synaptic vesicle-associated Ca2+ channel promotes endocytosis and couples exocytosis to endocytosis. *Cell,* 138**,** 947-960.

YASUDA, T. 2006. Cartilage destruction by matrix degradation products. *Modern rheumatology,* 16**,** 197-205.

YOUNG, D. A., BARTER, M. J. & WILKINSON, D. J. 2019. Recent advances in understanding the regulation of metalloproteinases. *F1000Research,* 8.

YU, J., ZHANG, W., TANG, H., QIAN, H., YANG, J., ZHU, Z., REN, P. & LU, B. 2016. Septin 2 accelerates the progression of biliary tract cancer and is negatively regulated by mir-140-5p. *Gene,* 589**,** 20-26.

ZETSCHE, B., GOOTENBERG, J. S., ABUDAYYEH, O. O., SLAYMAKER, I. M., MAKAROVA, K. S., ESSLETZBICHLER, P., VOLZ, S. E., JOUNG, J., VAN DER OOST, J. & REGEV, A. 2015. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell,* 163**,** 759-771.

ZHEN, G., WEN, C., JIA, X., LI, Y., CRANE, J. L., MEARS, S. C., ASKIN, F. B., FRASSICA, F. J., CHANG, W. & YAO, J. 2013. Inhibition of TGF-β signaling in mesenchymal stem cells of subchondral bone attenuates osteoarthritis. *Nature medicine,* 19**,** 704.

ZHU, L., CHANALARIS, A., GROVES, K., FURNISS, D., WATT, F., GARDINER, M. & VINCENT, T. 2018. Polymorphic variants in ALDH1A2 determine the expression level of ALDH1A2 and CYP19A1 in the cartilage of patients undergoing trapeziectomy for severe thumb osteoarthritis. *Osteoarthritis and Cartilage,* 26**,** S157.

ZURIS, J. A., THOMPSON, D. B., SHU, Y., GUILINGER, J. P., BESSEN, J. L., HU, J. H., MAEDER, M. L., JOUNG, J. K., CHEN, Z.-Y. & LIU, D. R. 2015. Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nature biotechnology,* 33**,** 73.

Appendix

8

Chapter 8. Appendices

8.1 MiSeq methodology

The following are excerpts from "Illumina: And introduction to Next-Generation Sequencing Technology."[1]

In principle, the concept behind NGS technology is similar to CE sequencing. DNA polymerase catalyses the incorporation of fluorescently labelled deoxyribonucleotide triphosphates (dNTPs) into a DNA template strand during sequential cycles of DNA synthesis. During each cycle, at the point of incorporation, the nucleotides are identified by fluorophore excitation. The critical difference is that, instead of sequencing a single DNA fragment, NGS extends this process across millions of fragments in a massively parallel fashion. More than 90% of the world's sequencing data are generated by Illumina sequencing by synthesis (SBS) chemistry. It delivers high accuracy, a high yield of error-free reads, and a high percentage of base calls above Q30

Illumina NGS workflows include four basic steps:

1. Library Preparation – The sequencing library is prepared by random fragmentation of the DNA or cDNA sample, followed by 5′and 3′adapter ligation. Alternatively, "tagmentation" combines the fragmentation and ligation reactions into a single step that greatly increases the efficiency of the library preparation process Adapter-ligated fragments are then PCR amplified and gel purified.
2. Cluster Generation—For cluster generation, the library is loaded into a flow cell where fragments are captured on a lawn of surface-bound oligos complementary to the library adapters. Each fragment is then amplified into distinct, clonal clusters through bridge amplification. When cluster generation is complete, the templates are ready for sequencing.
3. Sequencing—Illumina SBS technology uses a proprietary reversible terminator–based method that detects single bases as they are incorporated into DNA template strands. As all four reversible terminator–bound dNTPs are present during each sequencing cycle, natural competition minimizes incorporation bias and greatly reduces raw error rates compared to other technologies. The result is highly accurate base-by-base sequencing that virtually eliminates sequence context–specific errors, even within repetitive sequence regions and homopolymers.
4. Data Analysis—During data analysis and alignment, the newly identified sequence reads are aligned to a reference genome. Following alignment, many variations of analysis are possible, such as single nucleotide polymorphism (SNP) or insertion-deletion (indel) identification, read counting for RNA methods, phylogenetic or metagenomic analysis, and more.

[1] https://www.illumina.com/documents/products/illumina_sequencing_introduction.pdf

A major advance in NGS technology occurred with the development of paired-end (PE) sequencing. PE sequencing involves sequencing both ends of the DNA fragments in a library and aligning the forward and reverse reads as read pairs. In addition to producing twice the number of reads for the same time and effort in library preparation, sequences aligned as read pairs enable more accurate read alignment and the ability to detect indels, which is not possible with single- read data. Analysis of differential read-pair spacing also allows removal of PCR duplicates, a common artefact resulting from PCR amplification during library preparation. Furthermore, PE sequencing produces a higher number of SNV calls following read-pair alignment. While some methods are best served by single-read sequencing, such as small RNA sequencing, most researchers currently use the paired-end approach.

8.2 Attended conferences

Cutting edge OA: June 2017; Oxford – 3 mins talk & poster presentation.

Genome Editing Congress: April 2018; Oxford

Gordon Research Conference, Cartilage Biology & Pathology: March 2019; Galveston, TX, USA – Poster presentation

BSMB: Stroma, Niche, and Repair; April 2019; Liverpool, UK – Poster presentation.

Cutting edge OA: June 2019; Oxford – 3 mins talk & poster presentation.

8.3 Awards

Best poster presentation at BSMB: Stroma, Niche, and Repair; April 2019; Liverpool, UK