

# **Identification and Characterisation of MicroRNAs Expressed in Articular Cartilage and Osteoarthritis**

**Natalie Elizabeth Crowe**

Thesis submitted for the degree of Doctor of Philosophy

University of East Anglia

School of Biological Sciences

Submitted June 2014

© This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that use of any information derived there from must be in accordance with current UK Copyright Law. In addition, any quotation or extract must include full attribution

## **Abstract**

Osteoarthritis (OA) is a debilitating disease suffered by millions of people around the world. It affects the whole joint but is defined in part by a degradation of cartilage which becomes more pronounced as the disease progresses. The initiation and progression of OA is still not fully characterised.

MicroRNAs are small, non-coding RNA molecules which are known to regulate gene expression, however, at the moment it is unclear exactly how many may be expressed in cartilage or how their expression may alter during osteoarthritis.

Using human OA cartilage and cells derived from it, we have optimised small RNA purification and have shown that there is a clear difference between microRNA expression in cartilage compared with cell isolates, with miR-140 most highly expressed in cells immediately after digestion. This may suggest a stress-type response.

RNA purified from chondrocytes immediately after digestion from cartilage was utilised for Illumina GIIIX deep sequencing in an attempt to identify novel microRNAs in addition to expressed known microRNAs. 990 previously documented and 1621 potential novel microRNAs were discovered and 16 novel candidates were chosen for further validation.

Validation was carried out using northern blotting and qRT-PCR on a number of samples and tissue types. In addition, the selected candidates were also examined in chondrogenesis models, and one candidate was probed in mouse embryo *in situ* hybridisation in an attempt to localise the candidate to cartilage tissue.

Three of the candidates with the most promising validation results were analysed for function using a whole genome microarray combined with computational target search analysis. A number of likely targets obtained from the array were further validated by subcloning and luciferase assays, resulting in the discovery of ITGA5 as a target of candidate novel 11.

The outcomes from this research provide us with an increased understanding of the miRNome of human OA cartilage and could impact upon future drug development and the use of microRNAs as biomarkers.

# Contents

Abstract .....	2
List of Figures .....	7
List of Tables.....	10
List of Appendices .....	12
Acknowledgements .....	13
Chapter 1 – Introduction .....	15
1.1 - Articular Cartilage And Its Properties .....	16
1.2 - Osteoarthritis.....	31
1.3 - Regulatory Mechanisms Of Osteoarthritis .....	36
1.4 – Biosynthesis Of MicroRNAs .....	40
1.5 - MicroRNA Involvement Within Osteoarthritis .....	51
1.6 – Aims And Objectives .....	57
Chapter 2 – Materials and Methods .....	58
2.1 Cell Lines .....	59
2.1.1 SW1353.....	59
2.1.2 DLD-1 .....	59
2.1.3 DF1.....	59
2.2 Human Tissue Panel RNA .....	60
2.3 Tissue Samples.....	60
2.4 Human Cartilage Dissection .....	60
2.5 Primary Chondrocyte Isolation .....	61
2.6 Cell Culture .....	63
2.7 RNA Extraction.....	63
2.7.1 RNA Extraction From Frozen Cartilage – Using mirVana™ miRNA Isolation Kit [Ambion (Glasgow)] .....	63
2.7.2 RNA Extraction From Monolayer Cells – Using mirVana™ miRNA Isolation Kit (Ambion).....	63
2.7.3 RNA Extraction From Post Digest Phase Of Primary Cell Isolation – Using mirVana™ miRNA Isolation Kit (Ambion) .....	64
2.7.4 RNA Enrichment For Small RNAs From Total RNA– Using mirVana™ miRNA Isolation Kit (Ambion) .....	64
2.8 Ethanol RNA Precipitation .....	64
2.9 cDNA Synthesis .....	64
2.9.1 cDNA Synthesis – Cells-to-cDNA II With MMLV .....	64

2.9.2 cDNA Synthesis - Reverse Transcription .....	65
2.9.3 cDNA Synthesis - miRCURY LNA™ Universal RT MicroRNA PCR Universal cDNA Synthesis Kit [Exiqon (Copenhagen)] .....	65
2.10 Quantitative Real Time PCR (qRT-PCR).....	65
2.10.1 qRT-PCR With Primers And Probes .....	65
2.10.2 qRT-PCR Using miRCURY LNA™ Universal RT MicroRNA PCR SYBR Green Master Mix Universal RT (Exiqon) .....	66
2.11 DNA Polyacrylamide Gel Electrophoresis .....	67
2.12 Northern Blot .....	67
2.13 LDH Cytotoxicity Assay (Promega).....	68
2.14 Deep Sequencing Of MicroRNAs.....	68
2.15 MicroRNA siRNA Transfection .....	69
2.16 Functional Analysis of Novels 2, 7 and 11 – Whole-Genome Expression Using Bead Arrays Microarray (Source Bioscience) .....	70
2.17 De-differentiation Assay of Human Articular Osteoarthritic Knee Primary Chondrocytes .....	70
2.18 In situ Hybridisation.....	71
2.18.1 Mouse Tissue In Situ Hybridisation Preparation .....	71
2.18.2 In Situ Hybridisation Of Novel MicroRNA In Mouse Embryo.....	71
2.19 Mouse Hip Avulsion Cartilage Wounding Assay.....	72
2.20 Sub-Cloning .....	73
2.20.1 Plasmids and gDNA.....	73
2.20.2 CaCl <sub>2</sub> Competent DH5α Cell Preparation.....	73
2.20.3 Polymerase Chain Reaction .....	74
2.20.4 Mini-Prep – QIAprep Spin Miniprep Kit [Qiagen (Manchester)] .....	74
2.20.5 Mini-Prep – QIAprep Spin Miniprep Kit (Without Filters) (Qiagen).....	74
2.20.6 Gel Purification [Zymoclean (Cambridge)] .....	74
2.20.7 Phenol: Chloroform Purification.....	75
2.20.8 Restriction Digests .....	75
2.20.8.1 SalI Digestion Of Plasmid And Inserts .....	75
2.20.8.2 Directional Digest .....	75
2.20.9 Vector De-Phosphorylation.....	75
2.20.10 GoTaq (Promega) dATP Extension Of Products .....	75
2.20.11 Ligation Using pGEM-T Easy Vector System (Promega).....	76
2.20.12 Transformation of dsDNA in pGEM-T Easy Vector System (Promega) ..	76

2.20.13 Ligation and Transformation Of Inserts With pmirGLO Vector .....	76
2.20.14 Source Bioscience Sequencing .....	76
2.20.15 Transformation of Constructs And Glycerol Stock Production .....	76
2.21 Transfection Of Cloned Constructs Into DF1 Cells .....	77
2.22 Luciferase Assay – Dual Luciferase Reporter Assay System (Promega) .....	77
2.23 Data Analysis, Statistics And Bioinformatics .....	78
2.23.1 Data Analysis Of qRT-PCR – Comparative C <sub>t</sub> Method .....	78
2.23.2 Data Analysis Of qRT-PCR – Standard Curve Method.....	78
2.23.3 Statistical Analysis Of qRT-PCR Data .....	78
2.23.4 NGS Deep Sequencing Bioinformatics.....	79
2.23.5 MicroRNA Target Search Using RStudios .....	80
2.23.6 Candidate novel microRNA species conservation analysis .....	80
2.23.7 Bioinformatics For Genome Array .....	81
Chapter 3 – MicroRNA Profiling of Osteoarthritic Cartilage.....	82
3.1 Introduction .....	83
3.2 Results .....	86
3.2.1 Expression of key microRNAs in primary chondrocytes .....	86
3.2.2 Profile of existing microRNAs in osteoarthritic primary chondrocytes .....	94
3.2.3 Candidate novel microRNA discovery in primary osteoarthritic chondrocytes .....	101
3.3 Discussion .....	106
Chapter 4 – Experimental Validation of Novel MicroRNAs.....	113
4.1 Introduction .....	114
4.2 Results .....	116
4.2.1 Candidate novel microRNA size and function validation.....	116
4.2.2 Candidate novel microRNA expression in human tissues .....	121
4.2.3 Candidate novel microRNA expression during chondrocyte differentiation and de-differentiation.....	133
4.2.4 Candidate novel microRNA expression in vivo.....	139
4.2.5 Genetic location and evolution of candidate novel microRNAs.....	142
4.3 Discussion .....	150
Chapter 5 – Novel MicroRNA Target Search Analysis.....	164
5.1 Introduction .....	165
5.2 Results .....	169
5.2.1 Experimental microRNA target search analysis .....	169

5.2.2 Bioinformatics microRNA target search analysis.....	182
5.2.3 MicroRNA target search analysis combining array data and bioinformatics .....	187
5.2.4 Experimental validation of microRNA targets .....	194
5.3 Discussion .....	205
Chapter 6 – General Discussion and Future Directions .....	216
6.1 – MicroRNA profiling .....	217
6.2 – MicroRNA validation.....	220
6.3 – MicroRNA functional analysis .....	222
6.4 – Final summary.....	224
Appendices .....	225
Bibliography.....	263

## List of Figures

Figure 1.1 Transcriptional regulation during chondrogenesis .....	19
Figure 1.2 Endochondral ossification .....	21
Figure 1.3 The zones of articular cartilage.....	24
Figure 1.4 Matrices formation within cartilage.....	27
Figure 1.5 The balance of anabolic and catabolic factors in healthy cartilage .....	30
Figure 1.6 MicroRNA synthesis process .....	44
Figure 1.7 Diagram of microRNA : mRNA interaction patterns.....	48
Figure 2.1 Images of primary chondrocyte isolation .....	62
Figure 3.1 MicroRNA expression in articular chondrocyte isolation and expansion. ....	89
Figure 3.2 Collagen gene expression in articular chondrocyte isolation and expansion.	90
Figure 3.3 Gene expression in articular chondrocyte isolation and expansion.....	92
Figure 3.4 Images of primary chondrocytes. ....	93
Figure 3.5 Known microRNAs detected from deep sequencing.....	96
Figure 3.6 Ten most highly expressed known microRNAs from deep sequencing.....	98
Figure 3.7 MicroRNA-140 expression in chondrocytes. ....	99
Figure 3.8 Expression of microRNA-140 on northern blots.....	100
Figure 3.9 Read numbers of candidate novel microRNAs from deep sequencing. ....	103
Figure 4.1 Assessment of candidate novel microRNA using northern blots. ....	118
Figure 4.2 Candidate novel microRNA expression in DLD-1 parental and DLD-1 Dicer null cell lines. ....	119
Figure 4.3 Gene and microRNA expression in DLD-1 parental and DLD-1 Dicer null cell lines.. ....	120
Figure 4.4 Candidate novel microRNA expression in HOA and NOF cartilage tissue.. ..	124
Figure 4.5 MicroRNA expression in HOA and NOF cartilage tissue .....	125
Figure 4.6 Gene expression in HOA and NOF cartilage tissue. ....	126
Figure 4.7 Candidate novel microRNA expression in KOA and NOF cartilage tissue. ....	127
Figure 4.8 MicroRNA expression in KOA and NOF cartilage tissue. ....	128
Figure 4.9 Candidate novel microRNA expression across a human tissue panel.....	129
Figure 4.10 MicroRNA expression across a human tissue panel.....	130
Figure 4.11 Candidate novel microRNA expression in a cartilage injury model .....	131
Figure 4.12 MicroRNA expression in a cartilage injury model.....	132

Figure 4.13 Candidate novel microRNA expression in a mesenchymal stem cell chondrogenesis assay. ....	135
Figure 4.14 Candidate novel microRNA expression in an ATDC5 chondrogenesis assay. ....	136
Figure 4.15 MicroRNA expression in an ATDC5 chondrogenesis assay. ....	137
Figure 4.16 MicroRNA expression in articular chondrocyte isolation and expansion. ....	138
Figure 4.17 Expression of candidate microRNA novel 11 during mouse development. ....	140
Figure 4.18 Expression of control microRNA miR-449 during mouse development. ....	141
Figure 4.19 Candidate microRNA novel 2 genetic location within its host gene. ....	144
Figure 4.20 Candidate microRNA novel 7 genetic location within its host gene. ....	145
Figure 4.21 Candidate microRNA novel 11 genetic location within its host gene. ....	146
Figure 4.22 Candidate microRNA novel 2 species evolutionary tree. ....	147
Figure 4.23 Candidate microRNA novel 7 species evolutionary tree. ....	148
Figure 4.24 Candidate microRNA novel 11 species evolutionary tree. ....	149
Figure 5.1 Gene expression change in response to novel microRNA mimics and inhibitors. ....	173
Figure 5.2 Gene expression fold change in response to novel microRNA mimics and inhibitors. ....	174
Figure 5.3 Candidate targets which responded as expected to microRNA mimic and inhibitor treatments. ....	181
Figure 5.4 Venn diagram of computational analysis combined with gain- and loss-of-function studies to identify microRNA gene targets. ....	189
Figure 5.5 Candidate targets which responded as expected to mimic and inhibitor treatments which also have at least one target site. ....	190
Figure 5.6 Expression of candidate targets of miR-6509-5p in original array samples by qRT-PCR. ....	198
Figure 5.7 Expression of candidate targets of miR-664b-3p in original array samples by qRT-PCR. ....	200
Figure 5.8 Expression of candidate targets of miR-3085-3p in original array samples by qRT-PCR. ....	201
Figure 5.9 Luciferase assays of cloned 3'UTR constructs of target genes in pmirGLO vector transfected into DF-1 cells and treated with microRNA mimics. ....	202



Figure 5.10 Luciferase assay of cloned 3'UTRconstructs of wild type ITGA5 and mutated ITGA5 constructs in pmirGLO vector transfected into DF-1 cells and treated with the miR-3085-3p mimic or a scrambled non-targeting control.....204

## List of Tables

Table 1.1 Collagen types in Cartilage Matrix .....	28
Table 1.2 Known microRNAs associated with osteoarthritis from the HMDD v2.0 database.....	54
Table 3.1 Candidate microRNAs summary.....	105
Table 5.1 Top 10 genes from microarray with highest fold change increase in response to miR-6509-5p mimic compared to a non-targeting control.....	175
Table 5.2 Top 10 genes from microarray with highest fold change decrease in response to miR-6509-5p mimic compared to a non-targeting control.....	175
Table 5.3 Top 10 genes from microarray with highest fold change increase in response to miR-6509-5p inhibitor compared to a non-targeting control.....	176
Table 5.4 Top 10 genes from microarray with highest fold change decrease in response to miR-6509-5p inhibitor compared to a non-targeting control.....	176
Table 5.5 Top 10 genes from microarray with highest fold change increase in response to miR-664b-3p mimic compared to a non-targeting control. ....	177
Table 5.6 Top 10 genes from microarray with highest fold change decrease in response to miR-664b-3p mimic compared to a non-targeting control. ....	177
Table 5.7 Top 10 genes from microarray with highest fold change increase in response to miR-664b-3p inhibitor compared to a non-targeting control.....	178
Table 5.8 Top 10 genes from microarray with highest fold change decrease in response to miR-664b-3p inhibitor compared to a non-targeting control.....	178
Table 5.9 Top 10 genes from microarray with highest fold change increase in response to miR-3085-3p mimic compared to a non-targeting control. ....	179
Table 5.10 Top 10 genes from microarray with highest fold change decrease in response to miR-3085-3p mimic compared to a non-targeting control. ....	179
Table 5.11 Top 10 genes from microarray with highest fold change increase in response to miR-3085-3p inhibitor compared to a non-targeting control.....	180
Table 5.12 Top 10 genes from microarray with highest fold change decrease in response to miR-3085-3p inhibitor compared to a non-targeting control.....	180
Table 5.13 Computational gene target search of miR-6509-5p using RStudios.....	184
Table 5.14 Computational gene target search of miR-664b-3p using RStudios.....	185
Table 5.15 Computational gene target search of miR-3085-3p using RStudios.....	186

Table 5.16 Table of the candidate gene targets of miR-6509-5p represented in figure 5.5 (A) above the 25% cut off point.....	191
Table 5.17 Table of the candidate gene targets of miR-664b-3p represented in figure 5.5 (B) above the 12.5% cut off point.....	192
Table 5.18 Table of the candidate gene targets of miR-3085-3p represented in figure 5.5 (C) above the 12.5% cut off point.....	193
Table 5.19 List of candidate targets of miR-6509-5p, miR-664b-3p and miR-3085-3p to experimentally validate..	197

## **List of Appendices**

**Appendix 1** – qRT-PCR and probe sequences

**Appendix 2** – Oligonucleotide sequences of northern blot probes for candidate novel and known microRNAs

**Appendix 3** – Primer sequences for sub-cloning. DNA sequences of forward and reverse primers used for PCR analysis

**Appendix 4** – Selected candidate novel microRNAs identified from next generation sequencing

**Appendix 5** – Read number of most highly expressed 100 known microRNAs

**Appendix 6** – Assessment of candidate novel microRNA using northern blots

**Appendix 7** – Candidate microRNA response to lack of Dicer

**Appendix 8** – Candidate novel microRNA expression in osteoarthritic cartilage

**Appendix 9** – Candidate novel microRNA expression across a human tissue panel

**Appendix 10** – Computational target search of miR-6509-5p

**Appendix 11** – Computational target search of miR-664b-3p

**Appendix 12** – Computational target search of miR-3085-3p

## **Acknowledgements**

*Firstly I would like to thank my supervisor Professor Ian Clark for his endless guidance and support throughout my PhD and his enthusiasm for the project. I have really enjoyed my time with the Clark lab and feel very lucky to have been a part of such a fantastic group of people over the past 4 years.*

*I would also like to thank the rest of my supervisory team: Professor Tamas Dalmay, for his invaluable knowledge of microRNAs and helpful advice, and Dr Tracey Swingler who taught me a huge amount, from cell culture and cell transfections, to sequencing DNA and sub-cloning.*

*The members of the Clark and Riley labs, both past and present, have proven both exceedingly supportive colleagues and an amazing lab family throughout my PhD. They have been fountains of knowledge; teaching me the ropes of the lab, and are very good friends with whom I was able to bounce a multitude of ideas off. Thanks go to Darren Ebreo, Ellie Jones, Ginny Carmont, Heather Bircher, Janine Wilkinson, Jonny Green, Kirsten Legerlotz, Orla Jupp, Sarah Gardener and Ursula Rodgers. I would also like to particularly thank Rose Davison for initiating me in her boundless knowledge of qRT-PCR. Linh Le for her help with all things 'R' and teaching me the use of R Studios, and Yvette Wormstone for her aid in the complex world of sub-cloning and bacteria.*

*Many thanks go to Carly Turnbull and Adam Hall who taught me how to run northern blots and work with radioactivity safely. To Guy Wheeler for teaching me in situ hybridisation, to Helio Pais who provided invaluable help in the initial analysis of the next generation sequencing part of this project, and Karim Sorefan who worked with me to produce the small RNA library for next generation sequencing.*

*A vital part to the project came from the collaboration with the Norfolk and Norwich University Hospital, specifically Professor Simon Donell and Tracey Potter who were instrumental in providing the tissue samples for the project.*

*I would also like to thank members of the Musculoskeletal Research Group at Newcastle University: David Young and Steven Woods for their advice and ideas, and Matt Barter for the kind use of samples from his chondrogenesis assay.*

*I couldn't have completed my PhD without the amazing support and encouragement of my partner Katie, who is always there to keep my spirits up whenever things seem too daunting. I am also very grateful for the support of my family and friends throughout.*

*Finally, I would like to thank UEA for hosting my project, and those that provided funding for my PhD. Firstly Arthritis Research UK who kindly funded my consumables budget, stipend, and provided expenses to allow me to attend the prestigious OARSI conference in Philadelphia, USA. And secondly, the John and Pamela Salter Trust, which with their generous grant, enabled me to purchase microRNA LNA probe sets for qRT-PCR*

# Chapter 1 – Introduction

## **1.1 - Articular Cartilage And Its Properties**

Osteoarthritis (OA) is a painful, degenerative disease which can affect all synovial joints within the body. All joint tissues are involved and affected during OA, however the main structural change which generally characterises OA is the breakdown and degradation of articular cartilage at moving joints such as the knee and hip, common locations of osteoarthritis (Goldring and Goldring, 2006).

As we become an increasingly ageing population with people living much longer, the incidence of OA is rising. Arthritis Care reported in 2012 that around 8.5 million people in the UK currently suffer from osteoarthritis, with 70% of those in constant pain and they predicted that by 2030, the number of people diagnosed with OA will have risen to over 17 million (Arthritis-Care, 2012). The peak health of joints is thought to be reached by the age of 30, specifically the tensile strength apparent in the superficial zone, whilst the tensile strength of cartilage in the deep zone decreases progressively with age (Kempson, 1982), this may explain the common occurrence of OA within the elderly, however this is still not well understood. It is also thought that the risks associated with age and osteoarthritis are due to changes in TGF $\beta$  signalling resulting in reduced proteoglycan synthesis (van der Kraan et al., 2010).

There are three main types of cartilage (one of many connective tissues) within the human body. The first is elastic cartilage, found in the outer ear, the eustachian tube of the ear and the epiglottis at the back of the throat. This type of cartilage is composed of elastin proteins and collagen fibres, and is therefore more flexible than the other forms of cartilage. Fibrocartilage is present in the meniscus of the knee and between the vertebral discs. Fibrocartilage functions to withstand shear stress and appears to be similar in structure to both hyaline cartilage and dense fibrous tissue (Hu et al., 2001). It contains aggrecan in addition to type I collagen and type II collagen, however, the proportions of type II collagen and aggrecan present are substantially lower than in hyaline cartilage. The cells within fibrocartilage resemble chondrocytes, however, they are more spread out and fewer in number than those found in hyaline cartilage (Freemont and Hoyland, 2006). Finally there is hyaline cartilage, also known as articular cartilage. This type of cartilage appears between ribs and sternum, in the trachea and at the ends of long bones at moving joints such as the knee and the hip (Bhosale and Richardson, 2008; Carter, 2007). Hyaline, or articular, cartilage is the most common form of cartilage found in the body (Umlauf et al., 2010).



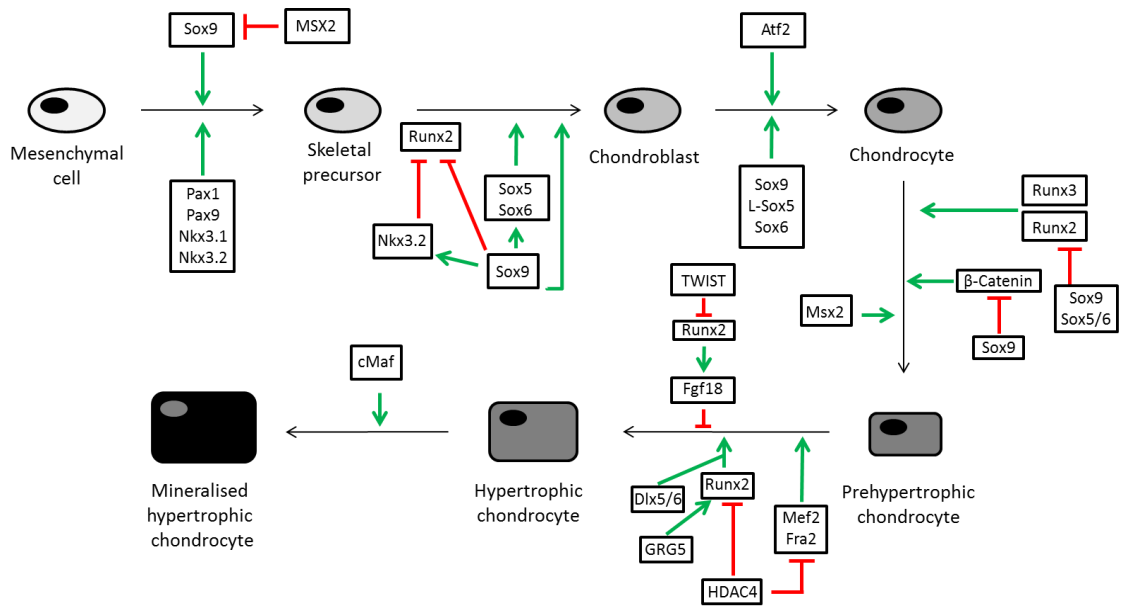
Articular cartilage develops through a process called chondrogenesis, where mesenchymal stem cells (MSC) progress through differentiation into the chondrocyte cells found in cartilage. This process is a vital part of endochondral ossification and the development of the long bones of the body.

Chondrogenesis begins with mesenchymal stem cells (originating from neural crest cells) migrating, proliferating and condensing as chondroprogenitor cells to form the cartilage anlagen scaffold. This involves cell-cell and cell-matrix interactions, and the production of type I and II collagen as well as hyaluronan (Olsen et al., 2000; Sandell et al., 1994). These chondroprogenitor cells develop, first into chondroblasts (small, flat, immature cartilage cells) and then finally to chondrocytes (mature cells) which are larger and more round in morphology than their immature counterparts. Chondrocytes progress through multiple differentiation events, sequentially: proliferation, hypertrophy and terminal differentiation. This is followed by mineralisation of the matrix, programmed cell death, and finally the invasion of osteoblasts and the formation of the long bone structures (Carter, 2007; Goldring et al., 2006; Umlauf et al., 2010).

During chondrogenesis, chondrocytes produce increasing amounts of extracellular matrix (ECM) molecules such as type II collagen and aggrecan (Carter, 2007). The chondrogenesis process can be defined specifically by the expression patterns of a number of ECM components. Firstly, type I collagen decreases as chondrogenesis progresses, whilst the expression of collagens II, IX and XI by chondrocytes begins to increase. Levels of other ECM components such as aggrecan, Gla protein (Luo et al., 1995) and link protein are also increased. Link protein is known to play a role in cross linkage and stabilisation between key ECM components of cartilage tissue. Furthermore, a fragment of link protein, known as LPP (Link protein N-terminal peptide) has been found to up-regulate the expression of the transcription factor SOX9 as well as cartilage components such as aggrecan and collagen type II in the matrix of cultured intervertebral discs (Wang et al., 2013).

A number of signalling molecules and pathways regulate chondrogenesis and endochondral ossification (Fig 1.1). Bone morphogenetic proteins (BMPs) and fibroblast growth factors (FGFs) have been linked to the control of chondrocyte proliferation during chondrogenesis (Minina et al., 2002). A number of Sox genes are also involved in chondrogenesis regulation. Sox9 is responsible for the expression of type II collagen, in addition to type XI collagen and cartilage-derived retinoic-acid-

sensitive protein (CD-RAP) by chondroprogenitor cells (Goldring et al., 2006). L-Sox5, Sox5 and Sox6 are further expressed alongside Sox9 during chondrocyte differentiation and regulate the production of aggrecan, type IX collagen, type II collagen and link protein (Ikeda et al., 2004; Smits et al., 2001). Pax1 and Pax9 (paired-box gene -1 and -9) are also involved in the induction of chondrogenesis, in part via their activation of Bapx1 [also known as NKX3-2 (NK3 Homeobox 2)] which in turn has been found to repress Runx2 (Runt-related transcription factor 2) and therefore enhance differentiation of chondrocytes (Lengner et al., 2005; Rodrigo et al., 2003). HDAC4 (Histone deacetylase 4) is also involved in the inhibition of Runx2 and subsequent prevention of chondrocyte hypertrophy (Vega et al., 2004).



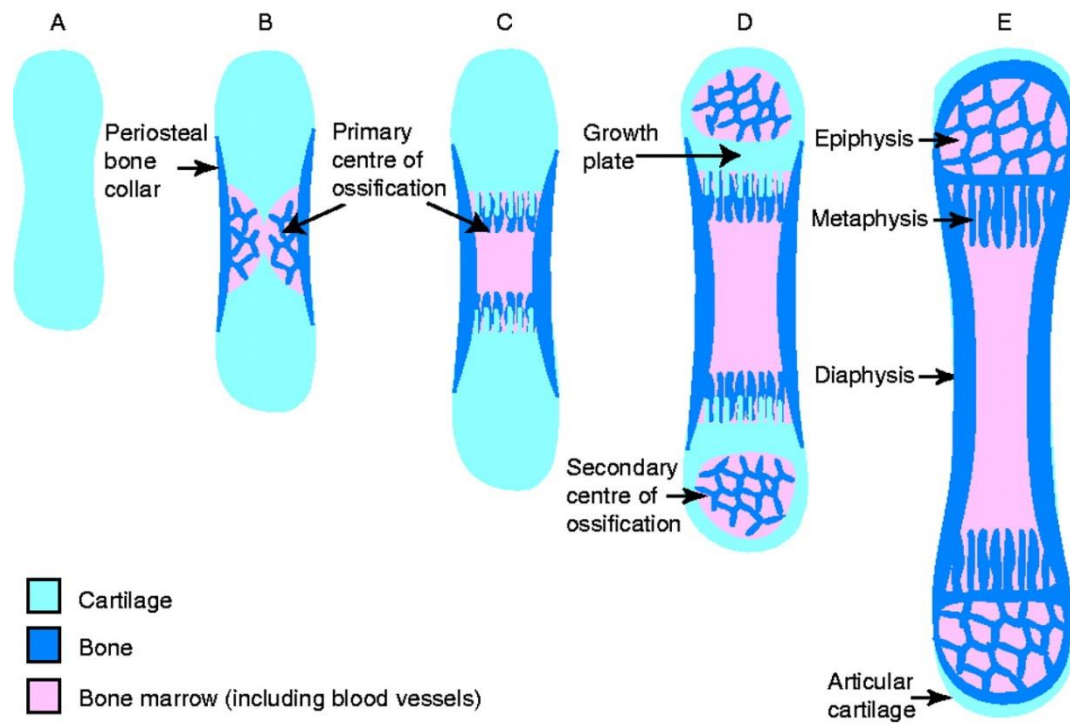
**Figure 1.1 Transcriptional regulation during chondrogenesis.** The regulation of a number of important transcription factors during chondrogenesis starting from a mesenchymal precursor cell. Green arrows represent positive factors, whilst red lines indicate negative factors. Diagram adapted from Hartmann (Hartmann, 2009).

During endochondral ossification (Fig. 1.2), cartilage tissue formed through chondrogenesis is replaced with bone tissue by osteoblasts. Mature chondrocytes firstly undergo hypertrophy in the centre of the prospective bone, and ossification is initiated by osteoblasts depositing a periosteal bone collar. Expression of Indian hedgehog (Ihh) is vital to the bone formation and is produced by chondrocytes prior to their hypertrophy (St-Jacques et al., 1999) along with Runx2 which is an important factor in regulating chondrocyte hypertrophy (Enomoto et al., 2000). Komari et al (Komori et al., 1997) show that the absence of Runx2 in mice results in the absence of chondrocyte hypertrophy.

The tissue is subsequently invaded by blood vessels, osteoclasts, osteoblasts and bone marrow precursor cells. Osteoclasts resorb the ECM of the cartilage tissue and bone tissue is deposited by osteoblasts in its place. Working in this manner, the expanse of bone tissue gradually grows outwards towards the ends of the cartilage model. In long bones, a growth plate of cartilage tissue can be seen between secondary ossification centres forming at the ends of the model. Within the growth plate, expression of Fibroblast growth factor receptor 3 (FGFR3) binding to the ligand FGF-18 is involved in inhibition of chondrocyte proliferation via the phosphorylation of Stat1, this results in an increase in p21 expression (Liu et al., 2002; Sahni et al., 1999).

The chondrocytes at the ends of long bones and their corresponding ECM remain, prevented from entering hypertrophy due to the induction of chondrocyte proliferation markers and repression of chondrocyte differentiation by Ihh (Indian hedgehog) and PTHrP (Parathyroid hormone-related protein) (Goldring et al., 2006). These chondrocytes form a layer of articular cartilage which is maintained by the chondrocytes themselves in a balance of production and degradation of ECM components (Carter, 2007; Mackie et al., 2011). The composition of the ECM in the areas of the model to be replaced by bone also changes; specifically as chondrocytes undergo hypertrophy, type X collagen is expressed (Umlauf et al., 2010).

The SMAD pathway (specifically SMAD3) is also responsible for maintaining the articular cartilage layer at the ends of long bones. Both TGF $\beta$  and SMAD3 inhibit the terminal hypertrophic differentiation of chondrocytes within the cartilage layer, therefore maintaining the tissue in its current state (Yang et al., 2001).



**Figure 1.2 Endochondral ossification.** Overview of endochondral ossification resulting in bone formation. (A) The cartilage model of future bone, (B) A periosteal bone collar forms and the primary centre of ossification begins to form, (C) The primary centre of ossification begins to expand towards the ends of the cartilage model. (D) The secondary centres of ossification form at each end of the bone, this leaves a cartilaginous growth plate between the primary and secondary ossification centres. (E) The mature bone has formed, with the growth plate cartilage being fully replaced by bone. The only cartilage which remains is a layer of articular cartilage at either end of the bone. (Mackie et al., 2011).

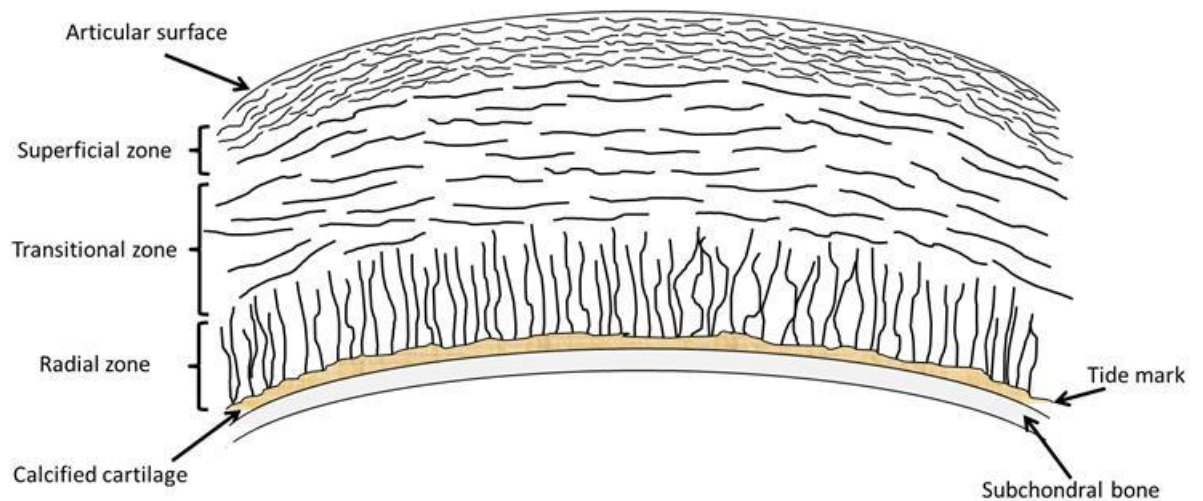
Articular cartilage as a tissue contains no nerve fibres or lymphatic vessels and is avascular. It uses very little oxygen for energy synthesis, the oxygen within the tissue ranges from <1% in the deeper zones to ~10% at the cartilage surface (Grimshaw and Mason, 2000). Instead articular cartilage relies on the Embden-Meyerhof-Parnas pathway of glycolysis to provide energy (Lee and Urban, 1997; Otte, 1991) via the proteins GLUT3 (Glucose transporter, type 3) and GLUT8 (Glucose transporter, type 8) (Mobasher et al., 2005). However, the presence of oxygen does appear to have a role within chondrocytes, with abnormal matrix production resulting from too little O<sub>2</sub> (Grimshaw and Mason, 2000; Zhou et al., 2004). It has been found in *in vitro* chondrocyte culture, that HIF-1 $\alpha$  (Hypoxia inducible factor 1 alpha) is up-regulated resulting in the induction in expression of numerous GLUT genes in response to the low oxygen levels. (Mobasher et al., 2005).

Articular cartilage is composed of four main zones (Fig. 1.3) and is made up of a combination of collagens (predominantly type II collagen, but also includes types I, III, VI, IX, X XI, XII, XIV and XXVII collagens), proteoglycans, specifically aggrecan; a large chondroitin sulphate proteoglycan, but also hyaluronan, syndecans, glypican, perlecan, epiphygan, decorin, lumican, fibromodulin, and biglycan. The glycoprotein lubricin is also found in cartilage, specifically the superficial zone and synovial fluid where it acts as a lubricant and functions to protect cartilage from friction-induced wear. The last component of cartilage is water (French, 2003; Rhee et al., 2005; Umlauf et al., 2010; Wardale and Duance, 1994).

The first zone of cartilage is the 'surface' or 'superficial' zone and is the thinnest of the four layers. It is in contact with the synovial fluid, and consists of fine collagen fibrils with chondrocytes aligned in parallel with the surface of the tissue (Martel-Pelletier et al., 2008); there is a low abundance of aggrecan at this level (Goldring and Marcu, 2009). The specific orientation of the collagen fibrils results in the tissue being able to withstand shearing forces placed on the joint by movement (French, 2003; Guilak et al., 1994). The centre layer is the 'middle zone' also known as the transitional zone. This thick zone consists of high concentrations of proteoglycans, such as aggrecan, and thick collagen fibres which are usually organised into layers or radial bundles. The chondrocytes at this level, unlike those at the surface layer, are more rounded and are present at lower concentration (Martel-Pelletier et al., 2008). The third layer is the deep zone (the radial zone) and this is closest to the bone layer. Out of the 3 zones, the deep zone contains the highest concentration of collagen fibres but less proteoglycans within

its extracellular matrix than the previous zone. The chondrocytes here are aligned perpendicularly to the surface layer, at a similar concentration to the middle layer (Martel-Pelletier et al., 2008; Umlauf et al., 2010). The very final zone, found immediately below the radial zone, is the calcified cartilage zone and is the area where the cartilage matrix interacts with the subchondral bone (Umlauf et al., 2010). It remains beyond the final closure of the growth plate and acts as a buffer between the bone and articular cartilage layers (Goldring and Marcu, 2009). All of the layers have distinctly different gene expression (Dunn et al., 2009).

Healthy articular cartilage functions to absorb impact to the joint, and cope with mechanical stresses such as shearing forces. The surface; or 'superficial' zone of cartilage is the area most responsible for managing mechanical strain and tensile forces placed upon the joint. The structure of the extracellular matrix (ECM) is in part the basis for the properties of articular cartilage in how it is able to withstand loads, and enable free movement of the joint, and consists mainly of aggrecan, collagen (in fibril form) and water. The degradation of articular cartilage ECM that occurs during osteoarthritis results in a loss of these functions and the development and progression of the disease (Martel-Pelletier et al., 2008).



**Figure 1.3 The zones of articular cartilage.** The superficial zone is the first zone beneath the articular surface and consists of fine collagen fibrils with chondrocytes aligned in parallel with the surface of the tissue. The centre layer is the transitional zone and contains thick collagen fibres which are usually organised into layers or radial bundles. The chondrocytes at this level are more rounded and are present at lower concentrations. The third layer is the radial zone and contains the highest concentration of collagen fibres. The chondrocytes here are aligned perpendicularly to the surface layer, at a similar concentration to the middle layer. The final zone, found immediately below the radial zone, is the calcified cartilage zone and interacts with the subchondral bone. Diagram adapted from Wiesler et al. (Martel-Pelletier et al., 2008; Umlauf et al., 2010; Wiesler et al., 2007).



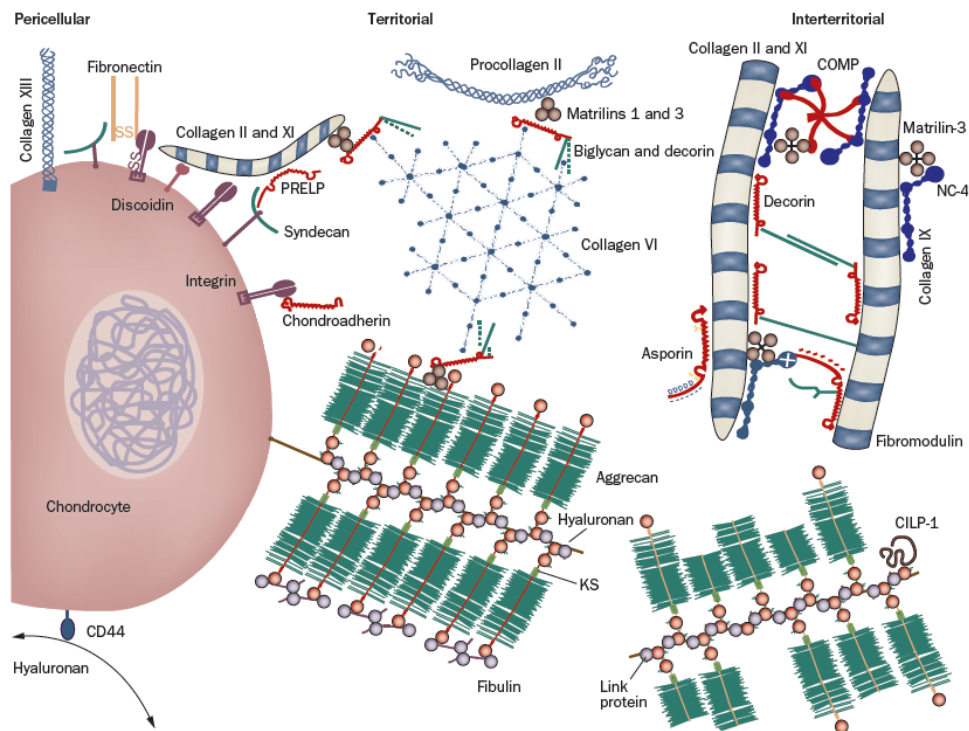
The only cell type present in cartilage, sparsely distributed, is the chondrocyte. Chondrocytes are found within chondrons, which are defined as the cell itself as well as the 1-5µm surrounding PCM (pericellular matrix). This in turn is surrounded by a fibrillar pericellular capsule (Poole, 1997). Chondrons are the mechanical units of hyaline cartilage and the pericellular matrix has been found to have an impact upon the mechanical properties of the chondrocyte, with chondrons thought to be responsible for absorbing mechanical load due to their ability to vertically compress when under strain and recover when unloaded (Nguyen et al., 2010; Poole, 1997).

Chondrocytes are dispersed in varying percentages between the zones of cartilage and often differ slightly in morphology between the layers. This is thought to be partially due to the mechanical environs found within each zone (Poole et al., 2007). It is these cells which (although they do not usually divide) produce extracellular matrix components such as collagens and proteoglycans and are responsible for maintaining this environment. Most matrix components are produced by early stage chondrocytes, with the main role of the mature cells being to maintain the homeostasis of this environment by interactions with integrin-mediated attachments (French, 2003). The matrix is also responsible for maintaining the phenotype of its surrounding chondrocytes (Aigner et al., 2006). The extracellular matrix, whilst mainly composed of collagen fibrils and proteoglycans, also contains a number of other molecules including cartilage oligomeric matrix protein (COMP) which acts as a catalyst in the formation of the fine collagen fibrils (fibrillogenesis) found in the matrix (Halasz et al., 2007) aided by perlecan (Kvist et al., 2006). Other components of the matrix include decorin, fibromodulin, biglycan and a number of matrilins (Goldring and Marcu, 2009). It is known that the maintenance and actual formation of articular cartilage matrix is greatly influenced by the interactions between these molecules, specifically type IX collagen, COMP and matrilin-3 (Leighton et al., 2007). These additional matrix components are also responsible for connecting different collagen types (such as VI and II) to each other and aggrecan (Wiberg et al., 2003).

The extracellular matrix is split into three matrices arranged in a specific order surrounding chondrocytes within the three main zones of cartilage. These are, in order of least distance from the chondrocytes: pericellular, territorial and interterritorial matrices (Fig. 1.4) (Heinegard and Saxne, 2011). Type II collagen, along with aggrecan, makes up the majority of the extracellular matrix of adult articular cartilage (Table 1.1)

and is involved in providing the tensile strength of the tissue as part of the interterritorial matrix (Martel-Pelletier et al., 2008). Specifically, compressive resistance is provided by aggrecan, a proteoglycan with a half-life of between 3 and 24 years (Goldring and Marcu, 2009). Type II collagen forms a framework for aggrecan which swells with water due to its negatively charged glycosaminoglycan side chains, creating large hydrodynamic domains. This combination of pressure restrained by a collagen grid gives cartilage its shape, tensile strength (Fosang and Beier, 2011) and elasticity (Aigner et al., 2006), and the proteoglycan network in turn protects the collagen network within the matrix. Collagen is known to have a half-life of over 100 years in cartilage if not exposed to deleterious conditions (Goldring and Marcu, 2009). The combination of the collagenous network and aggrecan is key to avoiding damage to cartilage during loading. They are able to dissipate the mechanical pressure resulting from loading and distribute it across a larger area (Aigner et al., 2006).

Type IX collagen is found at intersections within collagen fibrils and co-localizes with type II at the surface, it provides stability to the collagen fibrils (Muller-Glauser et al., 1986). Type XI collagen is also found throughout collagen fibrils however, unlike type II and IX, it cannot be detected by antibodies on intact fibrils, only at areas where fibril structure is incomplete or has been disrupted (Mendler et al., 1989). This is due to its function as a template for the fibril (Table 1.1). Both collagens type IX and XI are vital in disease processes, as mutations in either can result in chondrodysplasia (Eyre et al., 2006). All three collagens (II, IX and XI) are heavily involved in cartilage formation and form what is known as a heteropolymeric template (Eyre et al., 2006). In contrast to the interterritorial zone, the territorial matrix contains collagen type VI but very little fibrillar collagen, and is located closer to the chondrocyte cells. Collagen XXVII is also a member of the fibrillar collagen family and has a role in maintaining stability within the matrix (Goldring and Marcu, 2009). More recently it has been found that type XXVII Collagen plays a vital role in the structure of the pericellular extracellular matrix of the growth plate and is also necessary for the organisation of the proliferative zone of cartilage (Plumb et al., 2011).



**Figure 1.4 Matrices formation within cartilage.** Diagram of the placement of matrices within cartilage in relation to the chondrocyte [figure from (Heinegard and Saxne, 2011)]. Pericellular is adjacent to the cell and is primarily the location where cell signalling with cell surface receptors occurs, one example of this is of the glycosaminoglycan hyaluronan binding to the CD44 receptor. The subsequent matrix zone (territorial) is comprised predominantly of aggrecan HA (hyaluronan) complexes and collagen VI whilst the zone furthest from the cell surface (interterritorial) contains mainly types II and XI collagen. Abbreviations: CILP-1 (cartilage intermediate layer protein 1), COMP (cartilage oligomeric matrix protein), and KS (keratan sulphate) (Heinegard and Saxne, 2011).

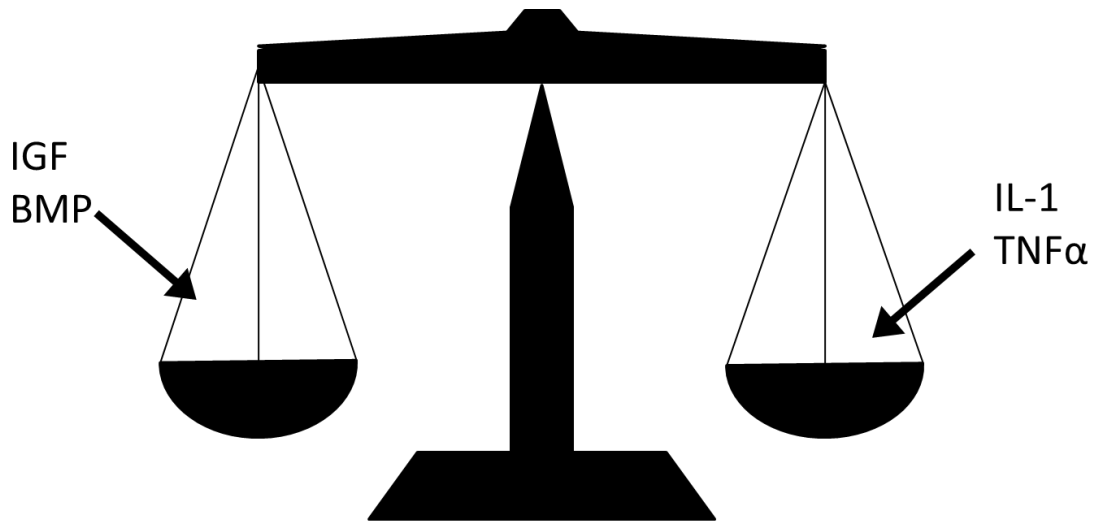
<b>Collagen Type</b>	<b>Percentage of total collagen</b>
Type I	Present in adult articular surface and pericellular
Type II	75% foetal, >90% adult
Type III	>10% in adult human articular cartilage
Type IX	Covalently fibril-associated collagen 10% foetal, 3% adult
Type X	Hypertrophic cartilage only
Type XI	Fibril template 10% foetal, 3% adult
Type VI	Chondron basket, microfilaments <1%
Type XII/XIV	Non-covalently fibril-associated collagens
Type XIII	Transmembrane
Type XXVII	Fibrillar collagen, role in maintaining stability

**Table 1.1 Collagen types in Cartilage Matrix** [table adapted from (Eyre et al., 2006; Plumb et al., 2011; Wardale and Duance, 1994)].

In healthy cartilage, equilibrium between the production and destruction of ECM components is maintained by the chondrocytes present and involves a number of both anabolic and catabolic factors (Fig. 1.5). Those involved in maintaining and producing ECM components include TIMPs, IGF-1 and members of the TGF-beta family. Whilst from the catabolic side, inflammatory cytokines, MMPs and aggrecanases are responsible for ECM breakdown (Aigner et al., 2006). If the balance between ECM synthesis and degradation becomes disrupted and begins to fall more on the catabolic side, as in OA, then the cartilage begins to degrade (Martel-Pelletier et al., 2008).

The release of proteins and other products from the breakdown of cartilage can further induce enzyme release from the surrounding macrophages and synovial cells via TLR (Toll-like receptor) activation; it can also result in other inflammatory events such as vascular hyperplasia in the neighbouring synovial membrane and invasion of immune cells such as B lymphocytes (Da et al., 2007). These processes can ultimately lead to positive feedback, with molecules such as tumour necrosis factor  $\alpha$  and interleukin-1 $\beta$  becoming induced and subsequently accentuating the degradation effect (Da et al., 2007).

Physical trauma can also result in inflammation in the early stages of OA, a symptom which was once thought to be a feature solely of rheumatoid arthritis (Saxne et al., 2003). Inflammation, through the production of cytokines can result in a switch to more catabolic processes by chondrocytes and other cells within the affected joint. As mentioned earlier this can cause more tissue damage and follow the increasing cycle of degradation, as well as meaning an increase in sensitivity of the joint to mechanical loading. Inflammation is also thought to be an important factor in the debilitating pain that most if not all patients of OA suffer (Heinegard and Saxne, 2011; Sellam and Berenbaum, 2010). However, the use of anti-inflammatory drugs often does not give the patient long-term respite against the pain, suggesting that OA may not be inflammatory in quite the same way as rheumatoid arthritis (Lin et al., 2004).



## Anabolism

Aggrecan  
 Collagen type VI  
 Collagen type IX  
 Link protein

## Catabolism

MMP-2,9 ADAMTS-1  
 MMP-3 ADAMTS-4  
 MMP-13 ADAMTS-5  
 MMP-14

**Figure 1.5 The balance of anabolic and catabolic factors in healthy cartilage.** In healthy tissue, anabolic and catabolic factors are in balance. All these factors are primarily produced by chondrocytes during autocrine and paracrine signalling. And it is when these factors become imbalanced that osteoarthritis occurs. (BMP: bone morphogenetic protein; IGF: insulin-like growth factor; IL-1 $\beta$ : interleukin 1 $\beta$ ; TNF: tumour necrosis factor). [Image adapted from (Aigner et al., 2006)].

## 1.2 - Osteoarthritis

Osteoarthritis is not just a disease of the cartilage, although this is usually the major area affected, and is used clinically to diagnose the disease. It is a disease encompassing the whole joint including the synovial tissue, tendons and surrounding bone. Specifically in OA there is a tendency for an increase in the thickness and remodelling of the subchondral plate beneath joint cartilage to occur. Symptoms also include other structural changes such as the overgrowth of bone (osteophytes) at the joint, which can press on nerve endings (Martel-Pelletier et al., 2008). The growth of osteophytes is thought by some to be an attempt by the body to repair the damage to the worn joint, but in effect it seems to lead to a misshapen joint and result in increased pain as well as adversely affecting mobility (Swales, 2010).

Although some attempt is seen by chondrocytes to repair cartilage matrix in early stage OA with an initial increased expression of type II collagen and aggrecan, this is also paired with increased production of matrix degrading enzymes (Goldring, 2000). The overall reason behind the poor ability of cartilage to repair itself may be the slow turnover of new cartilage due to mature chondrocytes functioning more to maintain the articular cartilage tissue rather than rapidly produce new matrix (Eyre et al., 2006). Cartilage tissue is also an avascular tissue and so has no ready access to a blood supply, hindering its repair (Oldershaw et al., 2010). Interestingly, it has been found that specific mouse models can rapidly repair damaged cartilage *in vivo*. Eltawil et al found that when DBA/1 and C57BL/6 mice underwent microsurgery in order to damage knee cartilage, mice of the DBA/1 strain showed a faster and more consistent repair of the tissue, whilst the C57BL/6 strain developed osteoarthritic characteristics. This shows a potential strain dependent outcome to induced osteoarthritis. Eltawil et al also showed this repair to be age dependent, with 8-week old DBA/1 mice showing more efficient repair than 8-month old DBA/1 which had undergone the same procedure (Eltawil et al., 2009).

Other symptoms, particularly of knee osteoarthritis, include, but are not limited to, structural changes within the menisci and cruciate ligaments (Englund et al., 2010). And at later stages of the disease, the structural changes increase further in severity, and chronic inflammation of the synovial membrane occurs (Martel-Pelletier et al., 2008). Some levels of synovitis are also thought to occur in early osteoarthritis but not to the same extensive degree; however Benito and colleagues saw an increase in inflammatory

markers at early stages of OA compared to the latter, advanced, stages (Benito et al., 2005). An inflamed synovial membrane is characterised by hypertrophy and hyperplasia of the membrane as well as the occurrence of lymphocytes invading the sub-lining tissue of the synovium (Benito et al., 2005). This was also seen in studies by Rollin et al who studied the expression of CD5 and CD69 and found an influx of lymphocytes into the synovium, however, it was reduced when compared to lymphocyte invasion in RA tissue (Rollin et al., 2008). The synovial inflammation triggers the release of more inflammatory mediators which further alters the homeostasis of the ECM in cartilage, accelerating degradation (Martel-Pelletier et al., 2008; Sellam and Berenbaum, 2010).

One of the main obstacles to dissecting the aetiology of this disease is the lack of tissue readily available from early stages of osteoarthritis to study. This is due to the fact that the clinical symptoms that diagnose OA, such as radiographic changes in cartilage thickness, generally only appear in the latter stages (Heinegard and Saxne, 2011; Rousseau and Garnero, 2012). Also, the radiographic scores do not always correspond to what is classed as the predominant clinical representation of OA: patient defined joint pain (Thakur et al., 2013). Rarely are early stage human OA tissue samples available to study since joints are only removed at the end point of the disease. However, a large area of arthritis research now lies around early stage markers to better diagnose the disease (Rousseau and Garnero, 2012). In addition to the problems of attaining early stage osteoarthritic cartilage, samples used experimentally as controls usually originate from trauma patients with fractured neck of femur. Generally this tissue is sourced from the elderly or those already predisposed to fragile bones such as those with osteoporosis. This means that although the tissue is non-arthritic, it is far from 'normal'. The ideal control would be tissue obtained via amputation, however obviously this tissue is harder to obtain and comes with its own problems, such as if the joint has been aberrantly loaded in its lifetime, and how long the joint has spent in an unloaded condition before the tissue is processed.

The progression and initiation of OA is still not completely understood and is compounded by the growing theory that there are in fact many subtypes of osteoarthritis and not, as once thought, merely primary and secondary OA. These subtypes are being defined by a number of factors, for example, joint site, presence of inflammation, structural aspects and the threshold of pain suffered by a patient (Bierma-Zeinstra and Verhagen, 2011).



Osteoarthritis is thought to be the result of a combination of many different factors including mechanical stress and enzymatic feedback loops (Martel-Pelletier et al., 2008) as well as a combination of both environmental and genetic factors. It has also been shown that OA can be more prevalent in certain joints in women more than men and vice versa, with one example being knee OA more prevalent in females (Dieppe and Lohmander, 2005). Most recently, initiation factors of OA are thought to include mutations within matrix proteins such as type II collagen (Jakkula et al., 2005; Kuivaniemi et al., 1997; Metsaranta et al., 1992), type IX collagen (Dreier et al., 2008; Fassler et al., 1994; Jakkula et al., 2005), and XI Collagen which correlate with an increased incidence of early onset primary osteoarthritis (Jakkula et al., 2005).

There have been a number of genome-wide association studies (GWAS) carried out to investigate osteoarthritis genetically. Although not totally successful in finding a conclusive genetic cause of OA, they have suggested a number of interesting genes with possible links to osteoarthritis. These include *DVWA*, a gene expressed specifically in cartilage (Miyamoto et al., 2008), and the *IL-1* gene cluster at 2q13 which has been associated with susceptibility to knee osteoarthritis (Loughlin et al., 2002). Other GWAS studies [including arcOGEN (Zeggini et al., 2012) and the TREAT-OA consortium (Kerkhof et al., 2011)] have identified other potential common loci for susceptibility to osteoarthritis such as *GLT8D1* (glycosyltransferase 8 domain containing 1), *ASTN2* (Astrotactin 2), *CHST1L* and *TP63*. These studies focussed on those patients who had joints replaced due to osteoarthritis and so the loci were also considered to be linked to the pain of OA sufferers since pain is the major reason behind replacement surgery. Although a large number of genes were identified within each GWAS, each observed only a small effect size, possibly suggesting a synergistic effect (Thakur et al., 2013).

Those most at risk of OA are thought to be the elderly, overweight or obese, and patients of trauma incidents (Attur et al., 2010), with incidents of osteoporosis thought to increase the chance of OA (Bellido et al., 2010). The link between osteoporosis leading to osteoarthritis, is however, refuted by a number of studies including one case study where patients with fractured hips were found to be at risk of getting hip OA in the same joint by one third compared to controls (Franklin et al., 2010).

The increased incidence, particularly seen in those overweight or obese patients, was traditionally thought to be due to the excess weight imposed upon the joints. However, there has been literature published which suggests that it is not necessarily the added weight, but the fat tissue itself which is responsible, or more precisely, inflammatory factors such as adipokines released from the adipose tissue (Aspden, 2011; Hotamisligil, 2006). The influence of adipose tissue upon OA is shown in part by the fact that obese patients also exhibit increased OA in finger joints which would not be affected by a larger mechanical load (Aspden, 2011). This is supported by the location and structure of adipocytes in relation to immune cells in that they are situated close to macrophages and can provide interactions between the immune system and metabolic responses (Hotamisligil, 2006). The infrapatellar fat pad is thought to be a potential source of adipokines in the knee and one study has suggested that cytokines, such as IL-6, produced from the infrapatellar fat pad in OA patients could be involved in inflammation in addition to being a factor in causing cartilage damage (Distel et al., 2009).

There is also thought to be a link between osteoarthritis and metabolic syndromes, to such an extent that metabolic osteoarthritis has now been described as a subtype of osteoarthritis. For example, hypertension has been linked to bone remodelling in OA due to its effect on the exchange of nutrients into articular cartilage, and dyslipidemia is thought to play a role in the initiation of OA by the dysregulation of lipid metabolism within the joint (Zhuo et al., 2012). Diabetes is also thought to be an additional independent risk factor of OA (Berenbaum, 2012), with those suffering from type 2 diabetes found to be at a much higher risk of developing severe OA than their non-diabetic counterparts (Schett et al., 2013). There has however been more recent research finding a more inconclusive link between OA and diabetes, suggesting that they are not causally related based on genetic risk scores for the diseases in a tested population (Hoeven et al., 2014). This could suggest instead that the link seen is due to common risk factors between the diseases such as increased weight and age.

There is also the possibility of cartilage degradation being aggravated, and the disease progressed, by an autoimmune response to the cartilage components (Nishioka, 2004). The concentrated presence of immune cells within the synovial tissue are thought to be involved with disease progression and development with one study showing an infiltration of B lymphocytes in half of the OA knee samples tested associated with an

increase in synovial inflammation (Da et al., 2007). There has also been shown to be a T-cell immune response to aggrecan specific peptides in osteoarthritis and rheumatoid arthritis patients (de Jong et al., 2010; Sakkas and Platsoucas, 2007), associated with an up-regulation in the cytokines and tumour necrosis factors (particularly IL- $\beta$  and TNF- $\alpha$  respectively) (Kobayashi et al., 2005; Malemud, 2010), and it is thought that these factors may have a role in the degradation of articular cartilage matrix proteins, as well as causing damage to subchondral bone. However this is still controversial as human trials utilising antagonists to these cytokines have so far proved unsuccessful (Malemud, 2010). Additional research has suggested a role for the innate immune response in osteoarthritis, with a focus on TLR (Toll-like receptors) interacting with hyaluronan within the joint (Scanzello et al., 2008).

Treatments for osteoarthritis are limited and are generally restricted to analgesics, non-steroidal anti-inflammatory drugs (NSAIDs) and corticosteroid injections (Arroll and Goodyear-Smith, 2004), although trials have shown that both anti-inflammatories and injections have a limited effect on pain from OA (Lin et al., 2004). There are as yet no effective DMOADs (disease modifying osteoarthritis drugs) and the last resort for OA patients is surgery with either partial (uni-compartmental) or full joint replacements. A knee replacement, for example, can last up to 20 years, but must eventually be replaced, especially in a younger patient (NHS, accessed 03/04/14). Alternative treatments include attempts to alter a patient's gait to treat the disease (Fregly, 2012).

There have also been investigations into the use of the compounds glucosamine sulphate and chondroitin sulphate as treatments for OA to use as both pain relief and to treat the structural symptoms of cartilage destruction. There is however differing reports as to the effectiveness of these treatments, with some studies finding no physiologically significant effect (Sawitzke et al., 2010), and others showing a significant reduction in OA symptoms when treated with either chondroitin sulphate or glucosamine (Martin et al., 2012; Reginster et al., 2012; Wildi et al., 2011).

### **1.3 - Regulatory Mechanisms Of Osteoarthritis**

Although it is observed that there is unlikely to be one sole cause of OA, much investigation has gone into mapping the various changes that occur during the disease. Primarily osteoarthritis develops when there is some form of dysregulation in the balance between matrix production and degradation.

Cartilage degradation is a key factor of osteoarthritis and usually begins with a loss of proteoglycans such as aggrecan due to the up-regulation of the aggrecanases; ADAMTS-4 and -5 (a disintegrin and metalloprotease domain with thrombospondin motifs). In the past, studies have shown that the degradation also corresponds with an up-regulation of MMPs (matrix metalloproteinases) which cleave aggrecan, although at different sites to the ADAMTS family (Nagase and Kashiwagi, 2003; Umlauf et al., 2010). However, more recent publications have noted that ADAMTSs are much more efficient at aggrecan degradation, meaning MMPs may have a smaller role in *in vivo* aggrecan degradation than once thought (Durigova et al., 2011). This theory could however be explained by research employing the use of neoepitope antibodies against aggrecan cleavage by both MMPs and aggrecanases, showing that MMPs are responsible for aggrecan cleavage at later stages of the disease compared with aggrecanases which appear responsible for the early stage loss of the protein (Fosang et al., 2010).

Aggrecan is responsible for the weight bearing properties of cartilage, and its progressive destruction results in structural instability of the ECM. There has been much research into which of the two aggrecanases: ADAMTS-4 and ADAMTS-5, is the key aggrecanase in osteoarthritis progression (Umlauf et al., 2010) and which specifically is most important in terms of that regulation. Although both are responsible for aggrecan degradation, ADAMTS-4 is stimulated by pro-inflammatory cytokines, and both ADAMTS-4 and 5 are expressed by chondrocytes. It is thought that ADAMTS-5 is the most active out of the two aggrecanases as seen from *in vitro* cartilage experiments, with activity 1000-fold higher than ADAMTS-4 (Gendron et al., 2007; Troeberg and Nagase, 2012). In mouse studies, ADAMTS-5 also appears to have the greatest effect in osteoarthritis. This is shown by ADAMTS-4 knock-out mice with surgically induced joint instability, which showed no halt in OA progression (Glasson et al., 2004), whereas ADAMTS-5 null mice show an abrogation of cartilage destruction in an inflammatory arthritis mouse model (Stanton et al., 2005).

It has been found that degradation of collagen fibrils begins after the loss of aggrecan, due in part to the theory that the presence of aggrecan in the matrix protects the collagen fibrils (Pratta et al., 2003; Troeberg and Nagase, 2012) and so collagen degradation is prevented whilst aggrecan is still intact.

Collagen degradation is due to the presence of collagenases such as MMP-13 which is known to be involved in maintaining the matrix equilibrium of cartilage as well as being implicated in chondrocyte differentiation (Goldring and Marcu, 2009; Troeberg and Nagase, 2012; Wu et al., 2002). Their involvement in osteoarthritis is shown by up-regulated levels of collagenase matrix metalloproteinases 1, 8 and 13 cleaving type II collagen in OA human cartilage compared to normal cartilage tissue (Billinghurst et al., 1997). This is combined with increased expression of MMP-3, MMP-10, MMP-11, MMP-2 and MMP-9 in osteoarthritic tissue (Umlauf et al., 2010). MMP-2 and MMP-14 have also been found to cleave collagen, although MMP-13 appears to be the predominant MMP with the highest specific activity against collagen involved in osteoarthritis (Kevorkian et al., 2004).

Collagen degradation can be partially inhibited by the introduction of MMP-8 and 13 inhibitors (Dahlberg et al., 2000; Piecha et al., 2010; Poole et al., 2002). This is supported by data from an MMP-13 knock out mouse model surgically induced through osteoarthritis which was found to have reduced cartilage destruction compared to a wild type control, suggesting that MMP-13 is key in the progression of osteoarthritis (Little et al., 2009). In addition, there is also evidence showing the effectiveness of MMP-13 inhibitors in the prevention of cartilage destruction in an osteoarthritic rat model (surgical induction of medial meniscus tear) (Baragi et al., 2009) as well as a decrease in collagen degradation recorded in articular cartilage from both human and bovine origins in response to MMP-13 inhibitors (Piecha et al., 2010).

Other factors, in addition to the aforementioned collagenases and aggrecanases, are also known to influence chondrocyte function during osteoarthritis. Pro-inflammatory cytokines such as IL-1 $\beta$  (interleukin 1 beta) and TNF- $\alpha$  (tumour necrosis factor alpha) are heavily involved in altering chondrocyte metabolism (as well as their phenotype) and therefore have a role in affecting matrix production and structure by activating catabolic pathways. They are known to induce the production of MMPs, including

MMP-1, MMP-3 and MMP-13, iNOS (inducible nitric oxide synthase) and COX-2 (cyclooxygenase 2), in addition to inhibiting TIMP production which are known MMP inhibitors (Umlauf et al., 2010). Both IL-1 $\beta$  and TNF- $\alpha$  are also known to activate a number of signalling pathways within chondrocytes, including JNK (c-Jun N-terminal kinases), p38, MAPK (Mitogen-activated protein kinases) and NF- $\kappa$ B (Nuclear factor kappa-light-chain-enhancer of activated B cells) (Goldring et al., 2011). MAPK, p38 and NF- $\kappa$ B are known to have roles within chondrocyte differentiation, with p38 showing expression in the growth plate of long bones (Bobick and Kulyk, 2008; Zhang et al., 2006b) and MAPK and NF- $\kappa$ B signalling implicated in MMP and ADAMTS induction (Goldring, 2012). The JNK pathway has been implicated in stress responses in chondrocytes and Kao et al describes its role in chondrocyte apoptosis after treatment with nitric oxide. They demonstrate a reduction in cell death with JNK inhibitor treatment (Kao et al., 2013).

Further, TNF- $\alpha$  has been implicated in disease as well as normal chondrocyte physiology, with research involving an inhibition of TNF- $\alpha$  showing a suppression in cartilage degradation in arthritis (Fernandes et al., 2002).

Additional cytokines and chemokines implicated in osteoarthritis include LIF (Leukaemia inhibitory factor), IL-6, IL-8, IL-15, IL-17, IL-18 and IL-21 (Kapoor et al., 2011). IL-6 has been found to up-regulate MMP-1 and MMP13 (Cawston et al., 1998; Rowan et al., 2001) and is detected in the synovial fluid of OA joints (Kaneko et al., 2000). IL-15 and IL-21 have also been detected at increased levels in OA synovial fluid (Scanzello et al., 2009). IL-18 on the other hand is expressed at higher levels in osteoarthritic chondrocytes compared to non-diseased cells and also appears to up-regulate the synthesis of MMPs as well as decrease expression of inhibitors of MMPs (Sandell and Aigner, 2001). IL-17 is a pro-inflammatory cytokine and has been found to down-regulate the levels of proteoglycans in OA (Lubberts et al., 2000), in addition to inducing the up-regulation of MMPs (Martel-Pelletier et al., 1999). The chemokine IL-8 has shown increased expression in OA tissue and is also thought to be involved in proteoglycan degradation (Alaaeddine et al., 1999). Finally LIF is another cytokine linked to OA, and has been detected at elevated levels in the synovial membrane and fluid of osteoarthritic joints, as well as being implicated in proteoglycan degradation (Alaaeddine et al., 1999).

A review by Tchetina *et al* comments on how some of the characteristic changes in OA that are shown by the cartilage cells themselves, resemble the changes that occur in development of the bone (endochondral ossification) (Tchetina, 2011). These include, for example, the expression of genes related to differentiation such as type X collagen which was found to be expressed more highly by OA chondrocytes than within a control and is a marker of hypertrophy in normal cell physiology (Girkontaite et al., 1996; von der Mark et al., 1992; Walker et al., 1995).

On a cellular level it has been found that apoptosis of terminally differentiated chondrocytes is also found in OA (Blanco et al., 1998), with up-regulated expression of Fas and Caspase 8 in OA compared to normal cartilage tissue (Robertson et al., 2006). Another paper noted in an animal OA model, the presence of apoptosis in the early stages of OA in both the superficial and middle zones of cartilage. This is thought to be due to the ongoing mechanical damage occurring in the joint. At later stages of the disease chondrocytes appear to undergo autophagy as well as apoptosis in the same two zones, shown by high levels of caspase 3 and LC3III within the chondrocytes. In the deep zone, only apoptosis is apparent, thought to be because of ossification of the subchondral bone (Almonte-Becerril et al., 2010).

## 1.4 – Biosynthesis Of MicroRNAs

MicroRNAs are small (20-22mer) double stranded RNAs. Other known small RNAs include repeat-associated short interfering RNAs (rasiRNAs), short interfering RNAs (siRNAs), piwi-interacting RNAs (piRNA), small nucleolar RNAs (snoRNA), transfer RNAs (T-RNAs), Y-RNAs, mirtrons, and endogenous short-hairpin RNAs (endo-shRNAs) (Kato et al., 2009; Kim et al., 2009; Meister and Tuschl, 2004; Nicolas et al., 2012).

These short RNAs, along with RNA silencing machinery are responsible for the suppression of target gene activity. In fact, the machinery involved was first discovered due to its role in the defence against RNA viruses (Waterhouse et al., 2001) as well as a preventative measure against the random integration of transposable elements into the genome.

PiRNAs are thought to play a key role in transposon silencing in the germline via their interaction with Piwi proteins (part of the Argonaute protein family) (Girard et al., 2006; Kato et al., 2009). RasiRNA are also thought to interact with Piwi proteins and are involved in the RNA interference pathway where they are responsible for silencing endogenous selfish genetic elements, for example repetitive sequences (Vagin et al., 2006).

siRNAs function in a similar manner to microRNAs, and they both undergo a similar biogenesis, however, whilst miRNAs are known to regulate endogenous genes, siRNAs are responsible for defending the genome against foreign nucleic acids, for instance viruses (Carthew and Sontheimer, 2009).

SnoRNAs exist in one of two groups, and are known to modify ribosomal RNA or aid in rRNA processing. The groups are known as box C/D and box H/ACA. Box C/D snoRNAs are mainly involved in methylating the 2'-hydroxyl groups of target nucleotides whilst the box H/ACA snoRNAs are responsible for converting uridine to pseudouridine (Kiss, 2002; Ono et al., 2010). Some snoRNAs are also thought to act as precursors for microRNAs (Scott et al., 2009). The small RNAs known as T-RNAs are involved in protein synthesis via their association with the ribosome complex (Fischer et al., 2010). Recently, Haussecker et al have identified a new form of small RNA derived from tRNA now known as tsRNA (tRNA-derived small RNAs). It is thought



that tsRNAs are found within the cytoplasm, associated with Argonaute proteins, and are involved in small RNA silencing, potentially having a regulatory role upon microRNA function (Haussecker et al., 2010).

Y-RNAs are involved in DNA replication. These Y-RNAs are longer, non-coding RNAs of around 100 nucleotides and have been linked to apoptosis, they bind to Ro60 and La (Nicolas et al., 2012). Mirtrons, on the other hand are thought to be a type of microRNA, or more specifically a new form of microRNA precursor and are generally found within the spliced out introns of host genes. These small RNAs follow an alternate route of biogenesis, utilising splicing rather than cleavage by the Drosha enzyme found in microRNA synthesis (see below). They do however function as microRNAs by regulating the expression of target genes (Westholm and Lai, 2011). Endogenous short-hairpin RNAs (endo-shRNAs) are thought to be another potential form of precursor for microRNAs, forming from Dicer cleavage of long hairpins (Babiarz et al., 2008; Matera et al., 2007).

MicroRNAs are responsible for the regulation of their target mRNA counterparts, and through this they are able to control gene expression. It is thought that they are involved in the translational regulation of at least 60% of all human genes (Akbari Moqadam et al., 2013; Ambros, 2004; Bartel, 2009; Friedman et al., 2009). There are over 2,500 published mature microRNA sequences described on the online database miRBase in humans alone, with over 30,000 mature sequences arising from more than 24,000 microRNA precursors across all species (Griffiths-Jones et al., 2006; Griffiths-Jones et al., 2008; Kozomara and Griffiths-Jones, 2011), proving that they are an important factor in genetics. There are many techniques utilised to establish microRNA function, including knock out mouse models, and to date there are around 50 published mouse lines of knockout microRNAs (Park et al., 2012).

The microRNA synthesis process (Fig. 1.6) begins with the transcription of the miRNA gene by RNA polymerase II and III in the nucleus into a long primary transcript (Borchert et al., 2006). MicroRNAs are generally transcribed from introns within other genes and are known as intronic. They can be co-transcribed with that gene, although some also possess their own promoters, and it is thought that they can be processed from both pre- and post- spliced introns (Kim and Kim, 2007; Ozsolak et al., 2008). Other microRNAs however have been found between gene transcripts (more than 1kb

away from known or predicted gene sequences) and are transcribed independently from their own promoters (Ozsolak et al., 2008). These are known as intergenic. (Hon and Zhang, 2007; Kim, 2005).

The RNase III enzyme Drosha combined with the double stranded RNA binding domain (dsRBD) protein DGCR8, cleaves the double stranded miRNA primary transcript (pri-miRNA) resulting in a stem loop precursor miRNA (pre-miRNA) (Gregory et al., 2004; Lee et al., 2003) with two overhangs, at the 5' end; a phosphate, and at the 3' end; a 2 nucleotide long segment (Borchert et al., 2006). It is known that Drosha can be regulated by certain transcription factors (such as Smad and p53), resulting in the maturation of specific microRNAs due to increased recruitment of the corresponding pri-miRNAs to the Drosha complex (Treiber et al., 2012).

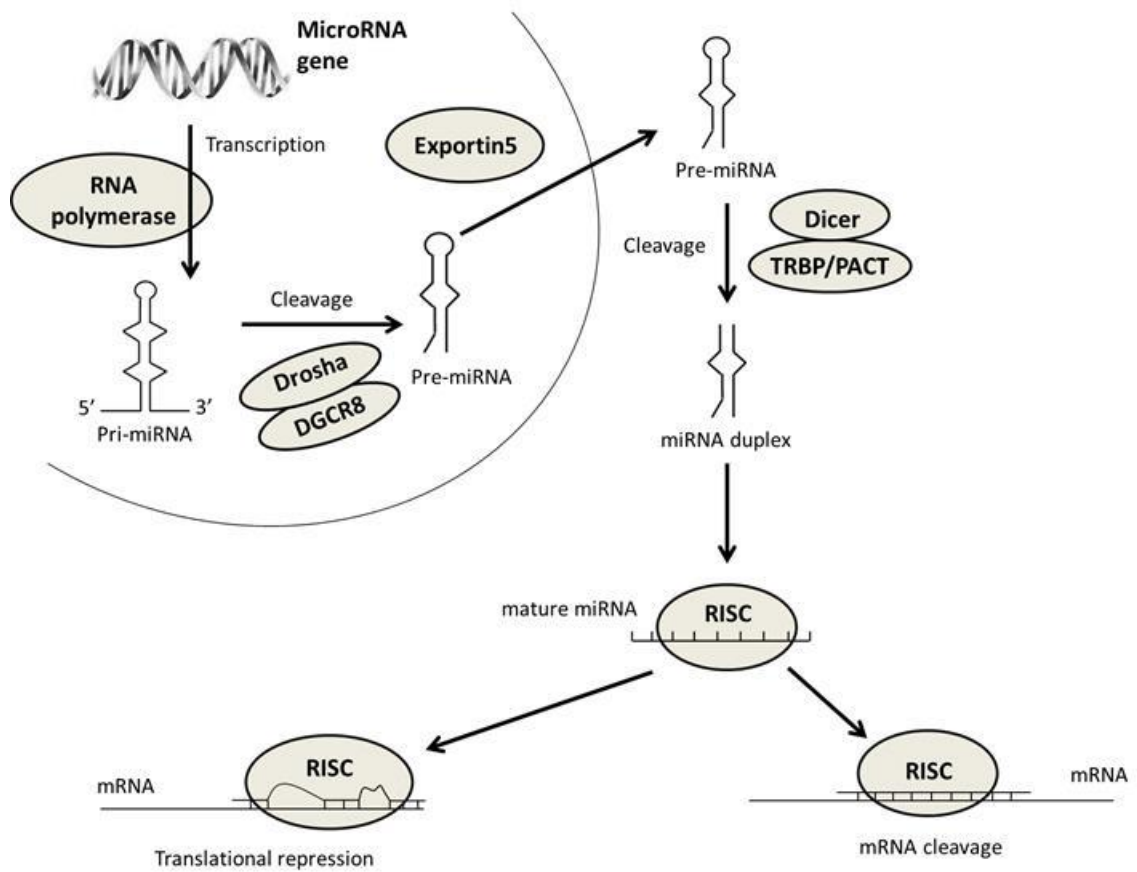
The precursor is subsequently transported by exportin-5 (in the presence of Ran-GTP cofactor), out of the nucleus and into the cytoplasm (Bohnsack et al., 2004; Yi et al., 2003). The pre-miRNA is then processed by Dicer (an endonuclease) to create a double stranded 20-22 nucleotide long RNA duplex intermediate, which has the same overhangs as the pre-miRNA (Zhang et al., 2002). As with Drosha, there are also factors which are known to regulate pre-miR processing by Dicer. Lin-28 is one such and is involved in the prevention of maturation of a number of let-7 microRNAs within embryonic stem cells, resulting in the increase in their target genes' expression (Treiber et al., 2012; Viswanathan et al., 2008).

The two strands of the miRNA-duplex-containing ribonucleoprotein particle (RNP) are separated into the guide and passenger strands. It is the guide strand which is incorporated into the RNA-induced silencing complex (RISC) (Gregory et al., 2005; Hammond et al., 2000). The RISC complex is comprised of an Argonaute protein (Ago2), Dicer, TRBP [Tar RNA binding protein (a double stranded RNA binding protein)], and PACT (protein activator of PKR) (Gregory et al., 2005; Kapinas and Delany, 2011). The mature microRNA guides the RISC complex to the complementary target site on the mRNA target sequence, resulting in the repression of the corresponding gene (Winter et al., 2009).

In plants, the guide strand is usually determined by the relative thermodynamic stability of the strands and is the strand with the weakest 5' base pair ending (Umate and Tuteja,

2010). The mature strand may be the 3' or 5' strand, and in some cases both may be used as the guide strand where necessary (Winter et al., 2009). In humans, in addition to the weakest base pairing rule applying to mature microRNA strand selection, where the strand with the lowest stable base pairing of the 2<sup>nd</sup> to 4<sup>th</sup> nucleotides from the 5' end is incorporated into the RISC complex (Schwarz et al., 2003), it is thought that the 5' strand is usually the guide strand due to the synthesis process. The 5' end of the 5' arm of the hairpin is cleaved by Drosha, whilst the 5' end of the 3' arm is cleaved by Dicer. Drosha is thought to cleave in a more specific manner and so this may explain why the 5' arm of microRNAs is often more highly expressed as it reaches its mature form with less errors (Hu et al., 2009).

It was thought that TRBP (a dsRBD protein) was mainly responsible for the inclusion of the guide strand into RISC and the stabilisation of Dicer (Gregory et al., 2005; Tomari et al., 2004), however more recent data suggests that Dicer may have a more important role in selection of the miRNA strand which is incorporated into the RISC complex. Dicer is thought to discriminate between strands based upon the 2-4nt 3' overhangs of the two strands of the microRNA duplex. This mediates the cleavage of the target mRNA (Peters and Meister, 2007; Sakurai et al., 2011), at the binding site [usually found at the 3'-untranslated region (UTR) of mRNAs] (Duursma et al., 2008).



**Figure 1.6 MicroRNA synthesis process.** The microRNA biosynthesis process, firstly the microRNA gene is transcribed to a pri-miRNA by RNA polymerase. This pri-miRNA is cleaved by Drosha to form a pre-miRNA which is exported out of the nucleus and into the cytoplasm where it is further cleaved by Dicer. The subsequent microRNA duplex is incorporated into the RISC complex and guides the complex to a microRNA binding site in the mRNA resulting in repression of the corresponding gene. Diagram adapted from (Howell et al., 2010).

Post biogenesis, microRNAs do not remain within cells indefinitely, they are eventually broken down and the turnover of microRNAs has been characterised. Some microRNAs are known to have long half-lives compared to others with the average being around 5 days. The panel of microRNAs analysed by Gantier et al showed a range between 4 and 9 days, but a half-life of up to 12 days for miR-208 has been seen by Van Rooij et al *in vivo*. MicroRNAs in general are also thought to be up to ten times more stable than mRNAs (Gantier et al., 2011; Ruegger and Grosshans, 2012; van Rooij et al., 2007; Zhang et al., 2012c). In human cultured cells microRNA degradation is controlled by ribosomal RNA processing protein 41 (RRP41), polynucleotide phosphorylase (PNPase) and XRN1. The first 2 enzymes degrade 3' to 5', but the 3<sup>rd</sup> degrades 5'-3' (Ruegger and Grosshans, 2012). Bail et al have also identified specific sequences found in certain microRNAs (specifically miR-382) outside the microRNA's seed site which may be responsible for how long a microRNA's half-life is, by determining its stability (Bail et al., 2010).

It is already known that microRNAs in general are vital for survival with a Dicer knock out mouse being embryonic lethal (Bernstein et al., 2003). This also shows that the Dicer enzyme itself is vital for microRNA synthesis. The Drosha/DGCR8 complex however may be dispensable, with *dgcr8* null mouse embryonic stem (ES) cells showing a less severe phenotype than *dicer* null mouse ES cells (Babiarz et al., 2008). This could suggest that there is an alternative route for microRNA production pre-Dicer cleavage. There has since, however, been subsequent research to suggest that Dicer might also be dispensable, but perhaps only with a sub group of microRNAs not vital for survival (Cheloufi et al., 2010).

The methods involved with miRNA function are multi-fold, depending on the organism in question. There are also some contradicting schools of thought as to which methods are generally more predominant. In plants, it's thought that microRNAs are able to bind directly and subsequently cleave a complementary target site of mRNA through the RISC complex, with non-complement matches not tolerated (Tang et al., 2003). Liu et al support this by showing that minor mismatches between a target site and microRNA seed site result in abrogated repression with most cases in plants. However they also discuss a small number of non-complementary target sites which result in less effective repression, rather than none, suggesting that these lower levels of repression may have been evolutionarily selected (Liu et al., 2014).

In mammals, microRNAs are more likely to have less complementarity to their target and so inhibit translation of the mRNA rather than cleave it directly, this also results in the preservation of the target mRNA structure (Ambros, 2004).

A number of microRNAs are produced in different cell types in different combinations and are thought to work co-operatively to regulate cell specific gene pathways. The target prediction algorithms currently available in addition to published research, show that a single microRNA may have hundreds of individual targets (Brennecke et al., 2005; Hendrickson et al., 2009), and that genes can be targeted by multiple different microRNAs or even have multiple target sites for the same microRNA. It is thought that these microRNAs work co-operatively to down-regulate gene expression, and thus can result in a stronger suppression (Akbari Moqadam et al., 2013; Hon and Zhang, 2007).

There are thought to be a number of mechanisms by which microRNAs are able to suppress gene expression. The first is whereby the microRNA prevents the translation initiation complex from forming by binding to its target. Next, the microRNA may bind to an area around, for example a stop codon, resulting in an interruption at a late initiation step and causing an early termination of translation. A final suggested mechanism is the prevention of ribosome assembly by the binding microRNA, resulting in translation suppression (Zinovyev et al., 2010).

The machinery involved in repression of target gene expression is also different to the RISC complex used in cleavage reactions. To this end complexes involved in repression are termed miRNPs (Mourelatos et al., 2002) and the repression complex is also thought to include the protein ago1, whereas cleavage complexes contain ago2 (Kapinas and Delany, 2011).

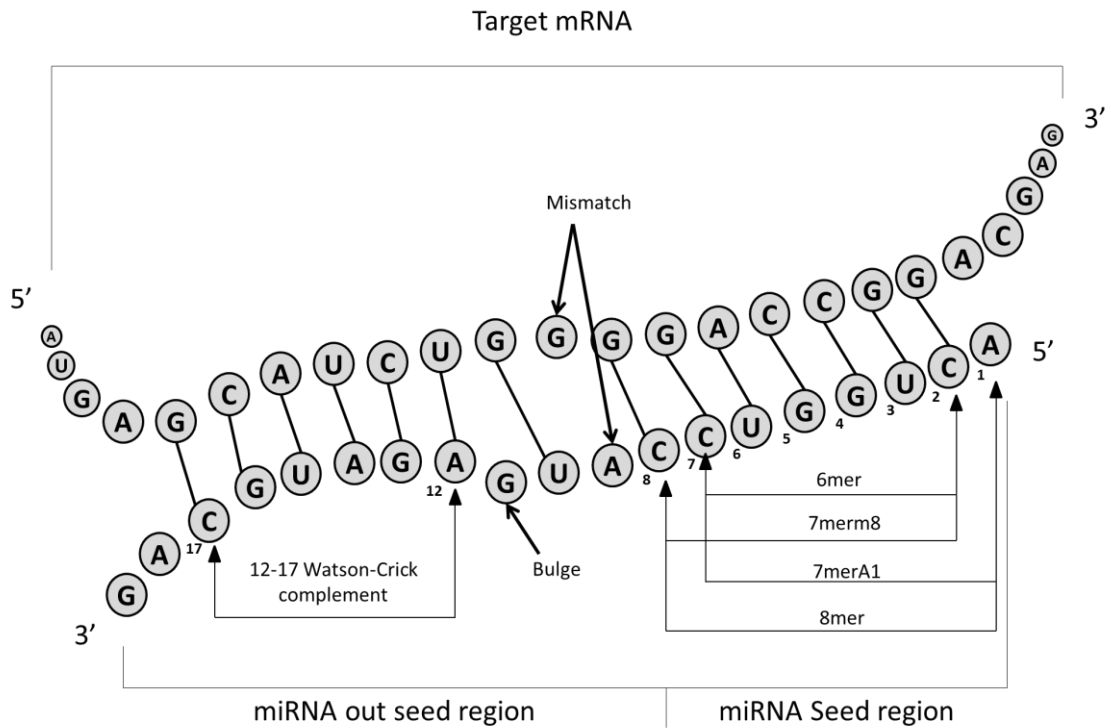
MicroRNAs bind to mRNA via specific conserved sites at the 5' end of the microRNA. These are known as seed sites which bind to corresponding complementary 'target sites' usually found at the 3' UTR of an mRNA sequence (Fig.1.7) (Akbari Moqadam et al., 2013). Target sites in the mRNA can also be found within the 5' UTR as well as coding regions of the mRNA (Lee et al., 2009). However, it is thought that microRNAs which bind at 3'UTRs cause a stronger or more efficient suppression of gene expression than if they were to bind to the 5'UTR, which could explain why targets in the 3'UTR are more prevalent (Moretti et al., 2010), or at least could explain why they are more

detectable. One microRNA (miR-605) has been found which targets both the 5'UTR and 3'UTR of the same gene, both resulting in down-regulation of the gene's expression (Lee et al., 2009).

In contrast to this, other research suggests that when a microRNA binds to a target site within the 5' UTR of an mRNA it actually results in the activation of gene expression rather than repression (Da Sacco and Masotti, 2012). This activation has been seen with miR-122 within the genome of the hepatitis C virus suggesting that perhaps the virus has evolved to exploit the microRNA process (Henke et al., 2008; Jangra et al., 2010; Jopling et al., 2005). It has also been found that miR-10a can induce the production of ribosomal proteins by binding to a site in the 5'UTR of the mRNA (Orom et al., 2008).

As already stated, microRNAs bind to mRNA sequences via specific sites at their 5' end and there are a number of seed sites of different sizes within the sequence of a microRNA which are thought to be responsible for binding to an mRNA target, and they start at a number of known locations within the microRNA (Fig. 1.7). The first is the 6mer, this is a site 6 nucleotides long which starts at the 2<sup>nd</sup> nucleotide and finishes at the 7<sup>th</sup> (inclusive). The next is the 7merm8. This site is 7 nucleotides long beginning at the 2<sup>nd</sup> nucleotide, and as the name suggests, ends at the 8<sup>th</sup> nucleotide. There is an additional site of the same length entitled 7merA1 however this site begins at the first nucleotide, finishing at the 7<sup>th</sup>. The final seed site is the 8mer, this site is 8 nucleotides long beginning at the first nucleotide and ending at the 8<sup>th</sup> (Mitra and Bandyopadhyay, 2011).

It's thought that although all 4 of these sites can result in binding sufficient to inhibit mRNA expression, the binding of the 8mer site results in a stronger down-regulation of the gene (Schnall-Levin et al., 2011). It is also theorised that in addition to the seed sites being the main locations of binding, the microRNA sequence can also bind to a mRNA sequence with its 3' end, specifically an area known as the Watson-Crick complement (Fig. 1.7) which runs from the 12<sup>th</sup> to the 17<sup>th</sup> nucleotide of the microRNA (Mitra and Bandyopadhyay, 2011).



**Figure 1.7 Diagram of microRNA : mRNA interaction patterns.** Diagram of all known potential binding areas between microRNA sequence and target mRNA. The miRNA seed region located at the 5' end of the microRNA contains the 4 most common seed sites: 6mer, 7merA1, 7merm8 and 8mer. Image adapted from Mitra et al (Mitra and Bandyopadhyay, 2011).



Aside from manual inspection of individual microRNA sequences to compare for target mRNA matches, there are a number of algorithms available online which attempt to predict the targets of the majority of microRNAs described on miRBase (Griffiths-Jones et al., 2006; Griffiths-Jones et al., 2008; Kozomara and Griffiths-Jones, 2011). These include TargetScan (Friedman et al., 2009), miRanda (Betel et al., 2008), PITA (Kertesz et al., 2007) and PicTar (Krek et al., 2005). However each database uses a different algorithm, and if results are compared they often do not overlap as can be seen in another online system which combines the results of a number of the databases: The microRNA body map (Mestdagh et al., 2011). One of the problems with using these online algorithms is that they all show high levels of false positive or false negative results (Mitra and Bandyopadhyay, 2011) and all targets must therefore be experimentally validated before determining them as definite targets, and microRNA target investigation must not solely rely on the current bioinformatics available.

The machinery involved in repression of target gene expression is also different to the RISC complex used in cleavage reactions. To this end complexes involved in repression are termed miRNPs (Mourelatos et al., 2002) and the repression complex is also thought to include the protein ago1, whereas cleavage complexes contain ago2 (Kapinas and Delany, 2011).

In addition to their functions, miRNAs can also be characterised into groups based on where they are 'clustered' in the genome (Houbaviy et al., 2003). The current theory is that these miRNAs may have arisen from the same mRNA molecule, and so understandably, miRNAs with similar sequences are found within the same cluster. In fact, not only are they transcribed from the same area, they are also often co-ordinately expressed (Cai et al., 2004) and share the same promoter (Treiber et al., 2012). Clusters of microRNAs are termed Intergenic clustered, Intronic clustered or Mirtrons (where the microRNA is part of the intron of a gene) (Kapinas and Delany, 2011). However, some miRNAs are clustered together with distinctly different sequences; it is thought that these particular miRNAs function to target a collection of genes in a coordinative manner, rather than being transcribed in a co-coordinative manner (Ambros, 2004). One example of this coordinated 'attack' on gene expression are the microRNAs let-7 and miR-125 in vertebrates. These two miRNAs are noticeably co-regulated in the developmental stage of the fruit fly (Bashirullah et al., 2003). This suggests that the miRNAs can be clustered together depending on their functional targets as well as their

sequence. Further data also supports the idea that clustered microRNAs work coordinately to target certain genes, such as the miR-290 cluster targeting Rbl2 (Benetti et al., 2008).

An example of genomic clustering within mammals includes the 6 microRNAs; miR-290, miR-291, miR-292, miR-293, miR-294, and miR-295 found expressed within mouse embryonic stem cells (ESC) (Claudo et al., 2009; Houbaviy et al., 2003). Additional research in humans by Altuvia et al (Altuvia et al., 2005) has discovered that 37% of the microRNAs studied appeared to cluster in at least groups of two or more and Seitz et al has computationally predicted 40 microRNAs clustered with the same 1 Mb of the human genome (Seitz et al., 2004).

MicroRNAs are already well-established as effectors of gene regulation in a number of diseases such as cancer, where some microRNAs can act as tumour suppressors, whilst others stimulate cancer development via their actions upon oncogenes (Raia, 2011). MicroRNAs are increasingly being seen as potential biomarkers of diseases and their progression, for example miR-155 in breast cancer (Bartels and Tsongalis, 2009). They are also involved in many physiological processes, from regulating embryo maturation in *Arabidopsis thaliana* (Willmann et al., 2011) to a possible role in the human circadian clock (for example miR-192 and miR-122) (Hansen et al., 2011).

A large number of microRNAs are documented in the Human microRNA Disease Database version 2.0 (HMDD v2.0) (Li et al., 2014). This database has recorded 572 known microRNAs which are known to be involved across 378 human diseases. These diseases include AIDS in which miR-198, miR-29a and miR-29b-1 are known to be involved, Alzheimer's disease in which 30 known microRNAs are involved and pulmonary disease, where 17 microRNAs are thought to play a role.

Further, microRNAs are no longer considered limited to acting just within cells, but are also detected within the blood. It is thought that these circulating microRNAs escape degradation by RNases found within the blood by the presence of Argonaute proteins (specifically Argonaute 2) which form a protective complex around the microRNA (Arroyo et al., 2011). In addition to their physiological role, these microRNAs are now being utilised as biomarkers of various diseases, with over 100 already identified as biomarkers (Weiland et al., 2012).

## 1.5 - MicroRNA Involvement Within Osteoarthritis

MicroRNAs are known to be involved in both chondrogenesis and various arthritic diseases including osteoarthritis, and have already been partially documented, but as we learn more about these molecules, the true depth of their involvement is only just being recognised.

The Dicer knockout mouse, as mentioned earlier, shows the general importance of all microRNAs for survival, with the mouse having a lethal phenotype (Bernstein et al., 2003). A Dicer knockout of zebrafish also shows the importance of Dicer and the microRNA population with zebrafish development halting at day 10 (Wienholds et al., 2003). Further, a conditional knockout of Dicer in early stage limb resulted in a shorter limb developing, suggesting that microRNAs also play a vital role in chondrogenesis (Harfe et al., 2005). This is supported by Dicer null chondrocytes which showed growth defects during chondrogenesis (Kobayashi et al., 2008).

A number of microRNAs are now associated with cartilage and chondrogenesis, including miR-199a, miR-124a and miR-96 (Suomi et al., 2008) with the first two found to be up-regulated in MSC derived chondrocytes, and miR-96 down-regulated. All three were found to target SOX-5 and HIF-1 $\alpha$ . The microRNAs miR-145 and miR-143 were also determined to increase in expression in chondrocytes derived from hMSCs (Karlsen et al., 2011; Yan et al., 2011), with miR-145 found to be a key element in early chondrogenic differentiation by its targeting of SOX9, resulting in the up-regulation of markers of hypertrophy (Yang et al., 2011a).

MiR-675 is also well studied, and is known to be influential in chondrogenesis and a part of the maintenance of cartilage homeostasis. The microRNA is regulated by SOX9 during chondrogenesis and indirectly results in an increase in expression of *COL2A1*. It was also found that miR-675 was able to combat type II collagen reduction in SOX9 inhibited chondrocytes (Dudek et al., 2010).

Expression of the miRNome of chondrocytes undergoing de-differentiation has also been investigated. The expression of five microRNAs were found to be higher at early chondrocyte passages compared to later stages, these were miR-30d, miR-140-3p, miR-210, miR-451 and miR-563 (Lin et al., 2011) suggesting chondrocyte specificity.

Comparisons carried out between osteoarthritic and normal cartilage cells have shown differential expression of a number of microRNAs. Jones and colleagues (2009) discovered 17 out of 157 microRNAs measured to be differentially expressed between late stage OA cartilage and normal cartilage, and 30 in OA bone compared to normal. These included miR-9 and miR-98 which were found to be upregulated in both OA cartilage and bone. These microRNAs are postulated to be involved in inflammation pathways and in particular, miR-9 appears to modulate the expression of *MMP13* (Jones et al., 2009). Another paper also suggested a link between *MMP13* expression in osteoarthritic chondrocytes and microRNAs, showing that miR-27b had a regulatory effect upon the expression of the metalloproteinase (Akhtar et al., 2010). This is supported by a further report which found miR-27a and miR-140 to be involved in the regulation of *MMP13* (Tardif et al., 2009).

Further publications describe the analysis of 723 microRNAs and the discovery of an increase in expression of miR-483-5p in OA chondrocytes whilst a further 6 microRNAs (miR-149-3p, miR-576-5p, miR-582-3p, miR-641, miR-634 and miR-1227) showed a decrease in expression in OA chondrocytes. These microRNAs are also thought to influence the expression of Wnt signalling and TGF- $\beta$ , suggesting a more in depth role in osteoarthritis (Diaz-Prado et al., 2012).

Additional profiling experiments have shown more microRNAs to be both up-regulated and down-regulated in diseased cartilage compared to normal. Iliopoulos et al looked at the expression of 365 microRNAs in cartilage tissue. They identified 9 to be up-regulated and 7 to be down-regulated in OA cartilage compared to normal. These microRNAs included miR-9, miR-98 and miR-146 which are all known to regulate IL-1 $\beta$  induced TNF- $\alpha$ , in addition to the targeting of *MMP13* by miR-9. This could suggest that these microRNAs could be involved in the prevention of OA or in the maintenance of healthy cartilage (Iliopoulos et al., 2008). MiR-22 was also examined functionally by the same group, and overexpression of the microRNA resulted in an increase in the expression of *MMP-13* and *IL-1 $\beta$* ; key factors in OA progression (Iliopoulos et al., 2008).

Additional research has been carried out upon miR-146a and its role in osteoarthritis, finding that it appears to be induced by joint instability in surgically induced OA models. This suggests that it could be involved in mechanotransduction of OA (Li et al.,

2012a) as well as its potential regulation of *MMP-13*, determined by an inverse pattern of expression between *MMP-13* and miR-146a (Yamasaki et al., 2009).

The online database HMDD v2.0 shows 8 recorded microRNAs to have published data on their influence in osteoarthritis, although this is not an exhaustive list (Table 1.2).

MicroRNA	Function	Reference
miR-140	Down-regulation of ADAMTS-5, and activated by Sox9 to target Sp1	(Miyaki et al., 2009) (Araldi and Schipani, 2010) (Yang et al., 2011b) (Zhang et al., 2012a)
miR-146a	Expressed in low-grade OA cartilage, Linked to pain in OA and controls knee joint homeostasis	(Yamasaki et al., 2009) (Li et al., 2011) (Okuhara et al., 2012)
miR-155	Highly expressed in OA cartilage	(Okuhara et al., 2012)
miR-181a-1	Highly expressed in OA cartilage	(Okuhara et al., 2012)
miR-181a-2	Highly expressed in OA cartilage	(Okuhara et al., 2012)
miR-223	Highly expressed in OA cartilage	(Okuhara et al., 2012)
miR-27b	Regulates MMP13 in OA chondrocytes	(Akhtar et al., 2010)
miR-675	Down-regulated in OA cartilage	(Steck et al., 2012)

**Table 1.2 Known microRNAs associated with osteoarthritis from the HMDD v2.0 database.** Table shows the 8 microRNAs described on the online Human microRNA Disease Database v2.0, how each is involved in osteoarthritis and the corresponding literature reference.

The microRNAs miR-140 and miR-455 have been studied extensively for their role in both osteoarthritis and chondrogenesis. MiR-140-3p, miR-140-5p, miR-455-3p and miR-455-5p have all been shown to have cartilage selective expression with miR-455-3p showing expression restricted to the cartilage of developing long bones in the chick and mouse (Miyaki et al., 2009; Miyaki et al., 2010; Swingler et al., 2012), and miR-140 expression specifically located to cartilage in the jaw, head and fins of zebrafish (Wienholds et al., 2005) as well as being identified within cartilage tissue of mouse embryo limbs at stage E11.5 and within digits of mouse limbs at stage E14.5 (Tuddenham et al., 2006). MiR-140 is also SOX-9 inducible; a factor heavily involved in chondrocyte development (Yang et al., 2011b).

A miR-140 null mouse developed by two separate research groups showed mild dwarfism with possible inhibition of chondrocyte differentiation. In addition, knockout mice showed a higher OA score compared to wild type controls and OA scores were significantly increased in knockout mice compared to wild type in a surgical OA susceptibility model (involving the resection of the medial meniscotibial ligament to induce joint instability). Further to this ADAMTS-5 and Dnpep (Aspartyl Aminopeptidase) were found to be direct targets of miR-140 (Miyaki et al., 2010; Nakamura et al., 2011) with *MMP-13* expression also thought to be mediated by the microRNA (Liang et al., 2012). In addition, an antigen-induced arthritis model with an overexpression of miR-140 in chondrocytes showed an apparent decrease in OA score suggesting that miR-140 could have a protective role (Miyaki et al., 2010). Further data shows miR-140 to be decreased in expression in both OA cartilage (Miyaki et al., 2009) and OA chondrocytes (Tardif et al., 2009) compared to normal tissue and cells. In contrast however, another study found miR-140 to increase in expression when measured in end stage hip OA cartilage tissue, corresponding with a decrease in ADAMTS-5 (a known target of miR-140) (Swingler et al., 2012) suggesting that the expression profile of miR-140 changes as the disease progresses, shifting away from the original protective role.

Further targets of miR-140 have subsequently been identified. IGFBP-5 (Insulin-like growth factor binding protein 5) is one such direct target and has been found to have an increased expression in osteoarthritic cartilage (Tardif et al., 2009).

Functional studies have been carried out on a range of microRNAs in association with cartilage, one study from 2010, showed miR-34a to be involved in chondrocyte apoptosis. Silencing of the microRNA resulted in reduced apoptosis of chondrocytes in a rat OA model, it was also shown to reduce interleukin-1 mediated down-regulation of type II collagen (Abouheif et al., 2010) suggesting a potential role for this microRNA in OA prevention.

In addition to microRNA profiling array studies, there have also been forays into the cartilage and osteoarthritic miRNome using next generation sequencing. One such study in rat cartilage using the Solexa platform identified 246 previously identified and 86 novel microRNAs (Sun et al., 2011b). These types of studies have also been used to investigate the miRNome of other tissues and diseases. For example, Wu et al identified 546 precursor microRNAs in breast cancer tissue using SOLiD sequencing in an investigation to identify potential cancer biomarkers (Wu et al., 2011). An additional study used next generation sequencing to identify 146 known microRNAs within tissue of the rat heart left ventricle as well as 68 novel microRNAs, which provides a further insight into cardiovascular diseases (Collins et al., 2013). On a more specific cell based level, Fehniger et al used a next generation sequencing project to identify the miRNome of natural killer cells. They identified 302 known microRNAs and 21 novel microRNAs within these cells, providing additional information to elucidate the function of microRNAs in natural killer cells and therefore their implicated roles within the immune system (Fehniger et al., 2010).

There are over 2500 mature human microRNAs described on miRBase (Griffiths-Jones et al., 2006; Griffiths-Jones et al., 2008; Kozomara and Griffiths-Jones, 2011) however only a fraction of them have been functionally analysed. An increasing number are becoming associated with cartilage and osteoarthritis, and it is likely, that as detection and identification technologies improve, soon there will be a comprehensive miRNome of osteoarthritis available. This will aid progression in both the diagnosis and hopefully prevention of osteoarthritis.



## **1.6 – Aims And Objectives**

The hypothesis of this project was that cartilage contains as yet unidentified microRNAs which may contribute to skeletal development or osteoarthritis. We aimed to address this hypothesis by the purification of small RNAs from primary chondrocytes, followed by the creation of a small RNA sequencing library for next generation Illumina sequencing to profile the miRNome of osteoarthritic cartilage. Novel microRNA candidates would undergo computational and experimental validation, and potential targets of a number of microRNAs would be uncovered through the use of a whole genome array combined with a computational study, and subsequently validated by sub-cloning and luciferase assays.

## **Chapter 2 – Materials and Methods**

## **2.1 Cell Lines**

### **2.1.1 SW1353**

SW1353 cells were obtained from the American Type Culture Collection [ATCC (Manassas, VA, USA)] (no. HTB-94). The line originates from a primary grade II chondrosarcoma and is a useful model for chondrocyte culture as there are some similarities in morphology and behaviour, although the phenotype is not completely comparable to primary chondrocytes (Gebauer et al., 2005). Cells were maintained in Dulbecco's Modified Eagle Medium (DMEM) + GlutaMAX™ [Life Technologies (Glasgow)] [+10% (v/v) heat inactivated Foetal calf serum (FCS) [PAA Laboratories (Somerset)] 100U/ml penicillin, 100µg/ml streptomycin [Gibco (Glasgow)]].

### **2.1.2 DLD-1**

DLD-1 Parental and DLD-1 Dicer null cell lines (kind gifts from Prof. Tamas Dalmay, University of East Anglia, Norwich Research Park, Norwich, UK) were developed at Horizon Discovery (Cambridge, UK) and originated from a colorectal adenocarcinoma. DLD-1 Dicer null is a homozygous knock out of Dicer and is assumed to be tumorigenic, produced through the homozygous disruption of the helicase domain in exon 5 of Dicer 1 (Cummins et al., 2006). The cell lines are isogenic apart from the Dicer expression. Cells were maintained in DMEM F12 + GlutaMAX™ (Life Technologies) [+10% inactivated FCS (PAA Laboratories), 40U/ml penicillin, 40µg/ml streptomycin (Gibco)].

### **2.1.3 DF1**

The DF1 cell line (a kind gift from Prof. Andrea Munsterberg, University of East Anglia, Norwich Research Park, Norwich, UK). The cell line originated from chicken embryo and is a spontaneously transformed fibroblast cell type (Foster and Foster, 1997). Cells were maintained in DMEM + GlutaMAX™ (Life Technologies) [+10% (v/v) heat inactivated FCS (PAA Laboratories) 100U/ml penicillin, 100µg/ml streptomycin (Gibco)].

## **2.2 Human Tissue Panel RNA**

RNA for 20 different human tissues (10µg RNA per tissue from a pool of 3 patients) was obtained from FirstChoice<sup>®</sup> Human Total RNA Survey Panel [Applied Biosystems (Glasgow)].

## **2.3 Tissue Samples**

Femoral head and knee cartilage were obtained from clinically diagnosed end stage osteoarthritic [HOA (Hip OA), KOA (Knee OA)] and trauma [NOF (Neck of femur)] patients undergoing total joint replacement surgery at the Norfolk and Norwich University Hospital in collaboration with Prof. Simon Donell (Consultant Orthopaedic Surgeon and Honorary Professor, University of East Anglia, UK). OA was diagnosed using clinical history and examination in addition to X-ray analysis. Fracture patients had no known history of joint disease and their cartilage was free of lesions at the time of removal. All subjects gave informed consent, and the study was performed with ethical approval. Tissue obtained for use in next generation sequencing included one female and two males of an age range of 56 – 71 years. Tissue obtained for use in the whole genome microarray included two females and one male of an age range of 61 – 81 years.

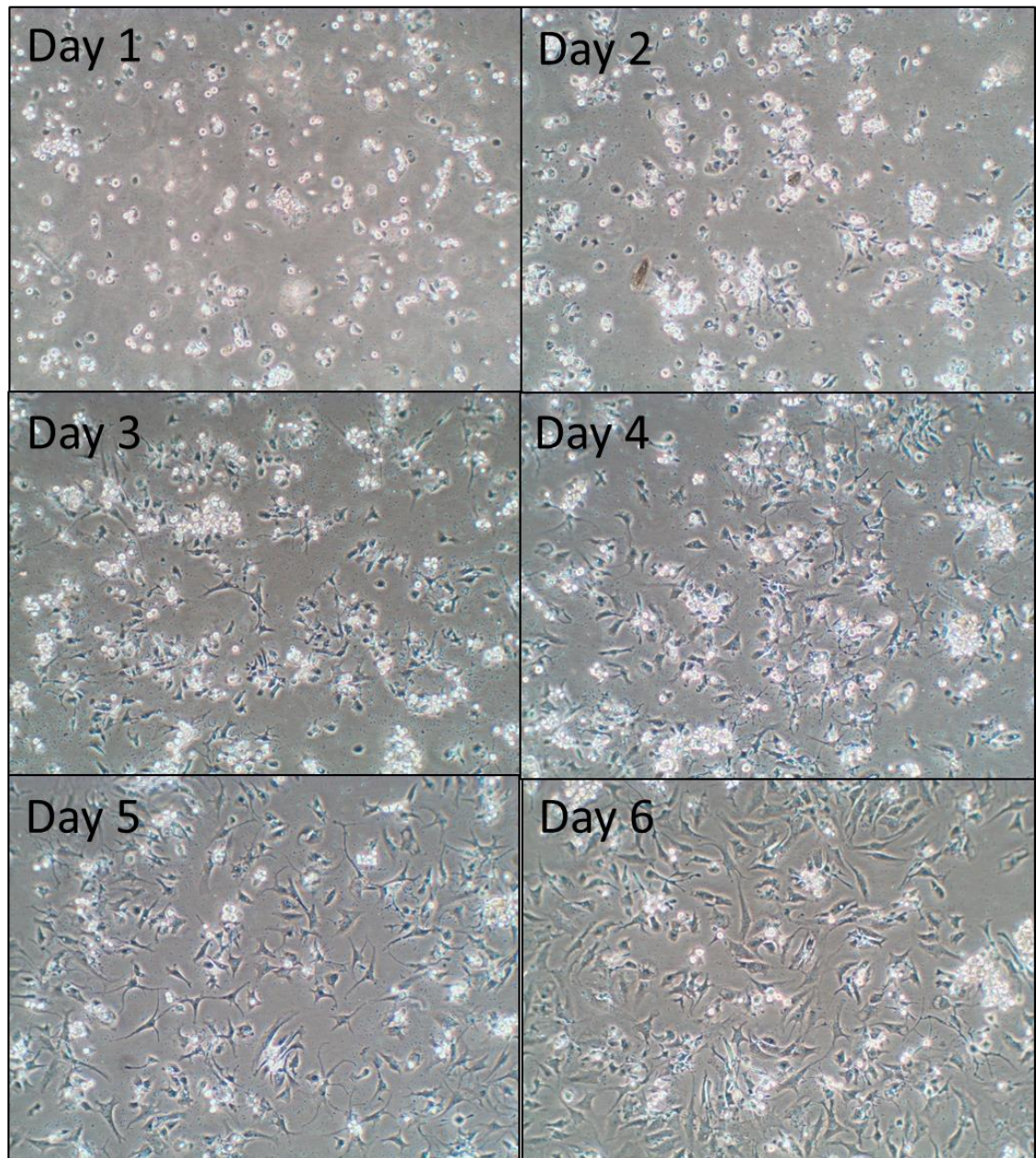
## **2.4 Human Cartilage Dissection**

Cartilage from KOA showing degeneration was removed from the entire surface of both tibial plateaus and femoral condyles including medial and lateral compartments. In the case of hip tissue, cartilage from HOA showing degeneration was removed from the entire surface of the femoral head of the hip and all tissue was removed from the femoral head of NOF. However any areas containing lesions observed on tissue designated NOF were not used.

Cartilage samples were removed from the joint using a sterile scalpel and washed in sterile phosphate buffered saline (PBS) solution (Life Technologies) containing 100U/ml penicillin, 100µg/ml streptomycin (Gibco) and cut into 2-5mm pieces. These sections were either snap frozen in liquid nitrogen-cooled n-hexane [Sigma-Aldrich (Dorset)] or processed to isolate primary chondrocytes (see section 2.5).

## 2.5 Primary Chondrocyte Isolation

Dissected cartilage sections were incubated at 37°C shaken at 180rpm ~18 hours in filter sterilised digestion media: [0.1% (w/v) collagenase type 1A (Sigma-Aldrich) and 0.4% (w/v) Hepes (Sigma-Aldrich) in DMEM + GlutaMAX™ (Life Technologies) containing 100U/ml penicillin, 100µg/ml streptomycin (Gibco) and 2.5µg/ml Fungizone (amphotericin B, Gibco) at 5ml/g of cartilage. The digestion mixture was filtered through 70µm nylon cell strainers (BD Falcon) and centrifuged at 180g for 5 minutes and the supernatant discarded. The cell pellet was washed in 5ml DMEM + GlutaMAX™ (Life Technologies) containing 100U/ml penicillin, 100µg/ml streptomycin (Gibco) and 2.5µg/ml Fungizone (amphotericin B, Gibco) and centrifuged at 180g for 5 minutes. The cell pellet was re-suspended in DMEM + GlutaMAX™ (Life Technologies) [+10% (v/v) heat inactivated FCS (PAA Laboratories) 100U/ml penicillin, 100µg/ml streptomycin (Gibco)]. Cells were seeded at  $4 \times 10^4$  cells/cm<sup>2</sup> and incubated at 37°C, 5% CO<sub>2</sub> and allowed to adhere for ~8 days. The media was changed following adherence, and at 80-90% confluency, cells were either frozen and stored in liquid nitrogen, or utilised in *in vitro* assays up to passage 1. Images of isolated and cultured primary chondrocytes for 6 days post digestion are shown in figure 2.1.



**Figure 2.1 Images of primary chondrocyte isolation.** Cells obtained from digestion of osteoarthritic knee cartilage, and cultured in monolayer. Images show cell adherence, and growth from day 1 to 6 post digest. At confluency seen at day 6 (>80%) cells are frozen and stored in liquid nitrogen.

## **2.6 Cell Culture**

Primary human chondrocytes, SW1353 and DF1 cells were cultured in DMEM + GlutaMAX<sup>TM</sup> (Life Technologies) [+10% (v/v) heat inactivated FCS (PAA Laboratories) 100U/ml penicillin, 100µg/ml streptomycin (Gibco)]. DLD-1 parental and Dicer null lines were cultured in DMEM F12 + GlutaMAX<sup>TM</sup> (Life Technologies) [+10% inactivated FCS (FCS) (PAA Laboratories), 40U/ml penicillin, 40µg/ml streptomycin (Gibco)]. Culturing conditions consisted of 37°C in 5% (v/v) CO<sub>2</sub>. Cells were passaged using 0.25% (w/v) trypsin-1mM EDTA (Life Technologies). Cells were frozen for storage by the following: trypsinisation, centrifugation at 180g for 5 minutes, resuspension in 90% (v/v) FCS, 10% (v/v) DMSO) and incubation in a cell freezing chamber (VWR) at -80°C overnight. Cells were stored long term in a liquid nitrogen dry store.

## **2.7 RNA Extraction**

### **2.7.1 RNA Extraction From Frozen Cartilage – Using mirVana<sup>TM</sup> miRNA Isolation Kit [Ambion (Glasgow)]**

All materials were RNase free. Snap frozen cartilage was crushed under liquid nitrogen into powder using a BioPulverizer (Biospec). 1.25ml Trizol (Life Technologies) was added to 50mg ground cartilage and vortexed for 30 seconds 4°C. 1/5 volume Trizol (Life Technologies) of chloroform (Fisher Scientific) was added, vortexed for 15 seconds and centrifuged at full speed for 10 minutes at 4°C. The aqueous phase was recovered and 1.25 x the aqueous volume of 100% (v/v) ethanol (Sigma-Aldrich) was added. The MirVana<sup>TM</sup> method continued from this point as per manufacturer's instructions with the following exceptions: at the elution stage, 50µl elution solution was put through the column twice. RNA was quantified using Nanodrop 2000 [Thermo Scientific (Loughborough)] and stored at -80°C.

### **2.7.2 RNA Extraction From Monolayer Cells – Using mirVana<sup>TM</sup> miRNA Isolation Kit (Ambion)**

The protocol was carried out according to manufacturer's instructions with the following exceptions: at the elution stage, 50µl elution solution was run through the column twice. RNA was quantified using Nanodrop 2000 (Thermo Scientific) and stored at -80°C.

### **2.7.3 RNA Extraction From Post Digest Phase Of Primary Cell Isolation – Using mirVana™ miRNA Isolation Kit (Ambion)**

A minimum of  $3 \times 10^6$  cells were obtained from digestion media during primary chondrocyte isolation (see section 2.5). Cells were centrifuged at 180g for 5 minutes RT, and washed with Hank's balanced salt solution (HBSS) (Life Technologies) before continuation of the mirVana™ isolation method as per manufacturer's instructions with the following exceptions: at the elution stage, 50µl elution solution was run through the column twice. RNA was quantified using Nanodrop 2000 (Thermo Scientific) and stored at -80°C.

### **2.7.4 RNA Enrichment For Small RNAs From Total RNA– Using mirVana™ miRNA Isolation Kit (Ambion)**

The protocol to enrich for small RNAs from total RNA sample was carried out according to manufacturer's instructions with the following exceptions: at the elution stage, 50µl elution solution was run through the column twice. RNA was quantified using Nanodrop 2000 (Thermo Scientific) and stored at -80°C.

## **2.8 Ethanol RNA Precipitation**

To improve RNA quality, to each RNA sample, 1/10 volume 3M sodium acetate pH4.5 (Fisher Scientific) and 2.5x volume 100% (v/v) ethanol (Sigma-Aldrich) was added. Sample was then incubated at -80°C overnight before being centrifuged at 4°C for 15 minutes at full speed to recover the RNA. Supernatant was discarded and 500µl ice cold 75% ethanol (v/v) (Sigma-Aldrich) added, the sample was centrifuged again at 4°C for 15 minutes at full speed. The supernatant was discarded and the pellet air dried and re-suspended in a smaller volume of sterile water than the original sample.

## **2.9 cDNA Synthesis**

### **2.9.1 cDNA Synthesis – Cells-to-cDNA II With MMLV**

Cells in a 96 well plate [Fisher Scientific (Loughborough)] format were washed twice using room temperature PBS, and 30µl 'Cells-to-cDNA' Lysis buffer (Ambion) was added. The lysate was transferred to a 96-well PCR plate (Applied Biosystems) and incubated at 75°C for 15 minutes. 2U DNase I (Ambion) and 10x DNase I buffer



(Ambion) were added to each well and incubated at 37°C for 15 minutes followed by 75°C for 5 minutes. 8µl from each well were transferred to a new PCR plate where 2.5mM dNTPs [Bioline (London)] and 200ng random hexamers [Invitrogen (Glasgow)] were added to each well and incubated at 70°C for 5 minutes. The samples were chilled on ice before 4µl 5x RT buffer (Invitrogen), 2µl DTT (0.1M) (Invitrogen), 100U MMLV reverse transcriptase (Invitrogen), and 20U RNasin ribonuclease inhibitor (Promega) were added to each well up to a total volume of 20µl in analytical grade H<sub>2</sub>O (Sigma-Aldrich). Samples incubated at 37°C for 50 minutes then 75°C for 15 minutes. To each well, 30µl analytical grade H<sub>2</sub>O (Sigma-Aldrich) was added and cDNA stored at -20°C.

### **2.9.2 cDNA Synthesis - Reverse Transcription**

400ng of random hexamers (Invitrogen), or specific microRNA RT primers [ABI (Glasgow)], were added to up to 1µg RNA (see sections 2.7) and incubated at 70°C for 10 minutes in a 11µl reaction. 4µl 5x RT buffer (Invitrogen), 2µl DTT (0.1M) (Invitrogen), 2.5mM dNTPs (Bioline), 200U Superscript (Invitrogen) and 40U RNase inhibitor [Promega (Southampton)] were added per reaction to a total of 20µl and incubated at 42°C for 1 hour, then at 70°C for 10 minutes. The cDNA was stored at -20°C.

### **2.9.3 cDNA Synthesis - miRCURY LNA<sup>TM</sup> Universal RT MicroRNA PCR Universal cDNA Synthesis Kit [Exiqon (Copenhagen)]**

The protocol followed the manufacturer's instructions with the following exceptions: half volumes of all reagents were used with the same concentration of total RNA extracted using mirVana<sup>TM</sup> miRNA Isolation Kit (Ambion) (see section 2.7).

## **2.10 Quantitative Real Time PCR (qRT-PCR)**

### **2.10.1 qRT-PCR With Primers And Probes**

Primers and probes for *MMP1* and *MMP13* were designed as described in (Nuttall et al., 2003). Primer and probe sets for microRNA-140, microRNA-455, *18S* and *ACAN* were designed by ABI. SNORA36A primers were designed by Exiqon. All other primers were designed using the Universal Probe Library Assay Design Center by Roche Applied Science and used with probes from the Universal Probe Library [Roche

Applied Science (West Sussex)]. Primer and probe sequences are documented in Appendix 1.

All genes were normalised to the *18S* RNA housekeeping gene to ensure there were no differences in mRNA concentration between samples. Any samples with *18S* RNA  $C_t$  levels over 1.5  $C_t$  from the median were excluded from subsequent analysis.  $C_t$  is the threshold cycle during a PCR run where the signal is statistically significantly above baseline and therefore detectable. *18S* has previously been established in our lab as the optimal housekeeping gene (out of a number of potential normalising genes tested) when using cartilage samples and chondrocyte cell lines.

Standard curves were prepared for each gene and microRNA to account for primer/probe set efficiency. *18S* RNA cDNA standards were diluted: 4, 2, 1, 0.5, 0.25, 0.125ng/10 $\mu$ l. The remaining genes' standards were diluted: 20, 10, 5, 2.5, 1.25, 0.625ng/10 $\mu$ l.

PCR was carried out using the Applied Biosystems 7500 Taqman system in MicroAmp optical 96-well plates (Applied Biosystems). *18S* RNA assays required 1ng cDNA and all other genes required 5ng cDNA per reaction. Per reaction, 2x Taqman Universal PCR Master Mix (Applied Biosystems), 100nM forward and reverse primers (Sigma-Aldrich) and 200nM universal probe (Roche) or specific probe (Sigma-Aldrich) were added and made up to a total reaction volume of 25 $\mu$ l with analytical grade H<sub>2</sub>O (Sigma-Aldrich). Optical adhesive covers (Applied Biosystems) were used to seal the plates. PCR cycling conditions were; 50°C for 2 minutes, 95°C for 10 minutes and 40 cycles of [95°C for 15 seconds followed by 60°C for 1 minute] Ramp- rate 1.6°C/s. Microsoft Excel was used to analyse the data.

### **2.10.2 qRT-PCR Using miRCURY LNA<sup>TM</sup> Universal RT MicroRNA PCR SYBR Green Master Mix Universal RT (Exiqon)**

Locked Nucleic Acid (LNA) primers for novel microRNAs were designed using the Exiqon web based assay tool (primer sequences not known). LNA primers for miR-140-3p, miR-140-5p, miR-455-3p, miR-455-5p, and *U6* RNA were designed by Exiqon (primer sequences not known). All genes were normalised to the *U6* RNA housekeeping gene to ensure there were no differences in miRNA concentration between samples.

Any samples with  $U6 C_t$  levels over  $1.5 C_t$  from the median were excluded from subsequent analysis.  $C_t$  is the threshold cycle during an RT-PCR run where the signal is above baseline and therefore detectable.

PCR was carried out as per manufacturer's instructions using the Applied Biosystems 7500 Taqman system with MicroAmp optical 96-well plates (Applied Biosystems). Optical adhesive covers (Applied Biosystems) were used to seal the plates. PCR cycling conditions were; 50°C for 2 minutes, 95°C for 10 minutes and 40 cycles of [95°C for 10 seconds followed by 60°C for 1 minute]. Ramp-rate was 1.6°C/s, followed by a melt curve analysis. The cycling conditions for the melt curve analysis were: 95°C for 15 seconds, 60°C for 1 minute, 95°C for 30 seconds and finally 60°C for 15 seconds. Ramp rate was 1.6°C/s. Melt curves ensure specificity of primers is known when using SYBR green. Microsoft Excel was used to analyse the data.

## **2.11 DNA Polyacrylamide Gel Electrophoresis**

qRT-PCR products from miRCURY LNA<sup>TM</sup> Universal RT microRNA PCR (Exiqon) (see section 2.10.2) on RNA isolated from NOF and KOA cartilage, were run on 15% (w/v) polyacrylamide gels in 0.5xTBE buffer to resolve small fragments (~21bp). The gel was stained with ethidium bromide (Sigma-Aldrich) and imaged under UV light. 10bp ladder (Invitrogen) run alongside and microRNA-140-3p was used as a size control.

## **2.12 Northern Blot**

12% (w/v) Polyacrylamide gels were prepared, for one gel: 3.15g urea (Sigma-Aldrich) heated with 1.5ml sterile water (Sigma-Aldrich). 2.8125ml (40%) (w/v) acrylamide [BIO-RAD (Hertfordshire)] was added followed by 0.75ml 5X TBE and mixed. 75µl 10% (w/v) ammonium persulfate (Sigma-Aldrich) and 3.75µl TEMED (BIO-RAD) were added and the gel poured rapidly and left to set for 30 minutes.

Total RNA extracted using mirVana<sup>TM</sup> miRNA Isolation Kit (Ambion) (see section 2.7.2) and RNA enriched for small RNAs (see section 2.7.4) were added to the same volume of sample buffer and heated at 70°C for 3 minutes. The wells were syringed and

the gel pre-run at 80V for 20 minutes. The wells were syringed a second time before the samples were added and the gel run at 100V for 2 hours.

The RNA was transferred from the gel to a Hybond-NX membrane [Amersham Biosciences (Buckinghamshire)] using a semi-dry transfer apparatus, run at 175V for 35 minutes. Whatmann filter paper was soaked in cross-linking solution [10ml sterile water (Fisher Scientific), 122.5µl 12.5M 1-methylimidazole (Sigma-Aldrich) and 20µl conc. hydrochloric acid to pH8, this was combined with 3.73g EDC (1-ethyl-3-dimethylaminopropyl carbodiimide) (Sigma-Aldrich) and made up to 12ml with sterile water]. The Hybond-NX membrane was placed on top of the soaked filter paper. This was wrapped in Saran wrap and baked at 60°C for 1-2 hours. The Hybond-NX membrane (Amersham Biosciences) was pre-hybridised for 1-2 hours at 37°C with 10ml ULTRAhyb-oligo buffer (Life Technologies). Simultaneously, the probe solution [1µl 10µM oligonucleotide (Sigma-Aldrich) without 5' phosphate, 4µl 5x forward buffer (Invitrogen), 12µl sterile water (Fisher Scientific), 2µl ATP [ $\gamma$ -<sup>32</sup>P] (3000Ci/mmol) [Perkin Elmer (Cambridgeshire)], 10U T4 polynucleotide kinase (Invitrogen) for hybridisation was incubated at 37°C for 1 hour. The probe solution was added to the Hybond-NX membrane (Amersham Biosciences) and the membrane was incubated in a hybridisation oven at 37°C for ~16 hours. The membrane was washed for 3x 20 minutes at 37°C using wash buffer (0.2x SSC; 0.1% (w/v) SDS). The final wash was poured away and the membrane wrapped in Saran wrap and placed in a cassette RNA side up with Fujifilm plate on top at 4°C to develop for 5 days.

The film was imaged on a Molecular Imager (BIO-RAD) using imaging software Quantity One. Membranes were stripped with boiling SSC (0.1% (w/v) SDS/5mM EDTA) and stored at -20°C. Oligonucleotide sequences are documented in Appendix 2.

### **2.13 LDH Cytotoxicity Assay (Promega)**

This assay was carried out as per manufacturer's instructions for 'CytoTox96 Non-Radioactive Cytotoxicity assay' (Promega).

### **2.14 Deep Sequencing Of MicroRNAs**

Total RNA was extracted from primary chondrocyte cells taken immediately after digestion from osteoarthritic human knee cartilage from three patients using mirVana™ miRNA Isolation Kit (Ambion) (see section 2.7.3). Small RNAs were enriched from

10µg total RNA using mirVana™ miRNA Isolation Kit (Ambion) (see section 2.7.4). The small RNA library was prepared using Illumina Small RNA V1.5 Sample Preparation Guide, however the sRNA adaptors were substituted with High Definition (HD) adapters (Sorefan et al., 2012).

200ng RNA enriched for small RNA was ligated to adenylated 3' HD adapter with truncated T4 RNA ligase 2 [New England Biolabs (Hertfordshire)]. The ligated fragment was then ligated to 5' HD adapter using T4 RNA ligase 1 (New England Biolabs). The ligated fragment was reverse transcribed as per protocol and size fractionated on an 8% (w/v) PAGE gel. A band corresponding to 100bp was gel purified and analysed on an Illumina Genome Analyzer IIX with 50nt read length [Baseclear (Netherlands)].

## **2.15 MicroRNA siRNA Transfection**

Primary osteoarthritic cells isolated from KOA cartilage tissue (see section 2.5) were plated at  $2.5 \times 10^5$  per well of a 6 well tissue culture plate (Fisher Scientific) and incubated overnight at 37°C 5% (v/v) CO<sub>2</sub> in DMEM + GlutaMAX™ (Life Technologies) [+10% (v/v) heat inactivated FCS (PAA Laboratories), 100U/ml penicillin, 100µg/ml streptomycin (Gibco)]. Two hours before treatment, the media was replaced with DMEM + GlutaMAX™ (Life Technologies) [+10% (v/v) heat inactivated FCS (PAA Laboratories)].

Cells were transfected in serum-free and antibiotic-free DMEM + GlutaMAX™ (Life Technologies) [using Lipofectamine 2000 (Invitrogen)] with novel microRNA mimics at 30nM (Qiagen), novel microRNA inhibitors at 50nM (Qiagen), or non-targeting controls [All Stars negative control at 30nM (Qiagen), miScript Inhibitor negative control at 50nM (Qiagen)]. Non-targeting siRNAs and mock transfections were used as further controls.

Cells were incubated for 6 hours at 37°C 5% (v/v) CO<sub>2</sub> in serum-free and antibiotic-free DMEM + GlutaMAX™ media (Life Technologies). The media was then replaced with DMEM + GlutaMAX™ (Life Technologies) [+10% (v/v) heat inactivated FCS (PAA Laboratories) 100U/ml penicillin, 100µg/ml streptomycin (Gibco)] and transfected cells were incubated for 48 hours at 37°C 5% (v/v) CO<sub>2</sub>.

RNA was extracted using the mirVana™ miRNA Isolation Kit (Ambion) (see section 2.7.2). Gene and microRNA expression analysis was carried out using quantitative real time PCR (see sections 2.10.1 and 2.10.2).

## **2.16 Functional Analysis of Novels 2, 7 and 11 – Whole-Genome Expression Using Bead Arrays Microarray (Source Bioscience)**

Primary chondrocyte cells isolated from KOA patients (see section 2.5) were seeded at passage 1 at  $2.5 \times 10^5$  per well of a 6 well tissue culture plate (Fisher Scientific) and incubated overnight at 37°C 5% (v/v) CO<sub>2</sub> in DMEM + GlutaMAX™ (Life Technologies) [+10% (v/v) heat inactivated FCS (PAA Laboratories), 100U/ml penicillin, 100µg/ml streptomycin (Gibco)]. Cells were transfected with siRNA mimics and inhibitors for novel microRNAs 2, 7 and 11 (see section 2.15).

RNA was extracted using the mirVana™ miRNA Isolation Kit (see section 2.7.2) and quantified with Nanodrop 2000 (Thermo Scientific). RNA samples were subsequently cleaned using an ethanol precipitation (see section 2.8), and analysed on the Experion™ system (BIO-RAD) with Experion RNA StdSens Analysis Kit (BIO-RAD) for RNA quality. 1µg of each transfected RNA sample with an RQI of over 7 were sent to Source Bioscience (Nottingham) for an Illumina Human HT12v4 whole genome microarray.

## **2.17 De-differentiation Assay of Human Articular Osteoarthritic Knee Primary Chondrocytes**

For this assay, RNA was extracted from 6 stages of primary chondrocytes from the same patient (see section 2.7). RNA was firstly extracted from KOA Cartilage, and KOA post digest cells. Cells digested from cartilage were then seeded at  $4 \times 10^4$  cells/cm<sup>2</sup> into 2x 25cm<sup>2</sup> tissue culture flasks (Fisher Scientific) in DMEM + GlutaMAX™ (Life Technologies) [+10% (v/v) heat inactivated FCS (PAA Laboratories) 100U/ml penicillin, 100µg/ml streptomycin (Gibco)], incubated at 37°C, 5% (v/v) CO<sub>2</sub> and allowed to adhere for ~8 days. This stage was labelled as passage 0 (P0). At ~80-90% confluency the cells from one flask were lysed and underwent RNA extraction (see section 2.7.2). The cells within the second flask were washed in HBSS (Life Technologies) and passaged with 0.25% (w/v) trypsin-1mM EDTA (Life

Technologies). Trypsinised cells were seeded at a 1/5 concentration into 2x 25cm<sup>2</sup> tissue culture flasks (Fisher Scientific) and labelled passage 1 (P1). This RNA extraction and passage stage was repeated until RNA from each stage P0 to P3 was obtained. RNA from each stage underwent reverse transcription (see sections 2.9.2 and 2.9.3) and qRT-PCR analysis (see sections 2.10.1 and 2.10.2).

## **2.18 In situ Hybridisation**

### **2.18.1 Mouse Tissue In Situ Hybridisation Preparation**

Strain C57BL/6 mouse embryos were obtained at stages E11.5, limbs at stage E14.5. Tissue was fixed with 4% (v/v) Paraformaldehyde (PFA) (Sigma-Aldrich) in PBS and dehydrated with increasing MeOH (Fisher Scientific) concentration washes from 25% to 100% made up in PBST [PBS with 0.1% (v/v) Tween 20 (Sigma-Aldrich)]. Samples were stored at -20°C.

### **2.18.2 In Situ Hybridisation Of Novel MicroRNA In Mouse Embryo**

Tissue was rehydrated with decreasing MeOH (Fisher Scientific) concentration washes from 100% (v/v) to 25% (v/v) made up in PBST. Proteinase K (Sigma-Aldrich) at 10µg/ml was used to digest the tissue for 30 mins followed by 2 PBST washes. Tissue was fixed in 4% (v/v) PFA (Sigma-Aldrich) for 20 minutes and a needle was used to prick the necks of whole embryos.

A pre-hybridisation stage was carried out at 57°C for 2.5 hours in hybridisation buffer consisting of 50% (v/v) formamide (Sigma-Aldrich), 5xSSC, 0.1% (v/v) Tween 20 (Sigma-Aldrich), 10mM citric acid pH6 (Sigma-Aldrich), 50µg/ml heparin (Sigma-Aldrich) and 100µg/ml tRNA (Sigma-Aldrich) made up in PBS. The prehybridisation buffer was replaced with fresh buffer and 20pmol Novel 11 LNA insitu hybridisation probe (Exiqon) and the tissue was incubated overnight at 57°C.

The hybridisation buffer was removed and the tissue washed at 57°C with a second hybridisation buffer containing 50% (v/v) formamide (Sigma-Aldrich), 5xSSC, 0.1% (v/v) Tween 20 (Sigma-Aldrich) and 10mM citric acid pH6 (Sigma-Aldrich). This was followed by washes in decreasing concentrations (75% - 25%) of hybridisation buffer in

2xSSC at 57°C, and then decreasing concentrations (75% - 25%) of 0.2xSSC in PBST at room temperature.

The tissue was washed in blocking solution [2% (v/v) Normal goat serum (Sigma-Aldrich), 2mg/ml BSA (Sigma-Aldrich), 0.1% (v/v) Triton X-100 (Sigma-Aldrich) and 0.05% (v/v) Tween 20 (Sigma-Aldrich) made up in PBS] for a minimum of 2.5 hours at room temperature. The tissue was subsequently incubated overnight at 4°C with Anti-Digoxigenin-AP Fab fragments antibody [Roche (West Sussex)] at a final dilution of 1:5000 in fresh blocking solution.

Over the next 2-3 days, multiple 2 hour washes in room temperature PBST were carried out to remove excess antibody. This was followed by 3-4 days of colour development: 25-50µl BCIP/NBT liquid substrate system (Sigma-Aldrich) in 3.5ml developing solution (100mM Tris-HCl pH9.5, 50mM MgCl<sub>2</sub>, 100mM NaCl, 0.1% (v/v) Tween 20, 2mM Levamisole).

To halt the colour reaction, the tissue was washed with 1mM EDTA (Life Technologies) pH8 in PBS for 15 minutes at room temperature followed by 3x washes with PBST. In order to post-fix the tissue, the samples were incubated in 4% (v/v) PFA (Sigma-Aldrich) for 30 mins, washes with PBST to remove the PFA, and then stored at 4°C in PBST.

## **2.19 Mouse Hip Avulsion Cartilage Wounding Assay**

Mouse femoral heads were dissected out of 3-5 week old C57BL/6 strain mice (beyond 5 weeks, the secondary centre of ossification in the hip epiphysis will have completely mineralised, and cartilage is unobtainable). The hip joint was disarticulated to expose the femoral head and the cartilage surface was twisted off using forceps (Stanton et al., 2011). The cartilage was washed in sterile PBS (Life Technologies) with 100U/ml penicillin, 100µg/ml streptomycin (Gibco). Cartilage samples were incubated in 250µl DMEM (Life Technologies) with 100U/ml penicillin, 100µg/ml streptomycin (Gibco)] at 37°C 5% (v/v) CO<sub>2</sub> in wells of a 96 well tissue culture plate (Fisher Scientific). Tissue was harvested into 500µl Trizol (Life Technologies) at the time points: 1, 3, 6, 12, 24 and 48 hours and stored at -80°C. The 0 hour time-point was collected immediately upon removal of cartilage from the mouse and stored in trizol (Life



Technologies) at -80°C. RNA was extracted (see section 2.7.1) and gene and microRNA expression analysed with qRT-PCR (see sections 2.10.1 and 2.10.2).

## **2.20 Sub-Cloning**

### **2.20.1 Plasmids and gDNA**

PmirGLO Dual-luciferase miRNA Target Expression Vector was purchased from Promega. Genomic DNA was used as a template for PCR amplification. The genomic DNA was extracted from SW1353 cells. SW1353 cells were harvested in 0.25% (w/v) trypsin-1mM EDTA (Life Technologies) and centrifuged at 180g for 5 minutes at room temperature. The cell pellet was washed in 200µl PBS (Life Technologies) and re-suspended in 150µl. 400µl of Lysis buffer [0.1M Tris-HCL (Sigma-Aldrich) pH 7.5, 1% (w/v) SDS], was added followed by 10µl proteinase K (20mg/ml) (Sigma-Aldrich) and the sample was incubated at 50°C for 2 hours. An volume equal to the sample total of phenol: chloroform: isoamyl alcohol (25:24:1) saturated with 10mM Tris, pH8.0, 1mM EDTA (Sigma-Aldrich) was added, this was followed by centrifugation at full speed for 10 minutes. The upper phase was transferred to a fresh centrifugation tube and an equal volume of chloroform (Fisher Scientific) was added, this was mixed by vortex and centrifuged at full speed for 10 minutes. The upper phase was removed and 2x volume of 100% ethanol (v/v) (Sigma-Aldrich) was added, this was centrifuged at full speed for 10 minutes. The resulting DNA pellet was washed with 2ml 70% ethanol (v/v) (Sigma-Aldrich) and centrifuged at full speed for 5 minutes. The pellet was dried at room temperature for 5 minutes and dissolved in 100µl nuclease free water (Sigma-Aldrich).

### **2.20.2 CaCl<sub>2</sub> Competent DH5α Cell Preparation**

DH5α bacterial colonies were grown from glycerol stock on LB/AGAR plate. One colony was inoculated into 5ml LB broth [1% (w/v) Tryptone (Fisher Scientific), 2% (w/v) Yeast Extract (Fisher Scientific) and 0.5% (w/v) Sodium Chloride (Fisher)] and incubated overnight at 37°C, 180rpm. 2x1ml of DH5α culture transferred to 2x 1l flasks and incubated at 37°C, 180rpm for 2.5 hours until the OD of the culture is between 0.4 and 0.6 OD<sub>600</sub>.

The culture was split into 50ml ice cold Falcon tubes (Fisher Scientific) and incubated on ice for 20 minutes. This was followed by centrifuging the cultures for 10 minutes at 1500g, 4°C. The pellet was gently re-suspended in 5ml 0.1M CaCl<sub>2</sub> (Fisher Scientific) and incubated on ice for 10 minutes before being centrifuged at 1500g, 4°C for 10

minutes. The pellet was re-suspended gently in 1ml ice cold 0.1M CaCl<sub>2</sub> (Fisher Scientific) and incubated on ice for 1 hour. 400µl glycerol (Fisher Scientific) was added to 1ml DH5α and snap frozen in liquid nitrogen and stored at -80°C.

### **2.20.3 Polymerase Chain Reaction**

AccuTaq LA 10x Buffer (Sigma-Aldrich), 5mM dNTPs (Bioline), 400pmol forward and reverse primers (Sigma-Aldrich), 100ng gDNA, 1µl DMSO (Sigma-Aldrich), 37µl sterile water (Fisher Scientific) and 5U AccuTaq<sup>TM</sup> LA DNA Polymerase (Sigma-Aldrich) were mixed on ice and cycled through the following cycling conditions: 1) 95°C: 2mins, 2) 95°C: 1min, 3) 50-60°C: 50sec, 4) 68°C: Extension time, 5) 68°C 10mins, 6) 4°C: ∞. Primer sequences are documented in Appendix 3.

### **2.20.4 Mini-Prep – QIAprep Spin Miniprep Kit [Qiagen (Manchester)]**

The protocol was followed as per manufacturer's instructions.

### **2.20.5 Mini-Prep – QIAprep Spin Miniprep Kit (Without Filters) (Qiagen)**

1.5ml of bacterial culture was pelleted at full speed in a centrifuge for 1 minute. The bacterial pellet was re-suspended in 100µl mini-prep buffer P1 (Qiagen), to this 200µl mini-prep buffer P2 (Qiagen) was added and mixed by inversion. This was followed by the addition of 150µl mini-prep buffer N3 (Qiagen) and mixed by inversion. The prep was centrifuged at full speed in a centrifuge at room temperature for 3 minutes and the supernatant was removed to a new tube and 1ml of 100% ethanol (Sigma-Aldrich) was added. This was vortexed and centrifuged at full speed for 10 minutes. The supernatant was removed and the pellet washed with 70% (v/v) ethanol (Sigma-Aldrich). The pellet was allowed to air dry and then re-suspended in 40µl sterile water.

### **2.20.6 Gel Purification [Zymoclean (Cambridge)]**

Gel purification of plasmids and cDNA was carried out as per manufacturer's instructions with the following exceptions; a third wash step was added (200µl DNA wash buffer), and the sample in the filter was centrifuged for 10 minutes. This was followed by a room temperature incubation for 15 minutes and elution with 12µl elution buffer.

### **2.20.7 Phenol: Chloroform Purification**

The DNA volume was made up to 200µl with sterile water. 200µl phenol: chloroform: isoamyl alcohol (25:24:1) saturated with 10mM Tris, pH8.0, 1mM EDTA (Sigma-Aldrich) was added and the sample was vortexed and centrifuged for 10mins at full speed at room temperature. The aqueous layer was removed to a fresh 1.5ml microcentrifuge tubes, 2.5x aqueous volume of 100% (v/v) ethanol (Sigma-Aldrich) and 0.1x aqueous volume 3M sodium acetate (pH5.2) were added. The tube was inverted to mix and centrifuged for 10mins at full speed. The supernatant was removed and the pellet washed with 70% (v/v) ethanol (Sigma-Aldrich) and centrifuged for 1min at full speed. The pellet was air dried and re-suspended in 20µl sterile water.

### **2.20.8 Restriction Digests**

#### **2.20.8.1 SalI Digestion Of Plasmid And Inserts**

Plasmids and subcloned cDNA were digested prior to ligation. 5µg plasmid DNA or 1µg subcloned DNA were mixed with 10x buffer D (Promega), and 20U SalI restriction enzyme (New England Biolabs) in a total 50µl reaction at 37°C for 1 hour.

#### **2.20.8.2 Directional Digest**

In order to ensure correct direction of insert in pmirGLO vector, site specific digestions were carried out. The reaction consisted of 10U restriction enzyme, 10x restriction buffer, 1µg/µl BSA (Promega), and 100ng construct made up to 10µl with sterile water. This was incubated at 37°C for 1 hour and then separated on a 1% (w/v) agarose gel to assess band size.

### **2.20.9 Vector De-Phosphorylation**

0.1x volume of 10x Antarctic Phosphatase reaction buffer (Biolabs) and 5U Antarctic Phosphatase (Biolabs) were added to SalIHF (high fidelity) (Biolabs) digested vector and incubated for 15 mins at 37°C.

### **2.20.10 GoTaq (Promega) dATP Extension Of Products**

After the PCR with Accutaq (Sigma-Aldrich), an A-tail extension was carried out upon the dsDNA. The reaction mix consisted of 5x clear GoTaq buffer, 1.25U GoTaq DNA Polymerase, 2mM dATP and DNA template (see section 2.20.6), made up to 50µl with sterile water. This was incubated at 72°C for 15 minutes.

#### **2.20.11 Ligation Using pGEM-T Easy Vector System (Promega)**

The protocol was followed as per manufacturer's instructions.

#### **2.20.12 Transformation of dsDNA in pGEM-T Easy Vector System (Promega)**

2µl ligation reaction (see section 2.20.11) was added to 80µl DH5α competent *E.coli* cells and incubated on ice for 20 minutes before being incubated for 30 seconds at 42°C. 950µl LB was added and incubated for 2 minutes on ice, followed by a 90 min incubation at 37°C in a shaking incubator. 100µl of the transformation was plated onto LB agar plates with 100µg/µl AMP, 0.003% (v/v) X-GAL, 0.1mM (v/v) IPTG and incubated at 37°C overnight.

#### **2.20.13 Ligation and Transformation Of Inserts With pmirGLO Vector**

Ligations of 3:1 ratio of DNA to plasmid were used. The ligation reaction consisted of 5x ligase buffer (Invitrogen), 1mmol ATP (Sigma-Aldrich), 2U T4 DNA ligase (Invitrogen) were mixed on ice in a 20µl reaction with 100ng digested vector and between 25-50ng insert. The ligations were incubated overnight at 14°C.

Half the ligation reaction was added to 50µl DH5α competent *E.coli* cells and incubated on ice for 20 minutes before being incubated for 30 seconds at 42°C. 300µl of LB was added and incubated for 2 minutes on ice, followed by a 90 min incubation at 37°C in a shaking incubator. Transfections were plated onto LB agar plates with 100µg/µl AMP and incubated at 37°C overnight.

#### **2.20.14 Source Bioscience Sequencing**

1µg of each plasmid was sent to Source Bioscience for sequencing using Source Bioscience stock primers T7 Reverse.

#### **2.20.15 Transformation of Constructs And Glycerol Stock Production**

100ng pmirGLO plasmid was added to 50µl DH5α competent *E.coli* cells and incubated on ice for 20 minutes before being heat shocked for 30 seconds at 42°C and finally

incubated on ice for 2 minutes. The transformation was plated onto LB agar plates with 100µg/µl AMP and incubated at 37°C overnight.

One colony was grown overnight in 5ml LB media with 100µg/µl AMP at 37°C, 180rpm. 500µl of culture was mixed with 500µl sterile glycerol (Fisher Scientific) and stored at -80°C.

## **2.21 Transfection Of Cloned Constructs Into DF1 Cells**

DF1 cells were plated onto a 96 well plate (Fisher Scientific) at a density of  $3.75 \times 10^4 / \text{cm}^2$  in DMEM + GlutaMAX™ (Life Technologies) [+10% (v/v) heat inactivated FCS (PAA Laboratories)] and incubated overnight at 37°C, 5% (v/v) CO<sub>2</sub>. The first transfection treatments were prepared, treatment one: containing (per well) 100ng of cloned DNA construct (in pmir-GLO) with 50nM of either microRNA mimic (Qiagen), inhibitor (Qiagen), non-targeting control [All Stars negative control (Qiagen), miScript Inhibitor negative control (Qiagen)] or a mock treatment and made up to 25µl with DMEM + GlutaMAX™ (Life technologies). Non-targeting siRNAs and mock treatments were used as negative controls. This treatment was incubated at room temperature for 5 mins.

The second transfection preparation; Lipofectamine 2000 (Invitrogen) made up to 25µl with DMEM + GlutaMAX™ (Life Technologies) was incubated at room temperature for 5 mins (per well).

25µl of the Lipofectamine 2000 preparation was added to each siRNA and mock preparation (per well) and incubated at room temperature for 20 minutes. The final treatment was added to the cells in a final volume of 100µl per well and incubated for 24 hours at 37°C, 5% (v/v) CO<sub>2</sub>. 6 hours after transfection, culture media was replaced with DMEM + GlutaMAX™ (Life Technologies) [+10% (v/v) heat inactivated FCS (PAA Laboratories)].

## **2.22 Luciferase Assay – Dual Luciferase Reporter Assay System (Promega)**

DF1 cells transfected with pmirGLO constructs (see section 2.21) were harvested as follows. Media was removed from wells of 96 well tissue culture plate (Fisher

Scientific), and cells washed with cold sterile PBS solution (Life Technologies). All liquid was subsequently removed before 70µl 1x cell lysis buffer (Promega) was added per well and left to gently shake for 30 minutes at room temperature. 10µl from each well was transferred to an opaque white 96 well plate, 50µl of re-suspended Luciferase Assay Substrate (Promega) was added per well and firefly luciferase activity measured on a spectrophotometer [EnVision 2103 Multilabel plate reader (Perkin Elmer)] (Luminescence 700 setting, 560nm) using Wallac EnVision Manager software. An additional 50µl Stop and Glow solution with 2% (v/v) substrate (Promega) was added per well and Renilla activity measured. Firefly luciferase relative light units were normalised to Renilla relative light units to account for the transfection efficiency.

## **2.23 Data Analysis, Statistics And Bioinformatics**

### **2.23.1 Data Analysis Of qRT-PCR – Comparative C<sub>t</sub> Method**

In order to compare across all genes and microRNAs profiled with qRT-PCR (see section 2.10.1 and 2.10.2), the efficiency of primers and probes were assumed to be equal and the data was transformed by  $\log_{10}2^{-\Delta C_t}$ , where:

$$\Delta C_t = \text{target gene } C_t - \text{housekeeping gene } C_t$$

(NB. *18S* RNA for mRNA or *U6* RNA for microRNA). Primer efficiency was checked with standard curve analysis prior to analysis using the comparative Ct method.

### **2.23.2 Data Analysis Of qRT-PCR – Standard Curve Method**

Relative gene expression was determined with standard curves generated from a positive control, to allow for efficiency of the primers and reaction. All genes measured were then normalised to *18S* RNA C<sub>t</sub>s for mRNA or *U6* RNA C<sub>t</sub>s for microRNAs. Relative quantification (RQ) of expression levels were determined by the following steps carried out for each gene and *18S* RNA.

- 1)  $RQ = (C_t - \text{slope}) / \text{intercept}$
- 2) Transforming to the inverse  $\log_{10}$
- 3) Normalising of the gene (step 2) to (step 2) for *18S* RNA

### **2.23.3 Statistical Analysis Of qRT-PCR Data**

The statistical methods used in the analysis of data included one-way ANOVA with a post hoc Tukey test to compare between more than two samples, using the statistical

software package 'PASW Statistics 18' (formerly SPSS) and Prism v.5, and the Student t-test which compares between two samples, using Prism v.5.

#### **2.23.4 NGS Deep Sequencing Bioinformatics**

The dataset obtained from NGS sequencing (see section 2.14) was initially analysed with miRCat online software [microRNA Categoriser - The UEA sRNA toolkit: animal version (Moxon et al., 2008)] by biomathematician Dr Helio Pais (School of Computing, UEA).

The MiRCat software maps the sequences obtained from the sequencing project to the genome of interest (in this case human). It then identifies regions with sRNAs containing at least 5 reads. Other criteria taken into account by miRCat are: the identified loci must not have more than four non-overlapping sRNAs, each sRNA must be 200nt or less away from another sRNA, and at least 90% of the identified sRNAs in the locus must have the same orientation.

The second stage carried out by miRCat analysis includes choosing the most abundant sRNA read at any locus as the most likely microRNA. It then obtains the sequences surrounding the sRNA to analyse with RNAfold to determine the sRNA secondary structure (of which the usual microRNA structure is a stable, partially double stranded stem-loop, often called a hairpin).

The secondary structure is subsequently trimmed and analysed to determine 1) the number of sequential mismatches between the microRNA complementary strands (miRNA and miRNA\*) of which there must be fewer than 3. 2) The number of paired nucleotides between the microRNA complementary strands of which there must be at least 17 around the ~25nt of the mature sequences. 3) The length of the hairpin structure must be at least 50nt. And 4) At least 50% of the bases within the hairpin structure must be paired.

Finally, the hairpin structure with the lowest adjusted minimum free energy is chosen to be the most likely precursor miRNA for the aforementioned miRNA candidate.

After the initial miRCat analysis of the sequencing dataset, which generated a large number of candidate microRNAs, further analysis was carried out to triage the dataset down to a manageable number to validate. Potential novel microRNA candidates were filtered based on their read number, presence in a DLD-1 cell line NGS dataset and fold-change in comparison to NGS dataset for Dicer knock out cell line (both datasets obtained by Professor T. Dalmay, UEA Norwich). In addition, the presence of an exact (or near exact) complementary miRNA strand in datasets obtained from miRCat (Moxon et al., 2008) were also used as criteria. In addition to triaging for novel microRNAs, the entire dataset were periodically run through the online microRNA database ‘miRBase’ (Griffiths-Jones et al., 2006; Griffiths-Jones et al., 2008; Kozomara and Griffiths-Jones, 2011) to check for any known microRNAs which may be present.

#### **2.23.5 MicroRNA Target Search Using RStudio**

The computer language R was utilised with open source RStudio (R Core Team, 2013) software and the open source package Biostrings\_2.28.0.tar.gz (Pages et al.).

To search for novel microRNA targets in 3’UTRs of the human genome as downloaded from Ensembl Biomart.

The R code used for this analysis is as follows:-

```
1) Utr3<-readDNASTringSet("Doc.")
2) a<-vcountPattern("seq",Utr3)
3) k<-strsplit(names(utr3),"\\ ")
4) m=c()
   for(i in k){
     m=rbind(m,i[1])
5) Targets<-cbind(a,m)
```

Genes determined to have novel microRNA target sites within their 3’UTR were ranked based on number and size of target match.

#### **2.23.6 Candidate novel microRNA species conservation analysis**

The candidate sequences novel 2, novel 7 and novel 11 were analysed for their presence in the genome of all species characterised using Ensembl (Flicek et al., 2014). Exact



sequence matches were labelled in green with all other matches being partial or in different genomic locations to the equivalent sequence in humans.

The estimated time of divergence between all species investigated was determined using the online software 'Time Tree: The Timescale of Life' (Hedges et al., 2006; Kumar and Hedges, 2011).

These data were combined into an evolutionary tree of all species containing an exact or partial match to the candidate novel microRNAs using Microsoft office PowerPoint.

### **2.23.7 Bioinformatics For Genome Array**

The dataset received from Source Bioscience for the whole genome microarray (see section 2.16) was normalised using RStudios with the Lumi package (R Core Team, 2013) by Miss Linh Le (School of Biological sciences, UEA). First quality assessment was carried out to identify any poor quality arrays by looking for unusual signal distributions. This was followed by background correction and normalisation by a between-array quantile method.

Normalised data was analysed to find fold change expression between each microRNA siRNA and the corresponding scrambled control. This was combined with the novel microRNA computational target search (see section 2.23.5). Potential gene targets were ranked based on their fold change expression (potential target expression increased with inhibitor and decreased with mimic) and number and size of novel microRNA target sites.

# **Chapter 3 – MicroRNA Profiling of Osteoarthritic Cartilage**

### **3.1 Introduction**

Chapter 1 described the importance of microRNAs in osteoarthritis and cartilage development in addition to their vital role in many other pathologies and developmental processes.

Online databases such as miRBase (Griffiths-Jones et al., 2006; Griffiths-Jones et al., 2008; Kozomara and Griffiths-Jones, 2011) describe the thousands of microRNAs which have already been discovered in humans as well as across many other species. However, the total population of microRNAs within humans and in any specific tissue is still largely unknown. MicroRNA expression and populations have been characterised in a number of tissues such as the thymus, testes and placenta (Barad et al., 2004), and microRNA arrays have mapped microRNA expression in others such as bone marrow, skeletal muscle, brain, breast, lung, spleen, heart, kidneys, liver, ovary and prostate, as well as a number of foetal tissues (Liu et al., 2004). The drawback with these studies however, is that a limited number of around 300 microRNAs are chosen for the custom microarrays. This is only 9.5% of the total number of human microRNAs registered on miRBase (Griffiths-Jones et al., 2006; Griffiths-Jones et al., 2008; Kozomara and Griffiths-Jones, 2011) and thus not a thorough analysis of what could potentially be present in these tissues.

To this end, in order to establish the entire microRNA population and relative expression of microRNAs within articular cartilage, a deep sequencing project was carried out. This does however bring its own problems in that the technology behind sequencing is in itself not perfect. Significant bias towards specific nucleotide sequences can be seen across existing platforms as described in a paper from Sorefan et al. (Sorefan et al., 2012). This is supported by further literature investigating bias detection in existing platforms (Hafner et al., 2008; Jayaprakash et al., 2011; Sun et al., 2011a).

The apparent strong bias results in a potentially large number of microRNAs being overlooked, as the system will detect only certain sequences. This is due to a preference of the RNA ligases used in small RNA cloning and library preparation, for certain secondary structures. The efficiency of RNA ligases has been shown to be dependent upon the secondary structure at the site of ligation (Nandakumar and Shuman, 2004; Yin et al., 2003). To support this, Sorefan et al. compared the minimum free energy of secondary structures of test small RNA libraries with detection rates from the

sequencing results and discovered the sequences with the most stable structures were those most commonly detected (Sorefan et al., 2012).

To measure bias, Sorefan et al. used a synthesised small RNA library composed of randomised oligonucleotide substrates where all sequences were of the same known concentration and therefore should all be detected at the same rate by the platform. When processed through an existing Illumina platform, only 42% of the presumed sequences were detected, this suggests that previous microRNA sequencing projects may have missed out on a large subset of the population. In addition Sorefan et al. discovered that the existing microRNAs reported in miRBase appear to be of a structure preferred by the 'biased' adapters.

To address this bias, Sorefan et al. developed modified, 'High Definition' (HD), degenerate adapters for use in the small RNA library preparation. These consisted of four random nucleotides added to the 5' end and 3' end of the existing 3' and 5' adapters respectively. When run through the same platform, they detected 81% of the sequences assumed to be present in the synthesised library. This is a substantial increase, and whilst there is obviously still some bias in sequence detection it is not as extreme as before. It appears that although the bias seen from the RNA ligases towards specific small RNA structures may have been resolved, there is still an issue surrounding the bias of detection which needs to be studied. Additional research from other labs has also investigated the issue of sequence bias in next generation sequencing and have proposed similar solutions to increase the detection rate of microRNAs (Jayaprakash et al., 2011; Sun et al., 2011a).

The new HD adapters developed by Sorefan et al. were utilised within the small RNA sequencing experiment on the Illumina platform for this project in an attempt to detect as many known and potentially novel microRNAs as possible within the still existing limitations of the method. The aim was to identify (as far as possible) the microRNA population of osteoarthritic cartilage with an aim of furthering the understanding of which microRNAs may be involved in the disease pathogenesis and progression, as well as chondrogenesis.

Small RNA was extracted from primary chondrocytes isolated from osteoarthritic knee cartilage obtained from total knee replacement surgery, in addition to extraction from cartilage tissue and primary chondrocytes grown in monolayer. A small RNA library from chondrocytes immediately after digestion from cartilage was created using the

Illumina method along with the aforementioned HD adapters. This was subjected to deep sequencing and the data were analysed using miRCat software (Moxon et al., 2008).

## 3.2 Results

### 3.2.1 Expression of key microRNAs in primary chondrocytes

In order to examine the expression of known key microRNAs in cartilage and isolated or cultured chondrocytes, both strands of the microRNAs miR-140 and miR-455 were measured using qRT-PCR from RNA purified from snap frozen cartilage tissue, chondrocytes isolated directly from the digestion media [termed Post Digest (PD)], passage 0 chondrocytes (P0) cultured in monolayer and allowed to reach confluency, and passage 1, 2 and 3 primary chondrocytes allowed to reach confluency. This de-differentiation assay showed a decreasing trend in expression from cartilage tissue to passage 3 for miR-140-3p (A), miR-455-3p (C) and miR-455-5p (D). For miR-140-5p expression in both cartilage and post digest chondrocytes was statistically significant from all other stages with  $p \leq 0.001$  (Fig. 3.1).

MiR-455-3p (C) expression in cartilage was significantly different from all other stages of chondrocytes apart from Post Digest, with  $p \leq 0.001$  between cartilage and cells at passage 0, 2 and 3, and a  $p \leq 0.01$  between cartilage and cells at passage 1. Chondrocytes from Post Digest showed expression significantly different from all subsequent passages with  $p \leq 0.001$  between PD and passage 3,  $p \leq 0.01$  between PD and passage 2 and  $p \leq 0.05$  between PD and passages 0 and 1 (Fig. 3.1).

MiR-455-5p (D) expression in cartilage showed a significant difference between all other stages of chondrocytes with  $p \leq 0.01$  between cartilage and Post Digest and  $p \leq 0.001$  between cartilage and all subsequent monolayer chondrocyte stages. Expression at the Post Digest level showed significance between passages with  $p \leq 0.01$  between PD and passage 0 and  $p \leq 0.001$  between PD and subsequent passages (Fig. 3.1).

In contrast to miR-140-3p, miR-455-3p and miR-455-5p which all showed a decreasing trend in expression from cartilage to passage 3, miR-140-5p (B) increases in expression from cartilage to the post digest stage before falling again through subsequent passages to passage 3 ( $p < 0.001$ ). Expression between PD and all other samples for miR-140-5p was significantly different with  $p \leq 0.001$ . The expression of miR-140-5p in cartilage also showed statistical significance compared to monolayer cells with  $p$  values of  $\leq 0.01$  between passages 1, 2 and 3 (Fig. 3.1).

The expression of seven human collagen genes was also measured across the de-differentiation assay (Fig. 3.2). COL1A1 (A) showed a statistically significant increase in expression from cartilage tissue to passage 3. A significant difference was seen between cartilage and passage 3 with  $p \leq 0.01$ . There was a significant difference of  $p \leq 0.05$  between PD and passage 2, and  $p \leq 0.001$  between PD and passage 3. A significant value of  $p \leq 0.05$  was also seen between passage 0 and passage 3.

The embryonic version of COL2A1 (VA) (B) did not show a specific trend in expression with a higher expression level seen at the cartilage as well as the passage 2 stage compared to all other chondrocyte passages, however none of these changes were significantly different. In comparison the adult version of collagen 2A1 (VB) (C) showed a strong decrease in expression from cartilage across all the passages. A significant difference was seen between cartilage tissue and all passaged chondrocyte stages with  $p \leq 0.001$  between cartilage and passages 0,  $p \leq 0.001$  between cartilage and passage 2 and  $p \leq 0.001$  between cartilage and passage 3. In addition, a significant difference of  $p \leq 0.01$  was seen between cartilage and passage 1.

Expression of COL3A1 (D) did not show a trend across de-differentiation with none of the changes showing statistical significance.

COL9A1 (E) showed a statistically significant decreasing pattern of expression through chondrocyte de-differentiation with  $p \leq 0.001$  between cartilage and all cultured chondrocytes. Similarly, COL10A1 (F) also had a pattern of decreasing expression from cartilage to passage 3, however none of the changes in expression were significantly different.

COL11A1 (G), showed a similar expression pattern to COL1A1 with expression increasing from cartilage to isolated and cultured chondrocytes up to passage 3. Statistics showed a significance between cartilage and passage 3 ( $p \leq 0.05$ ) and the Post Digest stage and passage 3 ( $p \leq 0.01$ ).

Finally COL27A1 (H) expression generally decreased through de-differentiation with the highest expression level seen in cartilage tissue. None of the changes in expression were statistically significantly different.

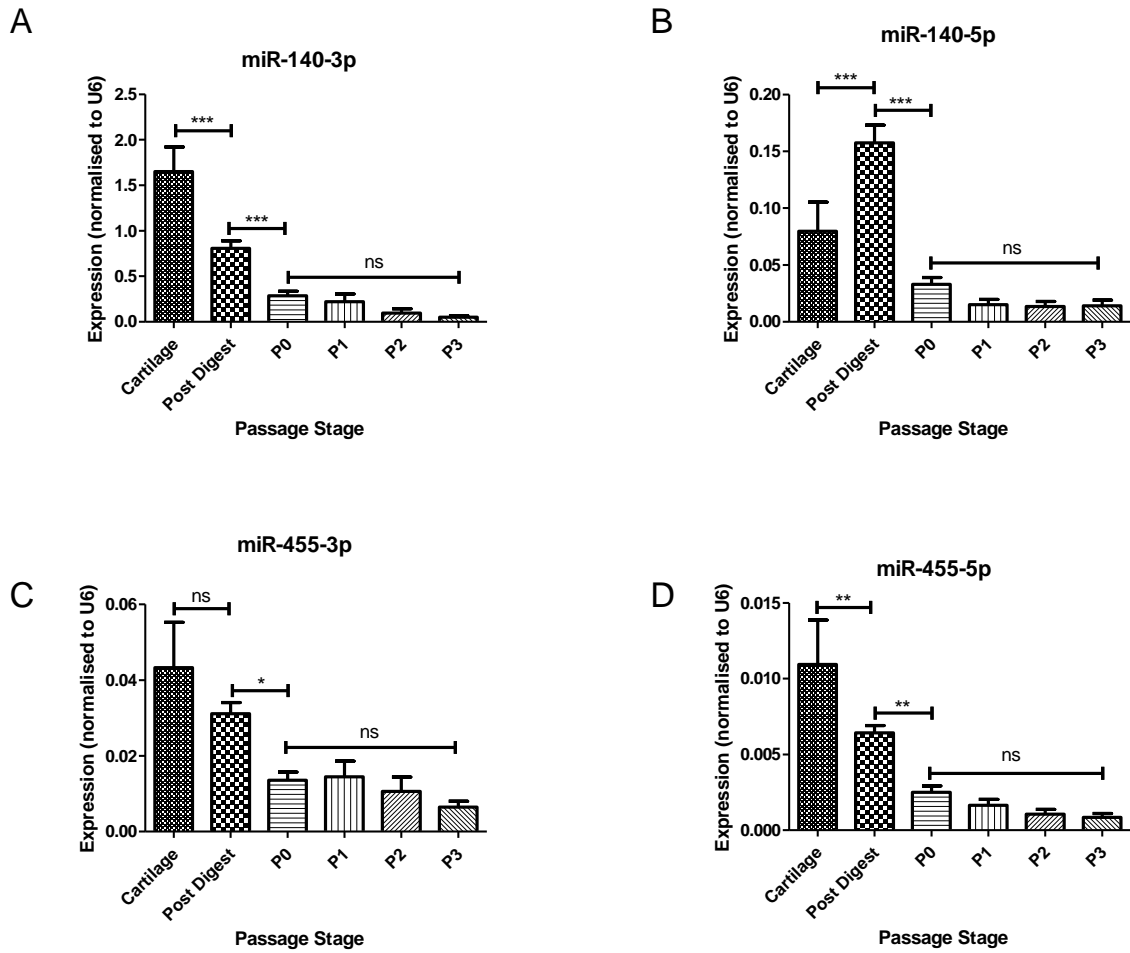
Figure 3.3 shows the expression of MMP1 (A), MMP13 (B) and aggrecan (C) in the de-differentiation assay. MMP1 showed low expression in cartilage tissue which then increased significantly at the post digest stage ( $p \leq 0.01$ ). Expression then steadily decreased across the subsequent passages. Significance is seen between cartilage tissue and passage 0 ( $p \leq 0.05$ ), post digest with passage 1 ( $p \leq 0.05$ ), between PD and passage 2 ( $p \leq 0.01$ ), and between PD and passage 3 ( $p \leq 0.001$ ).

MMP13 (B) shows a similar pattern to MMP1, with expression initially low in cartilage tissue and isolated cells before increasing at passage 0 and then decreasing through passage 1 to 3. There is a statistically significant difference seen between cartilage and passage 0 ( $p \leq 0.001$ ), post digest and passage 0 ( $p \leq 0.05$ ), and finally passage 0 with passage 2 ( $p \leq 0.001$ ) and 3 ( $p \leq 0.001$ ).

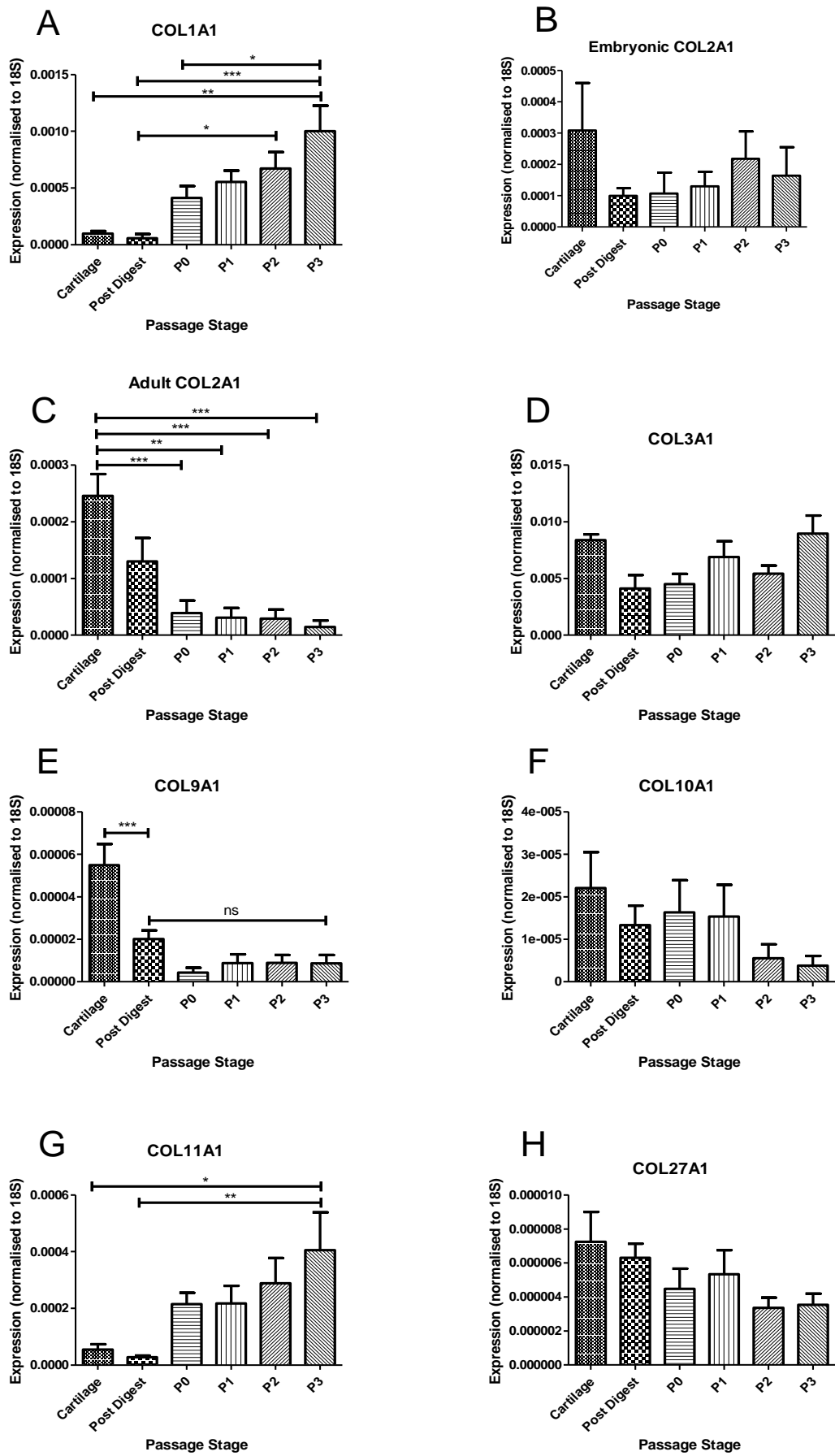
Aggrecan (C) expression decreased slightly from cartilage to passage 0 before increasing again to passage 3; however there was no statistical significance.

Morphology changes in the de-differentiation assay were shown in the images from (Fig. 3.4). Chondrocytes become more fibroblastic and elongated in shape from passage 0 to passage 3.



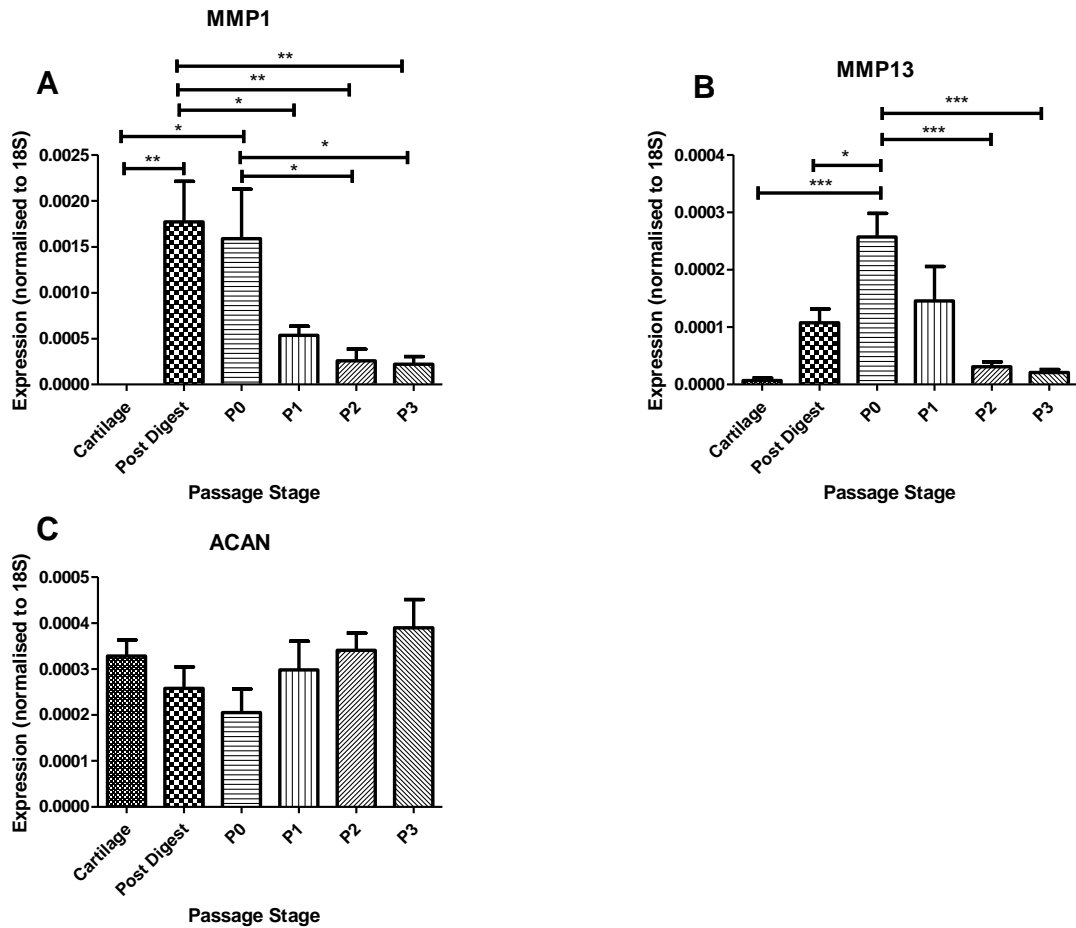


**Figure 3.1 MicroRNA expression in articular chondrocyte isolation and expansion.** Expression levels of (A) miR-140-3p, (B) miR-140-5p, (C) miR-455-3p, (D) miR-455-5p were measured by qRT-PCR from RNA isolated from: osteoarthritic knee cartilage tissue, isolated chondrocytes (post digest) and subsequent P0, P1, P2, and P3 passaged chondrocytes from monolayer culture. Data were normalised to *U6* RNA expression. (\* =  $p \leq 0.05$ , \*\* =  $p \leq 0.01$ , \*\*\* =  $p \leq 0.001$ . RNA was obtained from 8 patients for chondrocytes and 4 patients for tissue. A one way Anova analysis with a post hoc Tukey test was used to test for significance).

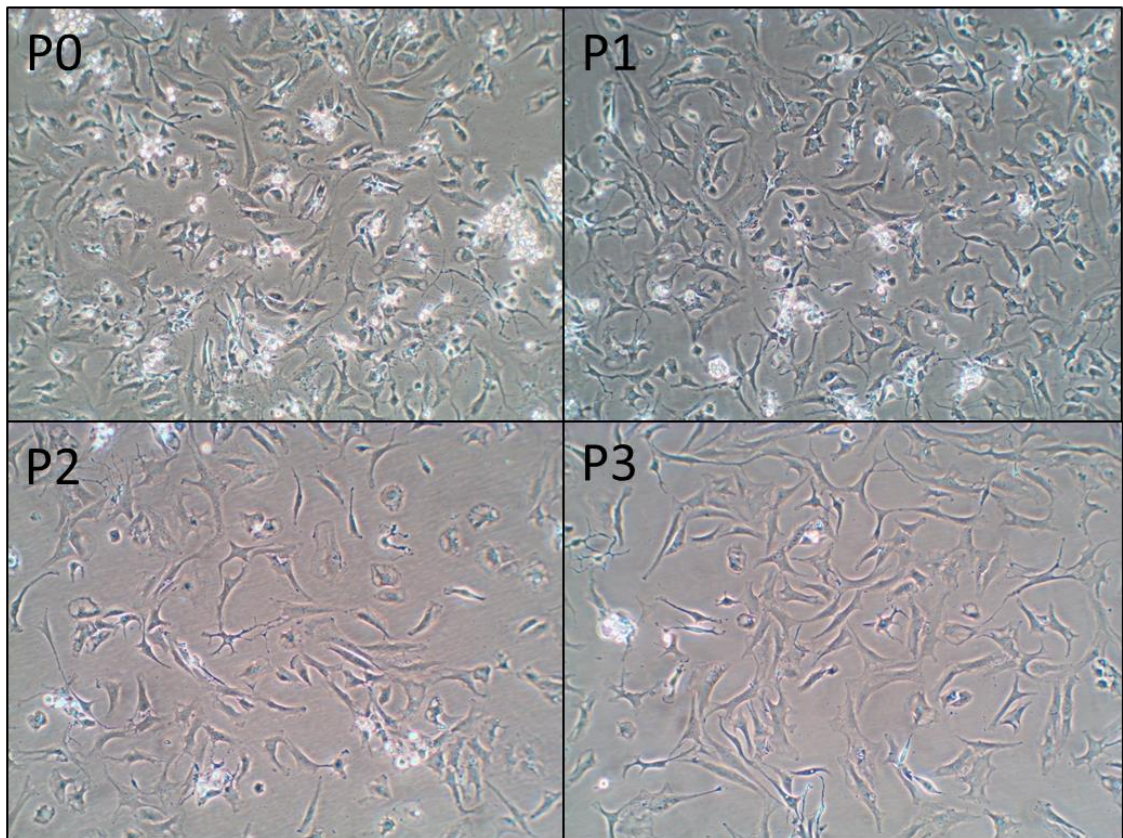


**Figure 3.2 Collagen gene expression in articular chondrocyte isolation and expansion.** Expression levels of (A) COL1A1, (B) COL2A1 VA, (C) COL2A1 VB, (D)

COL3A1, (E) COL9A1, (F) COL10A1, (G) COL11A1 and (H) COL27A1 were measured by qRT-PCR from RNA isolated from: osteoarthritic knee cartilage tissue, isolated chondrocytes (post digest) and subsequent P0, P1, P2, and P3 passaged chondrocytes. Data were normalised to *18S* RNA expression. (\* =  $p \leq 0.05$ , \*\* =  $p \leq 0.01$ , \*\*\* =  $p \leq 0.001$ . RNA was from 8 patients for chondrocytes and 4 patients for tissue. A one way Anova analysis with a post hoc Tukey test was used to test for significance).



**Figure 3.3 Gene expression in articular chondrocyte isolation and expansion.** Expression levels of (A) MMP1, (B) MMP13, (C) ACAN were measured by qRT-PCR from RNA isolated from: osteoarthritic knee cartilage tissue, isolated chondrocytes (post digest) and subsequent P0, P1, P2, and P3 passaged chondrocytes. Data were normalised to *18S* RNA expression. (\* =  $p \leq 0.05$ , \*\* =  $p \leq 0.01$ , \*\*\* =  $p \leq 0.001$ . RNA was from 8 patients for chondrocytes and 4 patients for tissue. A one way Anova analysis with a post hoc Tukey test was used to test for significance).



**Figure 3.4 Images of primary chondrocytes.** Obtained from primary chondrocytes isolated from osteoarthritic knee cartilage tissue. Each image represents passages 0, 1, 2, and 3 in sequence and shows elongation of chondrocytes at passages 2 and 3 as well as slower growth.

### **3.2.2 Profile of existing microRNAs in osteoarthritic primary chondrocytes**

Due to the high level of miR-140-5p (a microRNA heavily implicated in osteoarthritis) (Araldi and Schipani, 2010; Swingler et al., 2012; Tardif et al., 2009) expression seen at the post digest stage in previous experiments compared to cartilage tissue and cultured chondrocytes, and the difficulty with obtaining high enough quality RNA directly from cartilage; RNA was extracted from chondrocytes isolated immediately after digestion to use with deep sequencing on the Illumina system with the addition of newly developed HD adapters (Sorefan et al., 2012). This was carried out with the aim to identify novel microRNAs with a potential role in osteoarthritis, in addition to establishing the entirety of the human osteoarthritic miRNome. A small RNA library was created following 'Illumina Small RNA V1.5 Sample Preparation' guidelines using RNA from the knee cartilage of 3 osteoarthritic individuals. HD adapters (Sorefan et al., 2012) were utilised in order to reduce sequence bias and the samples were run on an Illumina Genome Analyzer IIX.

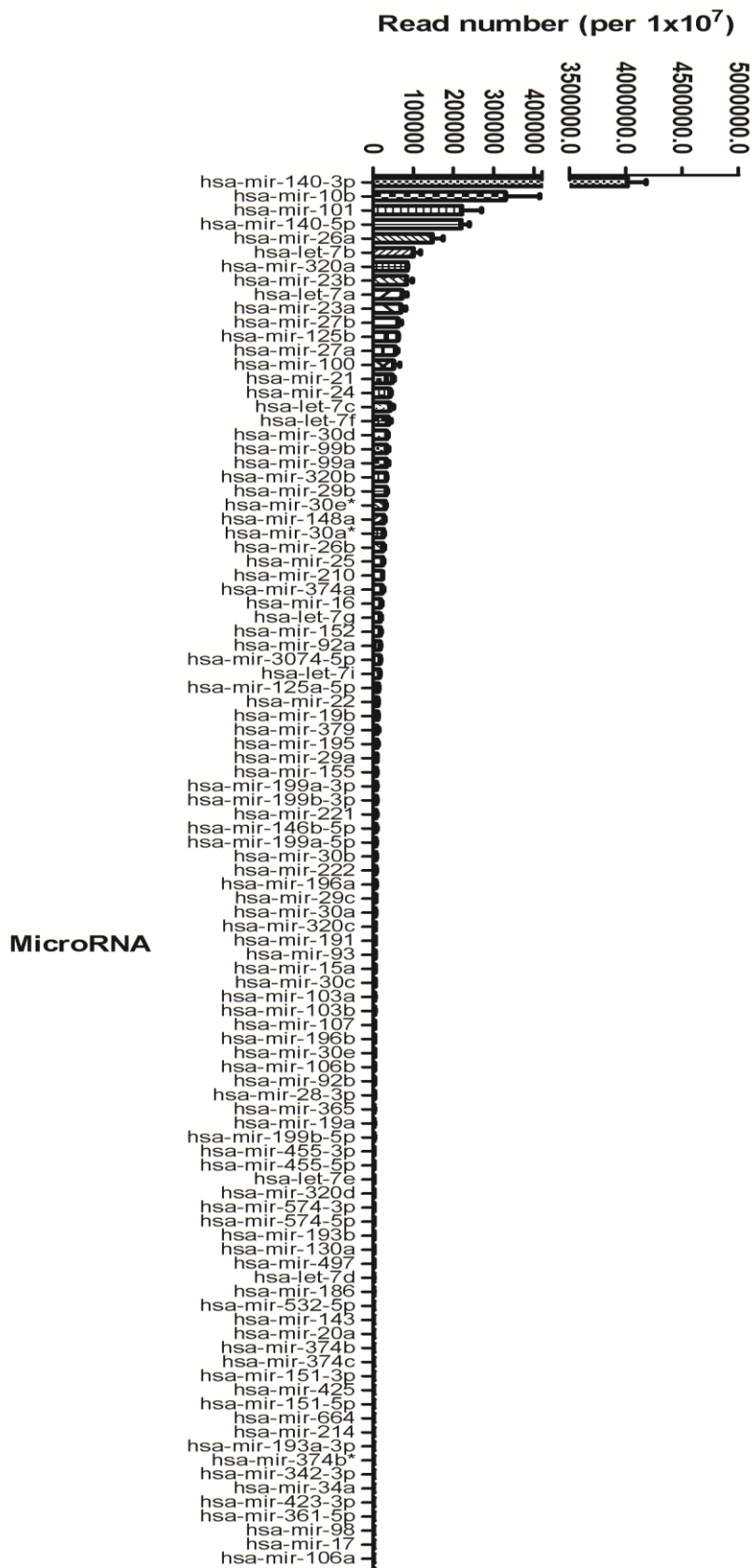
The initial dataset obtained from deep sequencing was analysed using miRCat software (Moxon et al., 2008) by bioinformatician Dr H. Pais (School of Computing, UEA). The miRCat algorithm maps the sequences obtained from the sequencing project to the human genome. It identifies potential microRNAs using a number of criteria. These include: the number of reads of the sRNA sequence in a given area, the orientation of the sequences, and the nucleotide distance between individual sRNAs. Additional identification stages involve choosing the most abundant sRNAs at any one genetic locus as candidates as well as calculating the predicted secondary structure to compare with typical structures of validated microRNAs. Finally the length of sequence and number of mismatches between the complementary strands of the candidate microRNAs were taken into account. This resulted in a preliminary list of candidate microRNAs which fulfilled the conditions of the miRCat algorithm.

Known microRNAs within this initial candidate list were confirmed using the online database miRBase which lists all published microRNA sequences (Griffiths-Jones, 2004; Griffiths-Jones et al., 2006; Griffiths-Jones et al., 2008; Kozomara and Griffiths-Jones, 2011). 990 individual known microRNAs were noted in the dataset of which 630 were present in all three patients, 151 in two patients and 209 in one patient. The most highly expressed 100 known microRNAs present in all three patients in terms of read number per  $10^7$  reads are shown in figure 3.5 (Read counts from each patient per microRNA with the top 100 are shown in Appendix 5). The most highly expressed ten

known microRNAs by read number are shown in figure 3.6. Ranked from highest read count to lowest are miR-140-3p, miR-10b, miR-101, miR-140-5p, miR-26a, let-7b, miR-320a, miR-23b, let-7a and miR-23a. The average read count of miR-140-3p showed a statistically significant difference from all other microRNAs detected with  $p \leq 0.001$ .

The total sequencing space of the sequencing performed was 20 million reads per sample. Figure 3.6 shows miR-140-3p (the most highly expressed microRNA in the dataset) accounted for 40% of that space with an average total read number of 8,031,127. In comparison, its complementary strand miR-140-5p, comprised only 2% of the total sequencing space (438,388 total reads), and hence in ranking, miR-140-5p was 4<sup>th</sup> overall in terms of read number. Figure 3.7 shows the expression (measured by qRT-PCR) of both strands of microRNA-140 in post digest stage chondrocytes (A) and in SW1353 cells in culture (B). There is a statistically significant increase of miR-140-3p expression of 5 fold and 2 fold seen in both models respectively ( $p \leq 0.001$  and  $p \leq 0.01$ ) compared to miR-140-5p.

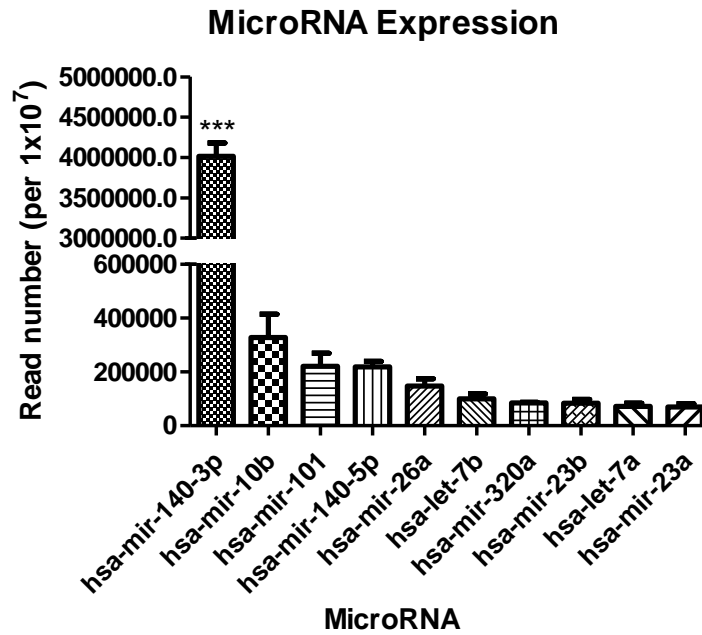
Expression of miR-140-3p and miR-140-5p was also measured by northern blot analysis in SW1353 RNA and quantified with standard curves (Fig. 3.8). Approximately 2.5 fmoles of miR-140-3p and less than 1 fmole of miR-140-5p were found in 10 $\mu$ g SW1353 RNA.



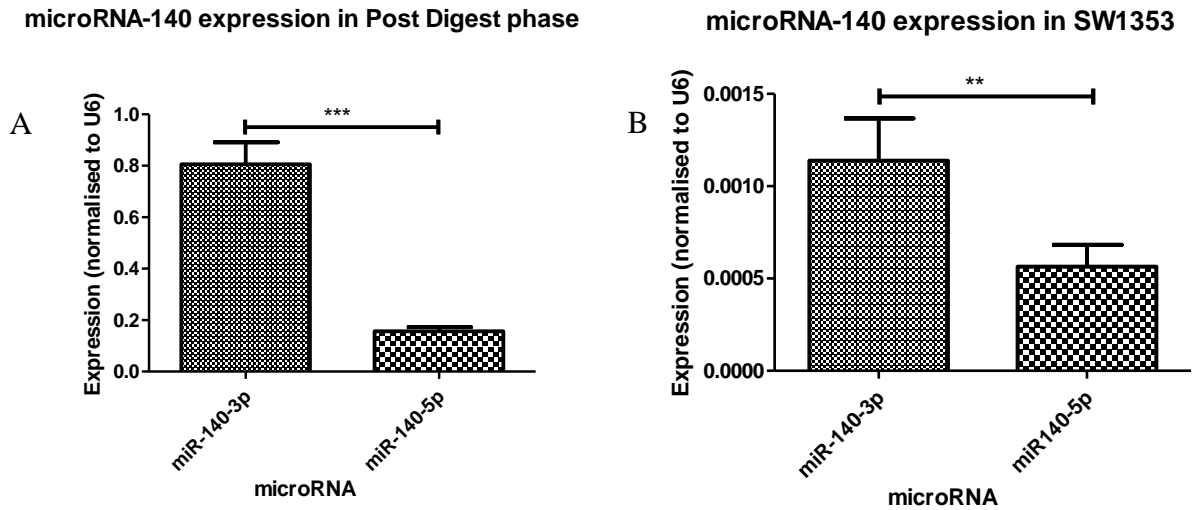
**Figure 3.5 Known microRNAs detected from deep sequencing.** Average read number of most highly expressed 100 known microRNAs shown in read number per  $10^7$  reads were obtained by deep sequencing (analysed with miRCat and miRBase software)



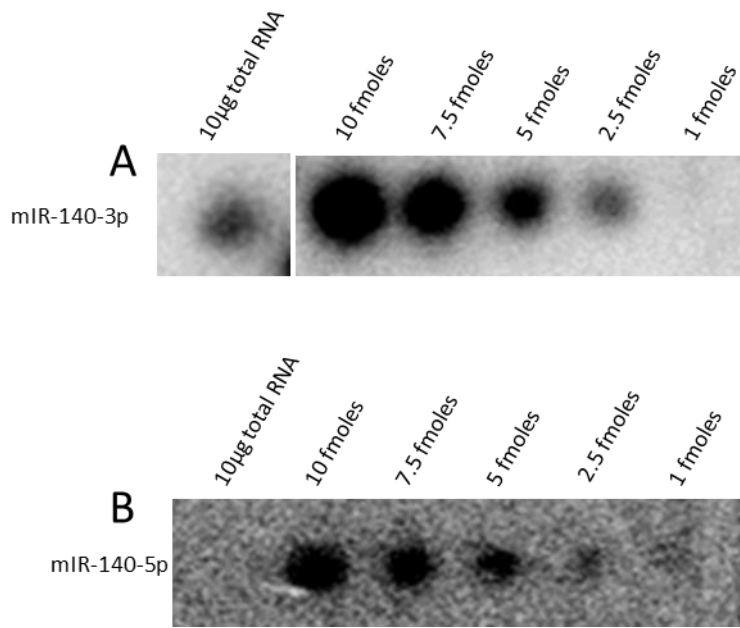
of cDNA transcribed from RNA obtained from chondrocytes of 3 patients isolated immediately after digestion.



**Figure 3.6 Ten most highly expressed known microRNAs from deep sequencing.** Average read number of most highly expressed 10 known microRNAs shown in read number per  $10^7$  reads were measured by deep sequencing of cDNA transcribed from RNA obtained from chondrocytes of 3 patients isolated immediately after digestion. (\*\*\*) =  $p \leq 0.001$  analyzed by one way Anova with post hoc Tukey test).



**Figure 3.7 MicroRNA-140 expression in chondrocytes.** (A) Expression of both strands of miR-140 were measured by qRT-PCR from RNA of chondrocytes isolated directly after digestion of cartilage tissue. RNA was obtained from the same samples of 3 patients used for deep sequencing (B) Expression of both strands of miR-140 were obtained by qRT-PCR from RNA isolated from the SW1353 cell line. Data were normalised to *U6* RNA expression. (n=3, \*\* =  $p \leq 0.01$ , \*\*\* =  $p \leq 0.001$  analyzed by a Student T-test).



**Figure 3.8 Expression of microRNA-140 on northern blots.** Expression of (A) miR140-3p and (B) miR140-5p measured by northern blot using 10µg SW1353 total RNA. Standard curves were created using DNA oligonucleotide positive controls at concentrations of 10 fmole to 1 fmole to quantify the level of microRNA expression present in 10µg SW1353 RNA.

### **3.2.3 Candidate novel microRNA discovery in primary osteoarthritic chondrocytes**

The dataset produced from the deep sequencing was analysed by Dr Helio Pais (School of Computing, UEA) utilising the MiRCat algorithm to identify candidate microRNA sequences, and miRBase V.17 (released in April 2011) database to determine how many were already published online on the database. 1621 candidate microRNAs were determined to be completely novel by this process. Out of the 1621 sequences, 61 were found across all three species (shown in figure 3.9). Average read counts per 10 million reads ranged from 328.91 to 0.49. Exact sequences and read counts of the 61 are shown in Appendix 4.

16 candidate microRNAs were chosen to study further. These are shown in table 3.1. The novels are labelled 1 to 16 and range in read count from 328.91 to 0.49. Sequence length ranges from 18 to 24 nucleotides.

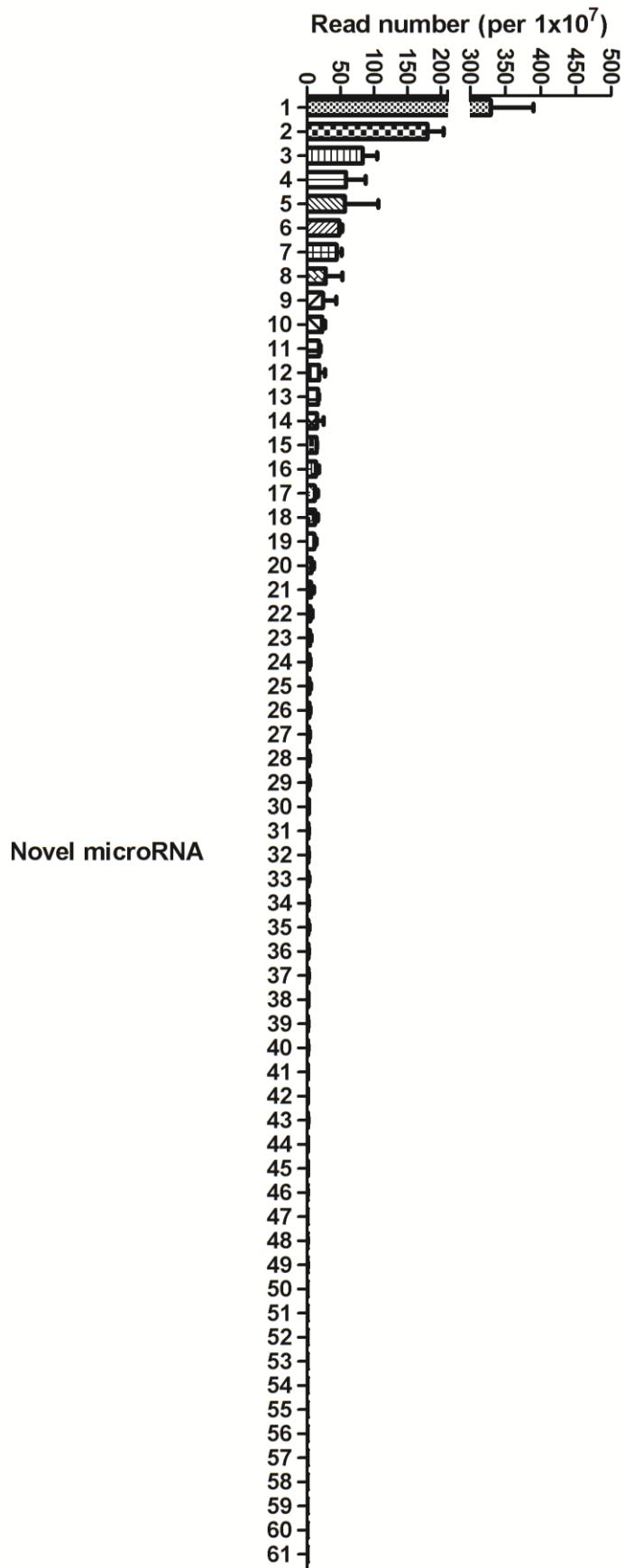
Triaging from the initial pool of 1621 candidate novel microRNAs to 16 was based on, first of all, whether they were present in all three patients, as presence in more than one patient increases the chance that the candidate is real. Secondly, the presence of the complementary strand of the candidate microRNA within the expression profile was also used as a positive selection. Both strands of a microRNA are more commonly being found to be expressed and functional as opposed to having just one dominant strand.

Finally, the candidates were cross-referenced with other sequencing projects obtained by Professor Tamas Dalmay's lab (School of Biological Sciences, UEA) to see if they appeared in other tissues or cell lines, this would provide support to their microRNA status if they were detected elsewhere, and suggest that they were not just artefacts of a specific experiment. The two cell lines compared were DLD-1 wild type and DLD-1 dicer knockdown. In addition to just comparing presence of the sequences between the projects, Dicer is an enzyme known to be vital to microRNA synthesis. If the candidate's expression was down-regulated in the Dicer knock down cell line compared to the wild type, then this added credibility to the candidacy, suggesting that the sequence was processed by the Dicer enzyme. However, this factor was only included in combination with the previously mentioned criteria, and a candidate was not discarded if it fulfilled the others but did not appear to be Dicer dependent.

At the time of writing miRBase has been updated a further 3 times since the initial sequencing analysis to V.20, and a number of microRNAs originally considered novel have now been described. Within the chosen top 16, 6 can now be found on miRBase.

Candidate novel 2 is described as hsa-miR-6509-5p (Joyce et al., 2011), candidate novel 7 is known as hsa-miR-664b-3p (Friedlander et al., 2012), candidate novel 8 is a truncated form of hsa-miR-1277-5p (Morin et al., 2008). Candidate novel 10 is described as hsa-miR-487a-5p in humans (Fu et al., 2005) and a similar sequence can be found as oar-miR-487a-5p in sheep (Caiment et al., 2010). Candidate novel 11, although still not annotated in humans can be found in mouse and rat as mmu-miR-3085-3p and rno-miR-3085 respectively (Chiang et al., 2010; Linsen et al., 2010). Finally candidate novel 12 can be found in humans as hsa-miR-539-3p (Sewer et al., 2005), with marginally altered sequences found in mouse, rat and sheep as mmu-miR-539-3p, rno-miR-539-3p and oar-miR-539-3p respectively (Caiment et al., 2010; Landgraf et al., 2007; Sewer et al., 2005). Additionally, candidate novel 15 was present in miRBase V.19 but has since been removed and annotated as a fragment of tRNA. Although these microRNAs have now been described on miRBase, none have been experimentally validated, with identification coming only from sequencing data or bioinformatics analysis.

The literature identifying candidate novels 8, 10 and 12 as microRNAs were published before the version of miRBase used during the initial data analysis; however those microRNAs were not identified in the analysis carried out for this project. This could be due to an error within the data analysis, or potentially the data from the literature referenced was not added to the miRBase database until after V.17 and hence not available for inclusion in the initial results.



**Figure 3.9 Read numbers of candidate novel microRNAs from deep sequencing.**

Average read number of the 61 candidate microRNA sequences (found in all three

patients) shown in read number per  $10^7$  reads were measured by deep sequencing (analysed with miRCat and miRBase software) of cDNA transcribed from RNA obtained from chondrocytes of 3 patients isolated immediately after digestion.



Novel ID	Sequence (5' - 3')	No. bases	Read number per 10,000,000			
			KOA15/11	KOA16/11	KOA17/11	Average
Novel 1	AAGTTTCTCTGAACGTGTAGAGC	23	39.44366915	54.17209332	51.62952183	48.41509476
Novel 2	ATTAGGTAGTGGCAGTGGAAC	21	0.433446914	0.492473576	0.537807519	0.487909336
Novel 3	CTGAAGATCTAAAGGTCCTGGT	23	356.72681	418.1100657	211.8961625	328.9110127
Novel 4	GGGGGTGTAGCTCAGTGG	18	133.9350964	191.0797473	215.1230076	180.0459504
Novel 5	TGAATCCAGCGATCCGAGT	19	66.31737781	126.5657089	57.54540453	83.47616376
Novel 6	TCGAACTTGACTATCTAGAGGA	22	6.93515062	7.387103634	157.0397956	57.12068327
Novel 7	TTCATTTGCCTCCCAGCCTACA	22	14.30374815	15.75915442	13.44518798	14.50269685
Novel 8	TATATATATATGTACGTAT	21	30.34128396	56.63446119	46.25144664	44.40906393
Novel 9	GATGCCTGGGAGTTGCGATCTG	22	14.73719507	19.20646945	21.51230076	18.48532176
Novel 10	GTGGTTATCCCTGCTGTGTTCCG	22	8.668938275	35.45809744	10.75615038	18.29439537
Novel 11	TCTGGCTGCTATGGCCCCCTC	21	3.034128396	14.28173369	18.28545565	11.86710591
Novel 12	ATCATACAAGGACAATTTCTTT	22	2.600681483	14.77420727	1.613422557	6.329437102
Novel 13	AGCTGGACTGAAGACTATCTGGC	23	1.300340741	3.447315029	2.151230076	2.299628615
Novel 14	CGCCTGCCCTCTCGGTGTCGGA	22	0.433446914	0.492473576	0.537807519	0.487909336
Novel 15	GGAGATCCTGGTTTGAATCCC	22	16.03753581	15.75915442	18.82326317	16.8733178
Novel 16	GACCCAGAAAAGGTGTTGGTTGA	24	8.668938275	2.954841454	2.151230076	4.591669935

**Table 3.1 Candidate microRNAs summary.** DNA sequences, nucleotide length, patient read numbers and average read numbers of the 16 triaged candidate microRNAs taken from sequencing data, read number per  $10^7$  reads.

### 3.3 Discussion

Published profiles of microRNA expression have used a variety of tissue sources or methods of culturing chondrocytes. These include varying passage numbers, with little information on the comparison between microRNAs taken directly from osteoarthritic cartilage and isolated chondrocytes grown in culture. In order to establish the best source of RNA for the sequencing experiment and to enable the optimum discovery of new microRNAs, we compared the expression of two specific known microRNAs (miR-140 and miR-455) in RNA extracted directly from cartilage with that from isolated chondrocytes (post digest) and those cultured in monolayer and allowed to grow to confluency. We also investigated their expression at three subsequent passages.

The de-differentiation assay (Fig. 3.1) showed significant expression of all four microRNAs measured in cartilage and at initial isolation (PD) of primary chondrocytes. Interestingly miR-140-5p showed an increase in expression after initial cell extraction compared to expression within cartilage tissue, whereas 140-3p, 455-3p and 455-5p showed a decreasing trend from cartilage to passage 3. This cartilage specificity is supported in the literature (Miyaki et al., 2009; Miyaki et al., 2010; Swingler et al., 2012). In addition, Hong et al have also shown the activity of a number of microRNAs in a de-differentiating chondrocyte assay, and found miR-140 to dramatically decrease through chondrocyte passage (Hong and Reddi, 2013). The expression pattern shown by miR-140-5p could suggest a potential stress or wounding response of miR-140-5p to the digestion conditions which cells undergo to enable isolation from tissue. It has previously been shown that some microRNAs can respond to stressful conditions in plants such as drought, low temperatures and high salinity (Liu et al., 2008). The same group have also discovered stress-related elements in the promoter regions of the microRNAs which appeared to respond to stress. Leung et al discuss how microRNAs may play a larger role than once thought in a cell's response to stress, with microRNAs thought to be involved in the restoration of cell homeostasis post stress (Leung and Sharp, 2010).

Although there is little existing information regarding potential responses of microRNAs to wounding conditions in cartilage, there have been investigations into the microRNA response to mechanical loading. It has been shown that in tendon fibroblasts in adult rats, the expression of a number of microRNAs were shown to significantly alter in response to mechanical loading. These included miR-100, miR-378, miR-1,

miR-133a, miR-133a\*, miR-133b, miR-206, miR-140-5p, let-7a, let-7e, miR-338, miR381, and miR743a (Mendias et al., 2012). Interestingly, although the group examined the expression levels of both miR-140-5p and miR-140-3p, only miR-140-5p showed a significant decrease in expression when under loading conditions compared to sedentary. In contrast, miR-140-3p showed no statistically significant difference between the two conditions. This is similar to the results seen in figure 3.1, suggesting that miR-140-5p is more responsive to certain stress conditions than its complementary strand miR-140-3p.

A number of matrix genes as well as the collagenases *MMP1* and *MMP13* were measured through the stages of the de-differentiation assay. As chondrocytes undergo de-differentiation, they change both in morphology and gene expression pattern as they become more fibroblastic over time. (Stokes et al., 2001). Figure 3.4 shows the morphological changes of primary chondrocytes during de-differentiation from passage 0 to 3 in monolayer culture. Over time in monolayer culture, the chondrocytes become more elongated and similar in morphology to fibroblasts.

Figure 3.2 shows the expression of several collagen genes during de-differentiation. Several collagen genes known to be cartilage specific show an expected decrease in expression from cartilage tissue to passage 3. These are *COL2A1*, *COL9A1* and *COL27A1*, this is supported by literature investigating collagen expression in de-differentiation and re-differentiation (Stokes et al., 2001). These collagens are known to be cartilage specific, and so expression decreases as chondrocytes lose their phenotype (Boot-Handford et al., 2003; Stokes et al., 2002; Stokes et al., 2001). *COL11A1* is also thought to be cartilage specific, however Figure 3.2 shows an increasing trend in expression for *COL11A1* through the assay as chondrocytes begin to lose their phenotype. This data contradicts research by Stokes et al who found a decrease in *COL11A1* through subsequent passages in de-differentiating monolayer chondrocytes (Stokes et al., 2001).

Aggrecan, another ECM component, has been shown previously to decrease during de-differentiation (Stokes et al., 2001), as well as degrading in osteoarthritis (Lark et al., 1997). But in the current study, aggrecan expression shows no obvious trend. Expression decreases slightly from cartilage tissue to passage 0, before increasing up to passage 3. However these changes do not show statistical significance.

Other collagens known to be expressed by fibroblasts include collagens type I and III (Stokes et al., 2001), and these show an increasing expression trend from cartilage to passage 3 (Fig. 3.2) as the chondrocytes lose their phenotype and become more fibroblastic. This is pattern of expression is supported by the literature (Stokes et al., 2001).

The expression of collagenases *MMP1* and *MMP13* are known to be up regulated in osteoarthritis (Bluteau et al., 2001). In this study, looking at de-differentiation (Fig. 3.3) they appear to have the highest expression at the post digest stage for *MMP1* and the passage 0 stage for *MMP13* before gradually decreasing down the remaining passage stages. Expression in RNA extracted directly from cartilage tissue is significantly lower for both MMPs compared to isolated and cultured chondrocytes, however only *MMP1* shows a statistical significance between cartilage tissue and isolated (PD) chondrocytes ( $p \leq 0.01$ ). This is supported by existing literature which also showed that expression of both MMPs was up-regulated *in vitro* compared with *in vivo* osteoarthritic and normal tissue (Bau et al., 2002).

Several attempts were made to get high amounts of good quality RNA from cartilage tissue, but this proved difficult (data not shown). The relatively high expression of selected known microRNAs seen at the post digest stage combined with the high quality of RNA needed for deep sequencing meant that RNA isolated from post digest chondrocytes were used for the miRNome profiling. The post digest stage was chosen as it was thought that a stage which showed a high level of expression of chondrocyte specific microRNAs would be the ideal stage to search for the entire miRNome of the tissue, and miR-140-5p is the best characterised microRNA implicated in osteoarthritis to date (Miyaki et al., 2009; Miyaki et al., 2010; Swingler et al., 2012).

To identify the miRNome, an Illumina next generation sequencing platform was utilised due to the design of the bias-reducing HD adapters for this specific platform (Sorefan et al., 2012). The adapters were intended to reduce the RNA ligase bias seen in previous editions of the platform.

We were able to identify a list of 1621 candidate novel microRNA sequences (61 found in all three patients) in addition to 990 previously described microRNAs (630 found in all three patients). This total number is much larger than any previous study, for

example, one deep sequencing study in rat cartilage using the Solexa platform identified 246 previously identified and 86 novel microRNAs (Sun et al., 2011b).

The difference in microRNA detection between the three patients suggests a potential detection bias still present within the system, or potentially a true detection rate, with the differences seen actually due to different osteoarthritic mechanisms between the three patients resulting in slightly different miRNome profiles. The idea that osteoarthritis can be separated into disease subtypes is gaining more evidence and might explain the differential microRNA expression patterns seen between different patients. (Davis, 1988; Driban et al., 2010). This has also been explored in terms of mRNA profiling of OA disease subtypes, showing for example that primary OA (thought to be previously defined as the development of the disease without an obvious mechanistic cause) can actually be divided into subsets based on genetic and hormone expression, as well as age related factors (Herrero-Beaumont et al., 2009).

MiRCat software (Moxon et al., 2008) was used to analyse the sequences of small RNAs obtained from the deep sequencing to determine whether they fulfilled the criteria for microRNAs or not. Factors considered by the algorithm when designating microRNA status to a sequence included; location, abundance, secondary structure, the number of mismatches between the microRNA and its complementary strand, and hairpin length. These sequences were then run through miRBase (Griffiths-Jones et al., 2006; Griffiths-Jones et al., 2008; Kozomara and Griffiths-Jones, 2011) to check for previously discovered microRNAs, and the remaining novel microRNA candidates were triaged to a reasonable number for experimental validation. The initial pool of 1621 novel microRNAs was reduced to 16 based on, first of all, whether they were present in all three patients, as presence in more than one patient increases the chance that the candidate is real. Secondly, the presence of the complementary strand of the candidate microRNA within the expression profile was also used as a positive selection. Both strands of microRNAs are more commonly being found to be expressed and functional as opposed to having only one dominant strand, although often one strand is usually found to be expressed more highly in specific conditions or tissues.

Finally, the candidates were cross-referenced with other sequencing projects obtained by Professor Tamas Dalmay's lab (School of Biological Sciences, UEA) to see if they appeared in other tissues or cell lines, which would provide support to their microRNA status if they were detected elsewhere, and suggest that they were not just artefacts of a

specific experiment. The two cell lines compared were DLD-1 wild type and DLD-1 dicer knockdown. In addition to just comparing presence of the sequences between the projects, Dicer is an enzyme known to be vital to microRNA synthesis. If the candidate's expression was down-regulated in the Dicer knock down cell line compared to the wild type, then this added credibility to the candidacy, suggesting that the sequence (like existing validated microRNAs) was processed by the Dicer enzyme. However, this factor was only included in combination with the previously mentioned criteria, and a candidate was not discarded if it fulfilled the others but did not appear to be Dicer dependent. This is due to recent literature showing that not all validated microRNAs are Dicer dependent, and that some (for example miR-451) are still synthesised without the presence of the enzyme, utilising a redundancy system of Ago proteins (Betancur and Tomari, 2012; Cheloufi et al., 2010).

In addition the three novel candidates with the highest read number across all three patients were included (candidate microRNAs 3, 4 and 5), as expression is an important factor when considering downstream validation techniques (Appendix 4 shows the criteria fulfilled by each of the 61 novel microRNAs to be found across all three patients).

In addition to the large number of microRNAs (both known and potentially novel) detected in the osteoarthritic chondrocyte RNA, we also found that the most highly expressed microRNA was miR140-3p. Previous literature has always speculated miR140-5p to be the mature strand which is most used as the guide strand in cartilage, with the miR-140-3p strand as the passenger strand (McAlinden et al., 2013; Shin, 2008; Swingler et al., 2012). However, other data from rat cartilage (Sun et al., 2011b), as well as our own from sequencing in chondrocytes and subsequent validation experiments has shown this not to be the case and that in fact miR-140-3p is the dominant microRNA guide strand expressed in chondrocytes and osteoarthritis. It is also now shown in miRBase (Griffiths-Jones et al., 2006; Griffiths-Jones et al., 2008; Kozomara and Griffiths-Jones, 2011) that miR-140-3p appears to have more reads than miR-140-5p across the deep sequencing projects submitted, suggesting it is more often the guide strand than miR-140-5p. There is also literature to support the higher expression of miR-140-3p over miR-140-5p in cartilage tissue (Kobayashi et al., 2008; Pando et al., 2012).

This function of miR-140-3p as the dominant strand is also supported by the theory that the guide strand of miRNA duplexes usually has the most thermodynamically unstable 5' end (Manavella et al., 2012) combined with a stable seed-site duplex (Hibio et al., 2012) and often begins with a uridine base as it appears to be the base usually (although not always) preferred by the RISC complex (Kawamata et al., 2011). MiR-140-3p begins with a uridine base whereas miR-140-5p begins with a cytosine at the 5' end.

The sequencing experiment showed a 20-fold difference in detection between the two miR-140 strands, with miR-140-3p taking up 40% of the total sequencing space available compared to the 2% of miR-140-5p.

qRT-PCR analysis supports the findings of the deep sequencing on miR-140, and shows a much higher expression of miR-140-3p than miR-140-5p in primary osteoarthritic chondrocytes taken from the post digest phase as well as in cells of the chondrosarcoma cell line SW1353 (Fig. 3.7) (showing 4 and 2 fold up-regulations respectively).

SW1353 cells are a useful model for chondrocyte culture as there are some similarities in morphology and behaviour, although the phenotype is not completely comparable to primary chondrocytes (Gebauer et al., 2005).

Northern blots also support this, showing 2.5 fold higher expression of miR-140-3p than miR-140-5p (Fig. 3.8). However the difference in expression was not as extreme in these methods as in the sequencing. The experimental techniques of northern blotting and quantitative RT-PCR used are known to be more reliable quantitatively than deep sequencing which is more useful for profiling the many different microRNAs and their isoforms rather than determining exact expression values (Baker, 2010). This is shown further by an extensive study carried out by Git et al., where a comparison of microarray profiling, real-time PCR, and deep sequencing on microRNA detection was carried out and found there was a low overlap in microRNA detection between the different techniques (Git et al., 2010). This all suggests that there is still some bias within each system towards possibly specific structures or sequences of which miR-140-3p may fit the profile, supporting the fact that any deep sequencing project must be supported with follow up experimental evidence.

However, this study has shown a large number of microRNAs (both known and potentially novel) to form the miRNome of osteoarthritic chondrocytes. This is a

substantial increase on previous published works, and the next step is to confirm that these candidate novel sequences are true microRNAs.

Validation is the focus of chapter 4, which will present a combination of experimental validation techniques used to validate the candidate novel microRNAs and ascertain if they are true microRNAs or if the sequences prove to be either false positives or an alternate sRNA molecule.



## **Chapter 4 – Experimental Validation of Novel MicroRNAs**

## 4.1 Introduction

Chapter 3 described the experimental methodology employed to profile the miRNome of osteoarthritic primary chondrocytes obtained from human articular cartilage and the results of a deep sequencing experiment. Aside from the 990 known microRNAs, 1621 previously undescribed microRNAs were discovered, termed candidate novel microRNAs. The next step post-discovery was validation to ensure that the sequences determined to be microRNAs by the MiRCat software (Moxon et al., 2008) were true microRNAs and not other small RNAs, anomalies or products of degradation which happened to fulfil the identification criteria by chance.

As discussed previously, deep sequencing technology is not infallible, as seen in the strong bias observed in existing platforms (Sorefan et al., 2012) and the fact that there are substantial differences seen in detection between various platforms and technologies (Git et al., 2010). Although part of this bias (associated with RNA ligases) was negated with the use of HD adapters for this project, there is still some bias remaining, as shown by Sorefan et al (Sorefan et al., 2012) where the number of different sequences detected within a known pool was increased from 42 to 81%. Although this is an impressive increase, there still appears to be 19% of sequences completely undetected.

The fact that there is bias remaining is also shown in the results of Chapter 3 by the unusually high miR-140-3p read counts (Fig. 3.6) which do not correspond quantitatively to experimental validation using northern blots and qRT-PCR (Fig. 3.7 and 3.8). In addition to bias towards specific structures, false positives are also a possibility with some sequences designated as microRNAs by miRCat potentially being simply products of degradation which happen to fulfil the criteria imposed by the miRCat (Moxon et al., 2008) algorithm.

Due to the problem of false positives, an important aspect of miRNome profiling is experimental validation of these sequences. Experimental validation involves firstly, size validation using northern blots to determine the expected microRNA size of ~22nt sequence. This is also combined with a functional validation, investigating the candidate microRNA's response to the presence and absence of Dicer (an enzyme vital to microRNA synthesis). Finally, expression of the sequences was measured across a number of models, including chondrogenesis assays using both human mesenchymal stem cells and mouse ATDC5 cells. Expression of the chosen candidates was also examined in a de-differentiation assay where primary chondrocyte cells were isolated

from human articular cartilage and cultured to passage 3 in monolayer, losing chondrocyte phenotype and becoming more fibroblastic.

In order to establish tissue specificity of the candidates, expression was also measured by qRT-PCR across a human tissue panel consisting of over 20 different human tissue types including articular cartilage. Expression was measured between osteoarthritic and non-diseased cartilage tissue to determine if the novel microRNAs changed in expression during disease. In addition to measuring the candidate microRNA expression in osteoarthritic tissue, expression changes were also measured in a controlled cartilage explant injury model to discover if expression of the microRNAs responded to wounding.

All 16 of the candidate novel microRNAs documented in Chapter 3 (Table 3.1) underwent experimental validation, however only three (candidate novel microRNAs 2, 7 and 11) will be discussed in detail here. The detailed results from the remainder will be documented in Appendices 6, 7, 8 and 9.

## 4.2 Results

### 4.2.1 Candidate novel microRNA size and function validation

In order to validate the candidate novel microRNAs, a number of techniques were used. Northern blotting is considered to be the gold standard for sizing microRNAs, and the northern blots of candidate microRNAs (A) novel 2, (B) novel 7 and (C) novel 11 are shown in figure 4.1. Total RNA was obtained from SW1353 cells using the miRVana microRNA isolation method. MiR-140-3p and U6 were used as size and loading controls respectively (Fig. 3.8 A) and northern blots of the remaining 13 candidates are shown in the Appendix 6.

A band of the correct size (~21bp) is seen for candidate microRNA novel 7 suggesting that this candidate is indeed a true microRNA as it fulfils the size requirements.

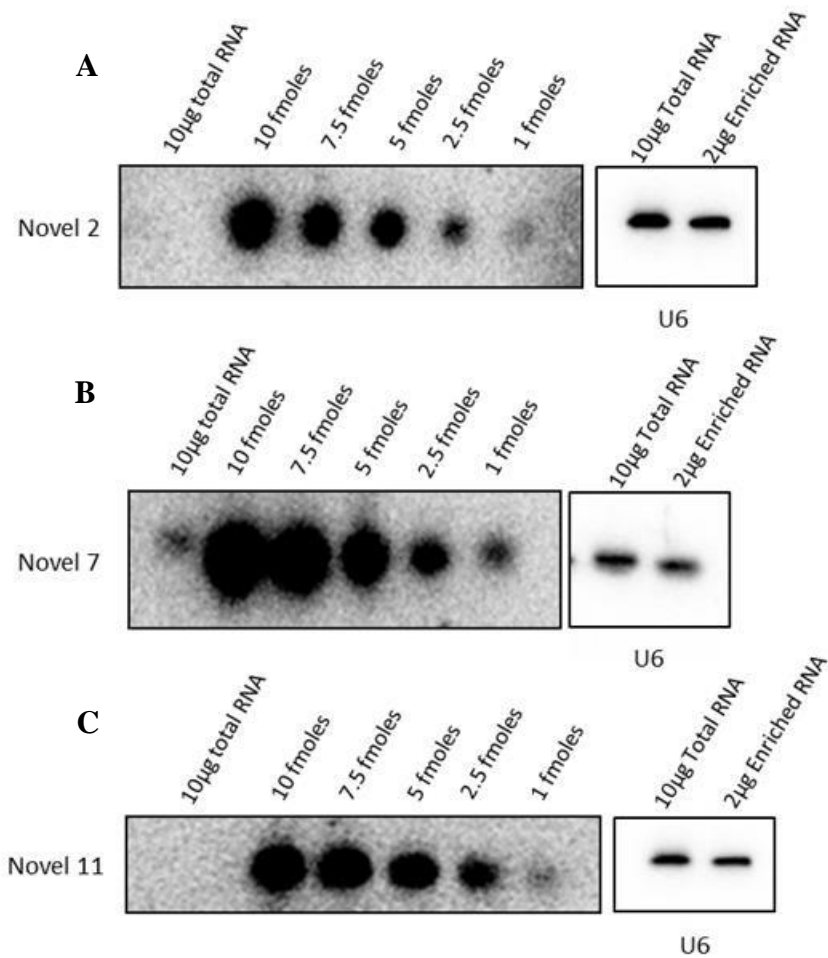
No signal was seen for either candidate novel microRNAs 2 and 11. This indicates that expression levels may be too low to be detected by this method but they may still however be microRNAs. Expression of the candidates was quantified with standard curves using DNA oligonucleotides of the complementary sequences; approximately 1 fmole of novel 7 was seen in 10µg SW1353 total RNA.

The Dicer enzyme is a vital part of the microRNA synthesis process and is involved in the production of the mature microRNA. If Dicer is not present, then the mature microRNA does not form. Expression of the candidate microRNAs was measured by qRT-PCR in RNA isolated from the wild type DLD-1 cell line in addition to the isogenic homozygous knock out of Dicer DLD-1 cell line, immature microRNAs are known to accumulate in the Dicer null cells with a reduction of mature microRNAs present (Cummins et al., 2006).

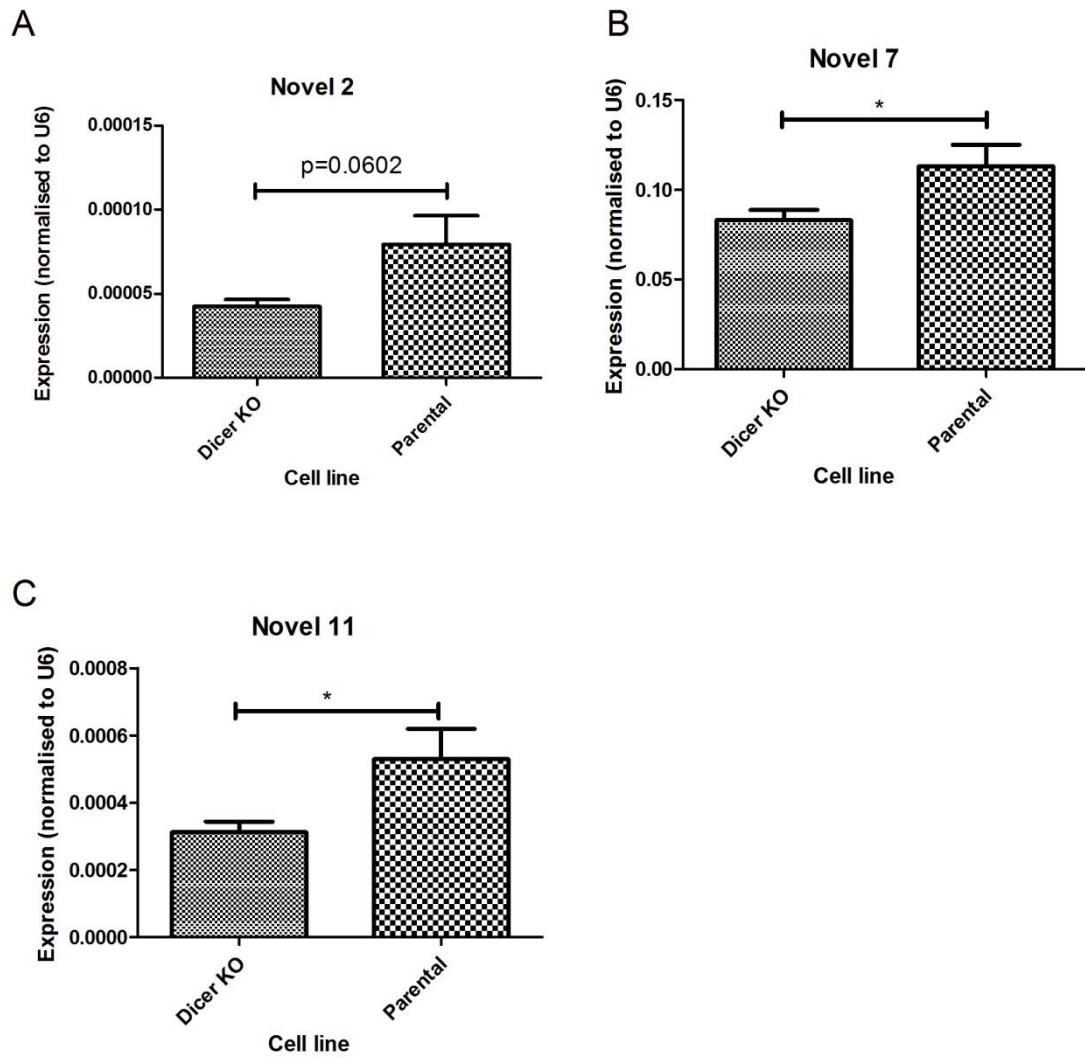
Figure 4.2 shows the expression of candidate microRNAs (A) novel 2, (B) novel 7, and (C) novel 11 in each cell line (the expression of the remaining candidates is shown in Appendix 7). All three candidate novel microRNAs showed a decrease in expression in the Dicer knockout cell line compared to the wild type. The down-regulation was statistically significant for both novel 7 and novel 11 with  $p \leq 0.05$ . The down-regulation of novel 2 approaches significance with  $p = 0.06$ .

In order to ensure the phenotype was maintained during the experiment, expression of *Dicer* RNA was also measured in both cell lines (Fig. 4.3 A). Expression decreased

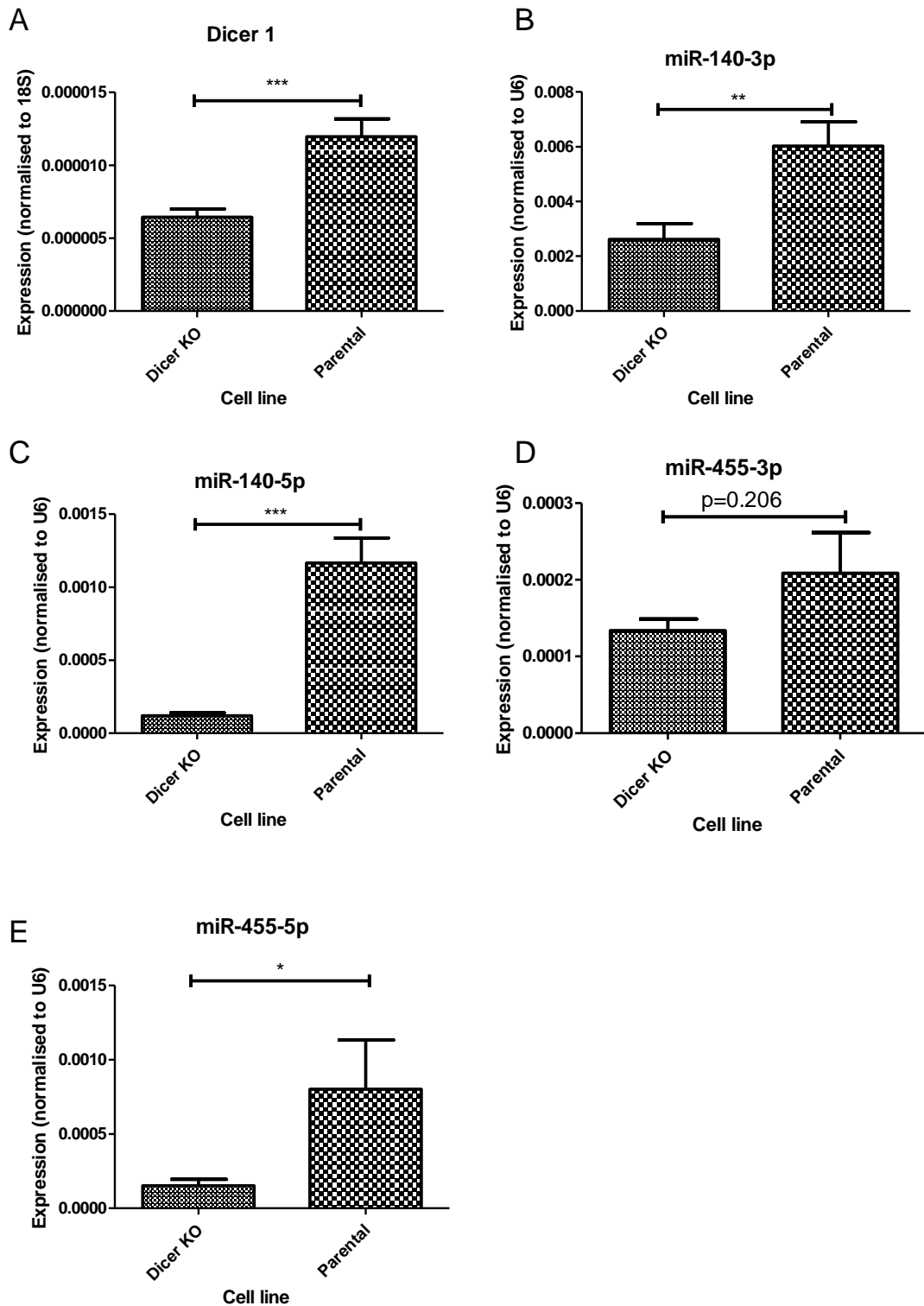
significantly with  $p \leq 0.001$  in the Dicer knock out line compared to the wild type by approximately 2-fold. As a further control, the expression of four known microRNAs was also measured in both cell lines to compare their expression pattern with that of the candidate novel microRNAs. Four microRNAs significantly decreased in expression in Dicer  $-/-$  cells compared with wild type, (B) miR-140-3p, (C) miR-140-5p, (D) miR-455-3p and (E) miR-455-5p. These changes were significantly different for miR-140-3p ( $p \leq 0.01$ ), miR-140-5p ( $p \leq 0.001$ ) and miR-455-5p ( $p \leq 0.05$ ) when analysed with a Student t-test.



**Figure 4.1 Assessment of candidate novel microRNA using northern blots.** Expression of candidate microRNA (A) novel 2, (B) novel 7 and (C) novel 11 was measured by northern blot using 10µg SW1353 total RNA. Standard curves were created using DNA oligonucleotide positive controls at concentrations of 10 fmole to 1 fmole to quantify the level of microRNA expression present in 10µg SW1353 total RNA. The small RNA *U6* was used as a standard reference for the northern blot.



**Figure 4.2 Candidate novel microRNA expression in DLD-1 parental and DLD-1 Dicer null cell lines.** Expression of candidate microRNA (A) novel 2, (B) novel 7 and (C) novel 11 were obtained by qRT-PCR from RNA purified from DLD-1 wild type (parental) and DLD-1 Dicer null cell lines. Data were normalised to *U6* RNA expression. Data shows mean  $\pm$  SEM,  $n = 12$ . (\*,  $p \leq 0.05$  analyzed by a Student T test).



**Figure 4.3 Gene and microRNA expression in DLD-1 parental and DLD-1 Dicer null cell lines.** Expression of (A) Dicer 1, (B) miR-140-3p, (C) miR-140-5p, (D) miR-455-3p, and (E) miR-455-5p were obtained by qRT-PCR from RNA of DLD-1 wild type (parental) and DLD-1 Dicer null cell lines. Data were normalised to *U6* RNA expression for microRNAs and *18S* RNA expression for Dicer 1. Data shows mean  $\pm$  SEM,  $n = 12$  (\*,  $p \leq 0.05$ , \*\*,  $p \leq 0.01$ , \*\*\*,  $p \leq 0.001$  analyzed by a Student t-test).



#### 4.2.2 Candidate novel microRNA expression in human tissues

The deep sequencing project established that the microRNA candidates were detected in osteoarthritic knee cartilage. The next step was to validate this using qRT-PCR. In order to ensure specific detection, locked-nucleic acid primers (Exiqon) were used. The products obtained from qRT-PCR were run on polyacrylamide gels to ensure specificity of the LNA primers (data not shown). The product from miR-140-3p was used as a size control. Additional minor products were shown in addition to the expected product for the primer sets of six of the sixteen candidate microRNAs. Candidate microRNAs novel 2, novel 7 and novel 11 all showed specific detection with only one product present.

Expression of the candidate microRNA sequences novel 2, novel 7 and novel 11 were measured in cartilage obtained from neck of femur (NOF) fractures, and osteoarthritic hip and knee total joint replacements (all remaining candidate microRNAs were measured and data are shown in Appendix 8). Known microRNAs miR-140-3p, 140-5p, 455-3p and 455-5p were also measured to explore the pattern of expression in these samples of microRNAs already known to be involved in osteoarthritis (Miyaki et al., 2009; Swingler et al., 2012).

Both candidate microRNA novels 2 and 11 [Figure 4.4 (A) and (C)] show a trend of increased expression in hip osteoarthritis compared to NOF; however these do not reach statistical significance when tested with a student t-test. Candidate microRNA novel 7 (B) appears to be expressed at equivalent levels between NOF and osteoarthritic hip cartilage suggesting it is not regulated in late stage disease.

Expression of miR-140-3p shows no apparent change between diseased and normal tissue [Fig. 4.5 (A)], however (B) miR-140-5p, (C) miR-455-3p and (D) miR-455-5p all show an increase in expression in osteoarthritic cartilage. Both strands of miR-455 show a statistically significant difference between the two conditions with  $p \leq 0.01$ . MiR-140-5p approaches significance with  $p = 0.0587$ .

Figure 4.6 shows expression of a number of marker genes in HOA cartilage and cartilage obtained from NOF. A disintegrin and metalloproteinase domain with thrombospondin motif 5 (ADMTS5) (A), matrix metalloproteinase 3 (MMP3) (B) and aggrecan (ACAN) (D) are increased in NOF cartilage compared with osteoarthritic hip cartilage ( $p \leq 0.05$ ,  $p \leq 0.05$  and  $p \leq 0.001$  respectively).

Matrix metalloproteinase 13 (MMP13) (C) shows a trend of increased expression in OA compared to NOF but this doesn't reach significance because of wide inter-patient variation. Type II collagen (COL2A1) (E) has an increased expression in osteoarthritic cartilage tissue compared to NOF ( $p \leq 0.001$ ).

Figure 4.7 shows the expression of candidate microRNAs (A) novel 2, (B) novel 7 and (C) novel 11 in osteoarthritic cartilage obtained from total knee joint replacement (KOA) compared to neck of femur fracture cartilage (NOF). All three showed decreased expression in KOA cartilage, with candidate microRNA novel 2 showing a statistically significant down-regulation ( $p \leq 0.01$ ).

The expression of 4 known microRNAs were also measured in KOA and NOF tissue (Fig. 4.8) and all showed a decreased expression in KOA cartilage compared to NOF. However, only (A) miR-140-3p showed a statistically significant difference ( $p \leq 0.05$ ).

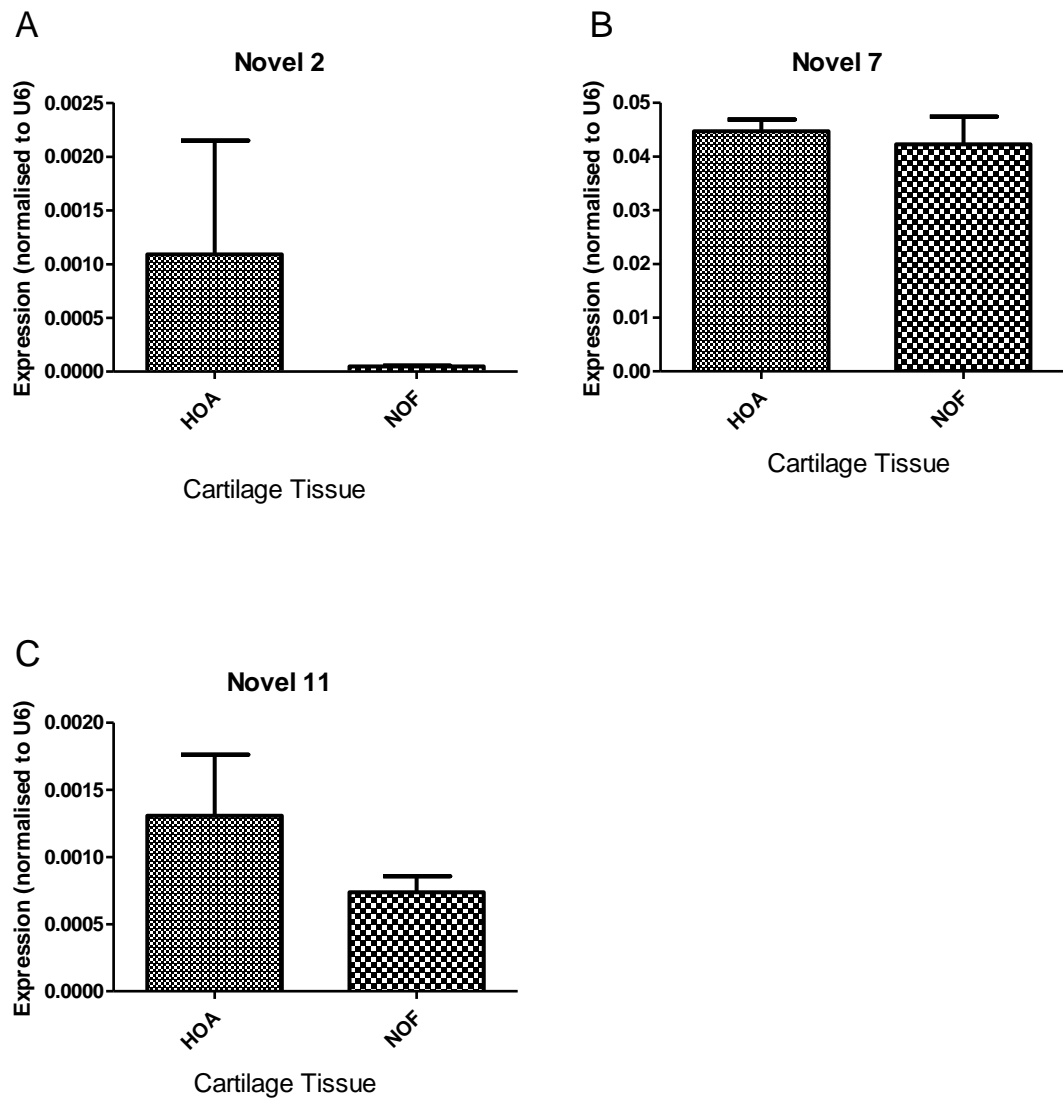
To assess whether the expression of these candidate novel microRNAs was cartilage specific or selective, their expression was measured using qRT-PCR across 23 different human tissue types (Fig. 4.9). These were adipose, bladder, brain, cervix, colon, oesophagus, heart, kidney, liver, lung, ovary, placenta, prostate, small intestine, skeletal muscle, spleen, testes, thymus, thyroid, trachea, osteoarthritic knee cartilage tissue (KOA), osteoarthritic hip cartilage (HOA) and neck of femur fracture cartilage (NOF). The cartilage tissue was obtained from the Norfolk and Norwich University Hospital from 9 patients (3 HOA, 3 NOF and 3 KOA). RNA from all remaining tissues was from a pool of 3 patients obtained from Applied Biosystems, the data for expression in all tissues other than cartilage shows  $n = 3$  technical replicates. Candidate microRNA novel 2 (A) showed highest expression values in RNA from cervix and skeletal muscle tissue. Candidate microRNA novel 7 (B) appeared to be most highly expressed in cervix and skeletal RNA, with expression in cartilage tissue relatively high as well. In contrast, Candidate microRNA novel 11 (C) showed the highest expression in cartilage RNA compared to all other tissues tested. The expression of the remaining candidate microRNAs is shown in Appendix 9.

Figure 4.10 shows the expression of miR-140-3p (A), miR-140-5p (B), miR-455-3p (C) and miR-455-5p (D) across the same 23 human tissues. Both miR-140-3p and miR-140-5p show highest expression in all the cartilage tissues, supporting the fact that miR-140 is a cartilage selective microRNA. MiR-455-3p appears to be most highly expressed in cervix tissue, with the expression in cartilage being comparable to the remaining tissues.

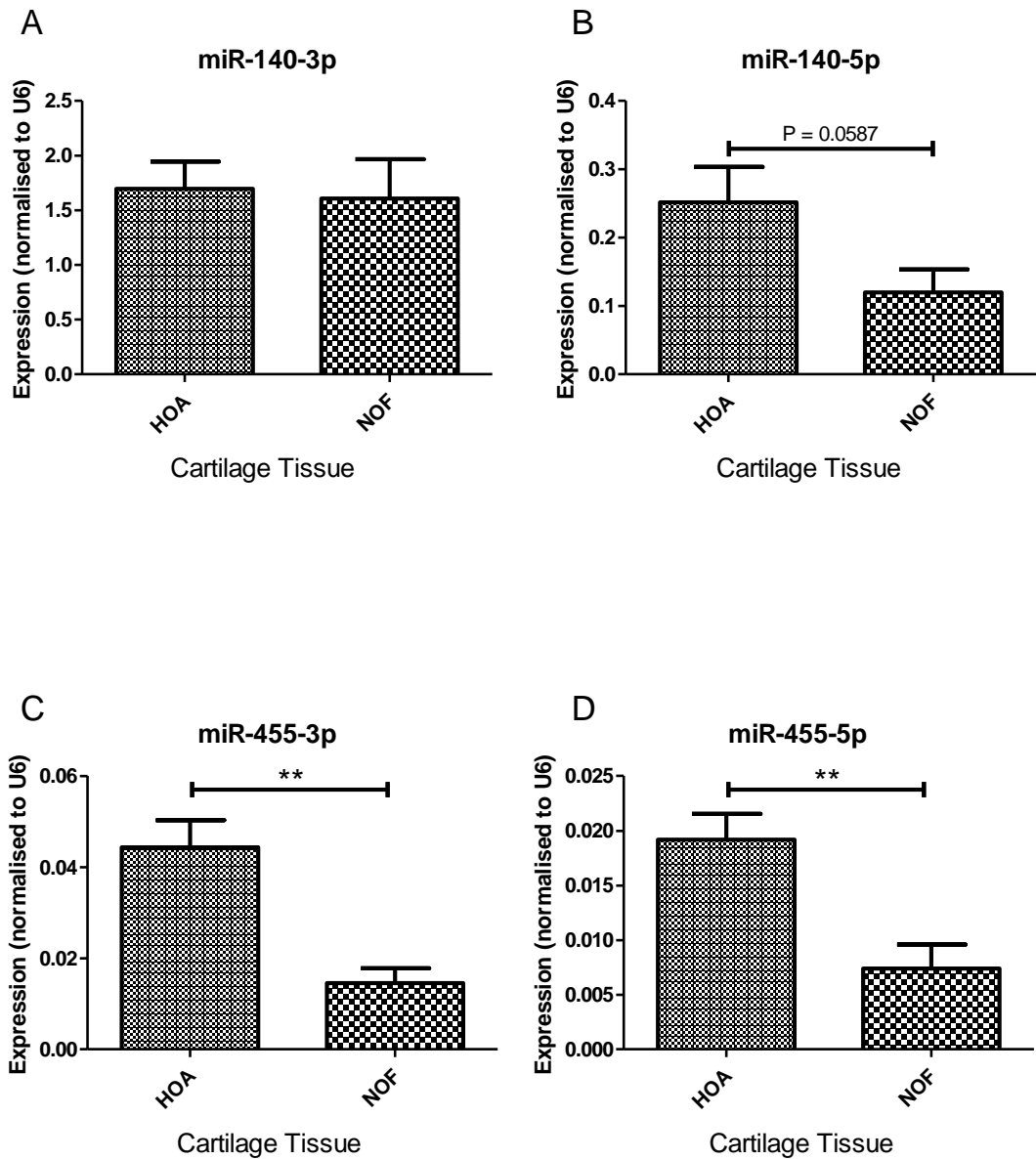
MiR-455-5p on the other hand shows the expression of the microRNA in osteoarthritic cartilage tissue (KOA and HOA) to be similar to cervix tissue, with expression in NOF and other tissues much lower.

In addition to analyzing the expression of candidate microRNAs in a number of tissues to assess their tissue and disease specificity, a cartilage injury assay was carried out using mouse hip avulsion (Chong et al., 2013; Stanton et al., 2011). (Fig. 11). The cartilage was removed from the hip joints of 3 week old mice and cultured over a period of 48 hours. Expression of the three candidate microRNAs; (A) novel 7 and (B) novel 11 was measured over set time points during the 48 hour period. Candidate microRNA novel 2 was not measured as it is not found within the mouse genome. Candidate microRNA novel 7 increased in expression from hour 0 to 48. A statistically significant difference was seen between 48 hours and all other time points of 0, 1, 3, 6, 12 and 24 hours ( $p \leq 0.001$ ,  $p \leq 0.001$ ,  $p \leq 0.001$ ,  $p \leq 0.05$ ,  $p \leq 0.001$ , and  $p \leq 0.05$  respectively). Candidate microRNA novel 11 appears to show a cyclic expression pattern with peaks in expression seen at 3 and 48 hours, however none of these changes reach statistical significance.

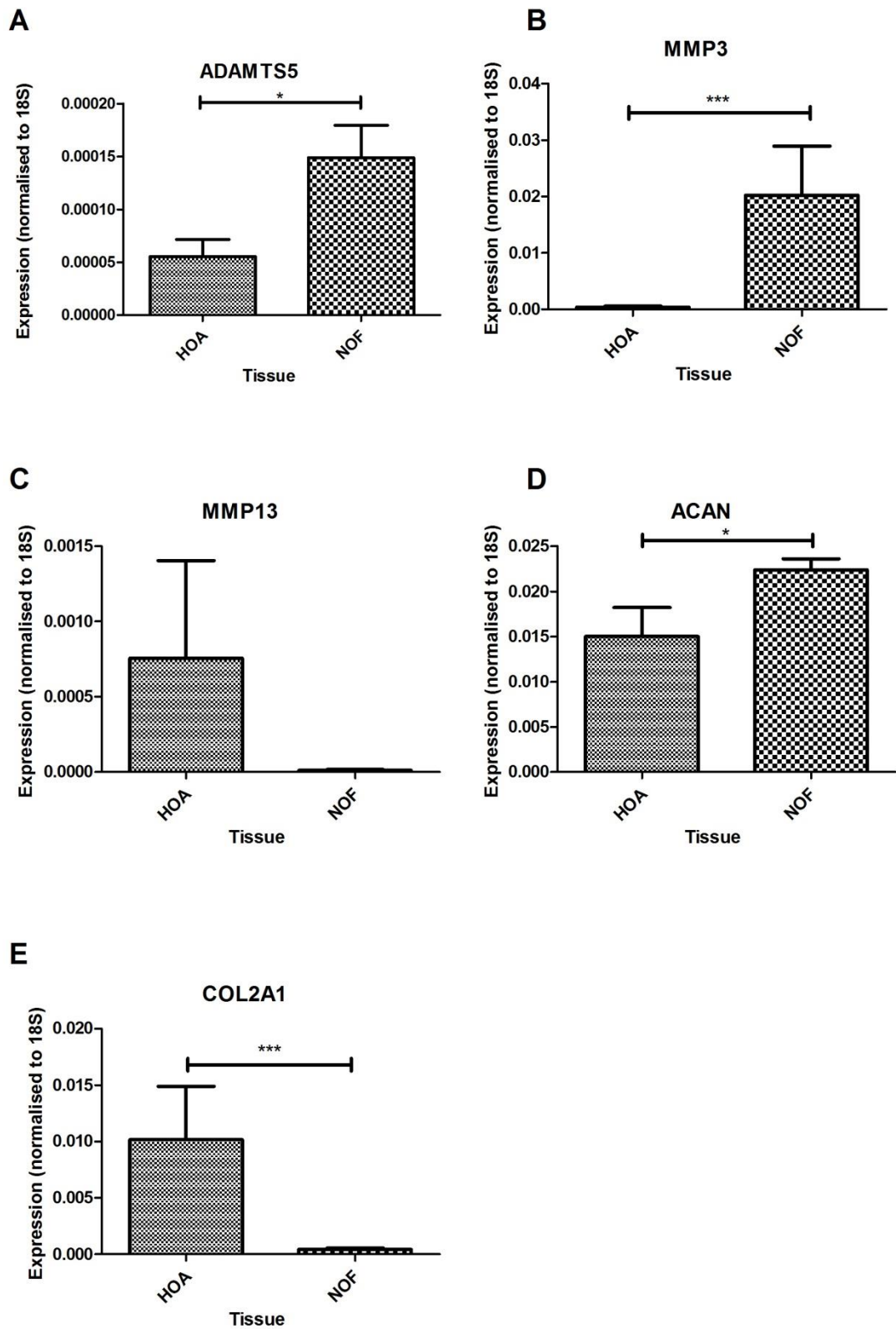
Figure 4.12 shows the expression of (A) miR-140-3p, (B) miR-140-5p and (C) miR-455-5p in the hip avulsion assay (the miR-455-3p sequence is not conserved between mouse and human and therefore not measured). All three microRNAs showed an increase in expression through the time points, peaking at 48 hours, with miR-140-3p and miR-140-5p reaching statistical significance.



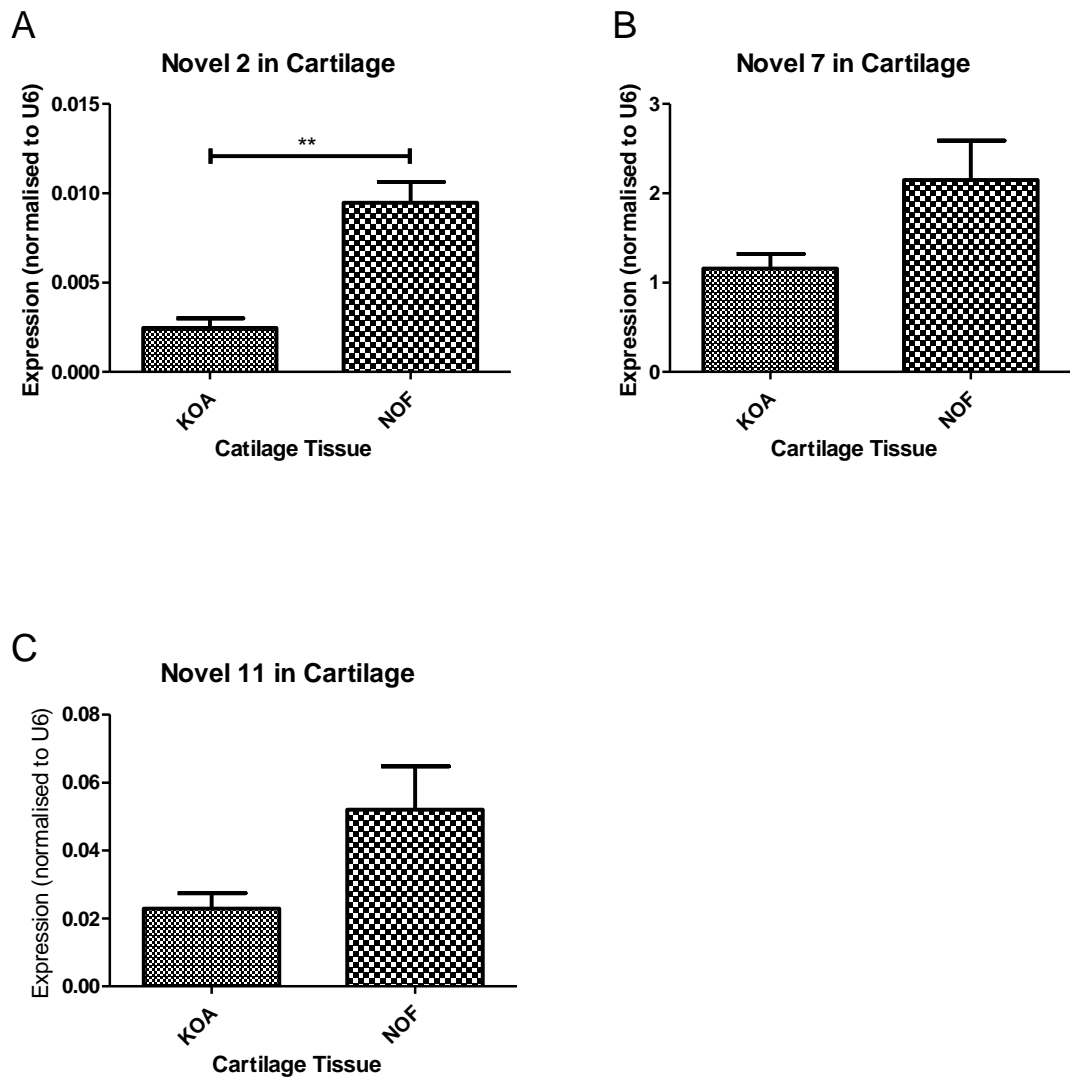
**Figure 4.4 Candidate novel microRNA expression in HOA and NOF cartilage tissue.** Expression levels of candidate microRNAs (A) novel 2, (B) novel 7 and (C) novel 11 were obtained by qRT-PCR from RNA isolated from: osteoarthritic hip cartilage tissue (HOA) and neck of femur fracture cartilage (NOF). Data were normalised to *U6* RNA expression. Data shows mean  $\pm$  SEM,  $n = 6$  (analyzed by a Students T test, no significance was seen).



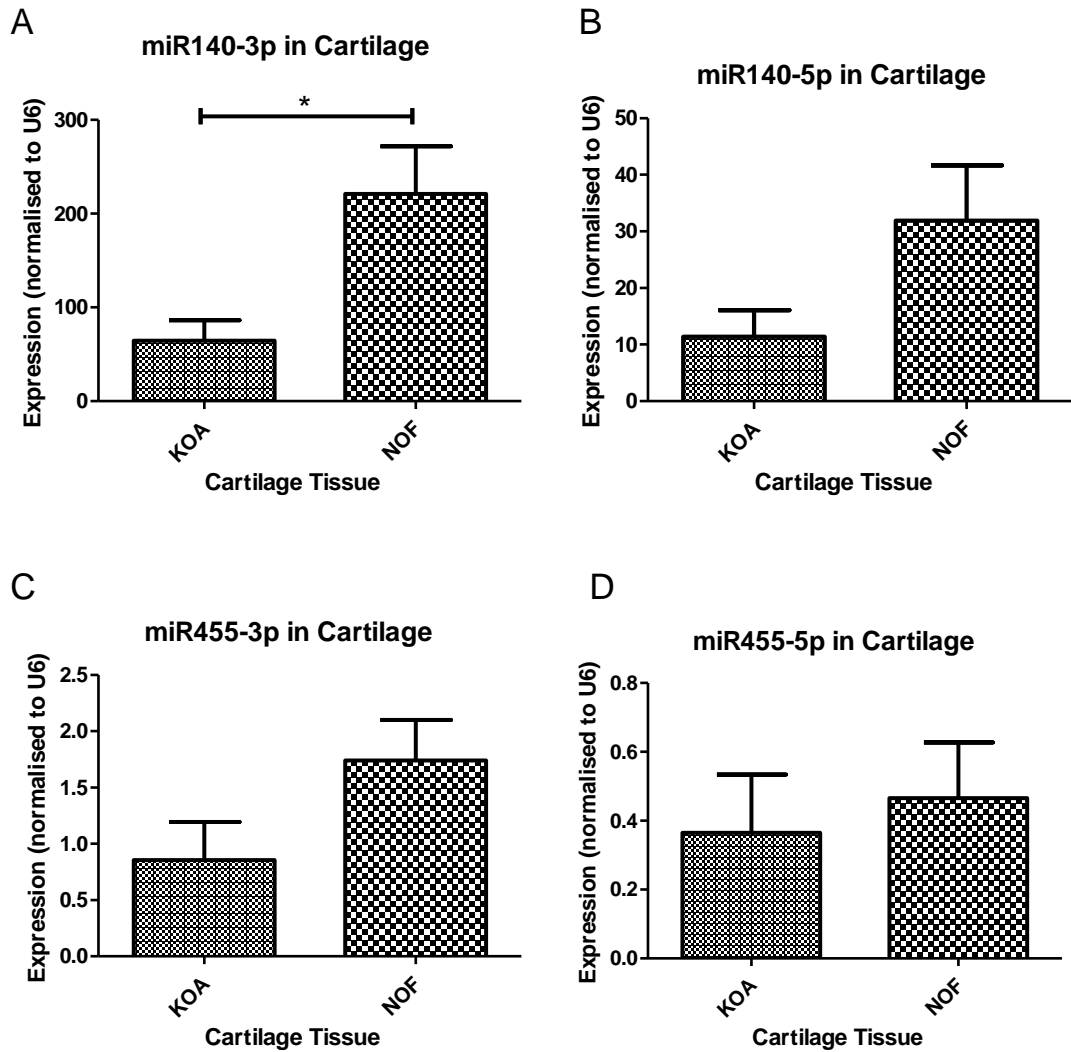
**Figure 4.5 MicroRNA expression in HOA and NOF cartilage tissue.** Expression levels of (A) miR-140-3p, (B) miR-140-5p, (C) miR-455-3p and (D) miR-455-5p were obtained by qRT-PCR from RNA isolated from: osteoarthritic hip cartilage tissue (HOA) and neck of femur fracture cartilage (NOF). Data were normalised to *U6* RNA expression. Data shows mean +/- SEM, n = 6 (\*\*,  $p \leq 0.01$  analyzed by Student t-test).



**Figure 4.6 Gene expression in HOA and NOF cartilage tissue.** Expression levels of (A) ADAMTS5, (B) MMP3, (C) MMP13, (D) ACAN, (E) COL2A1 were obtained by qRT-PCR from RNA isolated from: osteoarthritic hip cartilage tissue (HOA) and neck of femur fracture cartilage (NOF). Data normalised to *18S* RNA expression. Data shows mean  $\pm$  SEM,  $n = 6$ . (\*,  $p \leq 0.05$ , \*\*,  $p \leq 0.01$ , \*\*\*,  $p \leq 0.001$  analyzed by Student t-test).

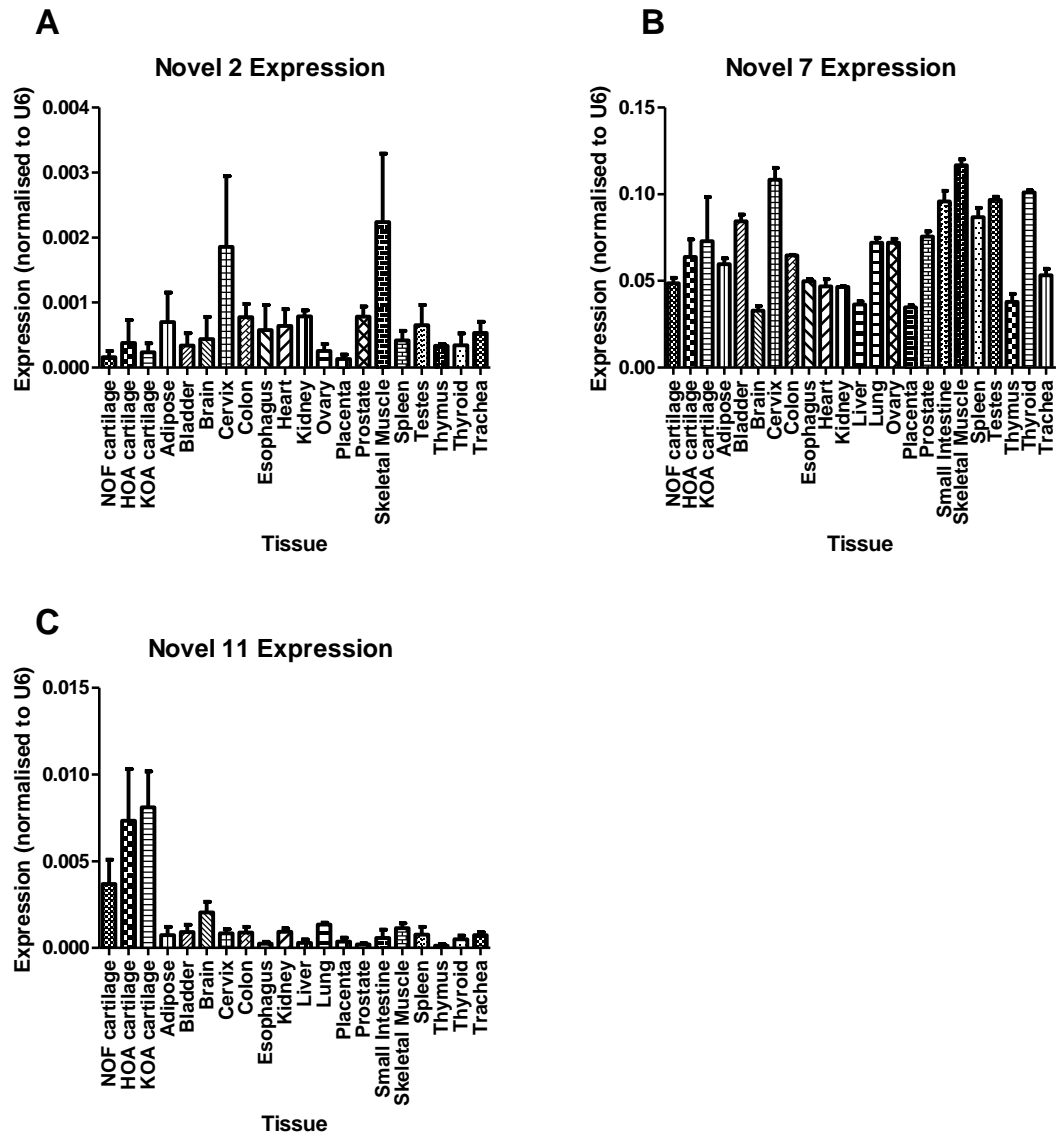


**Figure 4.7 Candidate novel microRNA expression in KOA and NOF cartilage tissue.** Expression levels of candidate microRNAs (A) novel 2, (B) novel 7 and (C) novel 11 were obtained by qRT-PCR from RNA isolated from: osteoarthritic knee cartilage tissue (KOA) and neck of femur fracture cartilage (NOF). Data were normalised to *U6* RNA expression. Data shows mean  $\pm$  SEM,  $n = 6$ . (\*\*,  $p \leq 0.01$  analyzed by Student t-test).

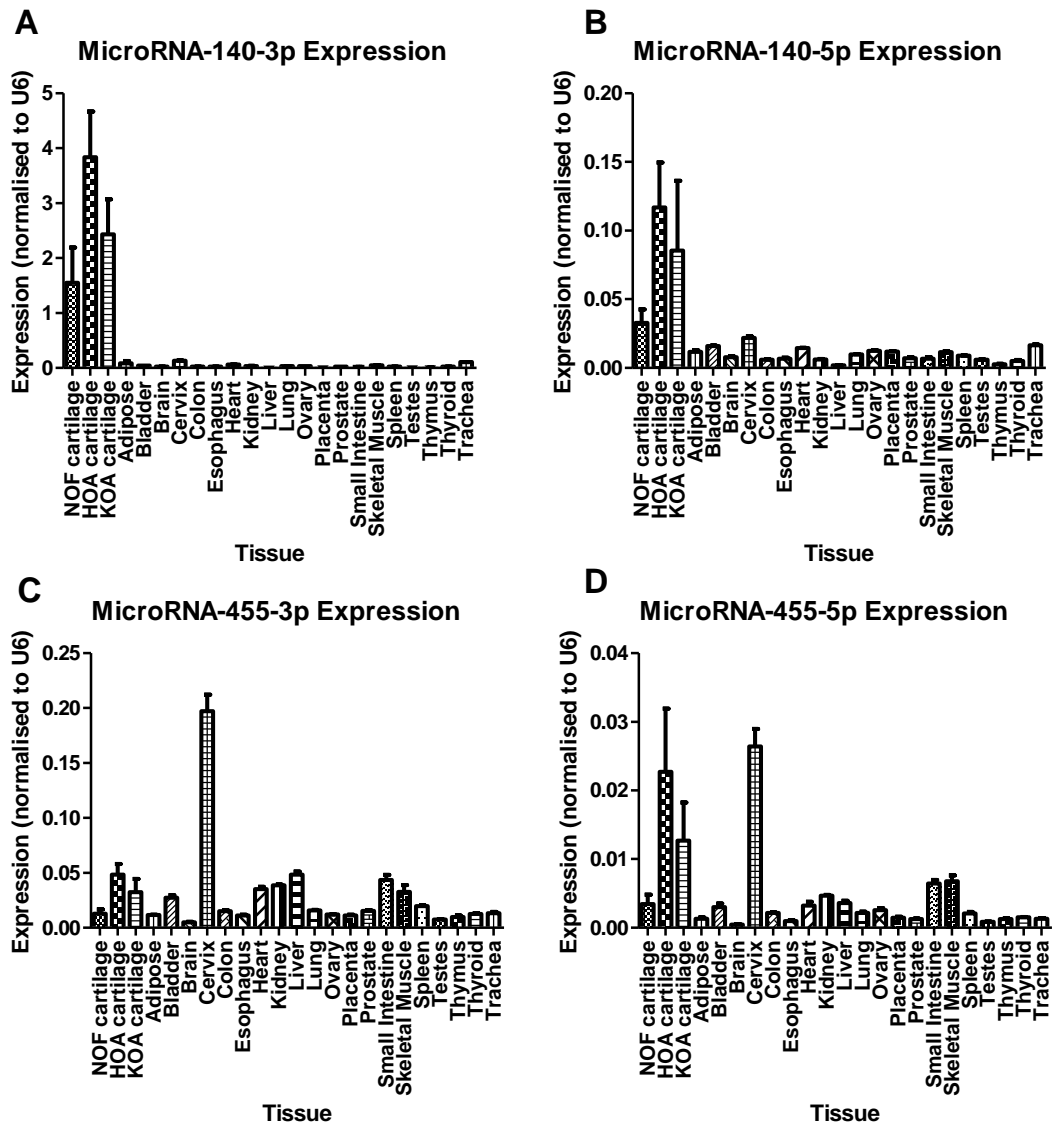


**Figure 4.8 MicroRNA expression in KOA and NOF cartilage tissue.** Expression levels of (A) miR-140-3p, (B) miR-140-5p, (C) miR-455-3p and (D) miR-455-5p were obtained by qRT-PCR from RNA isolated from: osteoarthritic knee cartilage tissue (KOA) and neck of femur fracture cartilage (NOF). Data were normalised to *U6* RNA expression. Data shows mean  $\pm$  SEM,  $n = 6$  (\*,  $p \leq 0.05$  analyzed by Student t-test).

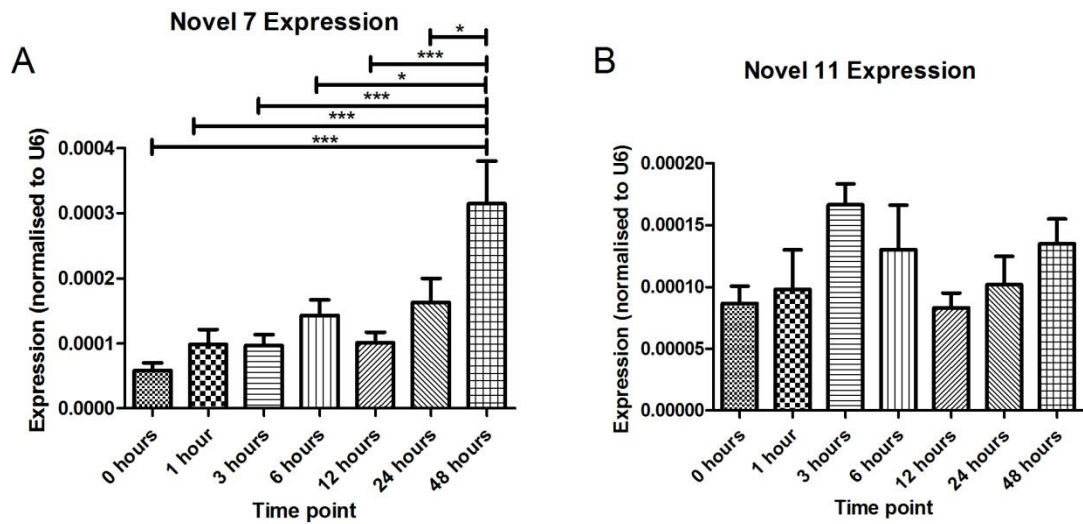




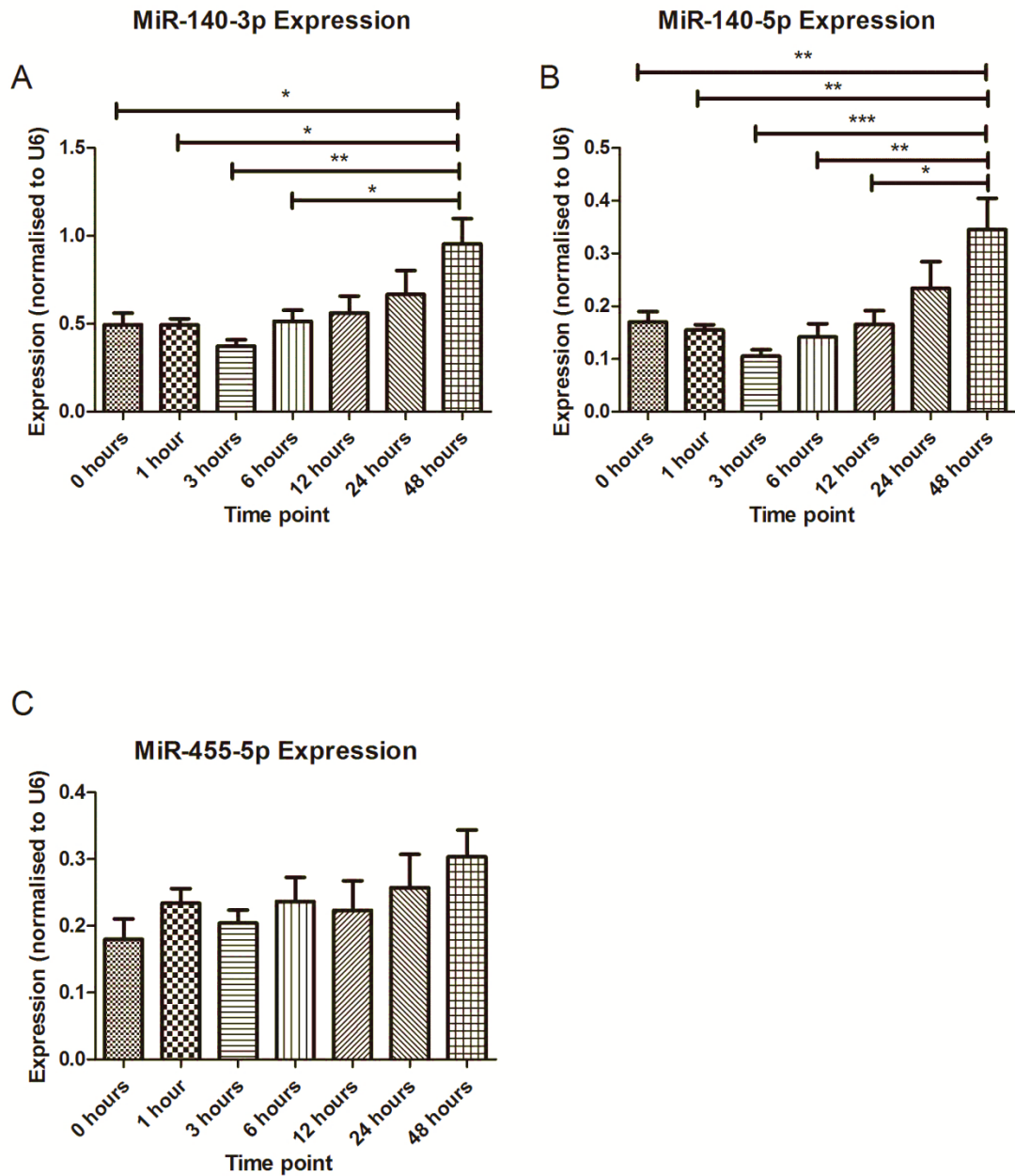
**Figure 4.9 Candidate novel microRNA expression across a human tissue panel.** Expression levels of candidate microRNAs (A) novel 2, (B) novel 7 and (C) novel 11 were obtained by qRT-PCR from RNA isolated from 23 human tissues including adipose, bladder, brain, cervix, colon, oesophagus, heart, kidney, liver, lung, ovary, placenta, prostate, small intestine, skeletal muscle, spleen, testes, thymus, thyroid, trachea, osteoarthritic knee cartilage tissue (KOA), osteoarthritic hip cartilage (HOA) and neck of femur fracture cartilage (NOF). Data were normalised to *U6* RNA expression. RNA of all cartilage samples was from 3 patients. RNA from all remaining tissues was from a pool of 3 patients with n= 3 technical replicates.



**Figure 4.10 MicroRNA expression across a human tissue panel.** Expression levels of (A) miR-140-3p, (B) miR-140-5p, (C) miR-455-3p and (D) miR-455-5p were obtained by qRT-PCR from RNA isolated from 23 human tissues including adipose, bladder, brain, cervix, colon, oesophagus, heart, kidney, liver, lung, ovary, placenta, prostate, small intestine, skeletal muscle, spleen, testes, thymus, thyroid, trachea, osteoarthritic knee cartilage tissue (KOA), osteoarthritic hip cartilage (HOA) and neck of femur fracture cartilage (NOF). Data were normalised to *U6* RNA expression. RNA of all cartilage samples was from 3 patients. RNA from all remaining tissues was from a pool of 3 patients with  $n = 3$  technical replicates.



**Figure 4.11 Candidate novel microRNA expression in a cartilage injury model.** Expression levels of candidate (A) novel 7 and (B) novel 11 were obtained by qRT-PCR from RNA isolated from hip cartilage removed from 3-5 week old mice and incubated in culture medium for 48 hours. RNA was isolated at the time points; 0, 1, 3, 6, 12, 24 and 48 hours. Data shows mean +/- SEM, n = 8. Data were normalised to *U6* RNA expression (\*,  $p \leq 0.05$ , \*\*\*,  $p \leq 0.001$  analyzed by a one way ANOVA with a post hoc Tukey test).



**Figure 4.12 MicroRNA expression in a cartilage injury model.** Expression levels of (A) miR-140-3p, (B) miR-140-5p and (C) miR-455-5p were obtained by qRT-PCR from RNA isolated from hip cartilage removed from 3 week old mice and incubated in culture medium for 48 hours. RNA was isolated at the time points; 0, 1, 3, 6, 12, 24 and 48 hours. Data shows mean  $\pm$  SEM,  $n = 8$ . Data were normalised to *U6* RNA expression (\*,  $p \leq 0.05$ , \*\*,  $p \leq 0.01$ , \*\*\*,  $p \leq 0.001$  analyzed by a one way ANOVA with a post hoc Tukey test).

### 4.2.3 Candidate novel microRNA expression during chondrocyte differentiation and de-differentiation

Further to exploring the expression of candidate microRNAs in a variety of human tissues, the expression of candidate microRNAs novel 2, novel 7 and novel 11 were measured in a number of differentiation and de-differentiation of chondrocyte assays.

Figure 4.13 shows a mesenchymal stem cell (MSC) chondrogenesis assay, where MSC cells are induced through chondrogenesis with TGF $\beta$ 3, ITS+L (Insulin-Transferrin-Selenium + L-ascorbic acid), ascorbic acid and dexamethasone to form chondrocytes, resulting in the formation of a cartilage disc at day 14. This assay was conducted by Dr M. Barter (Newcastle University, UK). RNA was isolated from the assay at set time points, and qRT-PCR was used to measure microRNA expression. Figure 4.13 shows candidate microRNA (A) novel 2, (B) novel 7 and (C) novel 11 expression at time points 0, 3, 7 and 14 days. Candidate microRNA novel 2 showed an increase in expression up to day 7 before decreasing again at day 14. A statistically significant difference was seen between day 0 and day 7 ( $p \leq 0.05$ ) and between day 7 and day 14 ( $p \leq 0.05$ ). Candidate microRNA novel 7 showed an increase in expression through all time points, peaking at day 14, becoming statistically significant, even by day 3 ( $p \leq 0.05$ ). Finally, candidate microRNA novel 11 expression was at its highest at day 0, and subsequently decreased at each time point to day 14, however none of these changes were statistically significant.

The expression of a number of known microRNAs was also measured by Dr M. Barter by microarray (data not shown). The data showed that miR-140-3p, miR-140-5p, miR-455-3p and miR-455-5p increased during chondrogenesis, with the highest expression of each seen in the cartilage disc at day 14.

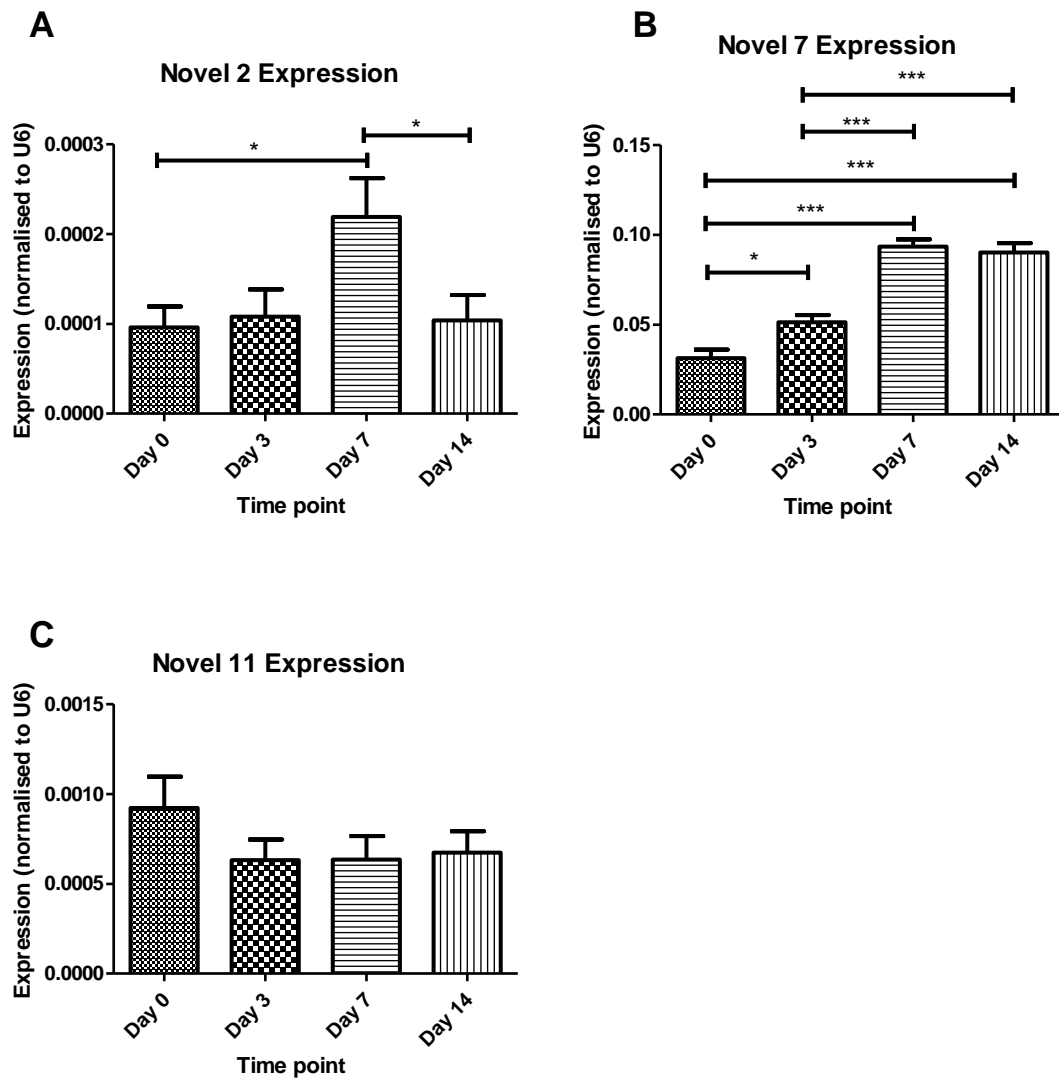
Figure 4.14 shows the expression of candidate microRNA (A) novel 7 and (B) novel 11 in an ATDC5 (mouse embryonic carcinoma cell) chondrogenesis assay, where ATDC5 cells are induced through chondrogenesis by stimulation with insulin, resulting in the development of chondrocytes. This assay was conducted by Dr T. Swingler (UEA). RNA was isolated from this assay at set time points, and qRT-PCR was used to measure microRNA expression. Figure 4.14 shows candidate microRNA (A) novel 7 and (B) novel 11 expression at time points 1, 21, and 42 days in both induced and control cells. Novel 2 was not measured as the sequence is not present in the mouse genome. The expression of candidate microRNA novel 7 increased through time in both control and

induced cells, peaking at day 42. Candidate microRNA Novel 11 expression decreased over time in both induced and control cells, apart from day 42 where there was a considerable decrease in expression in induced cells compared to control. However none of these changes were statistically significant.

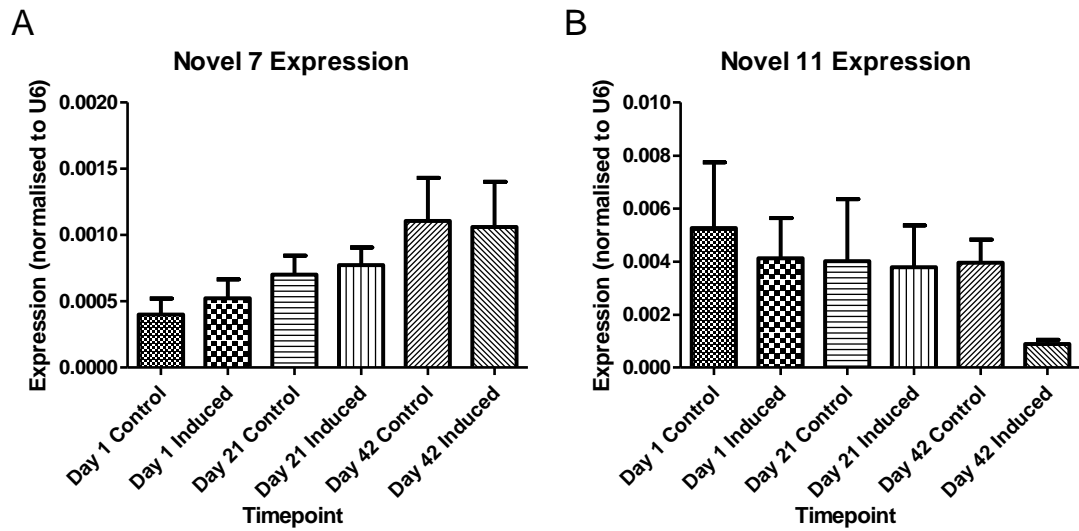
The known microRNAs miR-140-3p and miR-140-5p were also measured in the ATDC5 assay in figure 4.15 (A) and (B) respectively. The expression of both microRNAs is relatively low at all 3 time points in both induced and control cells except for induced cells at day 42. At this point there is a significant increase in expression with  $p \leq 0.001$  between induced cells at day 42 and every other time point.

As well as determining the expression of candidate novel microRNAs during chondrogenesis, expression was also measured during a chondrocyte de-differentiation model (figure 4.16), where expression is measured in RNA obtained from osteoarthritic cartilage, cells directly isolated from this cartilage, and primary chondrocytes passaged in monolayer culture.

Candidate microRNA novel 2 showed a decrease in expression from cartilage to passage 3; however this decrease was not statistically significant. Candidate microRNA novel 7 showed the highest expression in cartilage tissue, but this decreased significantly in cells isolated immediately after digestion (PD) ( $p \leq 0.001$ ) and expression decreased through all subsequent passaged cells. Candidate microRNA novel 11 showed a decreasing expression from cartilage to monolayer chondrocytes, and then a continued low expression across passage ( $p \leq 0.001$  between cartilage and post digest and  $p \leq 0.001$  between post digest and passage 0).

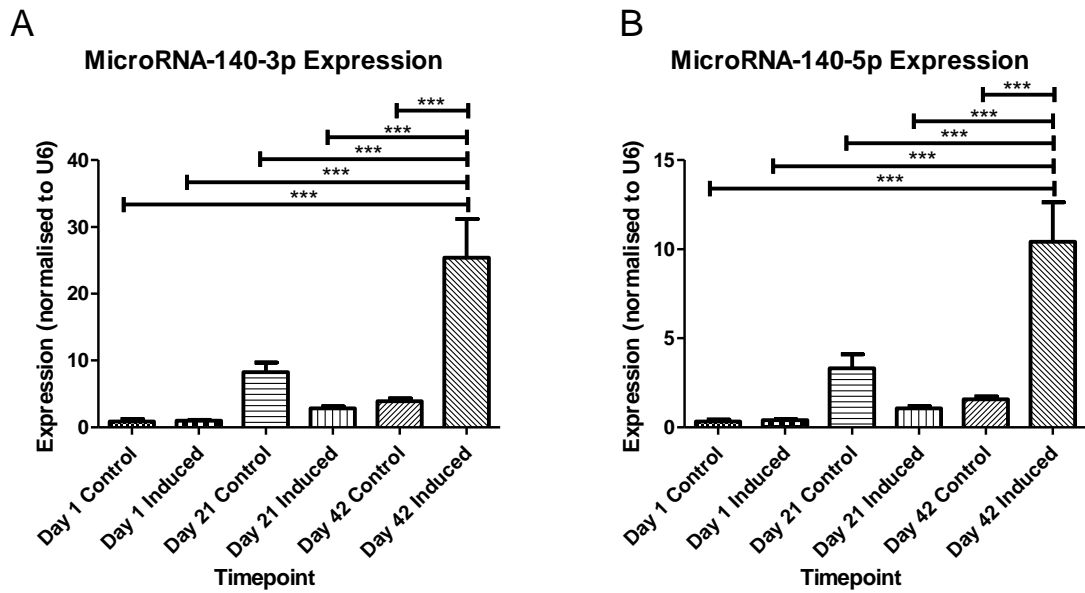


**Figure 4.13 Candidate novel microRNA expression in a mesenchymal stem cell chondrogenesis assay.** Expression levels of candidate microRNA (A) novel 2, (B) novel 7 and (C) novel 11 were measured by qRT-PCR from RNA obtained from mesenchymal stem cells induced through chondrogenesis to form cartilage discs. RNA samples were taken at time points during the assay: day 0, 3, 7, 14. Data shows mean  $\pm$  SEM,  $n = 3$ , data were normalised to *U6* RNA expression (\*,  $p \leq 0.05$ , and \*\*\*,  $p \leq 0.001$  analyzed by a one way ANOVA with a post hoc Tukey test).

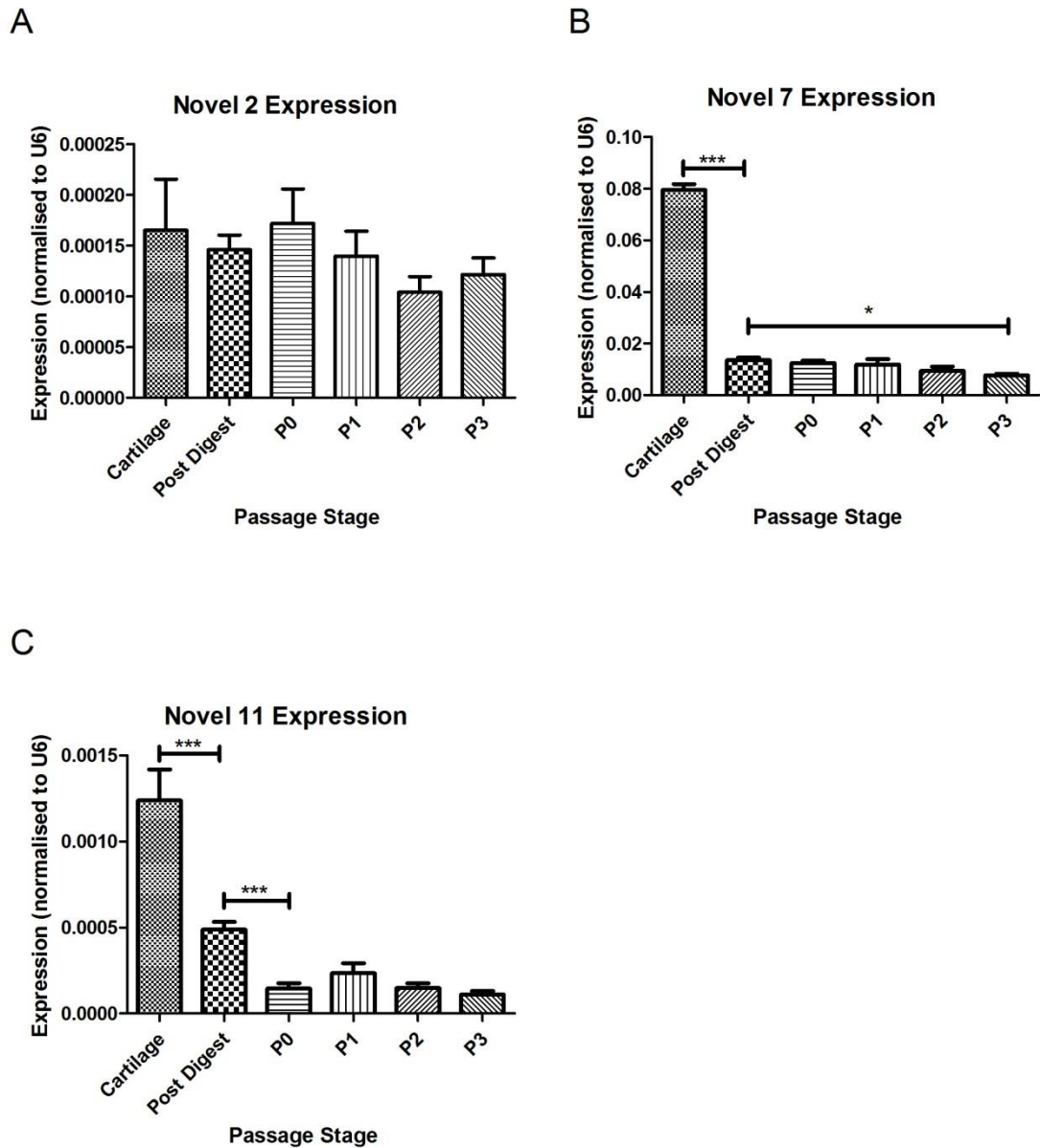


**Figure 4.14 Candidate novel microRNA expression in an ATDC5 chondrogenesis assay.** Expression levels of candidate microRNA (A) novel 7 and (B) novel 11 were measured by qRT-PCR from RNA obtained from ATDC5 mouse carcinoma cells induced through chondrogenesis. Control samples were not induced through chondrogenesis. RNA samples were taken at time points during the assay: day 1, 21, 42. Data shows mean  $\pm$  SEM,  $n = 3$ , data were normalised to *U6* RNA expression (from one experiment).





**Figure 4.15 MicroRNA expression in an ATDC5 chondrogenesis assay.** Expression levels of (A) miR-140-3p and (B) miR-140-5p were measured by qRT-PCR from RNA obtained from ATDC5 mouse carcinoma cells induced through chondrogenesis. Control samples were not induced through chondrogenesis. RNA samples were taken at time points during the assay: day 1, 21, 42. Data shows mean  $\pm$  SEM,  $n = 3$ , data were normalised to *U6* RNA expression (\*\*\*,  $p \leq 0.001$ , analyzed by a one way ANOVA with a post hoc Tukey test).



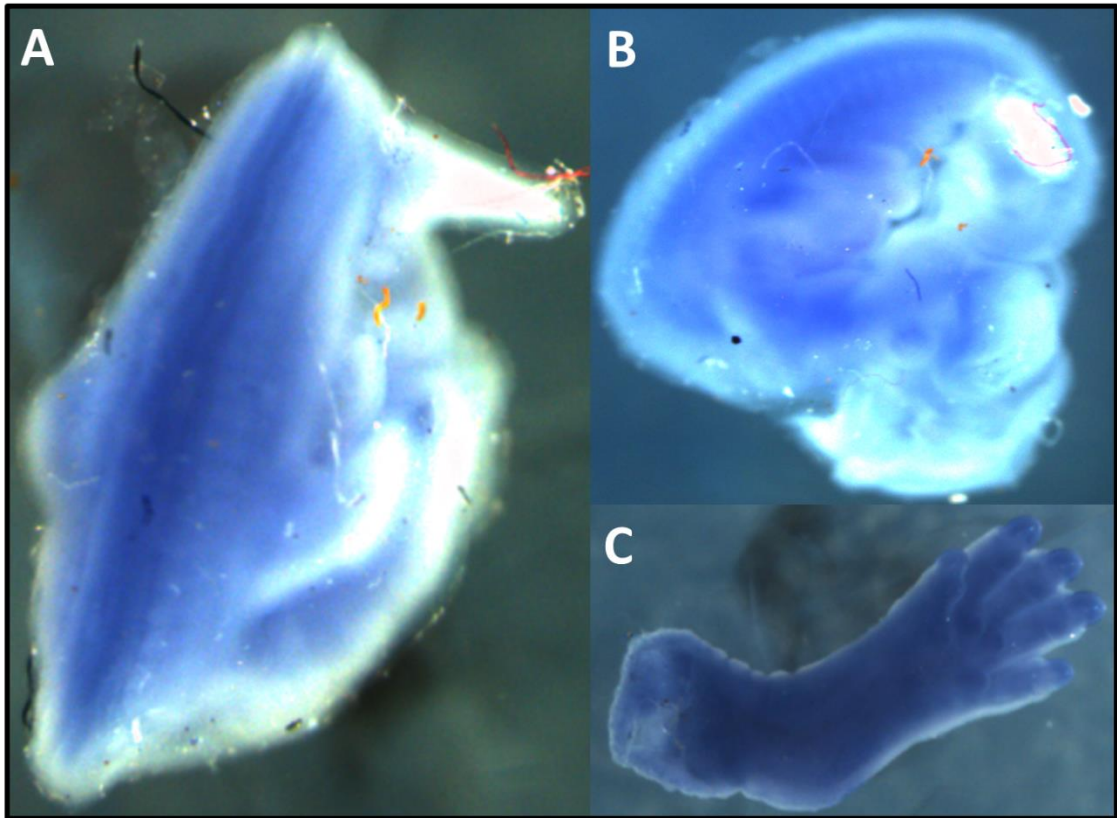
**Figure 4.16 MicroRNA expression in articular chondrocyte isolation and expansion.** Expression levels of candidate microRNA (A) novel 2, (B) novel 7 and (C) novel 11 were measured by qRT-PCR from RNA isolated from: osteoarthritic knee cartilage tissue, isolated chondrocytes (post digest) and subsequent P0, P1, P2, and P3 chondrocytes passaged through monolayer culture. Data were normalised to *U6* RNA expression. (\*,  $p \leq 0.05$ , and \*\*\*,  $p \leq 0.001$ . RNA was obtained from 8 patients for chondrocytes and 4 patients for tissue, data shows mean  $\pm$  SEM, a one way ANOVA analysis with a post hoc Tukey test was used to test for significance).

#### **4.2.4 Candidate novel microRNA expression in vivo**

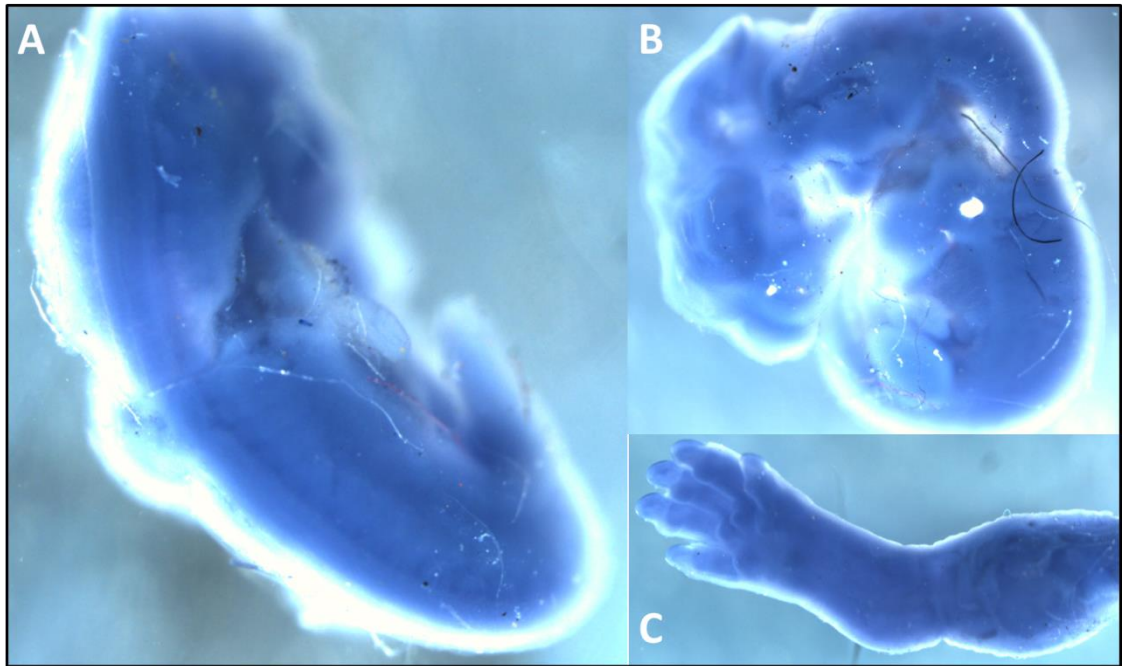
Expression of the candidate novel microRNAs was measured in a number of *in vitro* assays as described previously in this chapter. In order to complement that data, the expression of one of the candidates, novel 11, was also measured *in vivo* using *in situ* hybridisation.

Figure 4.17 shows the localisation of candidate microRNA novel 11 in the whole mouse embryo at stage E11 and limb at stage E14. Whilst no localised staining is noticeable in the limb (C), there appears to be some localisation along the spine (A) seen in the dorsal view of the whole E11 embryo. This expression cannot be seen in the lateral view of the whole embryo (B) and doesn't appear to be present between the spinal segments.

MiR-449 (a brain specific microRNA) was used as a control (Fig. 4.18) and expression was measured *in vivo* in whole E11 embryos and whole E14 limbs. No noticeable expression was seen in any area of the (A) dorsal or (B) lateral view of the whole embryo or the (C) whole limb.



**Figure 4.17 Expression of candidate microRNA novel 11 during mouse development.** *In situ* hybridization was carried out upon whole-mount mouse embryos (A, B) and dissected limbs (C). Expression of candidate novel 11 was not detected in limbs at E14 (C). Very faint staining was seen in the spinal cord of E11 embryo (A) but not between the vertebrae or in the E11 lateral view (B). Background staining on all tissue was relatively high.



**Figure 4.18 Expression of control microRNA miR-449 during mouse development.** *In situ* hybridization was carried out upon whole-mount mouse embryos (A, B) and dissected limbs (C). Expression of miR-449 was not detected in E11 whole embryos (A, B) or limbs at E14 (C).

#### 4.2.5 Genetic location and evolution of candidate novel microRNAs

Experimental validation is an important aspect of microRNA discovery. However a bioinformatics assessment is also useful. Figures 4.19, 4.20, 4.21 describe the location of the 3 candidate microRNAs investigated within their host genes.

Novel 2 is located within WD Repeat Domain 91 (WDR91) (Fig. 4.19). It is found within the 4<sup>th</sup> intron of each transcript of WDR91 except ENST00000485942, where it is found within the 1<sup>st</sup> intron.

Novel 7 is located within small nucleolar RNA, H/ACA box 36A (SNORA36A) (Fig. 4.20). Unlike novel 2 it is not an intronic microRNA, being located in the only exon of SNORA36A.

Novel 11 is located within cartilage acidic protein 1 (CRTAC1) (Fig. 4.21). It is found within the last intron of each transcript (the 11<sup>th</sup>, 14<sup>th</sup> and 13<sup>th</sup> respectively of ENST00000413387, ENST00000370597 and ENST00000298819). However the 2<sup>nd</sup> transcript is known to undergo splicing during transcription which would result in the removal of the microRNA sequence.

In order to predict the presence of the candidate microRNAs within multiple species aside from humans, an analysis was carried out using Ensembl ([www.ensembl.org](http://www.ensembl.org)) to search for each sequence within the species' genomes available. Figures 4.22, 4.23 and 4.24 show all species where the sequence was present (in either partial or complete form), and presents this in as an evolutionary tree, with species linked by time of divergence. The presence of an exact copy of the candidates is denoted by a green box.

The candidate microRNA novel 2 (Fig. 4.22) sequence is found in partial form in a number of species, but the exact sequence is found within *Homo sapiens*, *Pan troglodytes*, *Pongo abelii*, and *Callithrix jacchus*.

The candidate microRNA novel 7 (Fig. 4.23) sequence is found in partial form in a number of species, but the exact sequence is found within *Homo sapiens*, *Pan troglodytes*, *Gorilla gorilla*, *Pongo abelii*, *Nomascus leucogenys*, *Macaca mulatta* and *Callithrix jacchus*.

The candidate microRNA novel 11 (Fig. 4.24) sequence is found in partial form in a number of species, but the exact sequence is found within *Homo sapiens*, *Pan*

*troglodytes, Gorilla gorilla, Pongo abelii, Nomascus leucogenys, Macaca mulatta, Callithrix jacchus, Microcebus murinus, Mustela putorius furo, Ailuropoda melanoleuca, Canis familiaris, Felis catus, Erinaceus europaeus, Vicugna pacos, Equus caballus, Myotis lucifugis, Pteropus vampyras, Rattus norvegicus, Mus musculus, Loxodonta Africana, Procavia capensis, Ictidomys tridecemlineatus, Dipodomys ordii and Tupaia belangeri.*

## Novel 2 in WDR91

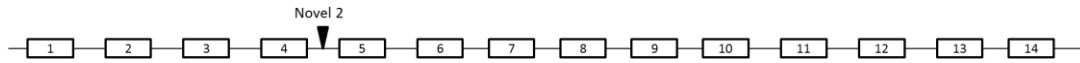
### Transcript

- ENST00000423565 – 4500bp – ENSP00000392555
- ENST00000354475 – 4588bp – ENSP00000346466
- ENST00000466182 – 3594bp – No protein product



### Transcript

- ENST00000344400 – 4506bp – ENSP00000340877



### Transcript

- ENST00000485942 – 582bp – No protein product



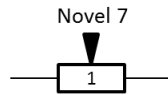
**Figure 4.19 Candidate microRNA novel 2 genetic location within its host gene.** The diagram shows the intron location of the candidate microRNA novel 2 within the multiple transcripts of its host gene, WD Repeat Domain 91. The boxes represent exons and the dashes represent introns.



## Novel 7 in SNORA36A

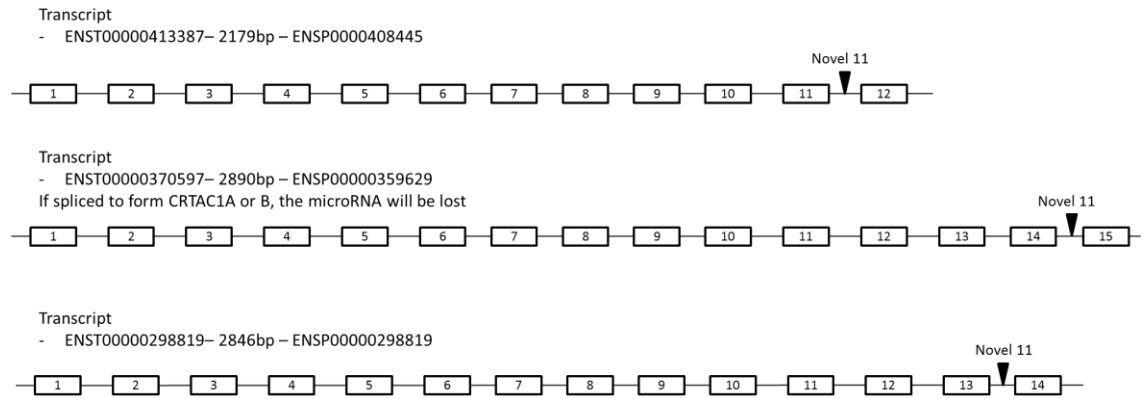
Transcript

- ENST00000384221- 132bp – No protein product



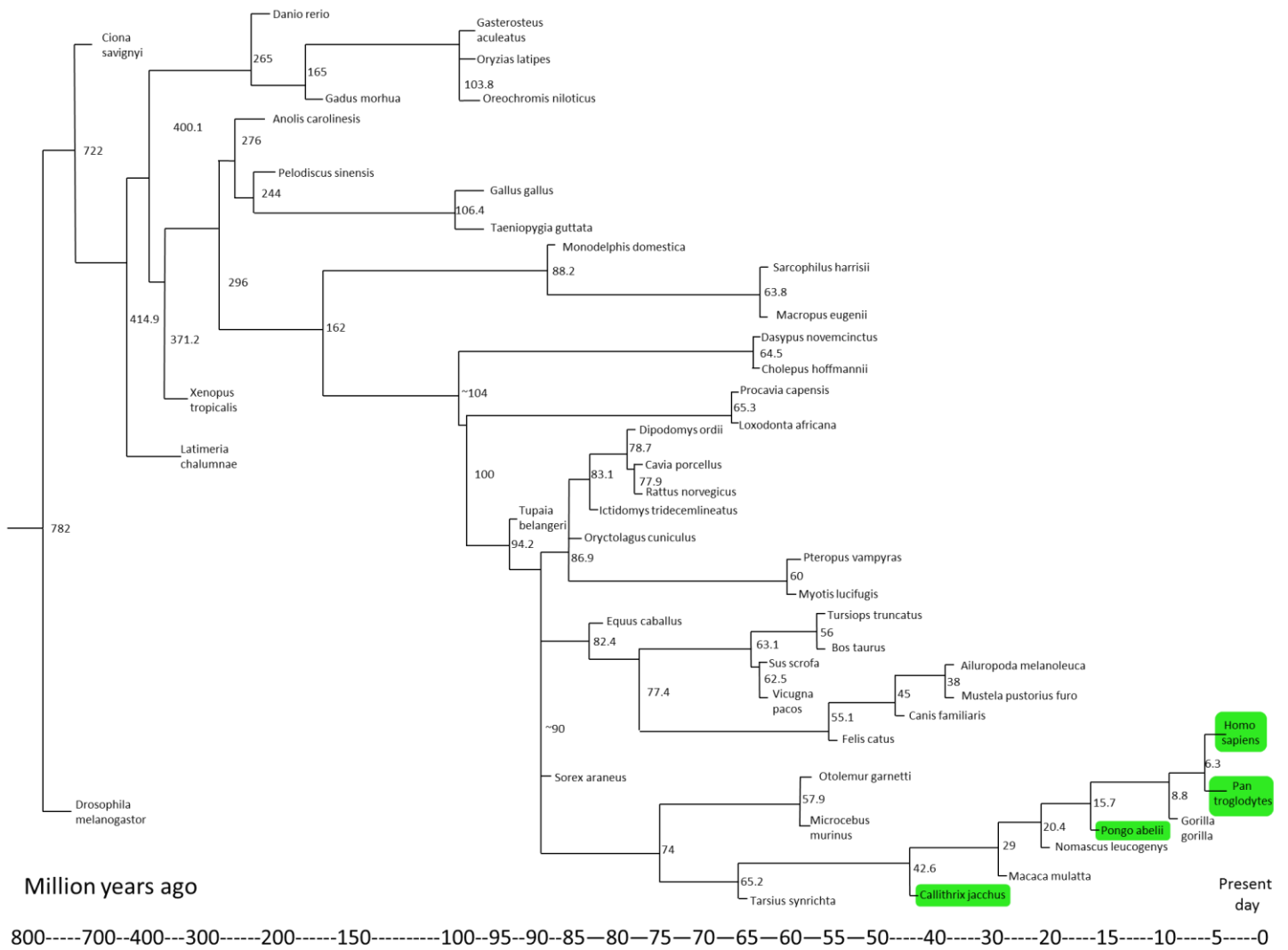
**Figure 4.20 Candidate microRNA novel 7 genetic location within its host gene.** The diagram shows the intron location of the candidate microRNA novel 7 within the transcript of its host gene, small nucleolar RNA, H/ACA box 36A. The box represents the exon and the dashes represent introns.

## Novel 11 in CRTAC1

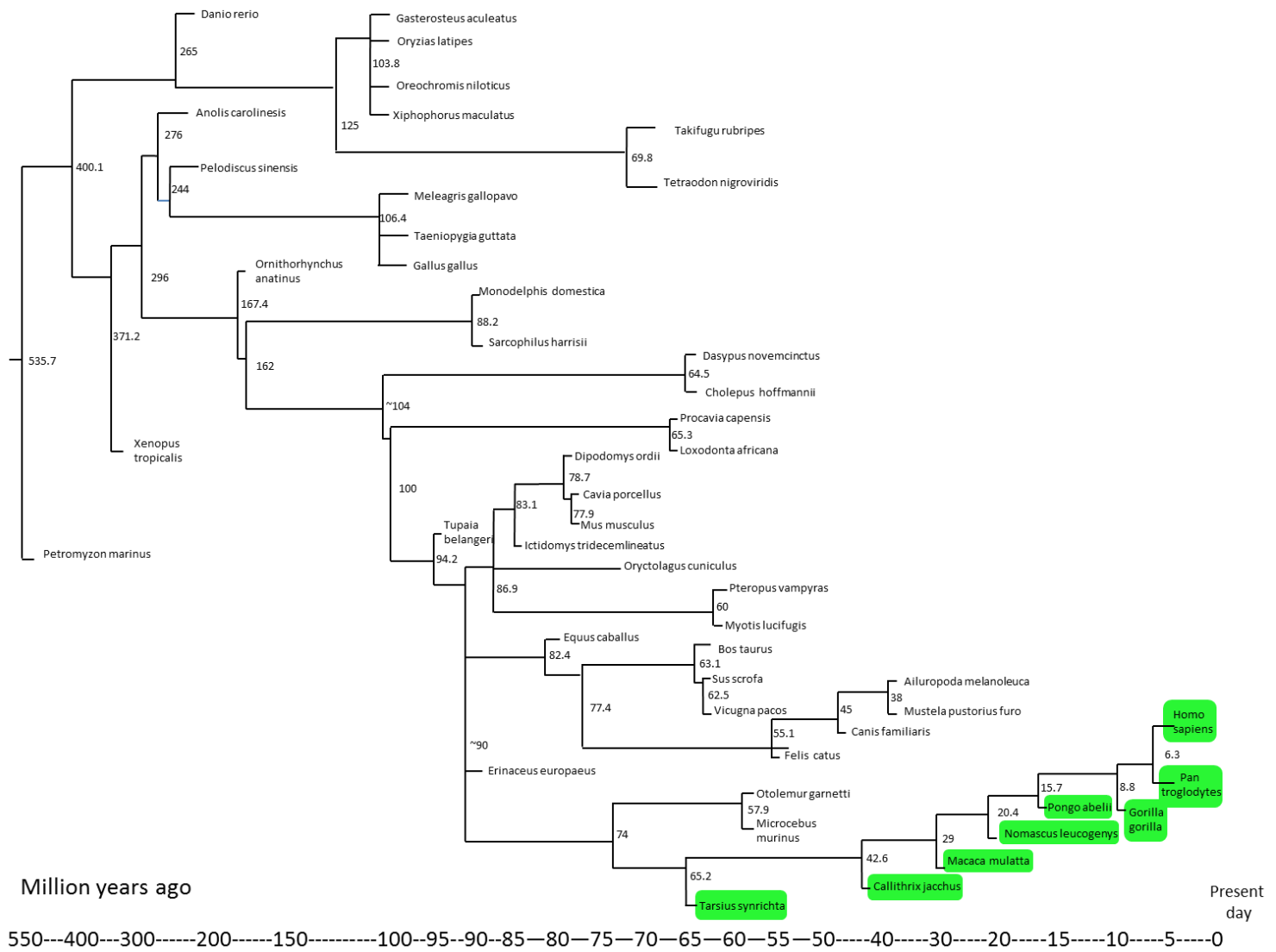


**Figure 4.21 Candidate microRNA novel 11 genetic location within its host gene.**

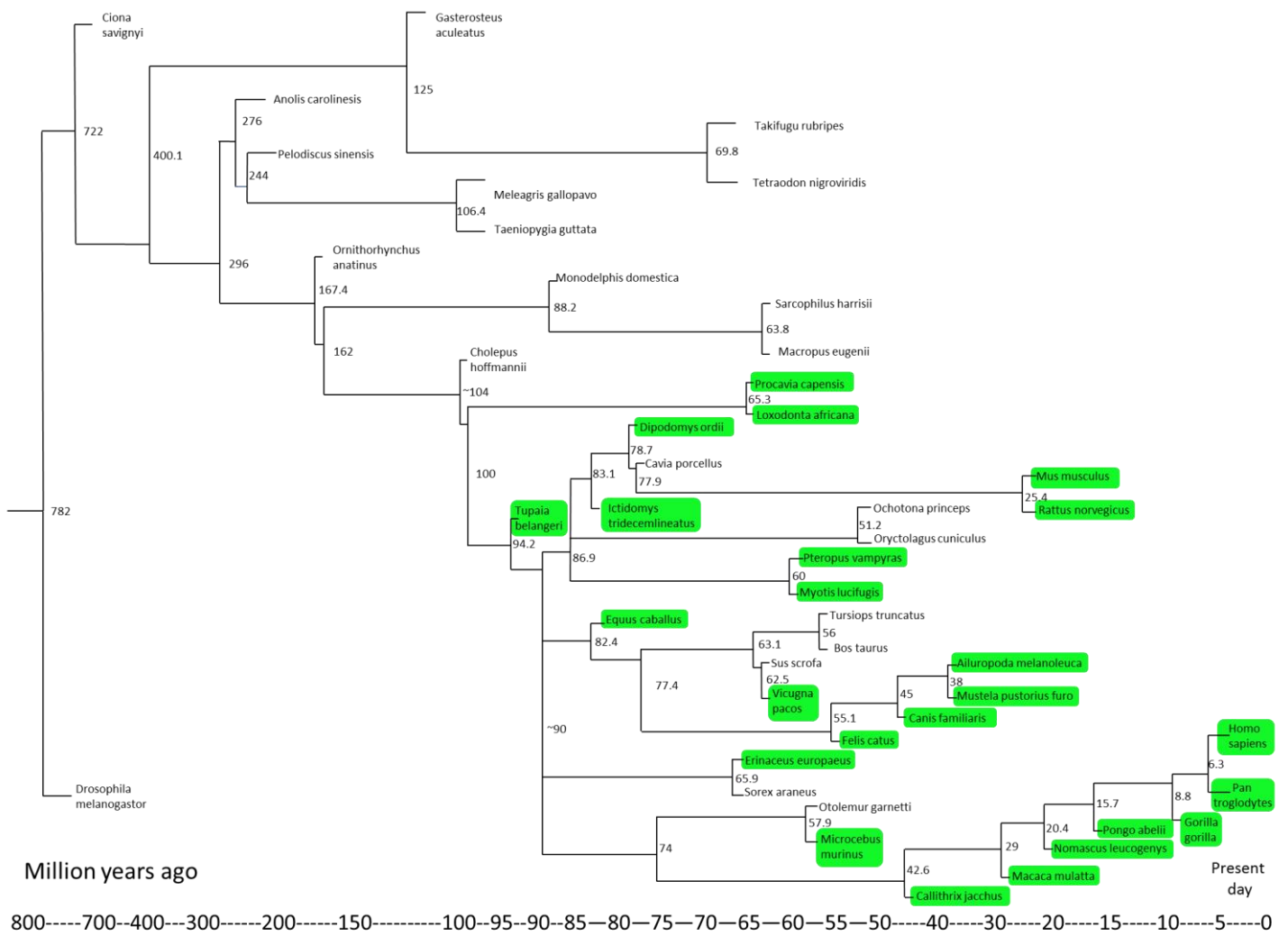
The diagram shows the intron location of the candidate microRNA novel 11 within the transcript of its host gene, cartilage acidic protein 1. The boxes represent exons and the dashes represent introns.



**Figure 4.22 Candidate microRNA novel 2 species evolutionary tree.** The diagram shows each species that the sequence is present in and the time of divergence between each. Green species indicates the presence of an exact sequence match. All other species contain a partial match. Time of divergence is measured in number of millions of years ago.



**Figure 4.23 Candidate microRNA novel 7 species evolutionary tree.** The diagram shows each species that the sequence is present in and the time of divergence between each. Green species indicates the presence of an exact sequence match. All other species contain a partial match. Time of divergence is measured in number of millions of years ago.



**Figure 4.24 Candidate microRNA novel 11 species evolutionary tree.** The diagram shows each species that the sequence is present in and the time of divergence between each. Green species indicates the presence of an exact sequence match. All other species contain a partial match. Time of divergence is measured in number of millions of years ago.

## 4.3 Discussion

The study of microRNAs is still a newly developing field, and as the technology progresses ever further, so too are more new sequences being discovered. This is shown by the large number of both new and known microRNAs found in cartilage during this project (shown in chapter 3).

However, in order to properly describe the miRNome of osteoarthritic cartilage, the microRNAs discovered must be proven to be true microRNAs by subsequent experimental validation, this is due to the problems of false positives in the detection process. False positives occur when the identification software falsely allocates a sequence to microRNA status. This could be due to degradation products fulfilling the identification criteria by chance, or perhaps a genuine small RNA sequence, however not a microRNA as they are currently described.

The first step was to identify the size of each candidate sequence using northern blotting. This is useful to identify if the sequence is the correct and expected size of a microRNA (~21bp). If the sequence is shown to be of a larger size by northern blot, then this suggests that the sequence discovered during the deep sequencing was indeed a product of degradation, and that the northern blot is detecting the original, entire sequence. Figure 4.1 shows the northern blots for 3 of the candidate microRNAs (northern blots were carried out for all 16 candidates and the remainder can be seen in Appendix 6). Candidate novel 7 (Fig. 4.1 B) showed a band of the usual microRNA size when compared to the size of miR-140-3p on northern blot. On the other hand, both candidate microRNAs 2 and 11 (Fig. 4.1 A and C) showed no detection by northern blot. Positive controls using synthesised DNA oligonucleotides of the complement sequences to the candidates worked well (as shown by the standard curves for each candidate) suggesting that the experiment itself and the primers used were working. The lack of detection could be explained by a very low expression value and so a band is not visible. In addition, RNA obtained from SW1353 cells was used for northern blotting due to the large amount of RNA needed, this is in contrast to the RNA obtained from primary chondrocytes used with the deep sequencing project, and could give another reason as to the lower expression of all known and candidate microRNAs detected with northern blot.

This low expression is supported by the pattern of expression seen from the deep sequencing project although this data came from primary chondrocytes rather than

SW1353 cells. Read numbers for all the potential candidates was relatively low compared to known described microRNAs. MiR-140-3p had an average read number of 8,031,127 in primary chondrocytes. Expression level of the same microRNA in 10µg of SW1353 chondrocyte cells was 2.5 fmoles when measured by northern blot. In comparison, miR-140-5p was barely detected by northern blot, yet had an average read number of 438,388 although this is high compared to other known microRNAs detected through deep sequencing (miR-140-5p was the 4<sup>th</sup> most highly detected microRNA). This suggests that, although northern blotting is the gold standard for sizing microRNAs, it could be less sensitive compared to other detection methods. This theory is supported in the literature by Fehr et al who showed higher sensitivity of detection with qRT-PCR than northern blot when measuring expression of *MMP3* mRNA (Fehr et al., 2000).

This apparent difference in detection efficiency between deep sequencing and northern blotting could explain why novel 2 and novel 11 remain undetected by northern blot (their respective average read numbers from sequencing were 0.488 and 11.867), but are likely to still be expressed at some level within chondrocytes and could still be classed as microRNAs. The remaining 14 candidate microRNAs were also investigated using northern blotting. 12 of the 14 were detectable by northern blot, with 1 being of correct size and 11 appearing to have products too large to be classified as microRNAs. Three of those showed multiple bands, all of which were too big for microRNAs. Since completing the deep sequencing project and validation process, a number of the candidates of interest were subsequently identified and recorded on miRBase (Griffiths-Jones et al., 2006; Griffiths-Jones et al., 2008; Kozomara and Griffiths-Jones, 2011). Candidate novel 8 has been identified as hsa-miR-1277-5p (Morin et al., 2008) although it has not been experimentally validated. The microRNA has also been implicated in laryngeal cancer (Xu et al., 2013). Candidate novel 10 is recorded as hsa-miR-487a-5p, whilst the -5p strand has only been identified via sequencing, the complementary -3p strand of miR-487a has been experimentally validated with northern blot (Fu et al., 2005). This contrasts with data from this project which observed multiple bands, all larger than a microRNA, for miR-487a-5p on a northern blot in RNA obtained from chondrocytes. The difference seen could be due to the cell type utilised, as Fu et al examined expression in human fetal liver RNA. MiR-487a has also been identified as a down-regulated microRNA in laryngeal carcinoma (Wang et al., 2010) and is thought to play a role in combatting the drug resistance of certain breast cancers (Ma et al., 2013).

There is also an entry of the same microRNA found in sheep (Caiment et al., 2010) however this sequence has a substituted base in its seed site. This could theoretically mean that the two versions of this microRNA may repress different targets, as they will be targeting slightly different mRNA sites. Candidate novel 12 has been described but not experimentally validated in both human and sheep as hsa-miR-539-3p (Landgraf et al., 2007) and oar-miR-539-3p (Caiment et al., 2010) respectively, although oar-miR-539-3p possesses an additional adenine base at its 5' end which would alter its seed site, resulting in a different set of gene targets to the human form. MiRBase also describes miR-539-3p in both rat and mouse; however, the sequences of these are slightly different to both the human and sheep forms at their 5' end and result in widely different seed sites.

Aside from using size as a validation attribute of microRNAs. There is also functional validation to investigate. The synthesis process of microRNAs has been described in detail (Bohnsack et al., 2004; Borchert et al., 2006; Gregory et al., 2005; Gregory et al., 2004; Lee et al., 2003) including a number of aspects of the process known to be vital in the production of mature microRNA sequences. One of these enzymes is Dicer, which is involved in cleaving the hairpin loop from the pre-microRNA structure, resulting in the formation of the mature miRNA duplex which subsequently splits into the guide and passenger strand (Gregory et al., 2005). If Dicer is missing from this process then mature microRNAs cannot be detected.

In order to investigate the response of the candidate microRNAs to the absence and presence of Dicer, firstly the enzyme was knocked down in SW1353 cells using a Dicer specific siRNA (data not shown). Expression of a number of known microRNAs as well as the *Dicer* RNA was measured by qRT-PCR at a number of time points including 48 hours, 72 hours and 1 week (data not shown). Although the expression of Dicer was seen to decrease in cells treated with the inhibitor compared to a control, no noticeable difference was seen in microRNA expression. Some microRNAs are known to have long half-lives with the average being around 5 days. The panel of microRNAs analysed by Gantier et al showed a range between 4 and 9 days, but a half-life of up to 12 days for miR-208 has been seen by Van Rooij et al *in vivo* (Gantier et al., 2011; van Rooij et al., 2007).

MicroRNAs in general are thought to be up to ten times more stable than mRNA (Gantier et al., 2011; Ruegger and Grosshans, 2012; Zhang et al., 2012c). This could



explain the lack of altered expression seen in the experiment, as although no new mature sequences were being produced, the existing mature microRNAs were still present within the system and had yet to degrade during the experimental timeframe.

To this end, the timeframe of the experiment was extended to over 1 week; however, Dicer knock down over this period was toxic to the cells. This suggests that although a longer time period was needed in order to see a measurable effect on microRNA expression, this was having a detrimental effect upon the health of the cells which was itself skewing the qRT-PCR results. This cell survival issue is supported by existing literature which looked at Dicer knock out mouse models (Bernstein et al., 2003), and shows that the knockout mouse is embryonic lethal.

A solution to this Dicer knock down problem was to utilise an existing Dicer null cell line, DLD-1. This is a commercially produced tumorigenic cell line which is able to survive and grow despite having Dicer knocked out. The expression of candidate microRNAs in addition to known microRNAs and *Dicer* RNA were measured by qRT-PCR and expression compared with that from the isogenic wild type DLD-1 line.

Figure 4.2 shows that each of the 3 candidate microRNAs (novel 2, novel 7 and novel 11) have a decreased expression in Dicer knockout cells compared to control suggesting that they are all dependent upon Dicer, indicating in turn that they are indeed microRNAs.

As a control, the expression of *Dicer 1* and a number of known microRNAs were also measured by qRT-PCR (figure 4.3). *Dicer 1* was found to have a significantly lower level of expression in the Dicer knock down cell-line compared to wild type, although there does appear to be some minimal expression which could explain the detection of mature microRNAs (albeit at lower levels than in wild type cells). The expression of miR-140-3p, miR-140-5p and miR-455-5p showed a statistically significant down-regulation in Dicer deficient cells. Although the expression of miR-455-3p also decreased in the Dicer knocked down cells, this was only a 1.5-fold difference compared to the 5-fold down regulation of miR-455-5p. These data show that these microRNAs are regulated by the Dicer enzyme; however different microRNAs appear to be affected at different rates, suggesting that perhaps there is an alternative microRNA synthesis pathway, or a form of redundancy within the system, allowing microRNAs to continue to be processed even without Dicer. This theory is supported in the literature, suggesting that not all microRNAs are Dicer dependent and that Ago

proteins may substitute to play a role in the manufacturing method. One of these microRNAs is miR-451, which although is processed by Drosha, appears to reach maturation without the activity of Dicer (Betancur and Tomari, 2012; Cheloufi et al., 2010). This then begs the question of classification: are these molecules still microRNAs but which can be synthesised through a different pathway, or are they a new subset of small RNAs with similar functions to microRNAs?

Kobayashi et al carried out a conditional knock out of Dicer in mouse cartilage and subsequently analysed, by microarray, the expression of a number of microRNAs in the RNA compared to a control. They discovered a reduction of 60-70% in expression of microRNAs from neonatal Dicer null growth plates in addition to decreased chondrocyte proliferation, accelerated chondrocyte hypertrophy, and a fold change decrease of 4.76 in *MMP13* expression. They postulate that the continued presence of these microRNAs (albeit at lower levels) is due to the inherent stability of microRNAs, and could therefore be a carryover of pre-existing miRNAs (Kobayashi et al., 2008). However, if Dicer was absent from birth, that suggests there were never any microRNAs to be 'carried over' in the first place, reinforcing the theory that there is at least a second microRNA maturation pathway in effect, although possibly not functioning with the same efficiency as Dicer (Cheloufi et al., 2010). Another possibility is the relatively recent discovery of circulating microRNAs in human plasma protected by AGO2 complexes (Arroyo et al., 2011). These could potentially be the source of the detected microRNAs in the Dicer null chondrocytes; however cartilage is a predominantly avascular tissue, rendering that theory less likely unless there was contamination from cells other than chondrocytes during the RNA extraction.

In addition to the initial validation methods, expression of the candidates was measured using qRT-PCR in a range of tissues. In order to determine if the candidates were altered in OA, expression was measured in both osteoarthritic hip cartilage and non-diseased neck of femur (NOF) fracture cartilage. NOF cartilage is used as a control for osteoarthritic cartilage. It is unfortunately not a perfect control as the cartilage is obtained from trauma patients, usually the elderly, and often the fractures result from other diseases such as osteoporosis. However, there may not be a perfect control tissue for cartilage (since even post-mortem cartilage is unloaded), but it is well validated in the literature (Kevorkian et al., 2004; Price et al., 2002).

Figure 4.4 showed that both candidate microRNA novel 2 and novel 11 show a trend to increase in OA, however the increased expression was not statistically significant compared to NOF cartilage. Candidate novel 7 showed a similar expression in OA compared to normal cartilage suggesting that expression of that candidate is not affected by the disease.

Expression was also measured in osteoarthritic knee cartilage to compare with expression in normal cartilage tissue from the hip (Fig. 4.7). In contrast to figure 4.4, expression of all 3 candidate microRNAs was highest in normal hip cartilage compared to osteoarthritic knee cartilage with candidate microRNA novel 2 showing a statistically significant increase of  $p \leq 0.01$ . This suggests that the baseline level of expression is actually higher to begin with in hip cartilage tissue than knee cartilage, and that a comparison using the same tissue type is more pertinent to detect differences in osteoarthritis. When comparing Ct values obtained from qRT-PCR analysis of the candidate microRNAs between hip and knee tissue, there are distinct differences. For each microRNA measured, Ct values were higher by one Ct or more for each candidate in knee tissue compared to both osteoarthritic and normal hip (the lower the Ct number, the higher the expression level).

In addition to assessing candidate microRNA expression in OA, the expression of a number of known microRNAs and mRNAs were measured. Figure 4.5 and 4.8 shows the expression of miR-140-3p, miR-140-5p, miR-455-3p and miR-455-5p in osteoarthritic hip (HOA) versus normal hip (NOF) cartilage and osteoarthritic knee (KOA) versus normal hip (NOF) cartilage respectively. When comparing expression within the same tissue type (Fig. 4.5) (cartilage from hip) expression of all 4 microRNAs is increased in osteoarthritic tissue with a statistical significance of  $p \leq 0.01$  for miR-455-3p and miR-455-5p, and a statistical value approaching significance of  $p = 0.0587$  for miR-140-5p. This supports the fact that these microRNAs are known to have altered in expression in OA with the expression of miR-140-5p and miR-455-3p often seen to increase in OA cartilage (Miyaki et al., 2009; Swingler et al., 2012). However, there is also literature observing a decreased expression of miR-140 in osteoarthritic chondrocytes (Miyaki et al., 2009; Tardif et al., 2009). This differing expression could be due to different stages of OA cartilage being analysed.

However, in figure 4.8, expression of all 4 microRNAs is higher in normal hip cartilage tissue than osteoarthritic knee cartilage with  $p \leq 0.05$  for miR-140-3p. This suggests that,

as with the candidate microRNAs, the baseline expression of a number of microRNAs is overall higher in hip cartilage than knee cartilage. This means it is always more accurate to compare between 2 tissue types of the same origin.

The gene expression examined in HOA versus NOF cartilage tissue (Fig. 4.6) included *ADAMTS5*, *MMP3*, *MMP13*, *ACAN* and *COL2A1*. *MMP3* showed significantly higher expression level in NOF cartilage ( $p \leq 0.001$ ) and *MMP13* appeared to increase in HOA tissue however it wasn't statistically significant. The expression seen from both of these MMPs is supported by TILDA data which saw the same expression pattern (Swingler et al., 2009). This is also reinforced by additional literature using qRT-PCR to quantify the expression of the MMPs in osteoarthritic cartilage compared with NOF (Kevorkian et al., 2004).

*ADAMTS5* showed a decreased expression in osteoarthritic cartilage ( $p \leq 0.05$ ), this does contradict some existing theories that *ADAMTS5* is one of the major aggrecanases involved in osteoarthritis (Stanton et al., 2005), although most of these studies have focussed on *in vivo* mouse experiments. Other research utilising RNA obtained from human cartilage of the hip (HOA and NOF) however agree with the data of this project in that *ADAMTS5* expression is significantly lower in OA cartilage (Kevorkian et al., 2004). This suggests that the expression of *ADAMTS5* could differ depending upon the stage of osteoarthritis at which it is measured. The data obtained by Kevorkian et al as well as from this project was obtained from end-stage OA patients.

*ACAN* expression is higher when measured in NOF cartilage compared to HOA ( $p \leq 0.05$ ), this makes sense as it has been shown that in late stage OA cartilage, *ACAN* production by chondrocytes is reduced (Goldring, 2000) in addition to increased degradation of the proteoglycan (Sandell and Aigner, 2001). This is supported by additional research by Brew et al who found a 4-fold decrease in *ACAN* expression in osteoarthritic cartilage compared to normal controls, measured by qRT-PCR (Brew et al., 2010).

Finally *COL2A1* expression was measured in osteoarthritic and normal hip cartilage, and showed a significant increase in OA tissue ( $p \leq 0.001$ ). Although matrix proteins are generally degraded during osteoarthritis, it is thought that some (like type II collagen) show high expression rates in late stage OA as a repair response, with a higher matrix remodelling rate (Aigner et al., 2001) this could explain the high levels of *COL2A1* seen during this project in osteoarthritic cartilage compared to NOF. It is also thought that

proteoglycans also increase their production rate during OA as part of a potential repair response, however, this production cannot compensate for the continued destruction (Aigner et al., 1997).

In addition to being detected within cartilage tissue by qRT-PCR, a human tissue panel (Fig. 4.9) was utilised to determine if each candidate microRNA was selectively expressed in cartilage or if it was present within a number of tissues at similar levels. Candidate microRNA novel 11 does appear to indeed be cartilage selective, with expression much higher in all articular cartilage samples than for any other tissue type. Candidate microRNA novel 2 on the other hand showed highest expression in RNA from the cervix and skeletal muscle tissue. Candidate microRNA novel 7 also showed high expression in cervix and skeletal muscle RNA, however, expression in cartilage tissue was also relatively high.

Figure 4.10 shows the expression of miR-140-3p, miR-140-5p, miR-455-3p and miR-455-5p across the same 23 human tissues. Both miR-140-3p and miR-140-5p showed highest expression in all the cartilage tissues, supporting the fact that miR-140 is a cartilage selective microRNA (Miyaki et al., 2009; Miyaki et al., 2010; Swingler et al., 2012). MiR-455-3p appears to be most highly expressed in RNA from cervix tissue, with the expression in cartilage being comparable to the rest of the panel. MiR-455-5p on the other hand shows the expression of the microRNA in osteoarthritic cartilage tissue (KOA and HOA) to be similar to cervix tissue, with expression in NOF much lower. Again, miR-455 is known to be an important microRNA in cartilage and osteoarthritis (Swingler et al., 2012; Zhang et al., 2012b).

A cartilage injury model in the form of a hip avulsion assay (Fig. 4.11) was carried out to investigate if the candidate microRNAs would respond to that type of stress as an alternative experiment to simply measuring expression between osteoarthritic and control cartilage.

This mouse hip avulsion model was developed by Dr. T. Vincent at the Kennedy Institute of Rheumatology. The femoral caps of dislocated hip joints of ~5 week old mice are removed and cultured *in situ*. RNA is subsequently extracted from the cartilage and gene expression analyzed by qRT-PCR. Vincent et al discovered an activation of the MAP kinase pathways in addition to the NF- $\kappa$ B pathway post injury. Furthermore, it was discovered that a number of the genes regulated by the surgical joint destabilization mouse model were also regulated by the mouse hip injury model, showing a link

between the pathways activated in both injury and surgically induced OA (Chong et al., 2013).

For this project the cartilage layer of a 3-5 week old C57BL/6 mouse femoral head was removed intact and incubated in cell culture medium for 2 days (Chong et al., 2013; Stanton et al., 2011). RNA was extracted directly from the cartilage at set time points (0, 1, 3, 6, 12, 24 and 48 hours) and analyzed by qRT-PCR.

Candidate microRNA novel 7 showed an increase in expression at each time point, peaking at 48 hours with statistical significance. This suggests that the candidate is responsive to injury even though there was no differential expression between osteoarthritic and normal cartilage tissue. This indicates that the injury model may activate different pathways to end stage osteoarthritis, or that the model could be more similar to early osteoarthritis and that novel 7 plays a more significant role in early OA rather than during the latter stages.

Candidate novel 11 showed a slight cyclical expression pattern unlike novel 7, with expression peaking at 3 and 48 hours (however without statistical significance). This again contrasts the differential expression seen between osteoarthritic and non-osteoarthritic tissue.

Three known microRNAs were also measured in the injury model (Fig. 4.12); miR-140-3p, miR-140-5p and miR-455-5p (miR-455-3p is not conserved between mouse and human and therefore was not measured). All three showed an increase in expression over time during the model, with the highest expression seen at 48 hours. This response to injury supports the earlier data described in chapter 3 (Fig. 3.1) showing miR-140-5p increasing in expression from cartilage to the post digest stage before decreasing again through passaged chondrocytes. This suggests that the earlier response was indeed due to injury rather than just the conditions of digestion. However, the other microRNAs measured also increased in expression across the injury model, yet did not increase at the post digest stage of isolated chondrocytes. This could suggest that although miR-140-5p responds to both injury and stress, the remaining microRNAs respond most strongly to injury, with stressed conditions having a lesser effect. Although this model is used as an injury response model, the effects seen in microRNA expression could also be due in part to unloading of the joint in combination with inflammation.

Measuring expression of microRNAs during chondrogenesis assays is another means to establish if these microRNAs are likely functional in chondrocytes. Figure 4.13 describes the data obtained from a mesenchymal stem cell chondrogenesis assay.

This assay was carried out by Dr M. Barter (Newcastle University, UK), who assessed the expression of a number of known microRNAs across the assay by microarray (data not shown). The data showed that miR-140-3p, miR-140-5p, miR-455-3p and miR-455-5p increased during chondrogenesis, with the highest expression of each seen in the cartilage disc at day 14. This supports previous research which show these microRNAs to be involved in chondrocyte development (Barter, 2012; Swingler et al., 2012).

During the MSC assay candidate microRNA novel 2 showed a stable expression throughout the assay except at day 7 where there was a spike in expression, suggesting that novel 2 is not regulated in this model. Candidate microRNA novel 7 shows an increase from day 0 through day 3 and 7, with expression at day 14 subsequently remaining constant. This suggests that candidate microRNA novel 7 is involved in chondrogenesis. Candidate microRNA novel 11 shows increased expression at day 0 of the assay, but this drops at day 3 and remains low for the remainder of the assay. This suggests that candidate microRNA novel 11 may not be involved in chondrogenesis although this contrasts with the expression of the candidate measured in the human tissue panel which suggested that novel 11 was cartilage selective. Interestingly, expression of novel 11 increases in SOX9 treated cells (data not shown), which is a marker of chondrogenesis (Ikeda et al., 2004).

An additional chondrogenesis assay was examined by qRT-PCR using cells obtained from a mouse cell line: ATDC5 [original assay carried out by Dr. T. Swingler (UEA)]. This is shown in figure 4.14 with candidate microRNA novel 7 showing an increase across time in both cells induced to undergo chondrogenesis and those acting as a control. Candidate microRNA novel 11 showed a comparable decrease in expression in both induced and control cells however at day 42, induced cells appear to decrease significantly (although not statistically) compared to control cells suggesting that any role novel 11 plays in chondrogenesis is minor. These data supports that seen within the MSC chondrogenesis assay suggesting novel 7 is implicated in chondrogenesis. However none of the changes seen here were statistically significant. In addition to measuring expression of candidate microRNAs, miR-140-3p and miR-140-5p were also measured by qRT-PCR (Fig. 4.15) and both microRNAs showed a significant increase

in expression in treated cells at day 42 compared to the previous time points (with  $p \leq 0.001$  between day 42 and all other time points) and compared to control this expression pattern is supported by the published report on this assay (Swingler et al., 2012).

To compare with and support the data from the chondrogenesis assays, candidate microRNA expression was also measured in an assay of human articular chondrocyte de-differentiation (Fig. 4.16). Many versions of this assay have been investigated with some extending past passage 3, and others only examining chondrocytes in culture. Hong et al observed miR-140 decreasing in expression during de-differentiating chondrocytes, but subsequently increase again during chondrocyte redifferentiation (Hong and Reddi, 2013). Lin et al on the other hand only examined chondrocytes after isolation from cartilage, however, their assay continued onto passage 5. Lin et al observed 13 microRNAs to be up-regulated across dedifferentiation and 12 down-regulated by microarray analysis. They also found miR-140-3p and miR-140-5p to be down-regulated in dedifferentiated chondrocytes compared to differentiated chondrocytes which supports the results seen across this project (Lin et al., 2011). A paper from Martinez-Sanchez et al also focusses on a model of chondrocyte dedifferentiation, looking specifically at miR-145 and how it regulates SOX9 (a key transcription factor in the synthesis of cartilage matrix) in dedifferentiated chondrocytes (Martinez-Sanchez et al., 2012). The same lab have also published data showing how SOX9 regulates another chondrocyte specific microRNA miR-675 by analysing expression across both dedifferentiation and redifferentiation (Dudek et al., 2010).

Although candidate microRNA novel 2 appeared to have similar expression at each stage from cartilage to passage 3, both novel 7 and novel 11 showed a decrease in expression from cartilage to isolated and expanded chondrocytes. Expression of candidate microRNA novel 7 decreased significantly from cartilage to PD but then remained stable, whereas novel 11 showed a decrease in expression at each subsequent stage between. As the chondrocytes progress through the assay they begin to lose their phenotype, becoming more fibroblastic. This suggests that both novel 7 and 11 are selectively expressed in chondrocytes due to their expression levels decreasing through the assay.

As candidate microRNA novel 11 has already been shown to be cartilage selective (Fig. 4.9 C) in the human tissue panel, and chondrocyte selective in the de-differentiation



model (Fig. 4.16 C); expression was measured in the development of mouse embryos using whole-mount *in situ hybridisation* (Fig. 4.17). Whole embryos of stage E11 and limbs from stage E14 were analysed [a developmental stage where a known cartilage selective microRNA is readily detected (Tuddenham et al., 2006)] . In whole mouse embryos, staining appears along the spine (A), suggesting it may be present within the cartilage of the spine. However unlike miR-455 which is a known cartilage expressed microRNA (Swingler et al., 2012), candidate microRNA novel 11 was not detected along inter digital regions of the E11 embryo (B) or E14 limb (C). A control (Fig. 4.18) was carried out with miR-449 [a known developmental brain specific microRNA (Saugstad, 2010)] and no specific staining was noted in any section of the embryo. However, it should be noted that this experiment was not an exhaustive assay across embryonic stages, and this would be informative since e.g. miR-140-5p is only expressed after E11.5 in mouse embryo limbs (Tuddenham et al., 2006).

In addition to functional validation, a bioinformatics assessment was also carried out upon the 3 candidate microRNAs novel 2, novel 7 and novel 11. Ensembl ([www.ensembl.org](http://www.ensembl.org)) (Flicek et al., 2014) and BLAST were utilised to identify the host gene of each candidate. Candidate microRNA novel 2 (Fig. 4.19) is found within an intron of WD Repeat Domain 91 (WDR91) (on chromosome 7) which is, as yet, an unvalidated protein encoding gene. We found WDR91 to have a similar expression pattern in HOA cartilage compared to NOF (data not shown), differing from the expression of candidate novel 2 which is increased in HOA cartilage, suggesting that in NOF cartilage, candidate novel 2 is not transcribed from its host gene at the same rate as in osteoarthritic cartilage.

Candidate microRNA novel 7 (Fig. 4.20) is found within the single exon of small nucleolar RNA, H/ACA box 36A (SNORA36A) which is in turn transcribed from chromosome 2. SnoRNAs are thought to guide RNAs as part of their function, and SNORA36A specifically is thought to be involved in the pseudouridylation of U105 and U1244 residues of 18S (Kiss et al., 2004; Lestrade and Weber, 2006). Moreover, further research has been carried out which determines that snoRNAs could in fact act as precursors of microRNAs, with the precursors of other known microRNAs such as miR-151, miR-140 and miR-605 possessing similar features to snoRNAs (Scott et al., 2009), or snoRNAs which function in a similar manner to microRNAs, for example the snoRNA ACA45 has been found to require cleavage by Dicer and is thought to have a role in gene silencing (Ender et al., 2008). We found comparable expression of

SNORA36A in cartilage from HOA and NOF (data not shown), suggesting it is not a marker of disease and follows the same expression pattern of candidate novel 7.

Candidate microRNA novel 11 (Fig. 4.21) is found in the last intron of each respective transcript of cartilage acidic protein 1 (CRTAC1) on chromosome 10. CRTAC1 [also known as CEP-68 (chondrocyte expressed protein 68)] interestingly encodes a glycosylated extracellular matrix protein found in the deep zone of cartilage within the interterritorial matrix (Steck et al., 2007). The protein is on occasion used as a marker to distinguish chondrocytes from other cell types in culture and the gene has been found to be an effective marker of chondrogenesis (Steck et al., 2001). CRTAC1 is also thought to be involved in certain cell-cell or cell-matrix interactions containing both integrin binding and calcium binding domains (Steck et al., 2006). We found CRTAC1 to be up-regulated in HOA cartilage compared to NOF (data not shown) in a similar manner to candidate novel 11, suggesting it could in addition be a marker of cartilage damage or disease. This is supported by Grgurevic et al who found CRTAC1 to be expressed in the plasma of patients with damaged cartilage. They posit that it could therefore be employed as a marker for cartilage damage and repair (Grgurevic et al., 2007).

In combination with identifying the host gene of each candidate, the presence of each sequence was also examined within the multiple species' genomes available within Ensembl to determine how conserved these candidates might be. The more species that a microRNA can be found in suggests that it is evolutionarily important and probably plays a vital role in common important biological pathways between those species.

Candidate novel 11 (Fig. 4.24) appeared to be the most highly conserved, with the sequence found in partial form in a number of species, but in its exact form within 24 different species in mammals which diverged approximately 94 million years ago. Candidate novel 7 (Fig. 4.23) was the second most highly conserved with the sequence found in partial form in a number of species, but in its exact form within 7 different species in mammals which diverged 65 million years ago. Candidate novel 2 (Fig. 4.22) sequence was found in partial form in a number of species, but the exact sequence was only found within 4 different species of mammals which diverged approximately 43 million years ago. Although miR-140 was not analysed in this manner during this project, data available on miRBase (Griffiths-Jones et al., 2006; Griffiths-Jones et al., 2008; Kozomara and Griffiths-Jones, 2011) shows that the microRNA can be found in

its mature form in at least 28 different species which is comparable to candidate novel 11 examined here.

These data suggest that candidate microRNA novel 11 may be the most evolutionarily conserved out of the 3 candidates examined. The conservation of microRNAs, and genetic elements in general across species is an indication of how important they are. Specifically with microRNAs, conservation of their sequence ensures conservation of binding site and therefore conservation of any potential gene targets. Conservation implies these sequences have been maintained despite speciation, which suggests a vital role in aspects of the organisms' function, survival or reproduction (Bejerano et al., 2004). As many microRNAs have been found to be conserved between species, as well as suggesting a potential importance to function, this has provided an additional strategy of identification. The sequences of microRNAs found in one species can be hunted across the sequencing data of many more (Zhang et al., 2006a).

Now that the candidate microRNAs have undergone significant functional analysis, we can assume from these data that novel 2, novel 7 and novel 11 are in fact microRNAs. In fact, since completion of this project each of the three candidate novel microRNAs has been published on miRBase (Griffiths-Jones et al., 2006; Griffiths-Jones et al., 2008; Kozomara and Griffiths-Jones, 2011). Candidate novel 2 is described in humans as hsa-miR-6509-5p, but has not been experimentally validated except for detection using Solexa sequencing (Li et al., 2012b). Candidate novel 7 is listed in humans as hsa-miR-664b-3p (Friedlander et al., 2012), but again there is no experimental validation, with identification coming from a novel computational analysis of existing deep sequencing data. Candidate novel 11 has been published on miRBase in both mouse and rat as mmu-miR-3085-3p (Chiang et al., 2010) and rno-miR-3085 (Linsen et al., 2010) respectively with no experimental validation beyond identification from deep sequencing projects. However, miR-3085 has not yet been described in humans.

The next step in this project was to investigate the function of the three chosen microRNAs: miR-6509-5p, miR-664b-3p and miR-3085-3p. This would hopefully enable the understanding of their potential importance within cartilage and osteoarthritis. Target search analysis was carried out using bioinformatics and microarray analysis, followed by functional validation of a number of the subsequently identified targets.

## **Chapter 5 – Novel MicroRNA Target Search Analysis**

## 5.1 Introduction

Chapter 3 and 4 described the discovery and subsequent validation of a number of microRNAs from osteoarthritic cartilage tissue obtained from total joint replacement surgery. However, it's not enough just to know these microRNAs exist, we need to know which genes they target in chondrocytes and why they have been evolutionarily conserved.

Chapter 5 depicts this next step which is to analyse these candidates functionally to discover genetic targets that may play a role in osteoarthritis or cartilage development.

The function of microRNAs is to bind to complementary targets within mRNA and suppress gene expression. In plants, it's thought that microRNAs are able to directly bind and subsequently cleave a complementary target site of mRNA through the RISC complex (Tang et al., 2003). In mammals, the microRNA is more likely to have less complementarity to their target and so inhibit translation of the mRNA rather than cleave it directly, this also results in the preservation of the target mRNA structure whilst still suppressing expression (Ambros, 2004).

MicroRNAs bind to mRNA via specific sites at the 5' end of the microRNA. These are known as seed sites which bind to corresponding complementary 'target sites' usually at the 3' UTR of an mRNA sequence (Fig.1.7), these are the most conserved areas of a microRNA (Akbari Moqadam et al., 2013). Target sites have also been found within the 5' UTR as well as coding regions of the mRNA (Lee et al., 2009). It is thought that microRNAs which bind at 3'UTRs cause a stronger or more efficient suppression of gene expression than if they were to bind to the 5'UTR, which could explain why targets in the 3'UTR are more prevalent (Moretti et al., 2010). A small number of microRNAs (including hsa-miR-1 and hsa-miR124) have been found which target both the 5'UTR and 3'UTR of the same gene, both resulting in down-regulation of gene expression (Lee et al., 2009).

There are a number of seed sites of different sizes within the sequence of a microRNA which are thought to be responsible for binding to an mRNA target and they start at a number of known locations within the microRNA. The first is the 6mer, this is a site which is 6 nucleotides long which starts at the 2<sup>nd</sup> nucleotide (from the 5' end) and finishes at the 7<sup>th</sup> (inclusive). The next is the 7mer8. This site is 7 nucleotides long beginning at the 2<sup>nd</sup> nucleotide, and as the name suggests, ends at the 8<sup>th</sup> nucleotide.

There is an additional site of the same length entitled 7merA1 however this site begins at the first nucleotide, finishing at the 7<sup>th</sup>. The final seed site is the 8mer, this site is 8 nucleotides long beginning at the first nucleotide and ending at the 8<sup>th</sup> (Mitra and Bandyopadhyay, 2011).

It's thought that although all 4 of these sites can result in binding sufficient to inhibit mRNA expression, the binding of the 8mer site results in a stronger down-regulation of the gene (Schnall-Levin et al., 2011). It is also theorised that in addition to the seed sites being the main locations of binding, the microRNA sequence can also bind to a mRNA sequence with its 3' end, specifically an area known as the Watson-Crick complement which runs from the 12<sup>th</sup> to the 17<sup>th</sup> nucleotide of the microRNA (Mitra and Bandyopadhyay, 2011).

The methods involved with miRNA function are multi-fold, depending on the organism in question. A number of microRNAs are produced in different cell types in different combinations and are thought to work co-operatively to regulate cell specific gene pathways. The target prediction algorithms show that a single microRNA can have hundreds of individual targets, and that genes can be targeted by multiple different microRNAs or even have multiple target sites for the same microRNA. It is thought that these microRNAs work co-operatively to down-regulate gene expression, and thus can result in a stronger suppression (Akbari Moqadam et al., 2013; Hon and Zhang, 2007).

There are a number of algorithms available online which attempt to predict the targets of the majority of microRNAs described on miRBase (Griffiths-Jones et al., 2006; Griffiths-Jones et al., 2008; Kozomara and Griffiths-Jones, 2011). These include TargetScan (Friedman et al., 2009), miRanda (Betel et al., 2008), PITA (Kertesz et al., 2007) and PicTar (Krek et al., 2005). However each database uses a different algorithm, and if results are compared they often do not overlap as can be seen in another online system which combines the results of a number of the databases: The microRNA body map (Mestdagh et al., 2011). One of the problems with using these online algorithms is that they all show high levels of false positive or false negative results (Mitra and Bandyopadhyay, 2011).

Identifying the target of any microRNA can be extremely difficult, especially in mammals, as gene regulation can occur by a number of microRNA : mRNA

interactions. Gene regulation is not dependent upon exact seed site complementarity; mismatches within the binding can still result in gene suppression as well as matching between other areas of the microRNA and the mRNA target. This is one of the problems with using prediction algorithms which mainly look for targets with strict seed site matches. This means that all functional analysis must include an experimental analysis. Even exact seed site matches may not result in an observed functional down-regulation. (Akbari Moqadam et al., 2013).

Three promising novel microRNAs (identified as part of the cartilage miRNome and validated in chapter 4) were chosen to investigate their potential functional role within osteoarthritis and cartilage development. These were candidate novels 2, 7 and 11, now described as miR-6509-5p miR-664b-3p and miR-3085-3p respectively. This was carried out using primary chondrocytes in monolayer at passage one, obtained from total joint replacement of osteoarthritic knees from three patients. miR-6509-5p, miR-664b-3p and miR-3085-3p were investigated using a gain- and loss-of-function method analysed with an Illumina Whole-Genome Expression BeadArrays Microarray encompassing over 47,000 genes, gene candidates, and splice variants. Qiagen microRNA mimics and inhibitors were designed for each microRNA and the genome array was analysed with R Studios (R Core Team, 2013) and Microsoft Excel.

Whilst the genome array is significantly more quantitative than, for example, a deep sequencing project, there was a limit to the number of genes and splice variants examined, whereas a deep sequencing analysis would examine the entire DNA sequence available.

The data obtained from the array were analysed and each gene was ranked based on its fold-change in expression from negative non-targeting siRNA controls when treated with a microRNA mimic or inhibitor. A positive gene target would theoretically be repressed with the microRNA mimic and increase in expression with the microRNA inhibitor.

This experimental analysis was combined with a computational microRNA target search in the genes of the human genome using the R programming language (R Core Team, 2013) and the Biostrings package (Pages et al., 2012).

To summarise, all genes measured by the array were assessed for their role as a target of one of the three newly discovered microRNAs by whether their expression levels

decreased with the microRNA mimic and increased with the microRNA inhibitor (when compared to a negative control) and whether they had at least one fully complementary target site. As previously explained, microRNAs can bind to a slightly mismatched target site on a gene and still cause an effect (although often not as extreme) and this analysis did not take into account any possible partial or 'mismatched' target sites. A more complex algorithm would be needed and there is still not enough data available to establish how many mismatches a microRNA can withstand and still successfully bind to a specific mRNA target.



## 5.2 Results

### 5.2.1 Experimental microRNA target search analysis

Primary chondrocytes from passage 1 isolated from osteoarthritic knee cartilage tissue were treated with either a microRNA mimic at 30nM (miR-6509-5p, miR-664b-3p and miR-3085-3p), inhibitor at 50nM (miR-6509-5p, miR-664b-3p and miR-3085-3p), or non-targeting controls for mimics and inhibitors at 30nM and 50nM respectively and incubated for 6 hours at 37°C 5% (v/v) CO<sub>2</sub>. RNA isolated from these samples was then sent to Source Bioscience to be analysed using a Whole-Genome Expression BeadArrays Microarray to measure the expression of over 47,000 genes, gene candidates and splice variants.

Figure 5.1 shows the normalised untransformed data obtained from the array. Each point represents a gene, gene candidate, or splice variant and how it responded to both the microRNA mimics and inhibitors in comparison to non-targeting controls of the microRNAs (A) miR-6509-5p (B) miR-664b-3p and (C) miR-3085-3p. Figure 5.2 shows the fold change in expression for each gene between a non-targeting control and the microRNA mimic and between the microRNA inhibitor and a non-targeting control of the microRNAs (A) miR-6509-5p (B) miR-664b-3p and (C) miR-3085-3p. The grey shading shows the area where genes are located if they are down-regulated by the microRNA mimic and up-regulated by the microRNA inhibitor. This indicates they are responding as expected for true microRNA targets.

The gene which appeared to increase the most (by 31.6 fold) when treated with the microRNA miR-6509-5p mimic was IFIT2 (Interferon-Induced Protein with Tetratricopeptide Repeats 2) (Table 5.1). This gene is part of the type 1 interferon response group of genes known to be involved in the immune response and interferon signalling. It has also been found to bind to the eIF-3C initiation factor and therefore inhibit protein synthesis (Berchtold et al., 2008; Terenzi et al., 2006). The gene which decreased in expression the most (by 2.4 fold) when treated with the miR6509-5p mimic was HNRPA1P4 (Heterogeneous Nuclear Ribonucleoprotein A1 Pseudogene) (Table 5.2) and as yet has no described function. The gene which increased the most (by 8.1 fold) when treated with the miR-6509-5p inhibitor was MX1 (Myxovirus Resistance 1, Homolog of Murine Interferon Inducible Protein P78) (Table 5.3). It encodes the protein MxA and is thought to be involved in the interferon-induced antiviral response of cells (Gao et al., 2011; Tazi-Ahnini et al., 2000) and is thought to target the

nucleoprotein of MxA-sensitive RNA viruses (Haller and Kochs, 2011). The gene which appeared to show the largest decrease in expression (by 1.9 fold) when treated with the miR-6509-5p inhibitor was MT1H (Metallothionein 1H) (Table 5.4), this gene codes a protein known to bind heavy metals (Stennard et al., 1994).

The gene which showed the highest expression in response to both the miR-664b-3p mimic and inhibitor was IFIT2 by 20.8 fold and 12.4 fold respectively (Tables 5.5. and 5.7). The gene which appeared to decrease in expression the most (by 2.7 fold) in response to the miR-664b-3p inhibitor was CPA4 (Carboxypeptidase A4) (Table 5.8), this is considered a metalloprotease and has potential links to histone hyperacetylation. It is also known to be linked to the aggressiveness of prostate cancer (Pallares et al., 2005; Tanco et al., 2010). STT3A (Subunit of the Oligosaccharyltransferase Complex, Homolog A) appeared to decrease in expression the most (by 2.3 fold) in response to the miR-664b-3p mimic (Table 5.6). It produces a protein which is the catalytic subunit of OST (N-oligosaccharyl transferase) and is involved in substrate specificity (Kelleher et al., 2003).

IFIT2 was also the most highly expressed gene (by 32.1 fold) when treated with the miR-3085-3p inhibitor (Table 5.11), whilst RPL14 (Ribosomal Protein L14) was the gene which appeared to decrease the most (by 2.5 fold) in response to the inhibitor (Table 5.12). RPL14 codes for a ribosomal protein (ribosomal protein L14) and is known to be involved in the RNA synthesis pathway (Uechi et al., 2001). The gene which showed the largest decrease in expression (by 2.2 fold) in response to the miR-3085-3p mimic was CSTB (Cystatin B) (Table 5.10) which functions as a thiol proteinase inhibitor (Laloti et al., 1997).

The array also measured the expression of a small number of microRNAs within the treated samples, and microRNA-1974 appeared to be the gene most highly expressed (by 2 fold) when the samples were treated with the miR-3085-3p mimic (Table 5.9). However, this microRNA has since been removed from miRBase as a false microRNA. It has been found to map to the mitochondrial genome and appears to overlap a mitochondrial tRNA sequence (Griffiths-Jones et al., 2006; Griffiths-Jones et al., 2008; Kozomara and Griffiths-Jones, 2011). To this end, the second most highly expressed gene in response to the miR-3085-3p mimic was CA2 (Carbonic Anhydrase 2) by 1.7 fold (Table 5.9). This codes for a zinc containing enzyme and is involved in catalysing the reversible hydration of carbon dioxide (Becker et al., 2011).

Figure 5.3 shows the ideal gene target area defined in figure 5.2 and the fold change in expression of each gene. All genes represented in these graphs increase in expression in response to the inhibitors and decrease in response to the mimics of the microRNAs (A) miR-6509-5p, (B) miR-664b-3p and (C) miR-3085-3p.

The gene which increased the most (by 1.9 fold) in response to the miR-6509-5p inhibitor whilst still showing a fold change decrease (by 1.1 fold) in response to the miR-6509-5p mimic was OAZ2 (Ornithine Decarboxylase Antizyme 2). OAZ2's functions include binding and destabilizing ornithine decarboxylase (Hayashi et al., 1996; Ramos-Molina et al., 2013). LOC646688 [which corresponds to RPL4P2 (Ribosomal Protein L4 Pseudogene)] was the gene which showed the largest decrease (by 2.1 fold) in response to the miR-6509-5p mimic whilst still showing a fold change increase (by 1.1 fold) in response to the miR-6509-5p inhibitor. RPL4P2 has as yet no known function.

The gene which showed the largest fold change increase (by 9.2 fold) in response to the miR-664b-3p inhibitor whilst showing a decrease (by 1.5 fold) in expression in response to the miR-664b-3p mimic was LOC100132564 (a hypothetical protein) and is uncharacterised. The 2<sup>nd</sup> most highly expressed gene (by 8 fold) in response to the miR-664b-3p inhibitor but still showing a marginal decrease (by 1.01 fold) in expression in response to the mimic was LOC100008589 which is a ribosomal RNA and encodes part of a 28S rRNA. The gene to decrease in expression the most (by 2.3 fold) in response to the miR-664b-3p mimic and still increase in expression (by 1.1 fold) in response to the miR-664b-3p inhibitor was STT3A (Subunit of the Oligosaccharyltransferase Complex, Homolog A). STT3A produces a protein involved in complex substrate specificity (Kelleher et al., 2003). The gene which appeared to increase the most (by 14 fold) when treated with the miR-3085-3p inhibitor but still decrease (by 1.1 fold) when treated by the miR-3085-3p mimic was RNU1G2 (also known as RNA, U1 Small Nuclear 4) which as yet has no described function. Finally, the gene which had the largest fold change decrease (by 2.1 fold) in response to the miR-3085-3p mimic but still increased in expression (by 1.9 fold) with the miR-3085-3p inhibitor was ITGA5 (Integrin alpha 5) which is a cell-surface receptor (Nagae et al., 2012; Takada et al., 2007).

The data obtained from the array suggested that, on the whole, the microRNA mimic siRNAs were more effective than the microRNA inhibitor siRNAs, especially when

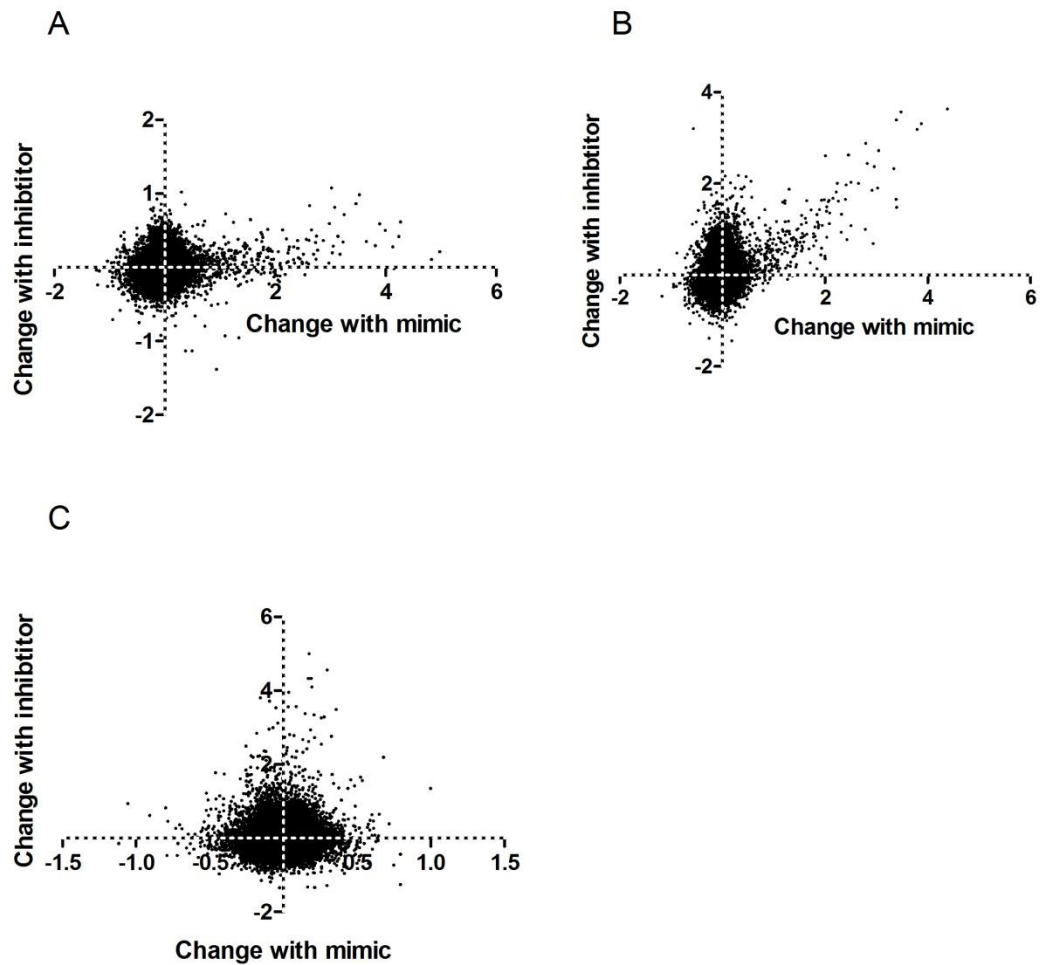
looking at gene targets which responded to the siRNAs in a manner expected of microRNA targets (increased expression with inhibitor and decreased expression with mimic) . This lead to the observation of genes showing, for example, a very large down-regulation when treated with the microRNA mimic but negligible increased expression in response to the microRNA inhibitor. To solve this problem, a series of cut off points were utilised in order to isolate genes which showed reasonable changes in expression to both the microRNA mimic and inhibitor siRNAs (Fig. 5.3).

The fold-change cut-off points were situated at, firstly, the point exactly halfway between the zero change point (fold change of exactly 1) and the most highly down-regulated gene (in response to the microRNA mimic) and the most up-regulated gene (in response to the microRNA inhibitor) and these lines were labelled in purple. No genes were found to be located beyond both of these lines, so new cut off points were created at both the 25% (red) and 12.5% (blue) from the zero change point.

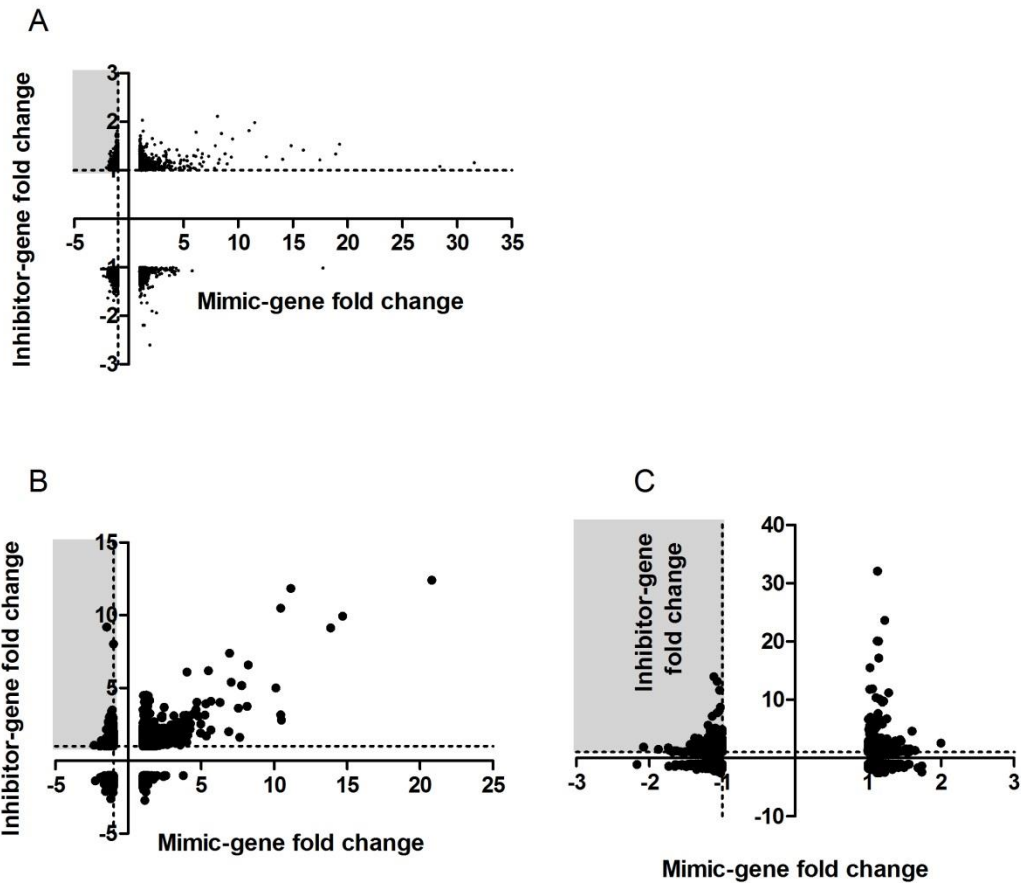
When these cut off points are taken into account 20 genes were found to be regulated in the correct manner by the miR-6509-5p mimic and inhibitor above the 25% cut off point and 452 genes were found to be regulated in the correct manner by the miR-6509-5p mimic and inhibitor above the 12.5% cut off point [Fig. 5.3 (A)].

One gene was found to be regulated in the correct manner by the miR-664b-3p mimic and inhibitor above the 25% cut off point and 11 genes were found to be regulated in the correct manner by the miR-664b-3p mimic and inhibitor above the 12.5% cut off point [Fig. 5.3 (B)].

No genes were found to be regulated in the correct manner by the miR-3085-3p mimic and inhibitor above the 25% cut off point but 18 genes were found to be regulated in the correct manner by the miR-3085-3p mimic and inhibitor above the 12.5% cut off point [Fig. 5.3 (C)].



**Figure 5.1 Gene expression change in response to novel microRNA mimics and inhibitors.** Untransformed data obtained from a Whole-Genome Expression BeadArrays Microarray measuring over 47,000 genes, gene candidates, and splice variants. Figure shows the gene expression results of RNA isolated from cells treated with (A) miR-6509-5p mimic and inhibitor, (B) miR-664b-3p mimic and inhibitor and (C) miR-3085-3p mimic and inhibitor. The x-axis represents a change in expression in response to the microRNA mimic and the y-axis represents a change in expression in response to the microRNA inhibitor with down-regulation corresponding to a negative value, and up-regulation corresponding to a positive value.



**Figure 5.2 Gene expression fold change in response to novel microRNA mimics and inhibitors.** Fold change in gene expression of genome array. Data obtained from a Whole-Genome Expression BeadArrays Microarray measuring over 47,000 genes, gene candidates, and splice variants. The figure shows the fold change in gene expression of RNA isolated from cells treated with (A) miR-6509-5p mimic and inhibitor, (B) miR-664b-3p mimic and inhibitor and (C) miR-3085-3p mimic and inhibitor. The x-axis represents a change in expression in response to the microRNA mimic and the y-axis represents a change in expression in response to the microRNA inhibitor. Down-regulation corresponding to a negative value, and up-regulation corresponding to a positive value. The shaded areas show the area corresponding to genes which decrease in expression in response to the mimic and increase in expression in response to the inhibitor.

Gene	Gene name	Mimic fold change	Inhibitor fold change
IFIT2	Interferon-induced protein with tetratricopeptide repeats 2	31.571	1.157
CCL5	Chemokine ligand 5	28.429	1.077
RSAD2	Radical S-adenosyl methionine domain containing 2	19.258	1.536
IFIT3	Interferon-induced protein with tetratricopeptide repeats 3	18.894	1.338
OASL	2'-5'-oligoadenylate synthetase-like	17.452	1.214
BST2	Bone marrow stromal cell antigen 2	14.054	1.227
IL18BP	Interleukin 18 binding protein	12.567	1.277
IFI44L	Interferon-induced protein 44-like	11.498	1.981
IFIT1	Interferon-induced protein with tetratricopeptide repeats 1	10.994	1.819
OAS2	2'-5'-oligoadenylate synthetase 2	9.487	1.642

**Table 5.1 Top 10 genes from microarray with highest fold change increase in response to miR-6509-5p mimic compared to a non-targeting control.** Gene expression fold change in response to miR-6509-5p mimic and inhibitor. Fold change in gene expression of genome array. Data obtained from a Whole-Genome Expression BeadArrays Microarray measuring over 47,000 genes, gene candidates, and splice variants. The table shows the fold change in gene expression of RNA isolated from cells treated with miR-6509-5p mimic and inhibitor.

Gene	Gene name	Mimic fold change	Inhibitor fold change
HNRPA1P4	Heterogeneous nuclear ribonucleoprotein A1 pseudogene 4	-2.355	-1.025
HNRPA1L-2	Heterogeneous nuclear ribonucleoprotein A1 pseudogene 10	-2.321	-1.179
ACO1	Aconitase 1, soluble	-2.272	-1.034
ICMT	Isoprenylcysteine carboxyl methyltransferase	-1.936	-1.141
TXNDC12	Thioredoxin domain containing 12	-1.927	-1.051
DANCR	Differentiation Antagonizing Non-Protein Coding RNA	-1.889	1.048
SWI5	SWI5 Recombination Repair Homolog	-1.836	1.100
ITGAE	Integrin, alpha E	-1.824	-1.022
FAP	Fibroblast activation protein, alpha	-1.815	1.182
RPS23	Ribosomal protein S23	-1.804	1.045

**Table 5.2 Top 10 genes from microarray with highest fold change decrease in response to miR-6509-5p mimic compared to a non-targeting control.** Gene expression fold change in response to miR-6509-5p mimic and inhibitor. Fold change in gene expression of genome array. Data obtained from a Whole-Genome Expression BeadArrays Microarray measuring over 47,000 genes, gene candidates, and splice variants. The table shows the fold change in gene expression of RNA isolated from cells treated with miR-6509-5p mimic and inhibitor.

Gene	Gene name	Mimic fold change	Inhibitor fold change
MX1	Myxovirus resistance 1, interferon-inducible protein p78	8.103	2.110
GPR1	G protein-coupled receptor 1	1.230	2.031
IFI44L	Interferon -induced protein 44-like	11.498	1.981
OAZ2	Ornithine decarboxylase antizyme 2	-1.102	1.897
IFIT1	Interferon -induced protein with tetratricopeptide repeats 1	10.994	1.819
PLIN5	Perilipin 5	-1.047	1.812
KIAA1751	KIAA1751	1.293	1.807
HERC6	HECT and RLD domain containing E3 ubiquitin protein ligase family member 6	6.139	1.788
ISG15	ISG15 ubiquitin-like modifier	8.454	1.759
NBPF8	Neuroblastoma breakpoint family, member 8	-1.194	1.727

**Table 5.3 Top 10 genes from microarray with highest fold change increase in response to miR-6509-5p inhibitor compared to a non-targeting control.** Gene expression fold change in response to miR-6509-5p mimic and inhibitor. Fold change in gene expression of genome array. Data obtained from a Whole-Genome Expression BeadArrays Microarray measuring over 47,000 genes, gene candidates, and splice variants. The table shows the fold change in gene expression of RNA isolated from cells treated with miR-6509-5p mimic and inhibitor.

Gene	Gene name	Mimic fold change	Inhibitor fold change
MT1H	Metallothionein 1H	1.912	-2.603
MT1X	Metallothionein 1X	1.408	-2.193
MT1F	Metallothionein 1F	1.296	-2.193
MT1G	Metallothionein 1G	2.533	-1.938
MT1M	Metallothionein 1M	2.130	-1.903
MT1E	Metallothionein 1E	1.653	-1.734
MEST	Mesoderm specific transcript	-1.380	-1.681
LOC148430	Ribosomal S2 Pseudogene	-1.922	-1.634
BLOC1S1	Biogenesis of lysosomal organelles complex-1, subunit 1	1.024	-1.606
PIGP	Phosphatidylinositol glycan anchor biosynthesis, class P	-1.142	-1.581

**Table 5.4 Top 10 genes from microarray with highest fold change decrease in response to miR-6509-5p inhibitor compared to a non-targeting control.** Gene expression fold change in response to miR-6509-5p mimic and inhibitor. Fold change in gene expression of genome array. Data obtained from a Whole-Genome Expression BeadArrays Microarray measuring over 47,000 genes, gene candidates, and splice variants. The table shows the fold change in gene expression of RNA isolated from cells treated with miR-6509-5p mimic and inhibitor.



Gene	Gene name	Mimic fold change	Inhibitor fold change
IFIT2	Interferon-induced protein with tetratricopeptide repeats 2	20.811	12.407
RSAD2	Radical S-adenylsyl methionine domain containing 2	14.697	9.949
IFIT3	Interferon-induced protein with tetratricopeptide repeats 3	13.879	9.123
BST2	Bone marrow stromal cell antigen 2	10.497	2.798
IFI44L	Interferon-induced protein 44-like	10.464	10.495
OASL	2'-5'-oligoadenylate synthetase-like	10.120	5.014
IFIT1	Interferon-induced protein with tetratricopeptide repeats 1	8.230	6.599
OAS1	2'-5'-oligoadenylate synthetase 1	8.141	3.741
ISG15	ISG15 ubiquitin-like modifier	7.769	5.176
CCL5	Chemokine ligand 5	7.646	1.609

**Table 5.5 Top 10 genes from microarray with highest fold change increase in response to miR-664b-3p mimic compared to a non-targeting control.** Gene expression fold change in response to miR-664b-3p mimic and inhibitor. Fold change in gene expression of genome array. Data obtained from a Whole-Genome Expression BeadArrays Microarray measuring over 47,000 genes, gene candidates, and splice variants. The table shows the fold change in gene expression of RNA isolated from cells treated with miR-664b-3p mimic and inhibitor.

Gene	Gene name	Mimic fold change	Inhibitor fold change
STT3A	Subunit of the oligosaccharyltransferase complex	-2.339	1.079
CPD	Carboxypeptidase D	-2.262	-1.380
CALM1	Calmodulin 1	-1.937	1.015
MAMDC2	MAM domain containing 2	-1.933	-1.149
CD151	CD151 molecule	-1.928	-1.198
PSAT1	Phosphoserine Aminotransferase	-1.766	-1.131
PLAU	Plasminogen activator, urokinase	-1.669	-1.151
GSTM1	Glutathione S-transferase mu 1	-1.663	-1.067
LEPRE1	Leucine proline-enriched proteoglycan 1	-1.659	-1.404
TXNDC12	Thioredoxin domain containing 12	-1.653	1.301

**Table 5.6 Top 10 genes from microarray with highest fold change decrease in response to miR-664b-3p mimic compared to a non-targeting control.** Gene expression fold change in response to miR-664b-3p mimic and inhibitor. Fold change in gene expression of genome array. Data obtained from a Whole-Genome Expression BeadArrays Microarray measuring over 47,000 genes, gene candidates, and splice variants. The table shows the fold change in gene expression of RNA isolated from cells treated with miR-664b-3p mimic and inhibitor.

Gene	Gene name	Mimic fold change	Inhibitor fold change
IFIT2	Interferon-induced protein with tetratricopeptide repeats 2	20.811	12.407
IFIT3	Interferon-induced protein with tetratricopeptide repeats 3	11.133	11.841
IFI44L	Interferon-induced protein 44-like	10.464	10.495
RSAD2	Radical S-adenylsyl methionine domain containing 2	14.697	9.949
MX1	Myxovirus resistance 1	6.936	7.385
IFIT1	Interferon-induced protein with tetratricopeptide repeats 1	8.230	6.599
HERC6	HECT and RLD domain containing E3 ubiquitin protein ligase family member 6	5.488	6.189
OAS2	2'-5'-oligoadenylate synthetase 2	7.054	5.413
ISG15	ISG15 ubiquitin-like modifier	7.769	5.176
OASL	2'-5'-oligoadenylate synthetase-like	10.120	5.014

**Table 5.7 Top 10 genes from microarray with highest fold change increase in response to miR-664b-3p inhibitor compared to a non-targeting control.** Gene expression fold change in response to miR-664b-3p mimic and inhibitor. Fold change in gene expression of genome array. Data obtained from a Whole-Genome Expression BeadArrays Microarray measuring over 47,000 genes, gene candidates, and splice variants. The table shows the fold change in gene expression of RNA isolated from cells treated with miR-664b-3p mimic and inhibitor.

Gene	Gene name	Mimic fold change	Inhibitor fold change
CPA4	Carboxypeptidase A4	1.138	-2.705
GLS	Glutaminase	-1.196	-2.609
RPS6KA2	Ribosomal protein S6 kinase, polypeptide 2	-1.040	-2.170
CA12	Carbonic anhydrase XII	1.201	-2.120
TFRC	Transferrin receptor	-1.145	-2.093
ACO1	Aconitase 1	-1.551	-2.089
ANGPTL5	Angiopoietin-like-5	-1.191	-2.076
EFEMP1	EGF containing fibulin-like extracellular matrix protein 1	-1.078	-2.064
IL7R	Interleukin 7 receptor	1.060	-1.994
BLOC1S1	Biogenesis of lysosomal organelles complex-1, subunit 1	-1.298	-1.959

**Table 5.8 Top 10 genes from microarray with highest fold change decrease in response to miR-664b-3p inhibitor compared to a non-targeting control.** Gene expression fold change in response to miR-664b-3p mimic and inhibitor. Fold change in gene expression of genome array. Data obtained from a Whole-Genome Expression BeadArrays Microarray measuring over 47,000 genes, gene candidates, and splice variants. The table shows the fold change in gene expression of RNA isolated from cells treated with miR-664b-3p mimic and inhibitor.

Gene	Gene name	Mimic fold change	Inhibitor fold change
MIR1974	MicroRNA-1974 (dead entry)	1.999	2.550
CA2	Carbonic anhydrase II	1.735	-1.351
CPA4	Carboxypeptidase A4	1.734	-2.390
MT1G	Metallothionein 1G	1.680	-1.656
SLC16A10	Solute carrier family 16, member 10	1.645	1.311
SNORA12	Small nucleolar RNA, H/ACA box 12	1.565	1.045
DESI2	Desumoylating Isopeptidase 2	1.560	-1.049
DAPK1	Death-associated protein kinase 1	1.546	-1.062
DICER1	Dicer 1, ribonuclease type III	1.537	1.058
ESM1	Endothelial cell-specific molecule 1	1.528	1.040

**Table 5.9 Top 10 genes from microarray with highest fold change increase in response to miR-3085-3p mimic compared to a non-targeting control.** Gene expression fold change in response to miR-3085-3p mimic and inhibitor. Fold change in gene expression of genome array. Data obtained from a Whole-Genome Expression BeadArrays Microarray measuring over 47,000 genes, gene candidates, and splice variants. The table shows the fold change in gene expression of RNA isolated from cells treated with miR-3085-3p mimic and inhibitor.

Gene	Gene name	Mimic fold change	Inhibitor fold change
CSTB	Cystatin B	-2.169	-1.103
ITGA5	Integrin, alpha 5	-2.079	1.915
CYTL1	Cytokine-like 1	-1.876	1.531
CSPG4	Chondroitin sulphate proteoglycan 4	-1.739	1.771
PIN1	Peptidylprolyl cis/trans isomerase, NIMA-interacting 1	-1.730	-1.384
MYADM	Myeloid-associated differentiation marker	-1.722	1.372
COPE	Coatomer protein complex, subunit epsilon	-1.683	1.006
CCND3	Cyclin D3	-1.660	1.208
HSPE1	Heat shock 10kDa protein 1	-1.649	-1.383
VPS26	Vacuolar protein sorting 26 homolog A	-1.643	-1.107

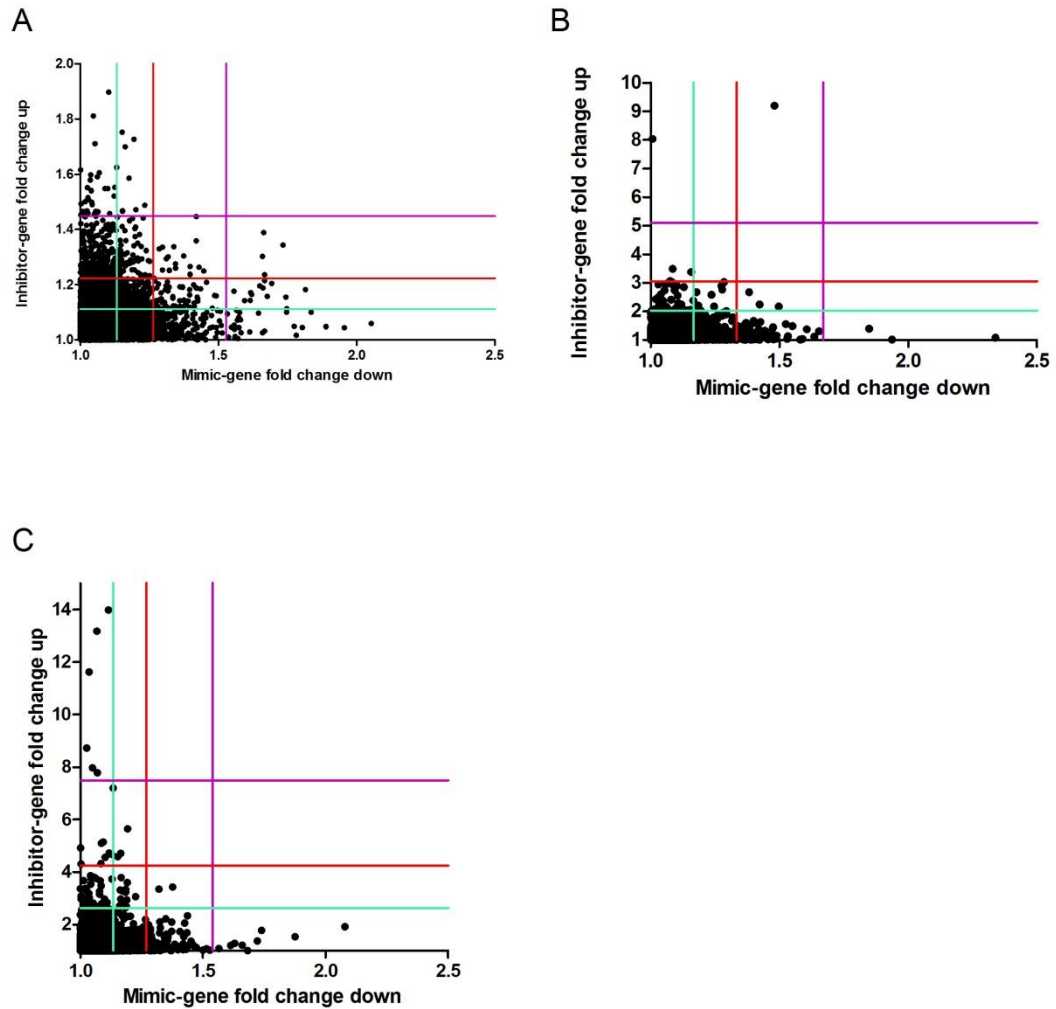
**Table 5.10 Top 10 genes from microarray with highest fold change decrease in response to miR-3085-3p mimic compared to a non-targeting control.** Gene expression fold change in response to miR-3085-3p mimic and inhibitor. Fold change in gene expression of genome array. Data obtained from a Whole-Genome Expression BeadArrays Microarray measuring over 47,000 genes, gene candidates, and splice variants. The table shows the fold change in gene expression of RNA isolated from cells treated with miR-3085-3p mimic and inhibitor.

Gene	Gene name	Mimic fold change	Inhibitor fold change
IFIT2	Interferon-induced protein with tetratricopeptide repeats 2	1.129	32.110
RSAD2	Radical S-adenylsyl methionine domain containing 2	1.230	23.662
IFI44L	Interferon-induced protein 44-like	1.125	20.085
IFIT3	Interferon-induced protein with tetratricopeptide repeats 3	1.140	20.071
RNVU1-18	RNA, Variant U1 small nuclear 18	1.025	15.509
RNU1-4	RNA, U1 small nuclear 4	-1.114	13.989
SNORD3D	Small nucleolar RNA, C/D box 3D	-1.068	13.174
RNU1-3	RNA, U1 small nuclear 3	1.063	11.897
RNU1-1	RNA, U1 small nuclear 1	-1.036	11.627
IFIT1	Interferon-induced protein with tetratricopeptide repeats 1	1.282	11.218

**Table 5.11 Top 10 genes from microarray with highest fold change increase in response to miR-3085-3p inhibitor compared to a non-targeting control.** Gene expression fold change in response to miR-3085-3p mimic and inhibitor. Fold change in gene expression of genome array. Data obtained from a Whole-Genome Expression BeadArrays Microarray measuring over 47,000 genes, gene candidates, and splice variants. The table shows the fold change in gene expression of RNA isolated from cells treated with miR-3085-3p mimic and inhibitor.

Gene	Gene name	Mimic fold change	Inhibitor fold change
RPL14	Ribosomal protein L14	-1.018	-2.543
GLS	Glutaminase	1.158	-2.521
ACO1	Aconitase 1	1.121	-2.504
MT1F	Metallothionein1F	1.227	-2.431
CPA4	Carboxypeptidase A4	1.734	-2.390
RPL23A	Ribosomal protein L23a	-1.062	-2.267
RPS6KA2	Ribosomal protein S6 kinase, polypeptide 2	-1.068	-2.081
SUMF1	Sulfatase modifying factor 1	-1.013	-2.008
FABP4	Fatty acid binding protein 4, adipocyte	-1.055	-1.972
BOLA3	BolA family member 3	1.120	-1.964

**Table 5.12 Top 10 genes from microarray with highest fold change decrease in response to miR-3085-3p inhibitor compared to a non-targeting control.** Gene expression fold change in response to miR-3085-3p mimic and inhibitor. Fold change in gene expression of genome array. Data obtained from a Whole-Genome Expression BeadArrays Microarray measuring over 47,000 genes, gene candidates, and splice variants. The table shows the fold change in gene expression of RNA isolated from cells treated with miR-3085-3p mimic and inhibitor.



**Figure 5.3 Candidate targets which responded as expected to microRNA mimic and inhibitor treatments.** Genes which decrease in expression in response to the microRNA mimic and increase in expression in response to the inhibitor as expected for targets of the novel microRNAs (A) miR-6509-5p – 11871 genes, (B) miR-664b-3p – 10800 genes and (C) miR-3085-3p – 10953 genes. The x-axis represents the fold change down in gene expression in response to the mimic and the y-axis represents the fold change up in gene expression in response to the inhibitor. The purple line represents a fold change cut off point at 50% of the largest fold changes in response to both mimic and inhibitors. The red and blue lines represent the 25% and 12.5% cut off points respectively.

### 5.2.2 Bioinformatics microRNA target search analysis

In combination with the functional array analysis to identify gene targets of the 3 microRNAs miR-6509-5p, miR-664b-3p and miR-3085-3p, an *in silico* analysis was also carried out. In order to suppress the function of a gene, a microRNA must bind to a specific site within the mRNA, either exactly or partially. To this end, the R programming language was used with RStudio (R Core Team, 2013) and the Biostrings package (Pages et al.) to identify all possible target sites (6mer, 7merA1, 7merm8 and 8mer) of each microRNA on the 3'UTR of each gene of the human genome as downloaded from the online database Ensembl (Flicek et al., 2014) using Biomart (Kasprzyk, 2011).

Four potential seed sites from each microRNA were used to produce the target site sequences in the search. For miR-6509-5p, the 6mer seed site is TTAGGT, the 7merA1 is ATTAGGT, the 7merm8 is TTAGGTA and the 8mer is ATTAGGTA. For miR-664b-3p, the 6mer seed site is TCATTT, the 7merA1 is TTCATTT, the 7merm8 is TCATTTG and the 8mer is TTCATTTG. For miR-3085-3p, the 6mer seed site is CTGGCT, the 7merA1 is TCTGGCT, the 7merm8 is CTGGCTG and the 8mer is TCTGGCTG. The complementary sequences to these sites were used as search criteria using the R programming language with RStudio (R Core Team, 2013) and the Biostrings package (Pages et al., 2012).

Tables 5.13, 5.14 and 5.15 show the top 20 genes identified as targets for each microRNA by this system, ranked firstly by the number of 8mer target sites, then by the total number of 7mer sites (including 7merA1 and 7merm8), and finally by the number of 6mer sites.

Table 5.13 shows the top 20 out of 4655 genes which possessed at least one 6mer target site within their 3'UTR for miR-6509-5p. The top 5 are as follows: LNPEP (Leiomodin 2) with 6 6mer sites, 8 7mer sites and 3 8mer sites. Next is KRAS (Karyopherin alpha 8) with 5 6mers, 6 7mers and 2 8mers. This is followed by CCNT2 (Cyclin B3) with 4 6mers, 6 7mers and 2 8mers. HELZ (Helicase with zinc finger) is in fourth place with 3 6mers, 5 7mers and 2 8mers. Finally, there is LPP (Lon peptidase 2, peroxisomal) with 3 6mers, 5 7mers and 2 8mers.

Table 5.14 shows the top 20 out of 9690 genes which possessed at least one 6mer target site within their 3'UTR for miR-664b-3p. The top 5 are as follows: ERAP2 (Endoplasmic reticulum aminopeptidase 2) with 6 6mer sites, 8 7mer sites and 4 8mer

sites. This is followed by AFF2 (AF4/FMR2 family member 2 with 15 6mers, 14 7mers and 3 8mers. Next is by APC (Adenomatous polyposis coli) with 10 6mers, 10 7mers and 3 8mers. MLL3 (Membrane bound O-acyltransferase domain containing 2) is in fourth place with 12 6mers, 9 7mers and 3 8mers. Finally, there is ATAD1 (ATPase family, AAA domain containing 1) with 7 6mers, 9 7mers and 3 8mers.

Table 5.15 shows the top 20 out of 8930 genes which possessed at least one 6mer target site within their 3'UTR for miR-3085-3p. The top 5 are as follows: TG (Thyroglobulin) with 8 6mer sites, 10 7mer sites and 4 8mer sites. Next is TANC2 (Tetratricopeptide repeat, ankyrin) with 12 6mers, 11 7mers and 3 8mers. This is followed by LMX1B (LIM homeobox transcription factor 1, beta) with 11 6mers, 11 7mers and 3 8mers. GFER (Growth factor, augments liver regeneration) is in fourth place with 9 6mers, 9 7mers and 3 8mers. Finally, there is PTPN18 (Protein tyrosine phosphatase, non-receptor type 18) with 6 6mers, 8 7mers and 3 8mers.

Target sites			Target information		
6mer	All 7mers	8mer	Ensemble gene ID	Gene	Gene name
6	8	3	ENSG00000113441	LNPEP	leiomodulin 2 (cardiac)
5	6	2	ENSG00000133703	KRAS	karyopherin alpha 3 (importin alpha 4)
4	6	2	ENSG00000082258	CCNT2	cyclin B3
3	5	2	ENSG00000198265	HELZ	helicase with zinc finger
3	5	2	ENSG00000145012	LPP	lipon peptidase 2, peroxisomal
3	5	2	ENSG00000146373	RNF217	ring finger protein 217
4	4	2	ENSG00000196632	WNK3	WNK lysine deficient protein kinase 3
3	4	2	ENSG00000081014	AP4E1	adaptor-related protein complex 4, epsilon 1 subunit
3	4	2	ENSG00000163545	NUAK2	NUAK family, SNF1-like kinase, 2
3	4	2	ENSG00000120948	TARDBP	TAR DNA binding protein
2	4	2	ENSG00000135932	CAB39	chromosome 9 open reading frame 72
2	4	2	ENSG00000241644	INMT	indolethylamine N-methyltransferase
2	4	2	ENSG00000131626	PPFIA1	protein tyrosine phosphatase, receptor type, f polypeptide, interacting protein, alpha 1
2	4	2	ENSG00000170234	PWWP2A	PWWP domain containing 2A
2	4	2	ENSG00000154114	TBCEL	tubulin folding cofactor E-like
2	4	2	ENSG00000118965	WDR35	WD repeat domain 35
6	5	1	ENSG00000163428	LRRC58	LIM domain containing preferred translocation partner in lipoma
5	5	1	ENSG00000149212	SESN3	sestrin 3
5	5	1	ENSG00000117500	TMED5	transmembrane emp24 protein transport domain containing 5
5	5	1	ENSG00000169905	TOR1AIP2	torsin A interacting protein 2

**Table 5.13 Computational gene target search of miR-6509-5p using RStudio.** All 4 target sites; 6mer, 7mer, 8mer, 7merA1 and 8mer were identified within the 3'UTR of all genes within the human genome as described on Ensembl, downloaded using Biomart. Target search was carried out with the R programming Biostrings package, and the top 20 genes of 4655 with target sites (ranked by size and number of target sites) is shown.



Target sites			Target information		
6mer	All 7mers	8mer	Ensemble gene ID	Gene	Gene name
6	8	4	ENSG00000164308	ERAP2	endoplasmic reticulum aminopeptidase 2
15	14	3	ENSG00000155966	AFF2	AF4/FMR2 family, member 2
10	10	3	ENSG00000134982	APC	adenomatous polyposis coli
12	9	3	ENSG00000055609	MLL3	membrane bound O-acyltransferase domain containing 2
7	9	3	ENSG00000138138	ATAD1	ATPase family, AAA domain containing 1
7	9	3	ENSG00000144824	PHLDB2	pleckstrin homology-like domain, family B, member 2
3	6	3	ENSG00000196166	C8orf86	chromosome 8 open reading frame 86
3	6	3	ENSG00000040199	PHLPP2	PH domain and leucine rich repeat protein phosphatase 2
15	12	2	ENSG00000168807	SNTB2	syntrophin, beta 2 (dystrophin-associated protein A1, 59kDa, basic component 2)
13	12	2	ENSG00000104290	FZD3	frizzled family receptor 3
10	10	2	ENSG00000166450	PRTG	protogenin
10	9	2	ENSG00000172795	DCP2	decapping mRNA 2
8	9	2	ENSG00000133739	LRRCC1	lysine (K)-specific methyltransferase 2C
12	8	2	ENSG00000152495	CAMK4	calcium/calmodulin-dependent protein kinase IV
11	8	2	ENSG00000010244	ZNF207	zinc finger protein 207
10	8	2	ENSG00000169213	RAB3B	RAB3B, member RAS oncogene family
9	8	2	ENSG00000154429	CCSAP	centriole, cilia and spindle-associated protein
8	8	2	ENSG00000036530	CYP46A1	cytochrome P450, family 46, subfamily A, polypeptide 1
8	8	2	ENSG00000143797	MBOAT2	leucine rich repeat and coiled-coil centrosomal protein 1
8	8	2	ENSG00000147862	NFIB	nuclear factor I/B

**Table 5.14 Computational gene target search of miR-664b-3p using RStudio.** All 4 target sites; 6mer, 7mer, 8mer, 7merA1 and 8mer were identified within the 3'UTR of all genes within the human genome as described on Ensembl, downloaded using Biomart. Target search was carried out with the R programming Biostrings package, and the top 20 genes of 9690 with target sites (ranked by size and number of target sites) is shown.

Target sites			Target information		
6mer	All 7mers	8mer	Ensemble gene ID	Gene	Gene name
8	10	4	ENSG00000042832	TG	Thyroglobulin
12	11	3	ENSG00000170921	TANC2	Tetratricopeptide repeat, ankyrin repeat and coiled-coil containing 2
11	11	3	ENSG00000136944	LMX1B	LIM homeobox transcription factor 1, beta
9	9	3	ENSG00000127554	GFER	Growth factor, augments liver regeneration
6	8	3	ENSG00000072135	PTPN18	Protein tyrosine phosphatase, non-receptor type 18
6	7	3	ENSG00000156639	ZFAND3	Zinc finger, AN1-type domain 3
5	6	3	ENSG00000110888	CAPRN2	Caprin family member 2
3	6	3	ENSG00000073331	ALPK1	Alpha-kinase 1
10	8	2	ENSG00000114735	HEMK1	HemK methyltransferase family member 1
10	8	2	ENSG00000130052	STARD8	StAR-related lipid transfer (START) domain containing 8
8	8	2	ENSG00000164742	ADCY1	adenylate cyclase 1
6	8	2	ENSG00000100106	TRIOBP	TRIO and F-actin binding protein
11	7	2	ENSG00000196092	PAX5	Paired box 5
9	7	2	ENSG00000136279	DBNL	Drebrin-like
9	7	2	ENSG00000174437	ATP2A2	ATPase, Ca <sup>++</sup> transporting, cardiac muscle, slow twitch 2
8	7	2	ENSG00000072501	SMC1A	Structural maintenance of chromosomes 1A
7	7	2	ENSG00000142765	SYTL1	Synaptotagmin-like 1
7	7	2	ENSG00000169118	CSNK1G1	Casein kinase 1, gamma 1
7	7	2	ENSG00000090905	TNRC6A	Trinucleotide repeat containing 6A
6	7	2	ENSG00000123643	SLC36A1	Solute carrier family 36 (proton/amino acid symporter)

**Table 5.15 Computational gene target search of miR-3085-3p using RStudio.** All 4 target sites; 6mer, 7mer, 8mer, 7merA1 and 8mer were identified within the 3'UTR of all genes within the human genome as described on ensembl, downloaded using Biomart. Target search was carried out with the R programming Biostrings package, and the top 20 genes of 8930 genes with target sites (ranked by size and number of target sites) is shown.

### 5.2.3 MicroRNA target search analysis combining array data and bioinformatics

In order to utilise both methods of gene target identification of the three chosen microRNAs: miR-6509-5p, miR-664b-3p and miR-3085-3p, the results of the two methods were combined (as shown in figure 5.4). All genes identified by the array analysis to show expression changes in the expected fashion of increasing when treated with the microRNA inhibitor and decreasing when treated with the mimic were cross-referenced with the list of human genes determined to have at least one target site. The resulting genes are presented in figure 5.5 for microRNAs (A) miR-6509-5p, (B) miR-664b-3p and (C) miR-3085-3p.

A system of fold-change cut off points were also utilised at the 50%, 25% and 12% marks, in an attempt to triage the possible targets to a number suitable for further experimental analysis.

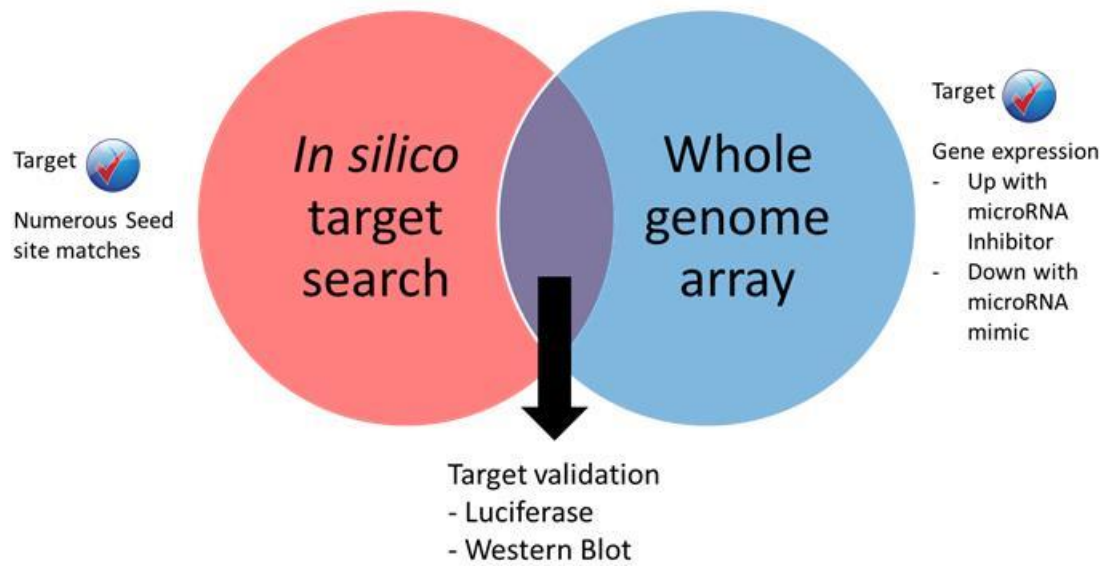
Table 5.16 shows the 23 genes (ranked by size and number of target sites) which fall above the 25% cut off point for miR-6509-5p. All these genes increase in expression when treated with the microRNA inhibitor and decrease in expression when treated with the mimic in comparison to non-targeting controls.

Table 5.17 shows the 53 genes (ranked by size and number of target sites) which fall above the 12.5% cut off point for miR-664b-3p. The cut off of 12.5% was used as there were only 4 genes which were found to be above the 25% point. All these genes increase in expression when treated with the microRNA inhibitor and decrease in expression when treated with the mimic in comparison to non-targeting controls.

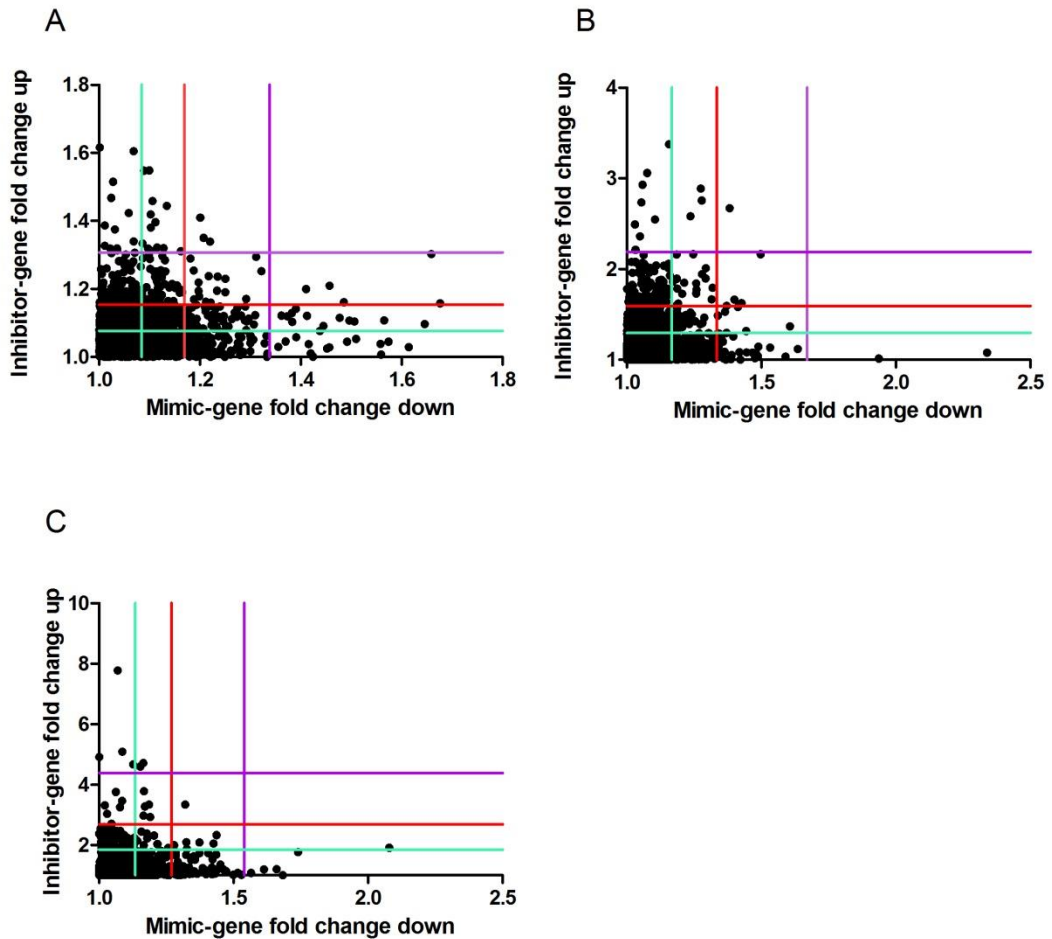
Table 5.18 shows the 21 genes (ranked by size and number of target sites) which fall above the 12.5% cut off point for miR-3085-3p. The cut off of 12.5% was used as there was only 1 gene which was found to be above the 25% point. All of the genes increase in expression when treated with the microRNA inhibitor and decrease in expression when treated with the mimic in comparison to non-targeting controls.

The gene for each microRNA which appeared to have the largest number of target sites within the subset of genes which increased with the inhibitor and decreased with the mimic were for miR-6509-5p; LRRC58 (Leucine rich repeat containing 58) with 6 6mers, 5 7mers and 1 8mer. SNTB2 (Syntrophin, beta 2) was the top gene for miR-664b-3p with 15 6mers, 12 7mers and 2 8mers. Finally the gene for miR-3085-3p was

ITPRIPL2 (Inositol 1, 4, 5-triphosphate receptor interacting protein-like 2) with 7  
6mers, 5 7mers and 1 8mer.



**Figure 5.4 Venn diagram of computational analysis combined with gain- and loss-of-function studies to identify microRNA gene targets.** The process undertaken to identify potential novel microRNA targets using data from both the computational analysis with RStudios and the whole genome microarray with microRNA mimics and inhibitors. Targets are considered promising if they contain at least one complementary target site, increase in expression when treated with the microRNA inhibitor, and decrease in expression when treated with the microRNA mimic.



**Figure 5.5 Candidate targets which responded as expected to mimic and inhibitor treatments which also have at least one target site.** Genes which decrease in expression in response to the microRNA mimic, and increase in expression in response to the inhibitor as expected for targets of the three microRNAs and which also contain at least 1 6mer target site, (A) miR-6509-5p – 1782 genes, (B) miR-664b-3p – 3245 genes and (C) miR-3085 - 3091 genes The x-axis represents the fold change down in gene expression in response to the mimic and the y-axis represents the fold change up in gene expression in response to the inhibitor. The purple line represents a fold change cut off point at 50% of the largest fold changes in response to both mimic and inhibitors. The red and blue lines represent the 25% and 12.5% cut off points respectively.

Gene	Gene name	Mimic fold	Inhibitor fold	N2 target sites		
		change down	change up	6mer	All 7mers	8mer
LRRC58	leucine rich repeat containing 58	1.292	1.171	6	5	1
RASSF6	Ras association (RalGDS/AF-6) domain family member 6, transcript variant 1	1.250	1.230	2	2	1
DCBLD2	discoidin, CUB and LCCL domain containing 2	1.207	1.350	2	2	1
PDGFD	platelet derived growth factor D, transcript variant 1	1.188	1.254	4	2	0
PLEKHA1	pleckstrin homology domain containing, family A (phosphoinositide binding specific) member 1, variant 1	1.410	1.200	2	2	0
KCNK2	potassium channel, subfamily K, member 2, transcript variant 3	1.322	1.252	2	2	0
PEG3	paternally expressed 3	1.199	1.178	2	1	0
GHR	growth hormone receptor	1.188	1.169	2	1	0
SFRP4	secreted frizzled-related protein 4	1.659	1.303	1	1	0
STEAP4	STEAP family member 4	1.486	1.161	1	1	0
CASB	carbonic anhydrase VB, mitochondrial, nuclear gene encoding mitochondrial protein	1.457	1.209	1	1	0
TROVE2	TROVE domain family, member 2, transcript variant 1	1.311	1.294	1	1	0
HERPUD2	HERPUD family member 2	1.234	1.195	1	1	0
AZIN1	antizyme inhibitor 1, transcript variant 1	1.220	1.339	1	1	0
RBM3	RNA binding motif (RNP1, RRM) protein 3	1.201	1.409	1	1	0
TGFBR3	transforming growth factor, beta receptor III	1.197	1.219	1	1	0
ENAH	enabled homolog (Drosophila), transcript variant 2	1.676	1.158	1	0	0
SLC2A3	solute carrier family 2 (facilitated glucose transporter), member 3	1.250	1.190	1	0	0
PDE5A	phosphodiesterase 5A, cGMP-specific, transcript variant 3	1.235	1.236	1	0	0
TTC3	tetratricopeptide repeat domain 3, transcript variant 1	1.222	1.160	1	0	0
HSPA12A	heat shock 70kDa protein 12A	1.217	1.238	1	0	0
RTN3	reticulon 3, transcript variant 1	1.212	1.163	1	0	0
SHROOM4	shroom family member 4	1.181	1.290	1	0	0

**Table 5.16 Table of the candidate gene targets of miR-6509-5p represented in figure 5.5 (A) above the 25% cut off point.** The ID and full name of all the gene targets of miR-6509-5p above the 25% cut off point which increase in expression when treated with the microRNA inhibitor and decrease in expression when treated with the microRNA mimic.

Gene	Gene name	Mimic fold	Inhibitor fold	N7 target sites		
		change down	change up	6mer	All 7mers	8mer
SNTB2	syntrophin, beta 2 (dystrophin-associated protein A1, 59kDa, basic component 2)	1.275	2.885	15	12	2
NFIB	nuclear factor I/B	1.241	1.502	8	8	2
DENND4C	DENN/MADD domain containing 4C	1.206	1.727	7	7	2
PVRL3	poliovirus receptor-related 3	1.497	2.163	5	6	2
HIPK2	homeodomain interacting protein kinase 2	1.236	2.581	6	5	2
MBTD1	mbt domain containing 1	1.232	1.393	5	4	2
BZW1	PREDICTED: basic leucine zipper and W2 domains 1	1.364	1.529	4	4	2
GPM6B	glycoprotein M6B, transcript variant 1, .	1.444	1.317	9	7	1
KLF5	Kruppel-like factor 5 (intestinal)	1.187	1.515	5	4	1
ABCA5	ATP-binding cassette, sub-family A, member 5, transcript variant 2	1.188	1.303	3	3	1
LRRFIP1	leucine rich repeat (in FLII) interacting protein 1	1.293	1.895	3	3	1
ARL5A	ADP-ribosylation factor-like 5A, transcript variant 3	1.207	1.402	2	3	1
RIOK3	RIO kinase 3 (yeast), transcript variant 1	1.190	1.487	2	3	1
VCAN	versican	1.316	1.666	6	4	0
SYAP1	synapse associated protein 1, SAP47 homolog (Drosophila)	1.219	1.493	5	4	0
FAM129A	family with sequence similarity 129, member A, transcript variant 2	1.370	1.595	4	3	0
NTAN1	N-terminal asparagine amidase	1.606	1.369	4	3	0
RAB27A	RAB27A, member RAS oncogene family, transcript variant 1	1.338	1.485	3	3	0
TIMP3	TIMP metalloproteinase inhibitor 3	1.382	2.671	3	3	0
PCYOX1	prenylcysteine oxidase 1	1.413	1.580	4	2	0
ZMAT3	zinc finger, matrin type 3, transcript variant 2	1.258	1.730	4	2	0
FBXL3	F-box and leucine-rich repeat protein 3	1.171	1.403	3	2	0
DAB2	disabled homolog 2, mitogen-responsive phosphoprotein (Drosophila)	1.192	1.619	2	2	0
RBPM52	RNA binding protein with multiple splicing 2	1.182	1.379	2	2	0
ALDH1A3	aldehyde dehydrogenase 1 family, member A3	1.191	1.442	3	1	0
CHRNA5	cholinergic receptor, nicotinic, alpha 5	1.236	1.522	2	1	0
DENR	density-regulated protein	1.259	1.771	2	1	0
ERRF1	ERBB receptor feedback inhibitor 1	1.288	1.463	2	1	0
FKBP14	FK506 binding protein 14, 22 kDa	1.279	1.943	2	1	0
LRRC1	leucine rich repeat containing 1	1.186	1.310	2	1	0
MEF2D	myocyte enhancer factor 2D	1.246	2.165	2	1	0
PRPF40A	PRP40 pre- processing factor 40 homolog A (S. cerevisiae)	1.183	1.593	2	1	0
SYNJ2BP	synaptojanin 2 binding protein	1.185	1.457	2	1	0
AGFG2	ArfGAP with FG repeats 2	1.171	1.503	1	1	0
CILP	cartilage intermediate layer protein, nucleotide pyrophosphohydrolase	1.273	1.392	1	1	0
GADD45A	growth arrest and DNA-damage-inducible, alpha	1.255	1.424	1	1	0
KRCC1	lysine-rich coiled-coil 1	1.319	1.793	1	1	0
LEO1	Leo1, Paf1/RNA polymerase II complex component, homolog (S. cerevisiae)	1.205	1.474	1	1	0
SF3B1	splicing factor 3b, subunit 1, 155kDa, transcript variant 1	1.212	1.438	1	1	0
ZNF483	zinc finger protein 483, transcript variant 2	1.258	1.842	1	1	0
CRCP	CGRP receptor component, transcript variant 1, .	1.185	2.159	2	0	0
IGF2BP2	insulin-like growth factor 2 binding protein 2, transcript variant 1, .	1.194	1.485	2	0	0
SLC4A5	solute carrier family 4, sodium bicarbonate cotransporter, member 5, transcript variant c	1.206	1.361	2	0	0
SOX4	SRY (sex determining region Y)-box 4	1.188	1.784	2	0	0
FGD5	FYVE, RhoGEF and PH domain containing 5	1.176	1.411	1	0	0
IVNS1ABP	influenza virus NS1A binding protein	1.294	2.009	1	0	0
MYO3B	myosin IIIB	1.171	1.404	1	0	0
PPA2	pyrophosphatase (inorganic) 2, nuclear gene encoding mitochondrial protein, variant 2	1.188	1.569	1	0	0
RHOBTB3	Rho-related BTB domain containing 3	1.185	1.893	1	0	0
SSPN	sarcospan (Kras oncogene-associated gene)	1.230	1.450	1	0	0
ZNF394	zinc finger protein 394	1.171	1.858	1	0	0
ZNF577	zinc finger protein 577	1.280	1.404	1	0	0
ZNF69	zinc finger protein 69	1.213	1.848	1	0	0

**Table 5.17** Table of the candidate gene targets of miR-664b-3p represented in figure 5.5 (B) above the 12.5% cut off point. The ID and full name of all the gene targets of miR-664b-3p above the 12.5% cut off point which increase in expression when treated with the microRNA inhibitor and decrease in expression when treated with the microRNA mimic.



Gene	Gene name	Mimic fold change down	Inhibitor fold change up	N11 target sites		
				6mer	All 7mers	8mer
ITPRIPL2	inositol 1,4,5-triphosphate receptor interacting protein-like 2	1.325	2.097	7	5	1
IL18BP	interleukin 18 binding protein, transcript variant A	1.153	4.593	1	2	1
PSMB8	proteasome (prosome, macropain) subunit, beta type, 8 (large multifunctional peptidase 7), variant 2	1.326	1.872	1	2	1
PODXL	podocalyxin-like, transcript variant 1	1.191	2.932	4	3	0
MYD88	myeloid differentiation primary response gene (88)	1.166	1.991	3	2	0
APOL3	apolipoprotein L, 3, transcript variant beta/a	1.177	2.383	2	2	0
TGFBR3	transforming growth factor, beta receptor III	1.321	3.342	2	2	0
TMEM140	transmembrane protein 140	1.165	2.975	2	2	0
KLF6	Kruppel-like factor 6	1.164	4.717	2	1	0
MSX1	msh homeobox 1	1.185	3.340	2	1	0
CD248	CD248 molecule, endosialin	1.217	2.032	1	1	0
CSNK1E	casein kinase 1, epsilon, transcript variant 1	1.179	2.217	1	1	0
LY6E	lymphocyte antigen 6 complex, locus E	1.205	2.313	1	1	0
FKBP14	FK506 binding protein 14, 22 kDa	1.142	2.142	3	0	0
ITGA5	integrin, alpha 5 (fibronectin receptor, alpha polypeptide)	2.079	1.915	2	0	0
DDIT4L	DNA-damage-inducible transcript 4-like	1.373	2.092	1	0	0
MXD1	MAX dimerization protein 1	1.158	2.445	1	0	0
PLIN5	perilipin 5	1.437	2.332	1	0	0
SCIN	scinderin	1.170	3.279	1	0	0
TSC22D1	TSC22 domain family, member 1, transcript variant 2	1.257	1.907	1	0	0
TSC22D3	TSC22 domain family, member 3, transcript variant 1	1.425	2.052	1	0	0

**Table 5.18 Table of the candidate gene targets of miR-3085-3p represented in figure 5.5 (C) above the 12.5% cut off point.** The ID and full name of all the gene targets of miR-3085-3p above the 12.5% cut off point which increase in expression when treated with the microRNA inhibitor and decrease in expression when treated with the microRNA mimic.

#### 5.2.4 Experimental validation of microRNA targets

In order to validate a number of targets of miR-6509-5p, miR-664b-3p and miR-3085-3p experimentally, a shortlist of candidates was created (Table 5.19). These genes were among those which showed the appropriate response to the siRNAs (decreased expression with the microRNA mimic and increased expression with the inhibitor) as well as having at least one target site which would exactly match one of the microRNAs' seed sites. These genes were also chosen based upon their function.

Potential targets of miR-6509-5p included PDGFD (Platelet derived growth factor D), SFRP4 (Secreted frizzled-related protein 4), CA5B (Carbonic anhydrase VB), RBM3 (RNA binding motif protein 3) and TGF $\beta$ R3 (Transforming growth factor, beta receptor III).

Targets of miR-664b-3p chosen were TIMP3 (tissue inhibitor of metalloproteinases 3), GPM6B (Glycoprotein M6B), VCAN (Versican), and CILP (Cartilage intermediate layer protein).

Possible targets of miR-3085-3p were TGF $\beta$ R3 (Transforming growth factor, beta receptor III), MYD88 [Myeloid differentiation primary response gene (88)], CSNK1E (Casein kinase 1, epsilon) and ITGA5 (Integrin, alpha 5).

To narrow this selection down even further, the expression of each gene was measured by qRT-PCR in the original RNA samples sent for the array analysis to determine if they followed the same rate of regulation as the array data showed. Figure 5.6 shows the qRT-PCR results for the genes theorised to be miR-6509-5p targets.

*PDGFD* (A) appeared to decrease in expression in response to the miR-6509-5p mimic but showed a similar expression between the miR-6509-5p inhibitor and non-targeting control. *SFRP4* (B) also showed a decrease in expression from non-targeting control to the microRNA mimic, but again remained the same in expression between the microRNA inhibitor and control. *CA5B* (C) decreased in expression significantly ( $p \leq 0.01$ ) with the microRNA mimic compared to the control. Expression of the gene appeared to decrease slightly between the microRNA inhibitor and the control, however there are large error bars. *RBM3* (D) decreased in expression from the control to the microRNA mimic which approached statistical significance at  $p = 0.0642$ , however expression also decreased from control to microRNA inhibitor. Finally the expression

of *TGFβR3* (E) decreased in response to the microRNA mimic but also appeared to decrease when treated with the inhibitor.

Figure 5.7 shows the qRT-PCR results for the genes predicted to be targets of miR-664b-3p. *GPM6B* (A) shows a slight decrease in expression between the control and microRNA mimic and remains fairly stable in expression between the control and microRNA inhibitor if error bars are taken into account. *VCAN* (B) decreases slightly when treated with the microRNA mimic, and increases when treated with the microRNA inhibitor with a p value approaching significance ( $p=0.0829$ ). *TIMP3* (C) increases significantly when treated with the microRNA inhibitor ( $p \leq 0.001$ ) but remained stable between the control and microRNA mimic. Finally *CILP* (D) increases in expression between the control and microRNA inhibitor in a manner approaching significance ( $p = 0.0533$ ) and decreases when treated with the microRNA mimic.

Figure 5.8 shows the qRT-PCR results for the genes predicted to be targets of miR-3085-3p. *MYD88* (A) shows an increase in expression when treated with the microRNA inhibitor, but similar expression rates between the microRNA mimic and control. *TGFβR3* (B) increases in expression in response to the microRNA inhibitor, however there is no obvious effect when treated with the microRNA mimic. The expression of *CSNK1E* (C) increases slightly when treated with both the microRNA inhibitor and mimic and finally *ITGA5* (D) shows a significant decrease with the microRNA mimic ( $p \leq 0.01$ ) and a slight increase when treated with the microRNA inhibitor.

Constructs of the 3'UTRs of *SFRP4*, *TIMP3*, *CILP*, *MYD88* and *ITGA5* were successfully sub-cloned into the pmirGLO vector. Constructs of the 3'UTRs of *RBM3*, *TGFβR3* and *VCAN* were also attempted but proved unsuccessful. The transformed vectors were transfected into DF-1 fibroblast cells and treated with the corresponding novel microRNA mimic. This then underwent a luciferase assay to determine if the microRNA mimic was actually binding the target sequence within the target gene (Figure 5.9). *SFRP4* (A) showed a decrease in relative light units produced when the transfected cells were treated with the miR-6509-5p mimic compared to a control, however the decrease was not significant and there was also a decrease in relative light units when cells were treated with a mimic for miR-29b. *TIMP3* (B) showed a significant decrease ( $p \leq 0.001$ ) in relative light units when transfected cells were treated with the miR-664b-3p mimic compared to a non-targeting control. There was also a significant difference ( $p \leq 0.05$ ) in relative light units between cells treated with miR-

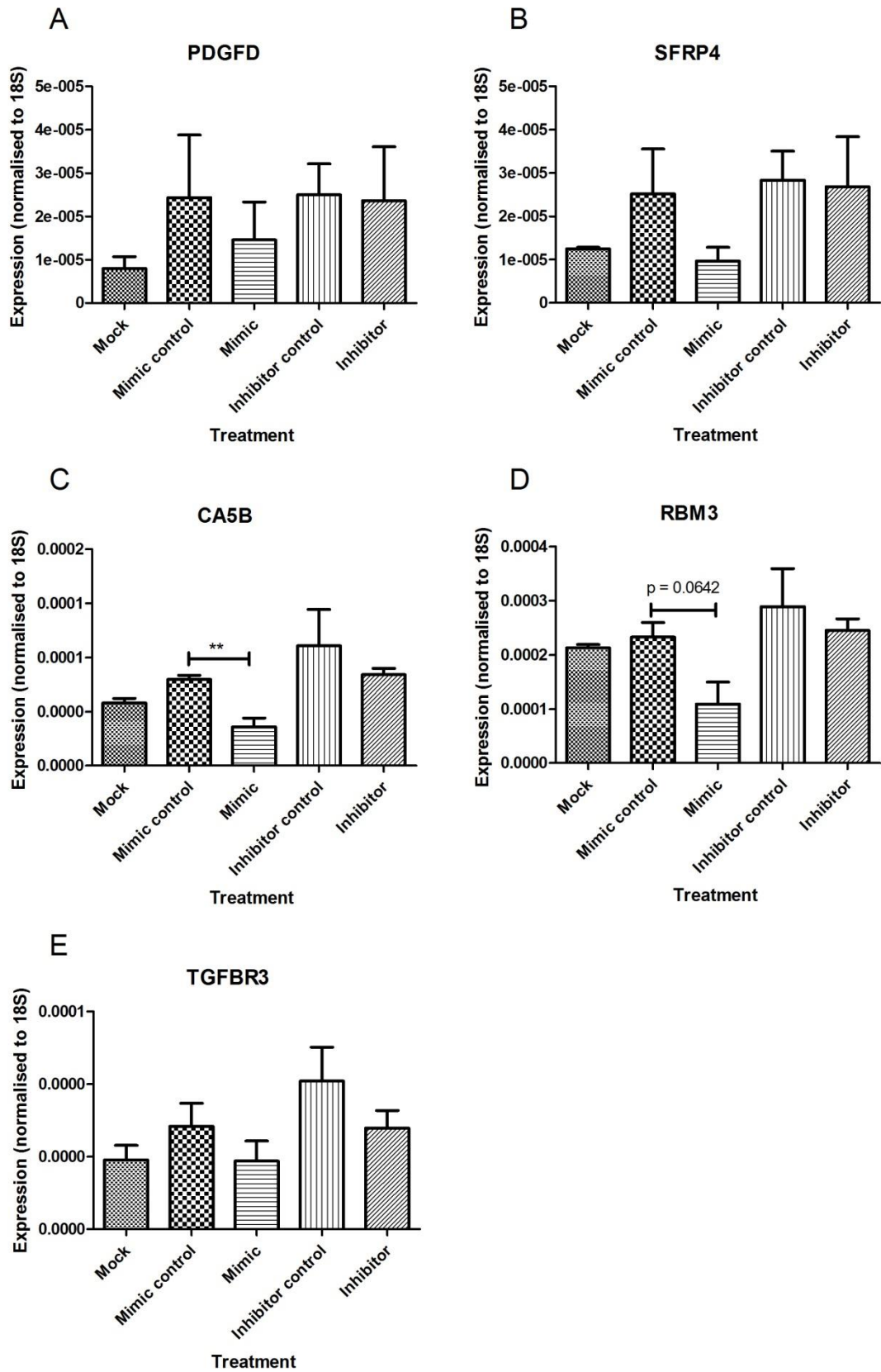
664b-3p and miR-29b mimics, with miR-664b-3p causing a significant decrease. CILP (C) showed the same relative light units when cells were treated with both a control and the miR-664b-3p mimic, with a decrease in relative light units actually being seen with cells treated with the miR-29b mimic ( $p \leq 0.01$ ). The expression of relative light units for cells transfected with the MYD88 (D) construct remained stable between the control, miR-3085-3p mimic and the miR-29b mimic. Finally ITGA5 (E) showed a statistically significant decrease in relative light units emitted by cells treated with the miR-3085-3p mimic compared to both the control and the miR-29b mimic ( $p \leq 0.001$ ).

Two subsequent constructs of ITGA5 were created with a mutation in one of the two target sites within the 3'UTR of the gene, and a further construct of ITGA5 with mutations in both target sites. These were used as further controls within the luciferase assay.

Figure 5.10 shows a significant decrease in emitted light units when cells transfected with the original ITGA5 construct were treated with the miR-3085-3p mimic ( $p \leq 0.001$ ) at 50nM compared to the non-targeting control at 50nM [as seen in figure 5.9 (E)]. When cells transfected with any of the three mutated constructs (two with mutations in only one seed site, one with mutations in both seed sites) were treated with the miR-3085-3p mimic, there was no statistical significance of emitted light units compared to treatments with the non-targeting control.

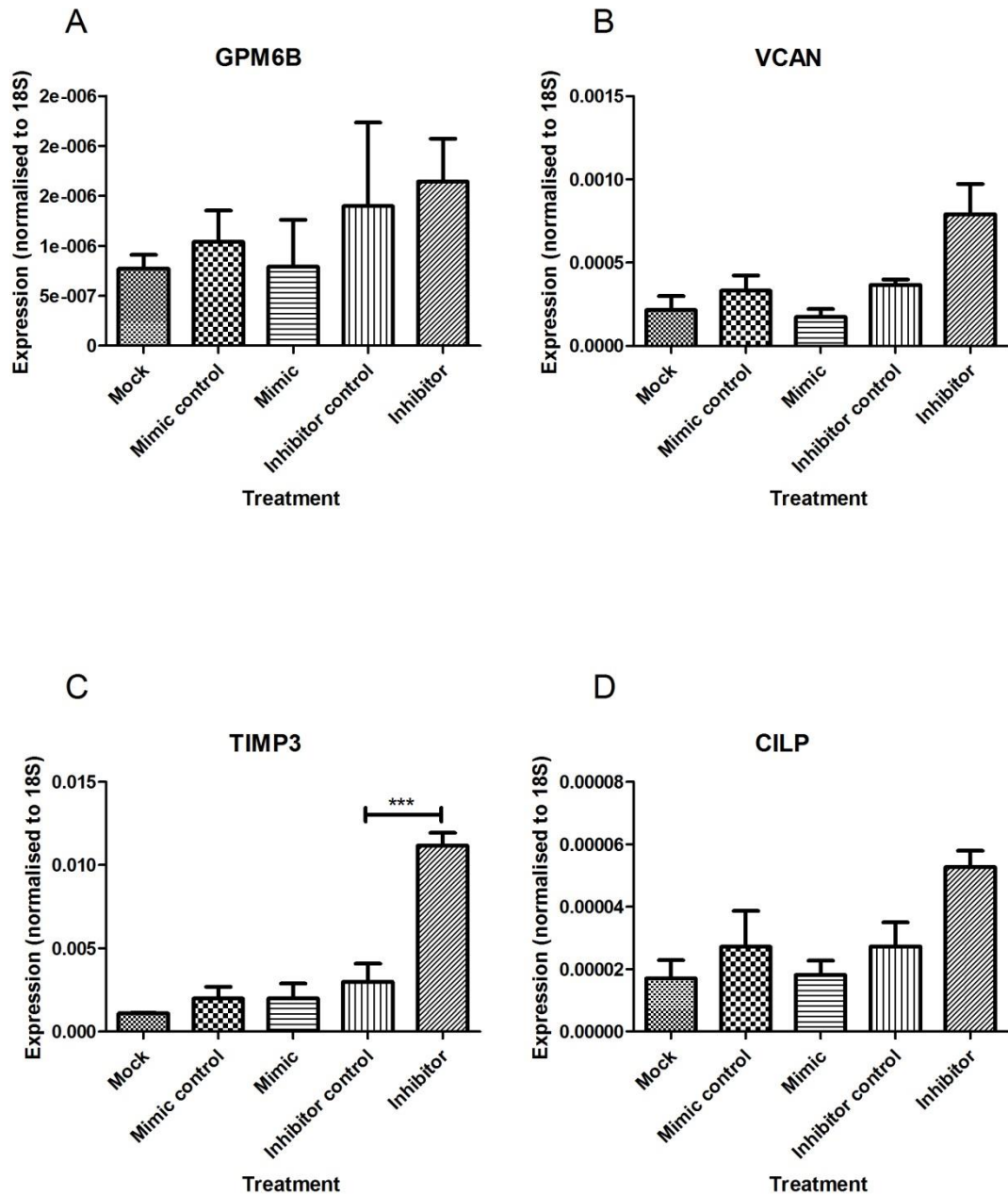
Novel microRNA	Ensembl gene ID	Gene	Gene name	Mimic fold	Inhibitor fold	N11 target sites		
				change down	change up	6mer	All 7mers	8mer
miR-6509-5p	ENSG00000170962	PDGFD	Platelet derived growth factor	1.188	1.254	4	2	0
	ENSG00000106483	SFRP4	Secreted frizzled related protein 4	1.659	1.303	1	1	0
	ENSG00000169239	CA5B	Carbonic anhydrase VB, mitochondrial	1.457	1.209	1	1	0
	ENSG00000102317	RBM3	RNA binding motif protein 3	1.201	1.409	1	1	0
	ENSG00000069702	TGFBR3	Transforming growth factor beta receptor 3	1.197	1.219	1	1	0
miR-664b-3p	ENSG00000046653	GPM6B	Glycoprotein m6b	1.444	1.317	9	7	1
	ENSG00000038427	VCAN	Versican	1.316	1.666	6	4	0
	ENSG00000100234	TIMP3	Tissue inhibitors of metalloproteinases 3	1.382	2.671	3	3	0
	ENSG00000138615	CILP	Cartilage intermediate layer protein	1.273	1.392	1	1	0
miR-3085-3p	ENSG00000172936	MYD88	Myeloid differentiation factor 88	1.166	1.991	3	2	0
	ENSG00000069702	TGFBR3	Transforming growth factor beta receptor 3	1.321	3.342	2	2	0
	ENSG00000213923	CSNK1E	Casein kinase 1-epsilon	1.179	2.217	1	1	0
	ENSG00000161638	ITGA5	Alpha 5 Integrin	2.079	1.915	2	0	0

**Table 5.19 List of candidate targets of miR-6509-5p, miR-664b-3p and miR-3085-3p to experimentally validate.** Candidate targets of the three microRNAs chosen to experimentally validate due to their expression pattern, target site content, and known or predicted function.



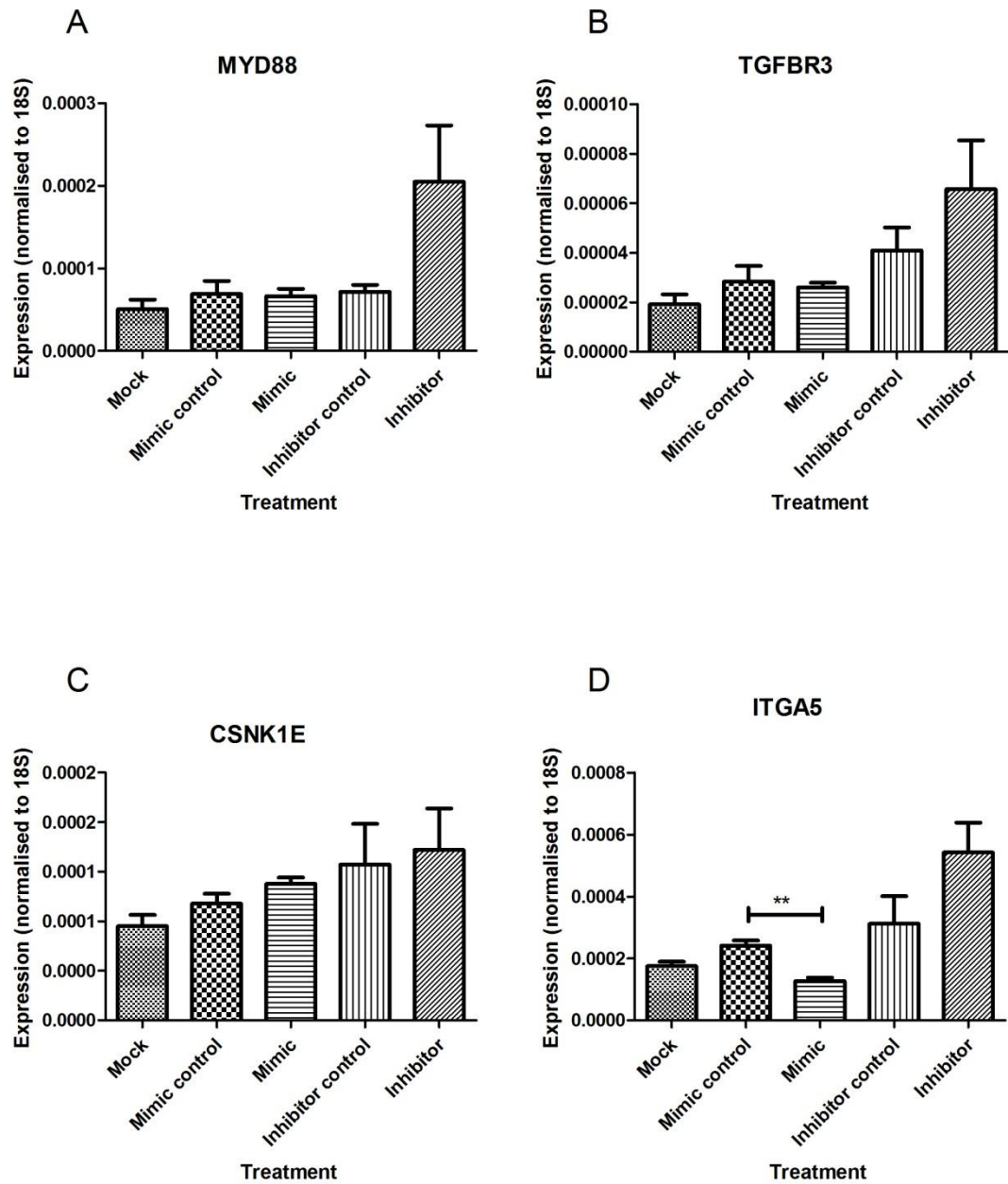
**Figure 5.6** Expression of candidate targets of miR-6509-5p in original array samples by qRT-PCR. Expression of (A) *PDGFD*, (B) *SFRP4*, (C) *CA5B*, (D) *RBM3* and (E) *TGFβR3* were obtained by qRT-PCR from RNA of primary chondrocytes

treated with miR-6509-5p mimic and inhibitor, and non-targeting controls. Data were normalised to *18S* RNA expression. Data are shown as mean +/- S.E.M, n = 3. (\*\* = p  $\leq$ 0.01 analyzed by a Student's t test).

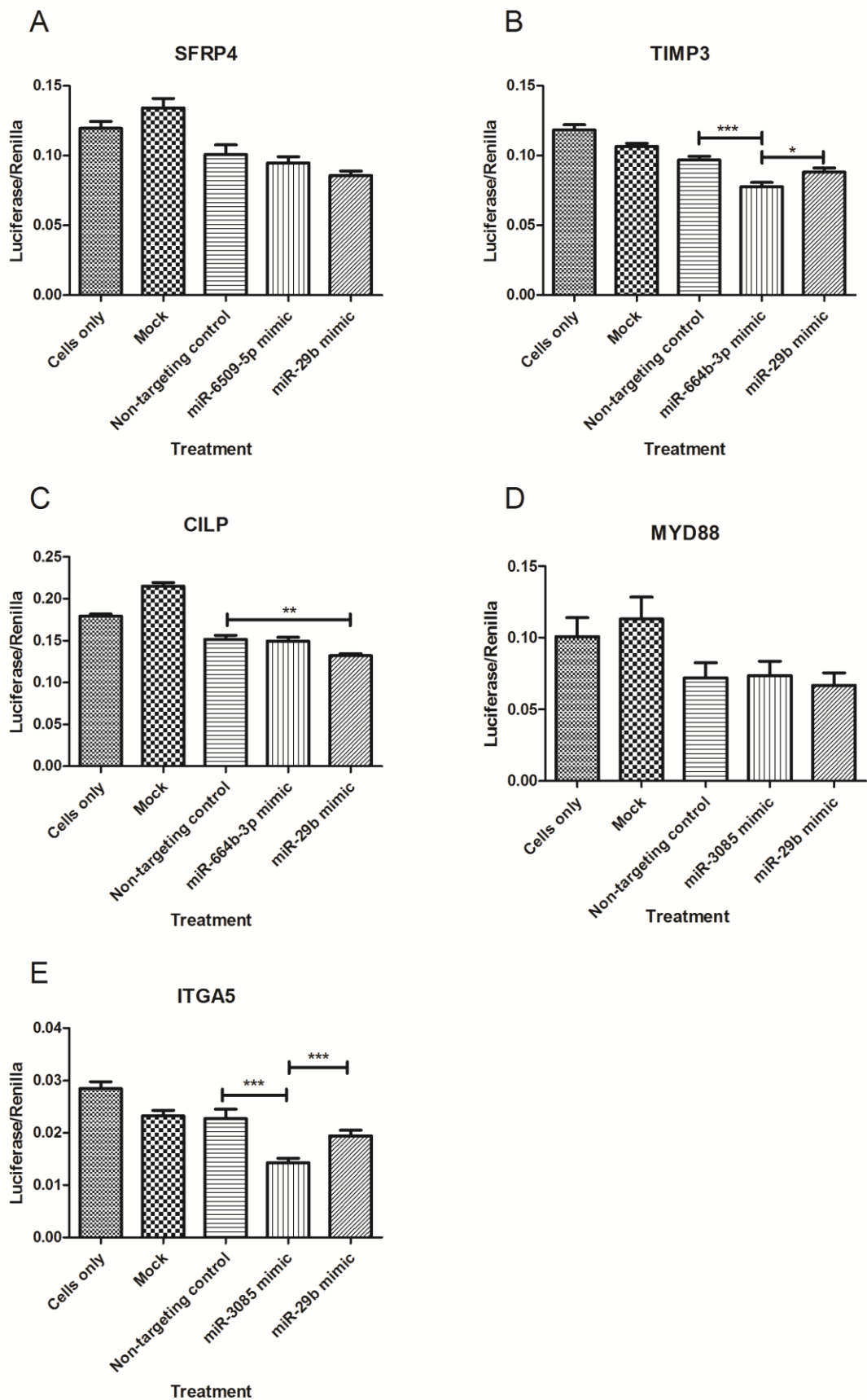


**Figure 5.7** Expression of candidate targets of miR-664b-3p in original array samples by qRT-PCR. Expression of (A) *GPM6B*, (B) *VCAN*, (C) *TIMP3* and (D) *CILP* were obtained by qRT-PCR from RNA of primary chondrocytes treated with miR-664b-3p mimic and inhibitor, and non-targeting controls. Data were normalised to *18S* RNA expression. Data are shown as mean +/- S.E.M., n = 3. (\*\*\*) =  $p \leq 0.001$  analyzed by a Student's t test).



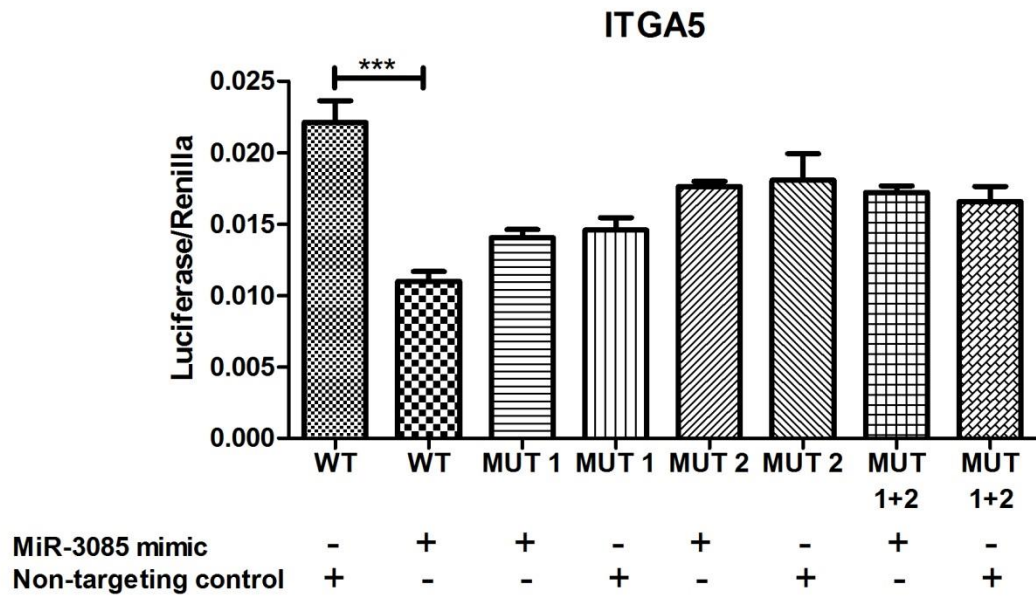


**Figure 5.8** Expression of candidate targets of miR-3085-3p in original array samples by qRT-PCR. Expression of (A) *MYD88*, (B) *TGFR $\beta$ 3*, (C) *CSNK1E* and (D) *ITGA5* were obtained by qRT-PCR from RNA of primary chondrocytes treated with miR-3085-3p mimic and inhibitor, and non-targeting controls. Data were normalised to *18S* RNA expression. Data are shown as mean  $\pm$  S.E.M., n = 3. (\*\* =  $p \leq 0.01$  analyzed by a Student's t test).



**Figure 5.9** Luciferase assays of cloned 3'UTR constructs of target genes in pmirGLO vector transfected into DF-1 cells and treated with microRNA mimics. 100ng of cloned DNA construct with 50nM of either microRNA mimic, inhibitor or a

non-targeting control siRNAs were transfected with 25% (v/v) Lipofectamine 2000 into DF1 cells for 24 hours at 37°C, 5% (v/v) CO<sub>2</sub>. Expression of firefly luciferase activity of (A) *SFRP4* treated with miR-6509-5p mimic, (B) *TIMP3* treated with miR-664b-3p mimic, (C) *CILP* treated with miR-664b-3p mimic, (D) *MYD88* treated with miR-3085-3p mimic and (E) *ITGA5* treated with miR-3085-3p mimic were obtained by firefly luciferase assay of the transfected DF-1 cells. Firefly luciferase relative light units were normalised to Renilla relative light units. Data are mean +/- S.E.M., n = 3. (\* = p ≤ 0.01, \*\* = p ≤ 0.01 and \*\*\* = p ≤ 0.001 analyzed by a Student's t test).



**Figure 5.10** Luciferase assay of cloned 3'UTRconstructs of wild type ITGA5 and mutated ITGA5 constructs in pmirGLO vector transfected into DF-1 cells and treated with the miR-3085-3p mimic or a scrambled non-targeting control. 100ng of cloned DNA construct with either 50nM of either microRNA mimic or non-targeting control siRNAs were transfected with 25% (v/v) Lipofectamine 2000 into DF1 cells for 24hours at 37°C, 5% (v/v) CO<sub>2</sub>. Expression of firefly luciferase activity when DF-1 cells were transfected with either an ITGA5 wild type (WT) construct, an ITGA5 construct with a mutation in one of two microRNA target sites (MUT 1, MUT 2), and an ITGA5 construct with a mutation in each microRNA target site (MUT 1+2). Firefly luciferase relative light units were normalised to Renilla relative light units. Data are mean +/- S.E.M., n = 3. (\*\*\*) = p ≤ 0.001 analyzed by a Student's t test).

### 5.3 Discussion

In order to capitalise upon the discovery of new microRNAs in osteoarthritic cartilage, the functions of these microRNAs were also explored. A number of techniques were used to identify possible targets of each of miR-6509-5p, miR-664b-3p and miR-3085-3p.

A functional analysis using the gain- and loss-of-function theory was carried out. This involved using custom-designed siRNAs of each microRNA. Both mimics and inhibitors of each microRNA were created and used to treat primary chondrocyte cells for a period of 48 hours, a non-targeting control for each of the mimics and inhibitors was also used. The resultant extracted RNA was sent for microarray analysis to measure the expression of over 47000 genes, gene candidates and splice variants when treated with the siRNAs compared to the non-targeting controls.

If a gene is a target of one of the microRNAs, then its expression should decrease when cells were treated with the mimic as this would introduce additional synthesised microRNA into the system which would increase suppression of the gene in question. When treated with the inhibitor the expression of the gene should increase, as the inhibitor would bind the existing microRNA within the system, reducing the amount of microRNA which is able to repress the same gene, resulting in an increase in gene expression. If a gene both decreases when treated with the mimic and increases when treated with the inhibitor, it is considered a potential target. This form of functional analysis is widely used to determine mRNA targets of microRNAs for example Cxcl12 [Chemokine (C-X-C Motif) Ligand 12] was validated as a target of miR-140 in this manner. (Nicolas et al., 2008).

However, the baseline level of microRNA already present in a cell can play a role in how effective the aforementioned methodology is. We found that the largest fold changes seen in the array were generally in response to the mimic siRNAs (Fig 5.2) however, the inhibitor of miR-3085-3p also seemed quite effective. The expression of each of the three microRNAs was measured in RNA obtained from siRNA treated cells prior to analysis by array. These data showed, for each microRNA, a dramatic increase in expression when treated with mimic compared to a non-targeting control, ranging from 40 to over 100 fold increases. In contrast, the expression of the microRNAs after treatment with a siRNA inhibitor, although decreased, showed much lower fold changes (data not shown). Overall, baseline expression of each microRNA within the primary

chondrocyte cells was relatively low, this is supported by the initial deep sequencing data in chapter 3 of primary chondrocytes and the subsequent northern blot validation using SW1353 RNA in chapter 4, meaning that the mimic siRNAs were more likely to have a larger effect than the siRNA inhibitors within the cells due to baseline levels initially being quite low.

We already know that for a microRNA to suppress the expression of a gene it must bind to a specific sequence usually found within the 3'UTR of the gene in question (Akbari Moqadam et al., 2013). This binding can be either partial or complete between the target site and a location on the microRNA known as the seed site.

For a candidate target discovered from the microarray to be a true target, there must be a form of binding site as well, otherwise this could represent a false positive result or indicate a downstream target. The potential seed sites of all 3 microRNAs were assessed and the exact target sequences they would bind to were hunted for amongst all the 3'UTRs of the genes within the human genome as downloaded from Ensembl.org (Flicek et al., 2014).

Four potential seed sites from each microRNA were used to produce the target sites sequences in the search: 6mer, 7merA1, 7merm8 and 8mer. The complementary sequences to these sites were used as search criteria using the R programming language with RStudios (R Core Team, 2013) and the Biostrings package (Pages et al., 2012). Unfortunately, the R code did not account for any partial target matches so it is likely that a number of potential gene targets were missed. It is still not known how many mismatches between a microRNA seed site and mRNA is acceptable to still function to repress the gene, meaning there are few gold standard negative gene : microRNA binding matches to work with to develop a perfect search algorithm (Akbari Moqadam et al., 2013). So to that end it was considered more productive in this project to simply look at exact microRNA seed site to mRNA target site matches.

There are existing online algorithms which are able to search for potential gene targets of microRNAs based upon their structure in addition to their seed site sequence, however, these only function with existing, described microRNAs currently available on miRBase (Griffiths-Jones et al., 2006; Griffiths-Jones et al., 2008; Kozomara and Griffiths-Jones, 2011). There is another algorithm which is able to assess target binding of any microRNA sequence, whether or not it is present on the miRBase database, and

this is RNAhybrid (Kruger and Rehmsmeier, 2006). However this algorithm has a much lower throughput compared to the R programming assessment which was used.

A large number of genes were found to have a least one 6mer target site within their 3'UTR for each of the three microRNAs. 4,655 genes had target site/s for miR-6509-5p, 9,690 genes had target site/s for miR-664b-3p and 8,930 genes had target sites for miR-3085-3p. The top 20 of each are presented in table 5.13, table 5.14 and table 5.15. These potential targets were ranked based on the size and number of their target sites. 8mer target sites are thought to cause a more significant repression than shorter matches (Schnall-Levin et al., 2011), so the target genes were ranked firstly based on their number of 8mer matches, followed by the total number of 7mer matches and then their 6mer matches.

Out of the 47,000 genes, gene candidates and splice variants tested (shown in Fig.5.2), those that decreased in expression with the microRNA mimic and increased in expression with the microRNA inhibitor compared to non-targeting controls included 11,871 for miR-6509-5p, 10,800 for miR-664b-3p and 10,953 for miR-3085 (Fig. 5.3). All of those genes were then cross-referenced with all of the genes discovered to have target sites with the computational analysis. This resulted in 1782 potential targets of miR-6509-5p, 3245 potential targets of miR-664b-3p and 3091 potential targets of miR-3085-3p (shown in Fig. 5.5).

Interestingly, when examining the array data without the addition of the computational target site analysis, a number of interferon-induced genes were found to be up-regulated when treated with the miR-6509-5p mimic and inhibitor, the miR-664b-3p mimic and inhibitor and the miR-3085 inhibitor suggesting an interferon response in the primary chondrocytes to the siRNAs. They were found to be within the top 10 genes showing regulation in this manner and included IFIT1 (Interferon-induced protein with tetratricopeptide repeats 1), IFIT2 (Interferon-induced protein with tetratricopeptide repeats 2), IFIT3 (Interferon-induced protein with tetratricopeptide repeats 3) and IFI44L (Interferon-induced protein 44-like). Interferon-induced genes are closely linked to the immune response to viral infection and IFIT1/2/3 are thought to target the FLUAV (influenza A virus), HPV (human papillomavirus), MHV (mouse hepatitis virus), RVFV (Rift Valley fever virus), SINV (Sindbis virus), VSV (vesicular stomatitis virus) and WNV (West Nile virus) viruses (Schoggins and Rice, 2011). The mechanism by which the IFIT genes respond to viral threats involves the formation of a

multiprotein complex which is able to sequester viral RNAs containing 5' triphosphates (Pichlmair et al., 2011; Schoggins and Rice, 2011) The IFIT genes are also induced by double stranded RNA and lipopolysaccharides in addition to interferons (Fensterl and Sen, 2011).

The induction seen from the array could therefore be the result of flooding the system with double stranded RNA in the form of synthesised microRNA mimics. However this does not account for the increased interferon presence when cells were treated with microRNA inhibitors as these were synthesised as single stranded RNA molecules.

Research has shown a common problem of siRNA technology in causing off target effects, especially in the form of an immune response. For example, it has been found that synthesised siRNAs can activate RIG-1 (retinoic acid-induced gene 1) and Toll-like receptors which are sensors of foreign RNA and can trigger an innate immune response, of which interferons play a role (Olejniczak et al., 2011). Aagaard et al also discuss the common problem of off-target effects from the therapeutic use of siRNAs (Aagaard and Rossi, 2007).

Although it was initially thought that short RNA sequences would not induce such a strong interferon response as the long dsRNA, it was found that specific siRNA sequences could activate an immune response through Toll-like receptors, resulting in the up-regulation of interferon-induced genes (Hornung et al., 2005; Marques and Williams, 2005). Further research from Reynolds et al also show that synthesised siRNAs can cause an immune response in transfected cells, and that although shorter siRNAs are thought to cause the smallest off target effects, this length threshold can differ greatly between cell types (Reynolds et al., 2006). Therefore it can be concluded that the apparent up-regulation in the array of many genes known to be interferon inducible is most likely a product of the off-targeting effects which siRNAs can cause, rather than a biologically relevant response to the microRNA mimics and inhibitors used in this project.

The large number of potential targets which responded appropriately to the microRNA siRNAs (showing repression in response to mimics and de-repression in response to inhibitors) in addition to possessing target sites for each microRNA in question were triaged further with the introduction of fold-change cut off points based upon their expression in response to both the microRNA mimics and inhibitors. This ensured that genes were chosen which had more of a balanced response to both the mimics and



inhibitors. The baseline expression of each microRNA was relatively low meaning the mimic of each microRNA would show a higher response than the inhibitor. The cut off points were set at 50%, 25% and 12.5% of the largest fold-change of each mimic and inhibitor (Fig. 5.5).

The triaged pool consisted of 23 genes for miR-6509-5p, 53 genes for miR-664b-3p, and 21 genes for miR-3085-3p. The three genes which contained the most microRNA target sites of miR-6509-5p, miR-664b-3p and miR-3085-3p and also increased in expression with the microRNA inhibitor and decreased in expression with the microRNA mimic were LRRC58, SNTB2 and ITPRIPL2 respectively.

LRRC58 is known to selectively interact with protein complexes (Kobe and Deisenhofer, 1994). SNTB2 codes for a protein which binds and organises the subcellular localisation of specific membrane proteins (Lumeng et al., 1999). Finally, ITPRIPL2 is known to be a protein coding gene linked to the lncRNA class, however its exact function has not been elucidated. ITPRIP is a paralog of ITPRIPL2 and is known to enhance the inhibition of ITPR (Inositol 1, 4, 5-trisphosphate receptor) calcium release (Mignery and Sudhof, 1990; van Rossum et al., 2006).

A number of the genes within these chosen subsets with known functions linked to cartilage, osteoarthritis, chondrogenesis or bone development were validated further using qRT-PCR which gives a greater sensitivity than microarray (Git et al., 2010).

The chosen targets of miR-6509-5p included PDGFD (Platelet derived growth factor D) which functions as a mitogenic factor for MSC derived cells as well as being involved in the development of kidneys, eyes and brain (Reigstad et al., 2005), SFRP4 (Secreted frizzled-related protein 4) which is a modulator of WNT signalling mediated by frizzled membrane receptors (Jones and Jomary, 2002). Also a target, was CA5B (Carbonic anhydrase VB) which is a member of the zinc metalloenzyme family and is involved in adipocyte development and metabolism. It is also known to be expressed in human articular cartilage (Schultz et al., 2011). RBM3 (RNA binding motif protein 3) was a third potential target and is a cold inducible protein which appears to inhibit microRNA function (Pilotte et al., 2011) and is thought to play a role in innate immune responses (Matsuda et al., 2011). Finally there is TGF $\beta$ R3 (Transforming growth factor, beta receptor III) which may be involved in the inhibition of TGF $\beta$  signalling (Moren et al., 1992; Santibanez et al., 2011).

Targets of miR-664b-3p chosen were TIMP3 (tissue inhibitor of metalloproteinases 3) which is known to inhibit the MMP function of many MMPs (both membrane bound and transmembrane) including MMP-1, MMP-2, MMP-3, MMP-7, MMP-9, MMP-13, MMP-14 and MMP-15 (Kashiwagi, Tortorella et al. 2001; Visse and Nagase 2003)(Apte et al., 1996). GPM6B (Glycoprotein M6B) was also chosen and is thought to be involved in osteoblast function and the development of bone (Drabek et al., 2011). VCAN (Versican) was the third gene chosen and encodes chondroitin sulphate proteoglycan which in turn is a cartilage ECM component (Sotoodehnejadmatalahi and Burke, 2013). Finally CILP (Cartilage intermediate layer protein) was the last chosen, which has a role in cartilage scaffolding and is also known to play a role in a number of cartilage diseases (Lorenzo et al., 1998; Wang et al., 2012).

The selected targets of miR-3085-3p were TGF $\beta$ R3 (Transforming growth factor, beta receptor III) and MYD88 [Myeloid differentiation primary response gene (88)] which is involved in the IL1 pathway as well as the activation of NF $\kappa$ B (Medzhitov et al., 1998). CSNK1E (Casein kinase 1, epsilon) was another and is part of the WNT pathway and responsible for phosphorylating both itself and other proteins including DVL1 (Dishevelled segment polarity protein 1) and a number of proteins involved in the regulation of circadian rhythm: PER1 (Period circadian clock 1) and PER2 (Period circadian clock 2). Through this it is thought to play a role in the negative regulation of circadian rhythms (Hino et al., 2003; Wager-Smith and Kay, 2000). The final selected target of miR-3085-3p is ITGA5 (Integrin Alpha 5). This gene codes an integrin, which functions as a transmembrane link between the actin microfilaments of the cell cytoskeleton and extracellular factors such as the ECM (Takada et al., 2007). ITGA5 encodes integrin alpha 5 which is a fibronectin receptor as well as an adhesion molecule (Nagae et al., 2012).

Figures 5.6, 5.7 and 5.8 show the expression of these genes measured in the original RNA samples (isolated from cells treated with microRNA mimics, inhibitors and non-targeting controls) sent to array.

The genes which showed the same pattern of expression with qRT-PCR as in the array (SFRP4, TIMP3, CILP, MYD88, ITGA5, RBM3, TGF $\beta$ R3 and VCAN) were selected for further experimental validation by a luciferase reporter assay. Sites within each gene's 3'UTR which contained the microRNA target site/s were sub-cloned into pmirGLO plasmid vectors with a firefly luciferase component. Constructs of SFRP4,

TIMP3, CILP, MYD88 and ITGA5 were successfully sub-cloned and the constructs were transfected into DF-1 chicken fibroblast cells. DF-1 cells are a commonly used cell line for transfections, and were chosen in this instance, due to unsuccessful transfection attempts into both SW1353 chondrosarcoma cells and primary chondrocytes. DF-1 cells, originating from chicken, should also not contain the human microRNAs under investigation at a baseline level, ruling out any non-specific effects.

The transfected cells were then treated with the relevant microRNA mimic. The emission of relative light units were measured and if they decreased when cells were treated with a microRNA mimic compared to the non-targeting control then it suggests that the microRNA was binding to the target site within the construct, decreasing the translation of luciferase.

Out of the 5 genes successfully cloned, 2 showed a statistically significant decrease in light units emitted when transfected cells were treated with the microRNA mimics compared to non-targeting controls suggesting a successful binding of microRNA to mRNA target. These were ITGA5 and TIMP3 (Fig. 5.9).

TIMP3 is known to inhibit the function of ADAMTS4, ADAMTS5 and TACE/ADAM17 (tumor necrosis factor- $\alpha$  converting enzyme) expression (Amour et al., 1998; Brew and Nagase, 2010; Kashiwagi et al., 2001; Visse and Nagase, 2003). In addition to lung and cardiac deficiencies, (Fedak et al., 2004; Leco et al., 2001) the TIMP3 null mouse shows an increase in collagen and aggrecan degradation similar to that seen during osteoarthritis. Showing that TIMP3 plays an important role in metalloproteinase balance in cartilage homeostasis (Sahebjam et al., 2007). Mahmoodi et al have also shown an increased inflammatory response in TIMP3<sup>-/-</sup> mice induced through arthritis by intra-articular injection, with an increased level of TNF- $\alpha$  detected. The research suggests that TIMP3 plays a role in combating TNF- $\alpha$  induced inflammation in inflammatory arthritis by the targeting of ADAM17 (Mahmoodi et al., 2005). In addition to its role in arthritis, mutations in TIMP3 have also been linked to other diseases such as Sorby's fundus dystrophy which is a degenerative eye disease (Weber et al., 1994).

TIMP3 is also known to inhibit many MMPs, a number of which are involved in the progression of osteoarthritis, specifically MMP1, 8 and 13 which are known to be up-regulated in osteoarthritic cartilage compared to normal (Billingham et al., 1997; Kevorkian et al., 2004). It has also been found that the degradation of collagen by

MMPs can be reduced by the introduction of inhibitors of these collagenases (Dahlberg et al., 2000; Poole et al., 2002). This suggests that TIMP-3 could be a useful target of miR-664b-3p to study further, and that miR-664b-3p could play an interesting role within the progression of osteoarthritis.

Additionally, mutant constructs of ITGA5 were also created to see if the effect of the microRNA mimic seen with the original construct could be prevented (Fig. 5.10). Mutant constructs of TIMP3 were also attempted but sub-cloning proved unsuccessful.

Two constructs of ITGA5 contained mutations in one of the two target sites present, whilst the third contained two mutated target sites. The luciferase assay shows the emission of light units significantly decreasing in cells transfected with the original construct and the miR-3085-3p mimic compared to a non-targeting control showing a decrease in luciferase translation meaning that the miR-3085-3p mimic was indeed binding to the sub-cloned target site. In comparison, the light units emitted in cells transfected with each of the three mutated constructs showed no significant change when treated with the miR-3085-3p mimic compared to the non-targeting control, showing that luciferase translation was unaffected due to the abrogation of microRNA : mRNA binding.

ITGA5 codes an integrin, which functions as a transmembrane link between the actin microfilaments of the cell cytoskeleton and extracellular factors such as the ECM (Takada et al., 2007). ITGA5 encodes integrin alpha 5 which is a fibronectin receptor as well as an adhesion molecule (Nagae et al., 2012). The integrin alpha 5 chain undergoes post-translational cleavage resulting in disulphide-linked light and heavy chains which are able to bind with integrin beta 1 and form the known integrin alpha5beta1 receptor which binds to fibronectin and fibrinogen (Nagae et al., 2012). Fibronectin is a glycoprotein found in the ECM of normal articular cartilage and is thought to play a role in the organisation of the ECM. It has been found to increase in expression in osteoarthritic cartilage due to increased production by chondrocytes, and this is thought to result in a potential role in the changing phenotype of chondrocytes during osteoarthritis (Chevalier, 1993).

Integrin alpha 5 beta 1 is also known to mediate chondrocyte attachment via COMP (cartilage oligomeric matrix protein) (Chen et al., 2005). Knockout mice of  $\alpha 5$  show a lethal phenotype with aberrant brain and intestine development and function, whilst  $\beta 1$

knock out mice shows a severe phenotype with defects within the vascular system.(Schaffner et al., 2013).

In addition, it has been found that Integrin alpha5beta1 has a role as a mechanoreceptor in both normal and OA chondrocytes. Millward-Sadler et al discovered that when subjected to pressure-induced strain, OA chondrocytes responded differently to their non-OA counterparts, with the cell membranes showing a depolarisation or hyperpolarisation response respectively. Though use of alpha 5 antibodies, Millward-Sadler and colleagues have determined that alpha5beta1 plays a vital role in this mechanotransduction event (Millward-Sadler et al., 2000).

Integrin alpha 5 beta 1 is obviously multifunctional and important in the cellular connections to ECM components. Using miR-3085-3p to manipulate gene expression could result in an impact upon both osteoarthritis progression and mechanotransduction pathways in the disease, but is also likely to affect other vital functions, unless the affect can be directed specifically to cartilage tissue.

Species conservation of microRNAs is a common method used to identify 'novel' microRNAs between species, with the often correct theory that the same microRNAs are often found within the genome of related species (Zhang et al., 2006a). The same process can be used when identifying targets of microRNAs which are found within multiple species. If a microRNA is conserved, then it is likely that its seed site is also conserved and it will bind to a common gene or homologue. For example, ITGA5 was found to be a target of miR-3085-3p. ITGA5 transcripts can be found in 50 species' genomes in addition to humans (Ensembl.org). Out of those 51 species, the 3'UTR sequence was only available for 17. And of those, 9 species (Chimpanzee, Dog, Ferret, Gorilla, Human, Marmoset, Mouse, Orangutan and Rat) were found to contain 6mer target sites for miR-3085-3p (data not shown) similar to that seen in humans, in addition to those species' genomes containing the miR-3085-3p sequence itself.

There are many experimental methods available to validate gene targets of microRNAs, including qRT-PCR, luciferase reporter assays and western blotting. QRT-PCR examines the co-expression of the microRNA and target mRNA, whilst western blots describe effects at the protein level. Reporter assays in turn show direct miRNA: mRNA interactions (Kuhn et al., 2008; Thomson et al., 2011). There are of course certain drawbacks with all experimental validation methodology, with successful sub-cloning

vital for reporter assays, and effective antibodies and primer probe sets necessary for western blotting and qRT-PCR respectively.

Targets which are regulated at the mRNA level can be detected by microarray, qRT-PCR and luciferase reporter assays and it is thought that the majority of microRNA targets can be detected by changes at the mRNA level due to the destabilisation caused by the microRNA: mRNA interaction (Guo et al., 2010). This suggests that the main mechanism of gene suppression by microRNAs is through mRNA destabilisation rather than affecting the efficiency of translation, in turn meaning that experimental validation looking at mRNA changes would be more efficient than examining protein regulation by western blot. Guo et al compared ribosome profiling with mRNA detection and discovered that over 84% of the protein reduction observed could be accounted for by a corresponding mRNA expression decrease (Guo et al., 2010; van Rooij, 2011). However recent data has shown that there are microRNA targets which can only be detected at the protein level such as Smad3 as a miR-140 target (Pais et al., 2010), this suggests that proteome analysis would be necessary in cases where a change in mRNA levels of a likely candidate target is not detected.

Methods of functional analysis of microRNAs more specific to cartilage and chondrocyte biology include the use of chondrogenesis assays such as cartilage disc models. Human mesenchymal stem cells can be induced through chondrogenesis to form cartilage discs with expression of known cartilage matrix genes (Murdoch et al., 2007). Dr. Barter from Newcastle University has shown an inhibition of *Dicer* within the chondrocytes of these discs results in malformation of the disc (Dr Matt Barter, Newcastle University, Personal Communication), showing that the microRNA population is vital for aspects of cartilage development. The general effect of one or multiple microRNAs can be assessed in the same way via overexpression or inhibition within these cartilage discs to determine if the microRNAs play a role in chondrocyte and cartilage formation.

Additional methods of experimental target validation include utilising the interaction between a microRNA and its target via the Ago (Argonaute) protein complex. Pull-down methods can be employed to detect the mRNA targets bound to these complexes alongside the microRNA (Beitzinger et al., 2007). Further techniques involve the use of *in vivo* experimentation; however there are many known issues with the delivery system of *in vivo* applications. For example, siRNAs are degraded easily by enzymes and

RNAses within the blood and tissue, and it has been found that accumulation of the siRNA in the area of interest can be problematically low, due mainly to poor delivery methods.

There have been a number of approaches to address these problems, including the use of electroporation and use of a carrier for the siRNA, such as liposomal bubbles (Higuchi et al., 2010). MicroRNAs have also been successfully injected into knee joints using an atelocollagen complex as a carrier. Atelocollagen consists of a purified type 1 collagen. Nagata et al used a microRNA-atelocollagen complex to effectively deliver miR-15a to the knee joints of mice (Nagata et al., 2009). This substance appears to be very effective in siRNA delivery and has also been used in the delivery of siRNAs to various cancer tissues (Mu et al., 2009).

This microRNA project has utilised a number of functional and target analysis methodologies to examine candidate targets of the three microRNAs validated in chapter 4. These combined data have determined miR-3085-3p to have at least one fully experimentally validated target which is known to play a role in osteoarthritis: ITGA5, with TIMP3 being a highly likely target of miR-664b-3p. This means that the two microRNAs may be able to influence aspects of osteoarthritis if correctly manipulated, although obviously not all of their potential targets have, as yet, been investigated which could result in unforeseen consequences if they were to be used therapeutically.

## **Chapter 6 – General Discussion and Future Directions**



## **6.1 – MicroRNA profiling**

This project has addressed the hypothesis that cartilage contains as yet unidentified microRNAs which may contribute to skeletal development or osteoarthritis.

Osteoarthritis is a debilitating disease affecting over 8 million people in the UK today, a number which is thought to increase to over 17 million by the year 2030 (Arthritis-Care, 2012). It is most common in the elderly but is also associated with risk factors such as obesity and mechanical injury (Attur et al., 2010). It is disease which encompasses the entire joint, including the progressive degradation of cartilage (Martel-Pelletier et al., 2008) and is the focus of much research.

MicroRNAs are small RNA molecules which effect gene expression. Initially thought to be merely fine tuners of gene regulation, extensive publications have demonstrated that microRNAs are also vital to survival (shown by the embryonic lethal Dicer knockout mouse model (Bernstein et al., 2003)) and play important roles in a large range of diseases and developmental processes, in addition to their continued minor regulation of a number of genes. Previous research has already shown a role for microRNAs in cartilage development and pathology, such as miR-140 and miR-455, two microRNAs known to be cartilage specific with roles in development, and altered expression in osteoarthritis (Miyaki et al., 2009; Miyaki et al., 2010; Swingler et al., 2012).

This project used a targeted approach first to identify the existing miRNome of osteoarthritic cartilage in as great a detail as possible, there has been little previous exploration into identifying the entire microRNA population of cartilage with only one recent study by Sun et al discovering 286 known microRNAs and 86 novel microRNAs, however this was in rat cartilage rather than human (Sun et al., 2011b). A next generation deep sequencing approach of primary chondrocytes extracted immediately from the osteoarthritic cartilage of three patients was utilised on the Illumina platform. To enable the most accurate detection rate, newly developed HD adapters were also used to reduce sequence bias (Sorefan et al., 2012).

Chapter 3 describes the results from the sequencing analysis. 990 microRNAs described on miRBase were discovered, along with 1621 novel microRNA candidates. These sequences were ascribed microRNA status by Dr H. Pais (School of Computing, UEA) using the MiRCat software package (Moxon et al., 2008). Ideally, if the theme of this project were to be extended and awarded further funding, the deep sequencing would be replicated on further OA patients. It is already known that samples often show wide

variation between patients, potentially due to undefined subtypes (Bierma-Zeinstra and Verhagen, 2011). One way to control for this would be to carry out multiple replicates, allowing for the detection of the microRNA population to be compared between multiple patients to discover the most common miRNome for osteoarthritis. Subtypes of osteoarthritis could also be investigated to find a common miRNome and enable stratifying of patients. Using additional patients would also reinforce and verify the discovery of novel microRNAs, if the same candidate is detected in multiple patients it suggests that it is a genuine detection and not a false positive created by the presence of a degraded longer sequence which fulfils the criteria of a microRNA. Deep sequencing analysis could also be carried out on primary chondrocytes from non-OA control cartilage samples, to allow comparison of the miRNome between diseased and normal tissues.

However, carrying out next generation sequencing on so many multiple samples does incur greater cost. This could however be negated, as sequencing technologies improve, especially in microRNA detection, and become cheaper. One example of this is nanopore sequencing. This method works by the passing of a single stranded DNA polymer through a protein nanopore. The nucleotide bases are identified as they pass through the pore. The technology is also able to detect specific microRNAs; the current passing through the nuclear pore is measured, with changes in the current signifying a probe-bound microRNA passing through the pore (Schneider and Dekker, 2012).

The large number of new sequences discovered in this project made validation of all in the timeframe impossible. To this end, a smaller pool of promising candidates were chosen to investigate further based on a number of criteria.

Triaging from the initial pool of 1621 candidate novel microRNAs to 16 was based on, first of all, whether they were present in all three patients, as presence in more than one patient increases the chance that the candidate was real. Secondly, the presence of the complementary strand of the candidate microRNA within the expression profile was also used as a positive selection. Both strands of a microRNA are more commonly being found to be expressed and functional as opposed to having just one dominant strand (Winter et al., 2009).

Finally, the candidates were cross-referenced with other sequencing projects obtained by Professor Tamas Dalmay's lab (School of Biological Sciences, UEA) to see if they appeared in other tissues or cell lines, this would provide support to their microRNA

status if they were detected elsewhere, and suggest that they were not just artefacts of a specific experiment. The two cell lines compared were DLD-1 wild type and DLD-1 dicer knockdown. In addition to just comparing presence of the sequences between the projects, Dicer is an enzyme known to be vital to microRNA synthesis. If the candidate's expression was down-regulated in the Dicer knock down cell line compared to the wild type, then this added credibility to the candidacy, suggesting that the sequence was processed by the Dicer enzyme. However, this factor was only included in combination with the previously mentioned criteria, and a candidate was not discarded if it fulfilled the others but did not appear to be Dicer dependent, this is due to the growing idea that not all microRNAs are dependent upon Dicer for their synthesis (Cheloufi et al., 2010).

Future work could focus on the validation of additional candidate microRNAs, using data obtained from subsequent sequencing analyses of primary chondrocytes, to determine the most likely true candidates. In addition, new techniques adopted in the lab since the initial sequencing experiment include a more efficient method of cartilage crushing, this could provide cleaner RNA for use in future sequencing, and could be used in addition to primary chondrocytes to provide a fuller overview of microRNA expression in osteoarthritis and cartilage.

Within microRNA research, over 30,000 microRNAs have been discovered across all species, however in comparison only a small number have been experimentally validated and further investigated functionally. If this is also combined with the knowledge that each microRNA could have hundreds of possible targets (Brennecke et al., 2005; Hendrickson et al., 2009); then there is still a wealth of untapped information regarding microRNA biology. The next direction of general microRNA research will most likely focus more heavily in fully understanding the existing microRNAs in physiology and pathophysiology in addition to the discovery of new microRNAs as detection methods are increasingly improved. This could be addressed with a more systems biology based approach (Watanabe and Kanai, 2011). MicroRNA biology is complex, and to truly understand the role of microRNAs within a specific tissue or disease, then a broader look must be taken to incorporate the entirety of miRNA-mRNA interactions within the system as a whole. This would require a combination of both transcriptomic and proteomic analysis of target mRNAs and microRNAs, in addition to a computational analysis of the targets and biological pathways.

## 6.2 – MicroRNA validation

Chapter 4 is focussed on the validation of the select few candidate novel microRNAs described in chapter 3. 16 candidates were examined and designated the temporary IDs of Novel 1 – 16. Since the initial sequencing analysis and microRNA validation portion of this PhD, the miRBase database has been updated 3 times and 6 of the 16 candidate novel microRNAs have been subsequently identified by other research groups and added to the database.

Candidate novel 2 is described as hsa-miR-6509-5p (Joyce et al., 2011), candidate novel 7 is known as hsa-miR-664b-3p (Friedlander et al., 2012), candidate novel 8 is a truncated form of hsa-miR-1277-5p (Morin et al., 2008). Candidate novel 10 is described as hsa-miR-487a-5p in humans (Fu et al., 2005) and a similar sequence can be found as oar-miR-487a-5p in sheep (Caiment et al., 2010). Candidate novel 11, although still not annotated in humans can be found in mouse and rat as mmu-miR-3085-3p and rno-miR-3085 respectively (Chiang et al., 2010; Linsen et al., 2010). Finally candidate novel 12 can be found in humans as hsa-miR-539-3p (Sewer et al., 2005), with marginally altered sequences found in mouse, rat and sheep as mmu-miR-539-3p, rno-miR-539-3p and oar-miR-539-3p respectively (Caiment et al., 2010; Landgraf et al., 2007; Sewer et al., 2005). Additionally, candidate novel 15 was present in miRBase V.19 but has since been removed and annotated as a fragment of tRNA. Although these microRNAs have now been described on miRBase, none have been experimentally validated, with identification coming only from sequencing data or bioinformatics analysis.

The literature identifying candidate novels 8, 10 and 12 as microRNAs were published before the version of miRBase (V.17) used during the initial data analysis; however those microRNAs were not identified as existing microRNAs by the bioinformatics analysis carried out for this project. This could be due to an error within the data analysis, or potentially the data from the literature referenced was not added to the miRBase database until after V.17 and hence not available for inclusion in the initial results.

Validation of the novel candidates encompassed a myriad of techniques, including northern blotting using RNA obtained from the chondrosarcoma cell line SW1353, to establish the size of these sequences and ascertain if the candidate microRNAs detected during the deep sequencing were, in fact, fragments of larger sequences and hence not

microRNAs. QRT-PCR was also used to detect the candidate microRNAs in a number of tissues and cell types, and microRNA expression was also measured in chondrogenesis and de-differentiation models as well as a Dicer knockout cell line. *In situ* hybridisation was also utilised to localise a candidate microRNA to a specific location within mouse embryo. Through this, candidate novels 2, 7 and 11 (miR-6509-5p, miR-664b-3p and miR-3085-3p respectively) were assigned microRNA status.

Due to time constraints, only miR-3085-3p was investigated with *in situ* hybridisation, and only one age of embryo was used [a developmental stage where a known cartilage selective microRNA is readily detected (Tuddenham et al., 2006)]. Further experimentation could involve all three candidate microRNAs at the same time point to verify results, but also use a range of developmental stages of embryo. *In situ* hybridisation could also be carried out on human cartilage samples (both osteoarthritic and control) as well as sections of embryo in addition to whole mount. This would allow for a more precise localisation of the microRNA in mouse embryo tissue, and identify whether a similar pattern of expression can be seen within human cartilage.

With additional time, a larger number of the 1621 candidate novel microRNAs could be investigated to establish their true microRNA identity before inclusion in miRBase. This would then provide a more accurate map of the miRNome of osteoarthritic cartilage.

### **6.3 – MicroRNA functional analysis**

Chapter 5 continued the analysis of the three validated microRNAs in the form of functional analysis. This utilised a computational search with RStudios, searching for complementary target sites of the three microRNAs in the 3'UTRs of the genes of the human genome, including all exacting matches of 6mers, 7mer8s, 7merA1s and 8mers. A genome analysis was also carried out on primary chondrocytes treated with siRNAs of the candidate microRNAs, and the results correlated with the computational analysis. A gene was considered a potential target if it contained one or more of the four target sites, if its expression decreased in response to treatment with a microRNA mimic and if expression increased in response to microRNA inhibition.

This combined analysis produced a large number of potential targets; 1782 for miR-6509-5p, 3245 for miR-664b-3p and 3091 for miR-3085-3p. Fold change cut off points addressing gene expression were introduced to triage the results down to a manageable number and a selection for each microRNA were analysed by sub-cloning and luciferase assays. The most likely target obtained from this assessment was ITGA5 (a fibronectin receptor (Nagae et al., 2012), a target for miR-3085-3p.

Further study would allow for a more in depth functional analysis of each of the three candidates such as exploring the remaining targets from the computational and genome array analysis, and utilising additional techniques such as western blotting to determine if protein expression of the targets of interest are affected when the microRNA is inhibited or enhanced. The interaction between a microRNA and its target via the Ago protein complex can also be utilised in functional analysis. Pull-down methods can be employed to detect the mRNA targets bound to these complexes alongside the microRNA (Beitzinger et al., 2007).

Additionally, a computational search for the microRNA targets could be carried out within the 5'UTR of genes plus the coding region, as research has increasingly shown that microRNAs can bind at more than just the 3'UTR (Lee et al., 2009). This would give a larger pool of targets to compare with the genome array data. It is also known, that a microRNA: mRNA binding does not have to be an exact complementary match. Further analysis could take this into account, searching for variations in the target sequence which may still function.

The expression of the microRNA candidates was measured in a mesenchymal stem cell chondrogenesis assay as part of the validation process described in chapter 4. This could

be taken further in future functional studies. Each candidate microRNA could be inhibited or overexpressed, and the cartilage disc produced as the end point of the assay could be analysed for functional or physical changes which would suggest a substantial role of the microRNA in cartilage development. In addition, the candidate microRNAs could be investigated as to whether they play a role in determining cell fate from mesenchymal stem cells, resulting in the formation of chondrocytes, osteocytes or adipocytes. It has already been shown that microRNAs can play an important role in cell fate determination. For example, Cordes et al found that both miR-145 and miR-143 are heavily involved in the differentiation of smooth muscle cells from multipotent murine cardiac progenitors (Cordes et al., 2009).

A more long term goal to address the functional qualities of the candidate microRNAs, specifically miR-3085-3p which was found to be more cartilage specific than the other two, would be to create a miR-3085-3p transgenic mouse and analyse the phenotype. Conditional cartilage and skeletal knockouts could also be carried out to determine if there is a detrimental effect to cartilage development *in vivo*.

Finally, all experiments involved either primary chondrocytes or cartilage obtained from end stage patients. Future experiments should, if possible, examine a range of osteoarthritic grades, and subtypes (of which there is yet no exact number) (Bierma-Zeinstra and Verhagen, 2011) . This could provide a pattern of microRNA expression across osteoarthritis progression.

## **6.4 – Final summary**

The hypothesis of this PhD that cartilage contains as yet unidentified microRNAs which may contribute to skeletal development or osteoarthritis, has been proven, with a significant miRNome with multiple novel microRNAs discovered in primary chondrocytes obtained from osteoarthritic cartilage samples. A number of the candidate microRNAs have been investigated and identified as true microRNAs, however there are many more left to analyse.

So, whilst the project has substantially addressed its original guidelines, there are much data left to progress this area of study forward. We already know microRNAs to be very useful in terms of understanding disease and their use (particularly circulating microRNAs) as biomarkers (Weiland et al., 2012). Over 100 microRNAs have been identified as biomarkers of disease so far, including miR-155, miR-133a, and miR-202 which have been identified as biomarkers for early pancreatic neoplasia, coronary artery disease and early stage breast cancer respectively (Fichtlscherer et al., 2010; Habbe et al., 2009; Schrauder et al., 2012; Weiland et al., 2012). MicroRNAs have also been implicated as biomarkers of arthritic diseases. For example, recently, Beyer et al found let-7e to be a probable biomarker of osteoarthritis at the knee and hip (Beyer et al., 2014).

The data from this project will therefore contribute to the growing field of microRNA biology in regards to osteoarthritis, and will hopefully aid in the battle against this increasingly common disease.



# Appendices

## Appendix 1- qRT-PCR primers and probe sequences

Gene Symbol	Gene name	Primer/probe	Sequence
18S	18S Ribosomal RNA	Forward primer: Reverse primer: Probe:	5'-GCCGCTAGAGGTGAAATTCTTG-3' 5'-CATTCTTGGCAAATGCTTTCG-3' 5'-FAM-ACCGGCGCAAGACGGA-TAMRA-3'
ACAN	Aggrecan	Forward primer: Reverse primer: Probe:	5'-CCGCTACGACGCCATCTG-3' 5'-CCCCACTCCAAAGAAGTTTT-3' 5'-FAM-TACACAGGTGAAGACTTTGTGGACATCCCA-TAMRA-3'
CA5B	Carbonic anhydrase VB	Forward primer: Reverse primer: Probe:	5'-ACACCGTGGACAGCAAATG-3' 5'-CAAATCTGACTGCGTTCCAA-3' Probe #3 (Roche Universal probe library)
CILP	Cartilage intermediate layer protein	Forward primer: Reverse primer: Probe:	5'-ATATAAGGCCGGCACTGGA-3' 5'-GGTCCCTGGGAAGTTTCT-3' Probe #15 (Roche Universal probe library)
COL1A1	Type I collagen	Forward primer: Reverse primer: Probe:	5'-GGGATTCCCTGGACCTAAAG-3' 5'-GGAACACCTCGCTCTCCA-3' Probe #67 (Roche Universal probe library)
COL2A1	Type II collagen	Forward primer: Reverse primer: Probe:	5'-GTGAACCTGGTGTCTCTGGTC-3' 5'-TTCCAGGTTTTCCAGCTTC-3' Probe #9 (Roche Universal probe library)
COL3A1	Type III collagen	Forward primer: Reverse primer: Probe:	5'-CTGGACCCCAGGGTCTTC-3' 5'-CATCTGATCCAGGGTTTCCA-3' Probe #20 (Roche Universal probe library)
COL9A1	Type IX collagen	Forward primer: Reverse primer: Probe:	5'-AGATAATCCTCAAGTTTCTGTCCA-3' 5'-GCAGCTCATGGCAAGTTTCT-3' Probe #89 (Roche Universal probe library)

## Appendix 1 - qRT-PCR primers and probe sequences

Gene Symbol	Gene name	Primer/probe	Sequence
COL10A1	Type X collagen	Forward primer: Reverse primer: Probe:	5'-CACCTTCTGCACTGCTCATC-3' 5'-GGCAGCATATTCTCAGATGGA-3' Probe #6 (Roche Universal probe library)
COL11A1	Type XI collagen	Forward primer: Reverse primer: Probe:	5'-TGCTCAAGCTCAGGAACCTC-3' 5'-CCTCTGTTACACTTTCAGCCTCTT-3' Probe #38 (Roche Universal probe library)
COL27A1	Type XXVII collagen	Forward primer: Reverse primer: Probe:	5'-GGTCTATCCAATTGCAACAA-3' 5'-AGCCGGTCTGGATAGCTGTA-3' Probe #43 (Roche Universal probe library)
CRTAC1	Cartilage acidic protein 1	Forward primer: Reverse primer: Probe:	5'-GGAGTGTGGCCAAGGATTC-3' 5'-GATGCATTATTGGTGTCCA-3' Probe #18 (Roche Universal probe library)
CSNK1E	Casein kinase 1 epsilon	Forward primer: Reverse primer: Probe:	5'-TCCATCAACACGCACCTG-3' 5'-GGAGCCCAGGTTGAAGTACA-3' Probe #80 (Roche Universal probe library)
DICER1	Dicer1	Forward primer: Reverse primer: Probe:	5'-TCCGATGGTTCTCGAAGG-3' 5'-GCAAAGCAGGGCTTTTCAT-3' Probe #47 (Roche Universal probe library)
GPM6B	Glycoprotein M6B	Forward primer: Reverse primer: Probe:	5'-AGTACCATCCCGTGCCAAC-3' 5'-CTCCCAGACACTTGATGCAG-3' Probe #64 (Roche Universal probe library)
ITGA5	Integrin alpha 5	Forward primer: Reverse primer: Probe:	5'-CCCATTGAATTTGACAGCAA-3' 5'-TGCAAGGACTTGTACTCCACA-3' Probe #55 (Roche Universal probe library)

## Appendix 1 - qRT-PCR primers and probe sequences

Gene Symbol	Gene name	Primer/probe	Sequence
Hsa-miR-140-5p	Hsa-miR-140	Forward primer: Reverse primer: Probe:	Sequence not known Sequence not known Sequence not known
Hsa-miR-455-5p	Hsa-miR-455	Forward primer: Reverse primer: Probe:	Sequence not known Sequence not known Sequence not known
MMP1	Matrix metalloproteinase-1	Forward primer: Reverse primer: Probe:	5'-AAGATGAAAGGTGGACCAACAATT-3' 5'-CCAAGAGAATGGCCGAGTTC-3' 5'-FAM- CAGAGAGTACAACCTACATCGTGTTCGGGCTC- TAMRA-3'
MMP13	Matrix metalloproteinase-2	Forward primer: Reverse primer: Probe:	5'-AAATTATGGAGGAGATGCCATT-3' 5'-TCCTTGGAGTGGTCAAGACCTAA-3' 5'-FAM- CTACAACCTGTTTCTTGTGCTGCGCATGA- TAMRA-3'
MYD88	Myeloid differentiation primary response gene 88	Forward primer: Reverse primer: Probe:	5'-GGACCCAGCATTGAGGAG-3' 5'-ACAGCGGCCACCTGTAAA-3' Probe #18 (Roche Universal probe library)
PDGFD	Platelet derived growth factor D	Forward primer: Reverse primer: Probe:	5'-CCATGACCGGAAGTCAAAAG-3' 5'-ATCCTGGGAGTGCAACTGT Probe #85 (Roche Universal probe library)
RBM3	RNA binding motif protein 3	Forward primer: Reverse primer: Probe:	5'-GAGAGCCATGAACGGAGAGT-3' 5'-CCTCTAGAGTAGCTGCGACCA-3' Probe #10 (Roche Universal probe library)
MYD88	Myeloid differentiation primary response gene 88	Forward primer: Reverse primer: Probe:	5'-GGACCCAGCATTGAGGAG-3' 5'-ACAGCGGCCACCTGTAAA-3' Probe #18 (Roche Universal probe library)

## Appendix 1 - qRT-PCR primers and probe sequences

Gene Symbol	Gene name	Primer/probe	Sequence
PDGFD	Platelet derived growth factor D	Forward primer: Reverse primer: Probe:	5'-CCATGACCGGAAGTCAAAAG-3' 5'-ATTCCTGGGAGTGCAACTGT-3' Probe #85 (Roche Universal probe library)
RBM3	RNA binding motif protein 3	Forward primer: Reverse primer: Probe:	5'-GAGAGCCATGAACGGAGAGT-3' 5'-CCTCTAGAGTAGCTGCGACCA-3' Probe #10 (Roche Universal probe library)
SFRP4	Secreted frizzled-related protein 4	Forward primer: Reverse primer: Probe:	5'-GCCTGAAGCCATCGTCAC-3' 5'-CCATCATGTCTGGTGTGATGT-3' Probe #88 (Roche Universal probe library)
SNORA36A	Small nucleolar RNA, H/ACA box 36A	Forward primer: Reverse primer: Probe:	Sequence not known Sequence not known Sequence not known
TGFβ3	Transforming growth factor, beta receptor III	Forward primer: Reverse primer: Probe:	5'-GATTCATCTTCGGCTGAAA-3' 5'-GCTCAGGAGGAATAGTGTGGA-3' Probe #82 (Roche Universal probe library)
TIMP3	TIMP metalloproteinase inhibitor 3	Forward primer: Reverse primer: Probe:	5'-GTGCAACTTCGTGGAGAGGT-3' 5'-AGCAGGACTTGATCTTGCACT-3' Probe #14 (Roche Universal probe library)
VCAN	Versican	Forward primer: Reverse primer: Probe:	5'-GCACCTGTGTGCCAGGATA-3' 5'-CAGGGATTAGAGTGACATTCATCA-3' Probe #54 (Roche Universal probe library)
WDR91	WD repeat-containing 91	Forward primer: Reverse primer: Probe:	5'-CTGGGGGACTCGGAACTT-3' 5'-AGGAAGCCCACACGAGGT-3' Probe #14 (Roche Universal probe library)

## Appendix 2 - Oligonucleotide sequences of northern blot probes for candidate novel and known microRNAs

MicroRNA ID	Probe sequence	Positive control sequence
Hsa-miR-140-3p	5'-CCGTGGTTCTACCCTGTGGTA-3'	5'-TACCACAGGGTAGAACCACGG-3'
Hsa-miR-140-5p	5'-CTACCATAGGGTAAAACCACTG-3'	5'-CAGTGGTTTTACCCTATGGTAG-3'
Novel 1	5'-GCTCTACACGTTTCAGAGAACTT-3'	5'-AAGTTTCTCTGAACGTGTAGAGC-3'
Novel 2	5'-GTTCCACTGCCACTACCTAAT-3'	5'-ATTAGGTAGTGGCAGTGGAAC-3'
Novel 3	5'-ACCAGGGACCTTTAGATCTTCAG-3'	5'-CTGAAGATCTAAAGGTCCCTGGT-3'
Novel 4	5'-CCACTGAGCTACACCCC-3'	5'-GGGGGTGTAGCTCAGTGG-3'
Novel 5	5'-ACTCGGATCGCTGGATTCA-3'	5'-TGAATCCAGCGATCCGAGT-3'
Novel 6	5'-TCCTCTAGATAGTCAAGTTCGA-3'	5'-TCGAACTTGACTATCTAGAGGA-3'
Novel 7	5'-TGTAGGCTGGGAGGCAAATGAA-3'	5'-TTCATTTGCCTCCCAGCCTACA-3'
Novel 8	5'-ATACGTACATATATATATATA-3'	5'-TATATATATATATGTACGTAT-3'
Novel 9	5'-CAGATCGCAACTCCAGGCATC-3'	5'-GATGCCTGGGAGTTGCGATCTG-3'
Novel 10	5'-CGAACACAGCAGGGATAACCAC-3'	5'-GTGGTTATCCCTGCTGTGTTTCG-3'
Novel 11	5'-GAGGGGGCCATAGCAGCCAGA-3'	5'-TCTGGCTGCTATGGCCCCCTC-3'
Novel 12	5'-AAAGAAATTGTCCTTGTATGAT-3'	5'-ATCATACAAGGACAATTTCTTT-3'
Novel 13	5'-GCCAGATAGTCTTCAGTCCAGCT-3'	5'-AGCTGGACTGAAGACTATCTGGC-3'
Novel 14	5'-TCCGACACCGAGAGGGCAGGCG-3'	5'-CGCCTGCCCTCTCGGTGTCGGA-3'
Novel 15	5'-GGGATTCGAACCCAGGATCTCC-3'	5'-GGAGATCCTGGGTTTGAATCCC-3'
Novel 16	5'-TCAACCAACACCTTTTCTGGGGTC-3'	5'-GACCCCAGAAAAGGTGTTGGTTGA-3'

**Appendix 3 – Primer sequences for subcloning. DNA sequences of forward and reverse primers used for PCR analysis.**

Gene symbol	Novel micro-RNA	Primer	Sequence
SFRP4	2	Forward primer: Reverse primer:	5'-GTACGTCGACGACTTCCGACTTCCTTAC-3' 5'-GTACGTCGACGTCAGTGAAGAAACCACC-3'
TGFβR 3	2	Forward primer: Reverse primer:	5'-GTACGTCGACCAGAAGGGTATCAGAGTGG-3' 5'-GTACGTCGACCAGACAATCTGCCTTGGAG -3'
RBM3	2	Forward primer: Reverse primer:	5'-GTACGTCGACGTGGTGTACGCCTGTAATC-3' 5'-GTACGTCGACCACAAAGATGAGTTCAGTGC-3'
CILP	7	Forward primer: Reverse primer:	5'-GTACGTCGACGTGGTACTTCACCCTCTTC-3' 5'-GTACGTCGACCCAGGAGATGTTCTTTCC-3'
TIMP3 - 1	7	Forward primer: Reverse primer:	5'-GTACGTCGACGCCATGCCAGAAAGAATG-3' 5'-GTACGTCGACCTCTCTTGTAGTTCCTTC-3'
TIMP3 - 2	7	Forward primer: Reverse primer:	5'-GTACGTCGACGTAGCTGAGACACTTCTG-3' 5'-GTACGTCGACCATCTTGGGGTAGTCCAG-3'
VCAN	7	Forward primer: Reverse primer:	5'-GTACGTCGACCAGCCAAAGTCCTAACTTC-3' 5'-GTACGTCGACGTCATCTTCAGCAGTCAC-3'
ITGA5	11	Forward primer: Reverse primer:	5'-GTACGTCGACCAGACTCCCATTCTGAAG-3' 5'-GTACGTCGACCTAGTTCTGGTCAGTGGG-3'
MYD88	11	Forward primer: Reverse primer:	5'-GTACGTCGACCTGCCTCCTCTTTTCGTTG-3' 5'-GTACGTCGACGCATCCATTCCCTCCCAAG-3'
TGFβR 3	11	Forward primer: Reverse primer:	5'-GTACGTCGACCAAAGTGCAACATAGGGAAG-3' 5'-GTACGTCGACGCTTAGAAAGCCTCAAAGC-3'

## Appendix 4 – Selected candidate novel microRNAs identified from next generation sequencing

Novel ID	Sequence	Size	Average reads per 10 million	Complementary strand detected	Presence in DLD	Down regulation in Dicer knock out
1	AAGTTTCTCTGAACGTGTAGAGC	23	48.4	Yes	Yes	Yes
2	ATTAGGTAGTGGCAGTGGAAC	21	0.5	Yes	Yes	Yes
3	CTGAAGATCTAAAGGTCCCTGGT	23	328.9	No	Yes	No
4	GGGGGTGTAGCTCAGTGG	18	180.0	No	No	No
5	TGAATCCAGCGATCCGAGT	19	83.5	No	No	No
6	TCGAACTTGACTATCTAGAGGA	22	57.1	Yes	Yes	No
7	TTCATTTGCCTCCCAGCCTACA	22	14.5	Yes	Yes	No
8	TATATATATATATGTACGTAT	21	44.4	Yes	No	No
9	GATGCCTGGGAGTTGCGATCTG	22	18.5	Yes	No	No
10	GTGGTTATCCCTGCTGTGTTTCG	22	18.3	Yes	No	No
11	TCTGGCTGCTATGGCCCCCTC	21	11.9	Yes	No	No
12	ATCATACAAGGACAATTTCTTT	22	6.3	Yes	No	No
13	AGCTGGACTGAAGACTATCTGGC	23	2.3	Yes	No	No
14	CGCCTGCCCTCTCGGTGTCGGA	22	0.5	No	Yes	Yes
15	GGAGATCCTGGGTTCGAATCCC	22	16.9	No	No	No
16	GACCCAGAAAAGGTGTTGGTTGA	24	4.6	No	No	No
	GGATTTTTGGAGCAGGGAGAT	21	58.6	No	No	No
	GGGGGCGGCGGCGGGGGG	18	28.2	No	Yes	No
	GGCGCGCGCGCGCGCGCGT	19	23.9	No	Yes	No
	GGAGATCCTGGGTTCGAATCC	21	22.7	No	No	No
	GGCGCGCGGCGGCGGCGGCGGC	22	15.1	No	No	No
	GGGATGTATTCGTA CTGTCTGATG	24	13.3	No	No	No
	GGCGGAGGGGCCGCGGGCCGG	21	11.0	No	Yes	No
	GGCGCGGCGGCGGCGGCGGC	19	6.4	No	No	No
	GGCGCGGCGGCGGCGGCGCG	20	5.2	No	No	No
	ATTCGTGTGGACATTTTGGAGA	22	4.0	No	Yes	No
	GAGAGAGAGAGATCGGAA	18	4.0	No	No	No
	CTCCACTGCTTCACTTGACTAGT	24	3.5	No	Yes	No
	AGCTTTCGCGAGTGGCAG	18	3.3	No	No	No
	GCGCGCGGCGGCGGCGGCGGC	21	3.3	No	Yes	No
	GGAGGGCGGCGGCGGCGGCGG	21	2.8	No	No	No
	TCAGCCGTGTCTGTTCTCTGG	21	2.7	No	No	No
	CCTCACCACCCCTTCTGCCTGC	22	2.5	No	No	No
	TACAGATGCAGATTCTCTGACT	22	2.1	No	No	No
	AAAAGTTATTGTGGTTTTTGCT	22	2.1	No	Yes	No
	CCCCGCCCGCGCCCTCT	18	2.1	No	No	No
	GAGAGAGAGAGAGATCGGAA	20	2.1	No	Yes	No
	GGGGGGCTGATTCTGCTCAGAT	22	2.0	No	No	No
	TCACCACCCCTTCTGCCTGCA	21	1.9	No	No	No



## Appendix 4 – Selected candidate novel microRNAs identified from next generation sequencing

Novel ID	Sequence	Size	Average reads per 10 million	Complementary strand detected	Presence in DLD	Down regulation in Dicer knock out
	TTATGCCTGTGACTGCTGTAG	21	1.6	No	No	No
	CAGGGCCTGGAACCTCAATGGGA	23	1.5	No	No	No
	ATTTGCAGTAACAGGTGTGAGC	22	1.4	No	No	No
	TGAGGCCGAGAAGGCAACCGCGA	23	1.4	No	Yes	No
	TCAGGACACTTCTGAACTTGGGA	22	1.2	No	Yes	No
	TGAGGCTGTAAGTGAAGAAAAG	22	1.2	No	No	No
	AGCAATGTGTGATTAGGAAAAG	22	1.2	No	Yes	No
	GGGCGGCGGGCGGGCGGTGGCGG	24	1.0	No	No	No
	GAAAGTAATGGTGGGTTTTGCC	22	1.0	No	Yes	No
	GGGGGGGCTTCTCCAGGA	19	1.0	No	No	No
	GGGAGCGGGGGGGCGGGCGCG	21	0.8	No	No	No
	GGGGGCGGGCGCCGGGGGGC	19	0.8	No	No	No
	GGTGGTGGCGGTGGCGGGCGGG	24	0.8	No	No	No
	AGAGTCGAGAGTGGGAGAAGAG	22	0.7	Yes	No	No
	GGCGGTGGTGGCGGTGGCGGC	21	0.7	No	No	No
	TTGCCTGCCCTTCTCCAG	21	0.7	No	No	No
	TCTCGTCCGGCCCTGCCCGCT	22	0.6	No	No	No
	TCCGTCGTTCTCAGGAGGTGG	22	0.6	No	Yes	No
	CAGCAATTACTTTGCACCACT	22	0.5	No	No	No
	TTGAGGGGAGAATGAGGTGGAGA	23	0.5	No	No	No
	TCATTCTGTTCTGTGGCCTGT	23	0.5	No	Yes	No

**Table A4.1 - Candidate novel microRNAs identified from next generation sequencing in all three patients.** 61 candidate novel microRNAs were found in all three patients out of the initial 1621 sequences identified. The complementary strands of 11 of the 60 candidates were detected, 10 of which are found within the triaged 16 candidate novels. 20 sequences were also detected in a sequencing project examining cells of the DLD cell line. Of those 20, 3 appeared down-regulated in a sequencing project of a DICER knockout DLD cell line. The triaged 16 contained 6 candidate sequences found in DLD cells, including the 3 candidates down-regulated in response to a DICER down-regulation. Also included in the chosen 16 candidates were the 3 candidate novels which were the most highly expressed out of the 60 sequences: candidate novel 3, 4 and 5.

## Appendix 5 - Read number of most highly expressed 100 known microRNAs

MicroRNA	Reads per 10 million reads			
	Patient 1	Patient 2	Patient 3	Average
hsa-mir-140-3p	4325532	3958193	3762965	4015563
hsa-mir-10b	166734	361833	457263	328610
hsa-mir-101	142849	311361	207423	220545
hsa-mir-140-5p	230886	247074	179621	219194
hsa-mir-26a	182911	164578	94932	147474
hsa-let-7b	67707	131882	100569	100053
hsa-mir-320a	86270	87063	82286	85206
hsa-mir-23b	58259	107486	85430	83725
hsa-let-7a	58306	96245	62147	72233
hsa-mir-23a	45619	84229	78874	69574
hsa-mir-27b	49403	79242	57848	62164
hsa-mir-125b	58044	62879	64987	61970
hsa-mir-27a	47965	67599	50943	55502
hsa-mir-100	24277	50548	78862	51229
hsa-mir-21	38211	53157	53220	48196
hsa-mir-24	40517	48748	37715	42327
hsa-let-7c	34793	59699	31642	42045
hsa-let-7f	29436	52644	31836	37972
hsa-mir-30d	39840	35425	28376	34547
hsa-mir-99b	24049	32075	44565	33563
hsa-mir-99a	23137	45767	30967	33290
hsa-mir-320b	31751	34899	33120	33256
hsa-mir-29b	26830	33086	37982	32633
hsa-mir-30e*	26966	35790	29143	30633
hsa-mir-148a	21373	34813	24745	26977
hsa-mir-30a*	23331	29949	25798	26359
hsa-mir-26b	27819	30071	16929	24940
hsa-mir-25	18824	27408	26089	24107
hsa-mir-210	23653	23873	23280	23602
hsa-mir-374a	22407	30470	16745	23208
hsa-mir-16	23150	22767	12525	19481
hsa-let-7g	15091	24578	15963	18544
hsa-mir-152	14799	22524	18061	18461
hsa-mir-92a	17882	15808	20769	18153
hsa-mir-3074-5p	16660	20482	16703	17949
hsa-let-7i	15521	19965	13244	16243
hsa-mir-125a-5p	16673	11126	13974	13924
hsa-mir-22	14600	13361	10632	12864
hsa-mir-19b	8960	15632	12615	12402
hsa-mir-379	2806	16592	12003	10467

## Appendix 5 - Read number of most highly expressed 100 known microRNAs

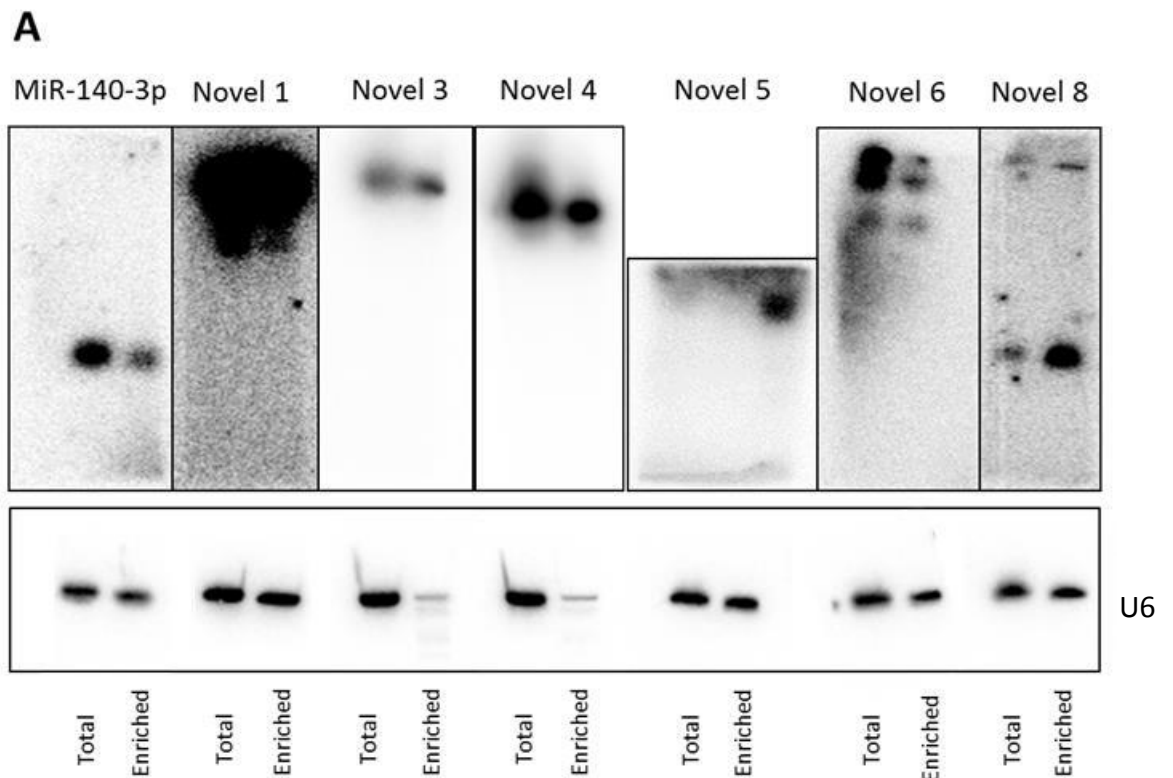
MicroRNA	Reads per 10 million reads			
	Patient 1	Patient 2	Patient 3	Average
hsa-mir-195	14893	9995	6218	10369
hsa-mir-29a	9303	11650	9516	10156
hsa-mir-155	6192	9589	11278	9020
hsa-mir-199a-3p	5504	11024	9640	8723
hsa-mir-199b-3p	5504	11024	9640	8723
hsa-mir-221	5352	10621	8910	8294
hsa-mir-146b-5p	4084	13371	5549	7668
hsa-mir-199a-5p	5408	8999	8189	7532
hsa-mir-30b	3938	9466	8688	7364
hsa-mir-222	5122	9203	7454	7259
hsa-mir-196a	4407	10558	5694	6886
hsa-mir-29c	5594	8104	6921	6873
hsa-mir-30a	6090	8592	5690	6791
hsa-mir-320c	5055	6780	6715	6183
hsa-mir-191	5287	7341	5769	6132
hsa-mir-93	4555	6328	6664	5849
hsa-mir-15a	5924	6662	3942	5509
hsa-mir-30c	4139	6520	5444	5368
hsa-mir-103a	4881	6024	4403	5103
hsa-mir-103b	4881	6024	4403	5103
hsa-mir-107	4448	5339	3941	4576
hsa-mir-196b	4572	5395	3744	4570
hsa-mir-30e	3909	5721	4001	4543
hsa-mir-106b	3170	5127	4598	4298
hsa-mir-92b	4143	3493	3981	3872
hsa-mir-28-3p	2804	4183	4575	3854
hsa-mir-365	4319	3619	3231	3723
hsa-mir-19a	2616	4348	3380	3448
hsa-mir-199b-5p	1317	4624	2861	2934
hsa-mir-455-3p	3652	3183	1906	2914
hsa-mir-455-5p	1682	3871	2489	2681
hsa-let-7e	3083	2891	2051	2675
hsa-mir-320d	1866	2957	3123	2649
hsa-mir-574-3p	2787	2205	2590	2527
hsa-mir-574-5p	1244	3069	3176	2496
hsa-mir-193b	1972	3095	2128	2398
hsa-mir-130a	1890	2973	2331	2398
hsa-mir-497	2009	2640	2181	2277
hsa-let-7d	2335	2256	1875	2155
hsa-mir-186	2048	2552	1688	2096

## Appendix 5 - Read number of most highly expressed 100 known microRNAs

MicroRNA	Reads per 10 million reads			
	Patient 1	Patient 2	Patient 3	Average
hsa-mir-532-5p	1368	2418	1831	1872
hsa-mir-143	1513	1613	2367	1831
hsa-mir-20a	1790	2027	1638	1818
hsa-mir-374b	1510	2216	1506	1744
hsa-mir-374c	1499	2215	1503	1739
hsa-mir-151-3p	1336	2134	1399	1623
hsa-mir-425	1043	2088	1670	1600
hsa-mir-151-5p	1710	1786	1120	1538
hsa-mir-664	1514	1611	1482	1535
hsa-mir-214	2002	1314	1237	1518
hsa-mir-193a-3p	1006	1680	1755	1480
hsa-mir-374b*	1176	1691	1243	1370
hsa-mir-342-3p	1395	1180	1207	1261
hsa-mir-34a	1505	937	1242	1228
hsa-mir-423-3p	969	1570	1008	1182
hsa-mir-361-5p	1219	1223	1026	1156
hsa-mir-98	1040	1383	884	1102
hsa-mir-17	1228	1141	911	1093
hsa-mir-106a	1225	1141	911	1092
hsa-mir-378	644	1240	1367	1083

**Table A5.1 - Read number of most highly expressed 100 known microRNAs** shown in read number per 10 million reads were obtained by deep sequencing (analysed with miRCat and miRBase software) of cDNA transcribed from RNA obtained from chondrocytes of 3 patients isolated immediately after digestion.

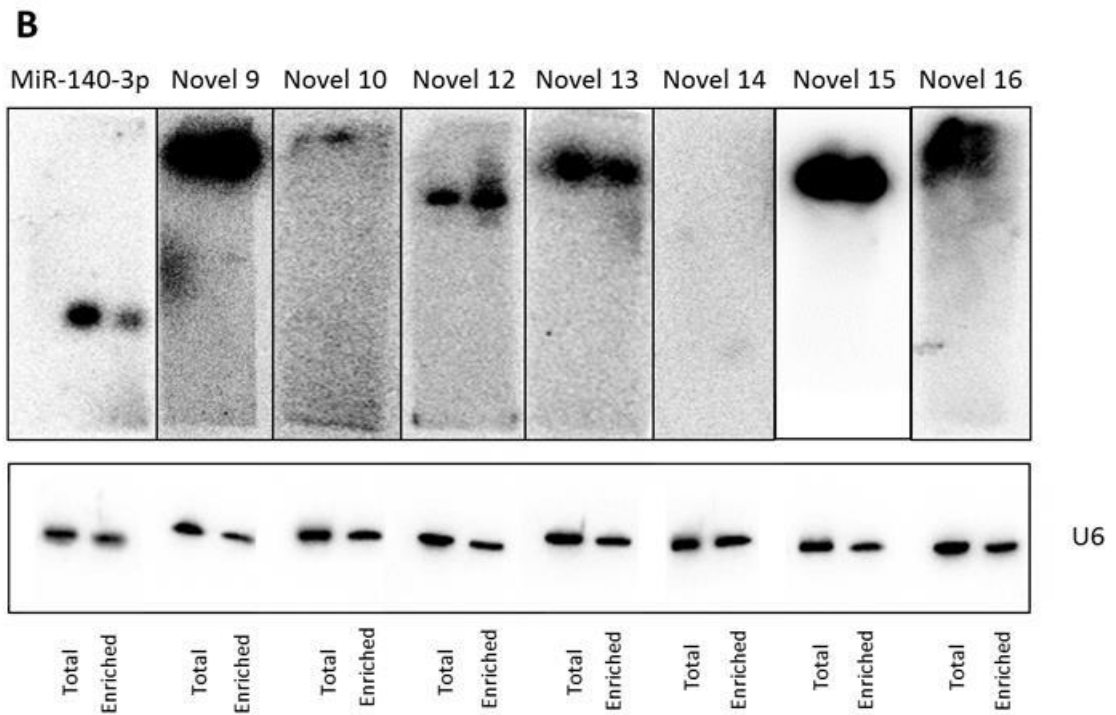
## Appendix 6 - Assessment of candidate novel microRNA using northern blots



**Figure A6.1 - Assessment of candidate novel microRNA using northern blots.**

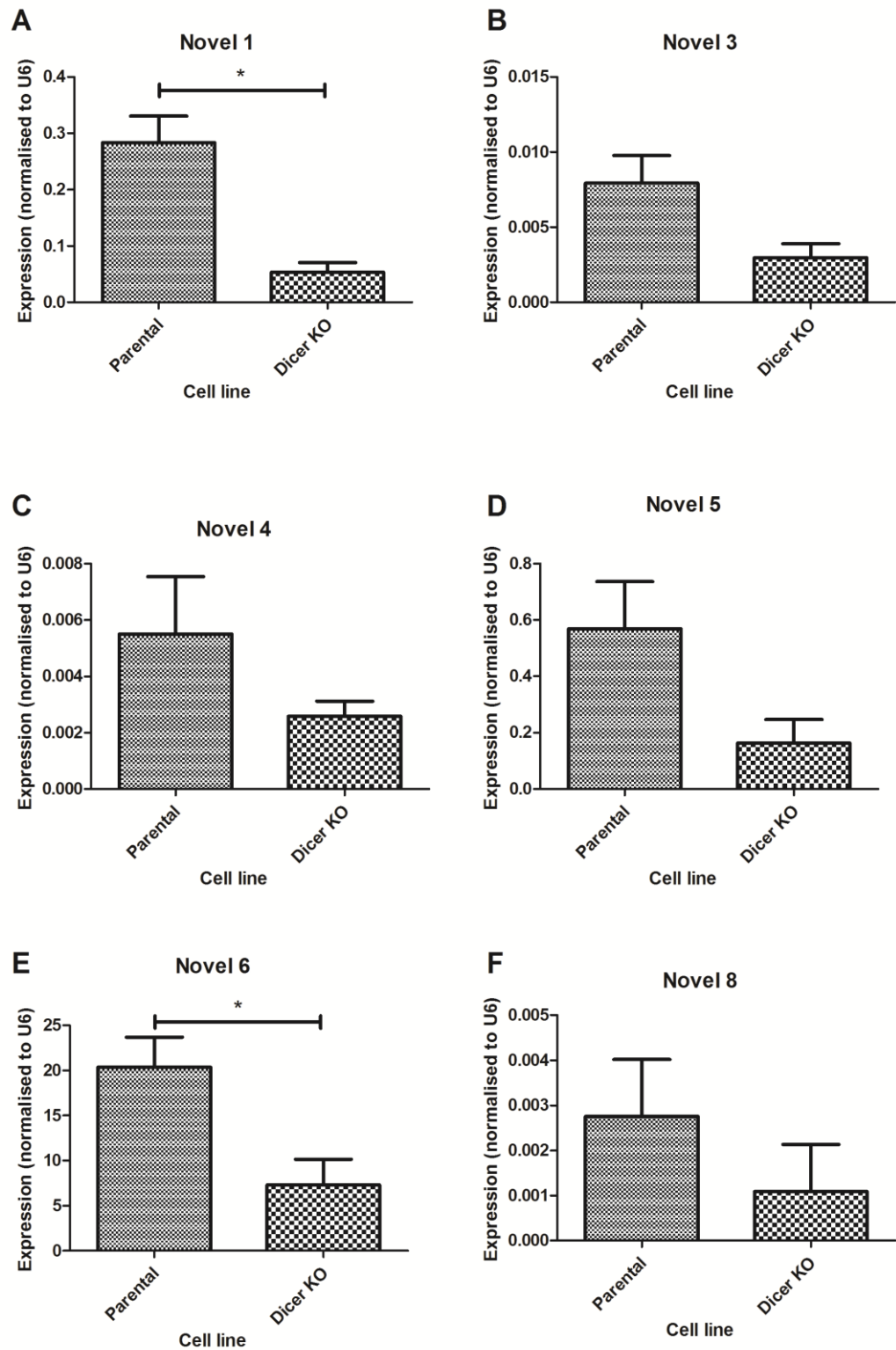
Expression of candidate novel microRNAs (A) novel 1, novel 3, novel 4, novel 5, novel 6, novel 8 was measured by northern blot using 10 $\mu$ g SW1353 total RNA or 2 $\mu$ g of RNA enriched for small RNAs. MiR-140-3p was used as a size control and the small RNA *U6* was used as a standard reference for the northern blot.

## Appendix 6 - Assessment of candidate novel microRNA using northern blots



**Figure A6.2 - Assessment of candidate novel microRNA using northern blots.** Expression of candidate novel microRNAs (B) novel 9, novel 10, novel 12, novel 13, novel 14, novel 15 and novel 16 was measured by northern blot using 10 $\mu$ g SW1353 total RNA or 2 $\mu$ g of RNA enriched for small RNAs. MiR-140-3p was used as a size control and the small RNA *U6* was used as a standard reference for the northern blot.

## Appendix 7 - Candidate microRNA response to lack of Dicer

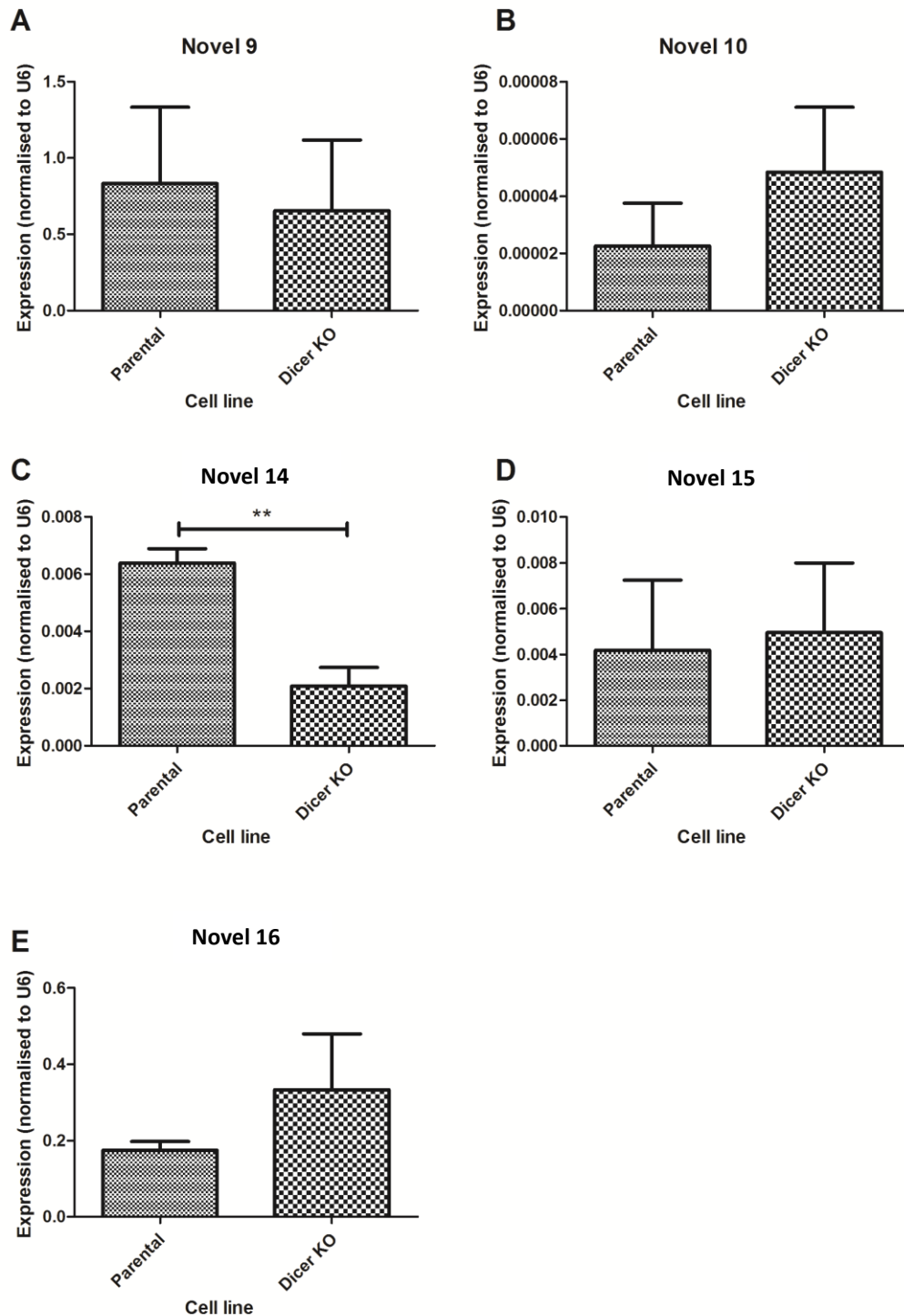


**Figure A7.1 - Candidate novel microRNA expression in DLD-1 parental and DLD-1 Dicer null cell lines.** Expression of candidate microRNA (A) novel 1, (B) novel 3, (C) novel 4, (D) novel 5, (E) novel 6 and (F) novel 8 were obtained by qRT-PCR from RNA

purified from DLD-1 wild type (parental) and DLD-1 Dicer null cell lines. Data were normalised to *U6* RNA expression. Data shows mean +/- SEM, n = 6. (\*,  $p \leq 0.05$  analyzed by a Student T test).



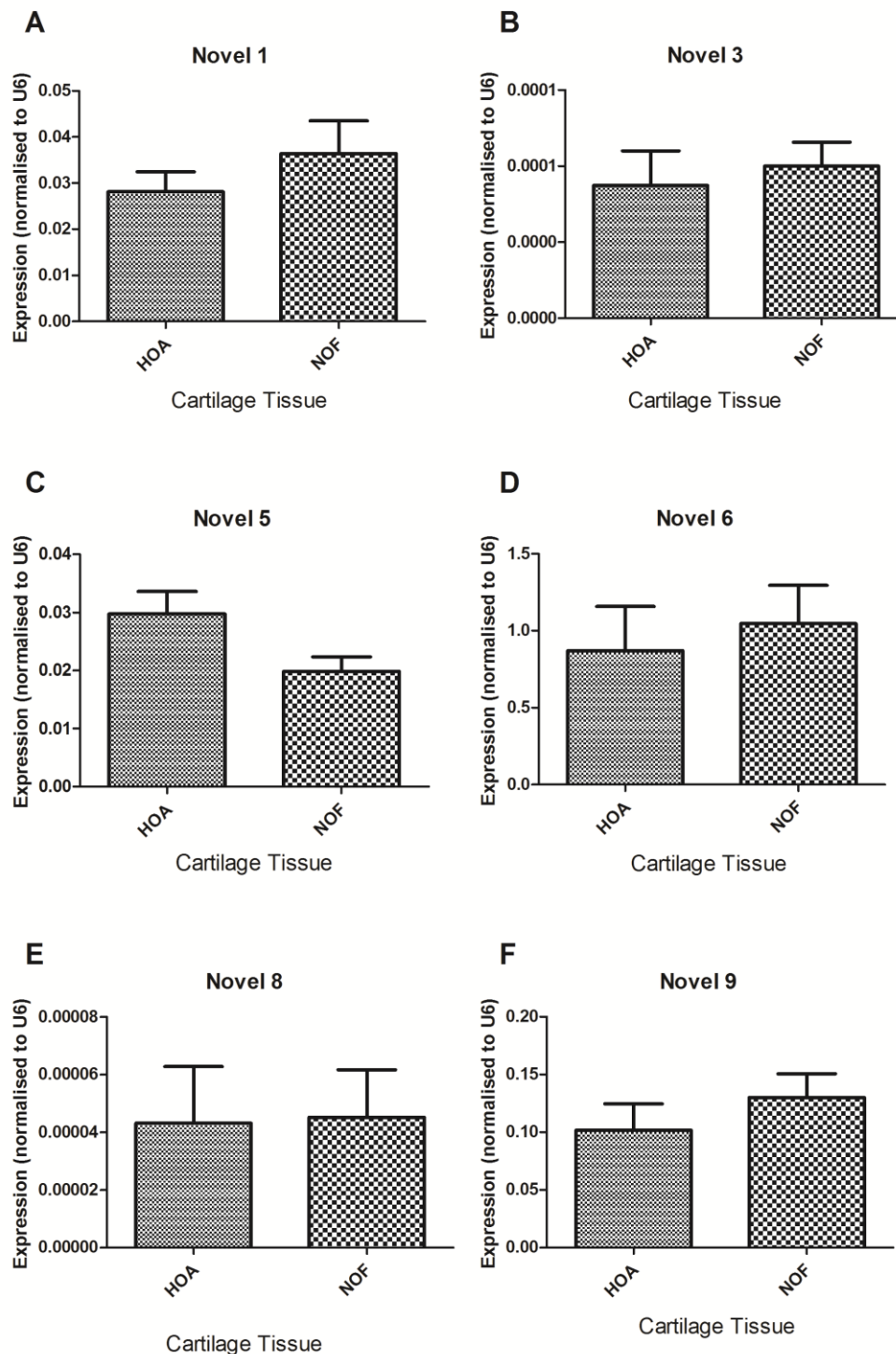
## Appendix 7 - Candidate microRNA response to lack of Dicer



**Figure A7.2 - Candidate novel microRNA expression in DLD-1 parental and DLD-1 Dicer null cell lines.** Expression of candidate microRNA (A) novel 9, (B) novel 10, (C) novel 14, (D) novel 15 and (E) novel 16 were obtained by qRT-PCR from RNA purified from DLD-1 wild type (parental) and DLD-1 Dicer null cell lines. Candidate

novels 12 and 13 were not detected in the DLD cell lines. Data were normalised to *U6* RNA expression. Data shows mean  $\pm$  SEM, n = 6. (\*\*,  $p \leq 0.005$  analyzed by a Student T test).

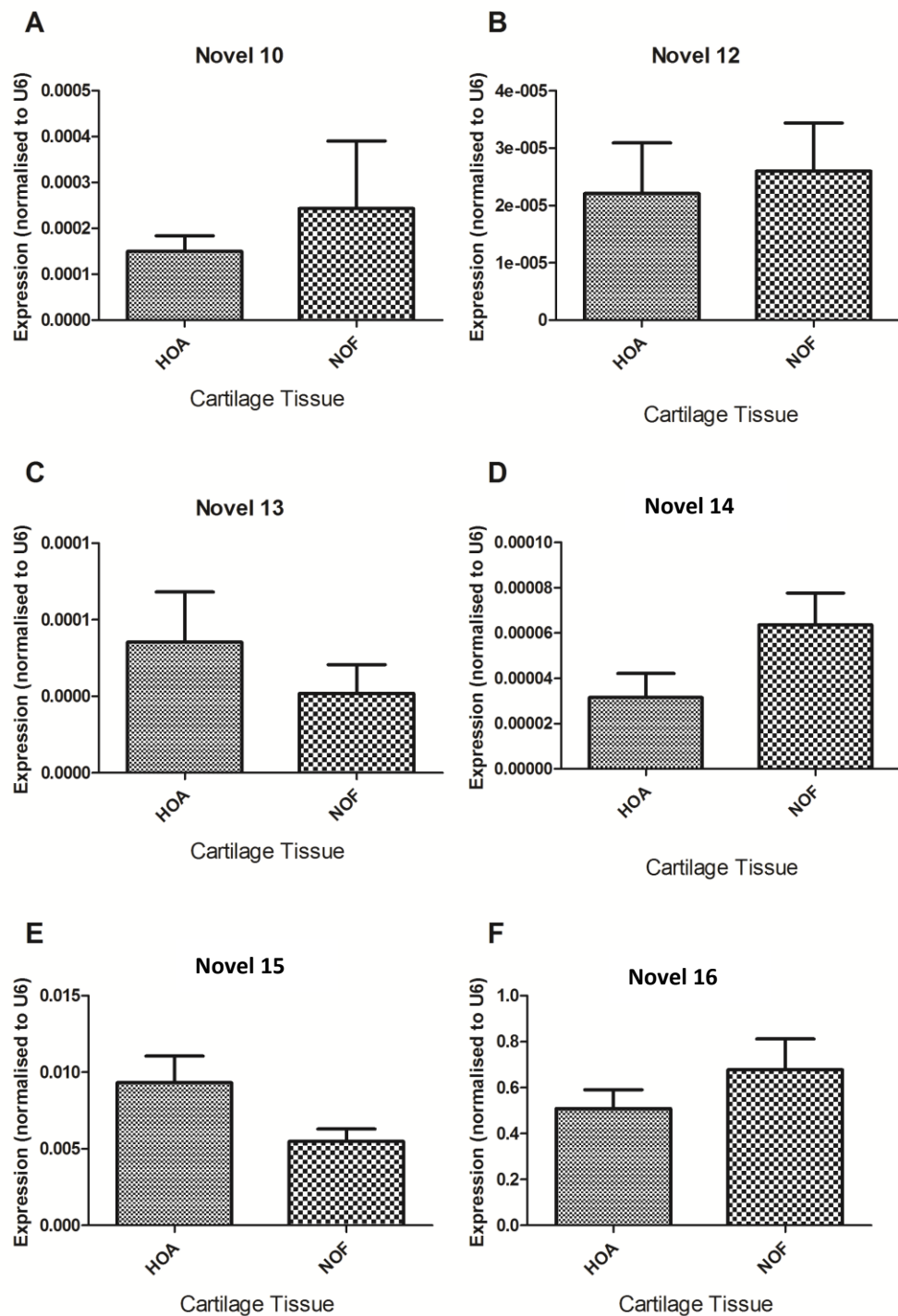
## Appendix 8 – Candidate novel microRNA expression in osteoarthritic cartilage



**Figure A8.1 - Candidate novel microRNA expression in HOA and NOF cartilage tissue.** Expression levels of candidate microRNAs (A) novel 1, (B) novel 3 and (C) novel 5, (D) novel 6, (E) novel 8 and (F) novel 9 were obtained by qRT-PCR from RNA isolated from: osteoarthritic hip cartilage tissue (HOA) and neck of femur fracture

cartilage (NOF). Candidate novel 4 was undetectable. Data were normalised to *U6* RNA expression. Data shows mean  $\pm$  SEM, n = 6 (analyzed by a Students T test, no significance was seen).

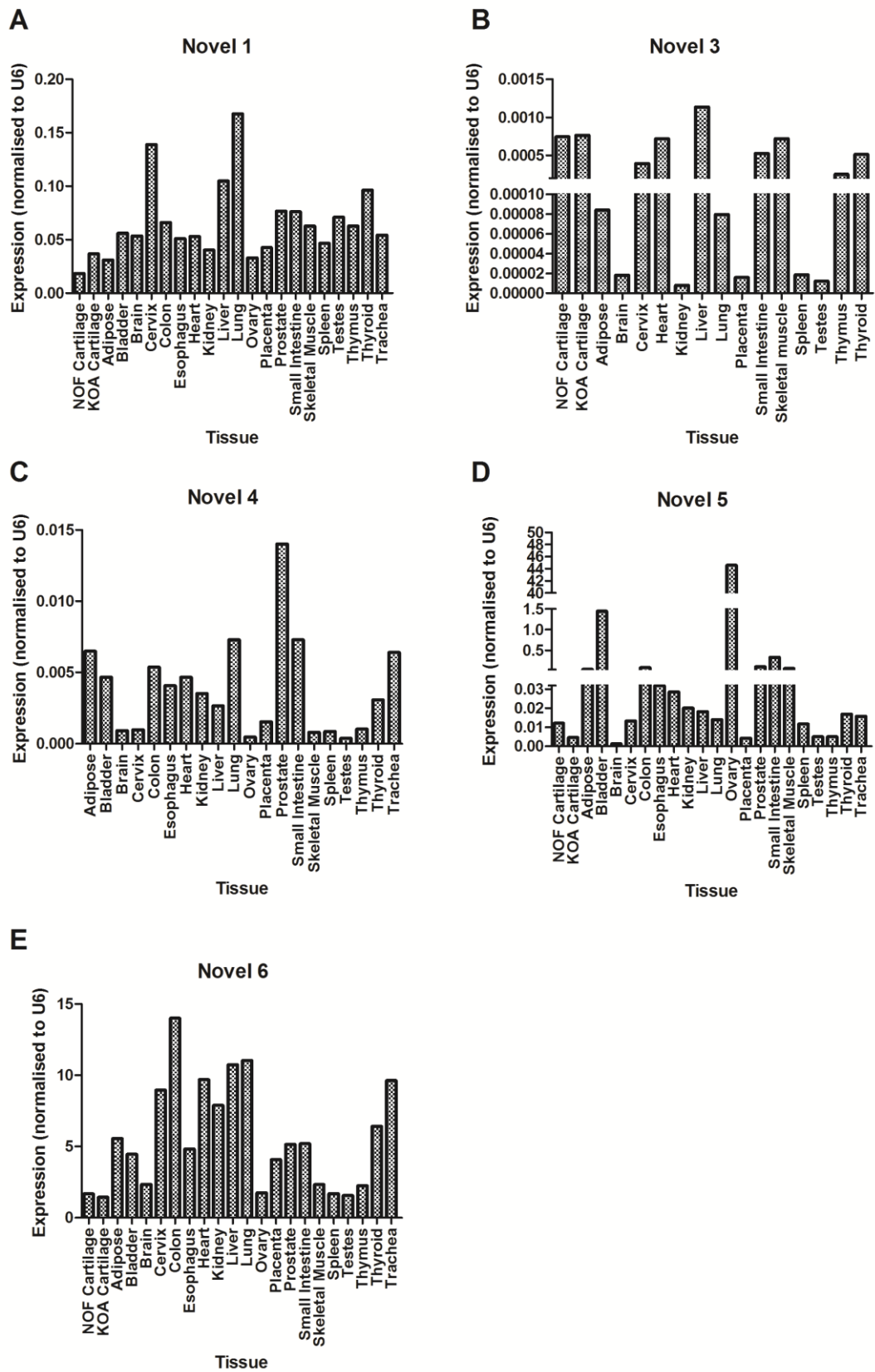
## Appendix 8 – Candidate novel microRNA expression in osteoarthritic cartilage



**Figure A8.2 - Candidate novel microRNA expression in HOA and NOF cartilage tissue.** Expression levels of candidate microRNAs (A) novel 10, (B) novel 12 and (C) novel 13, (D) novel 14, (E) novel 15 and (F) novel 16 were obtained by qRT-PCR from RNA isolated from: osteoarthritic hip cartilage tissue (HOA) and neck of femur fracture

cartilage (NOF). Data were normalised to *U6* RNA expression. Data shows mean +/- SEM, n = 6 (analyzed by a Students T test, no significance was seen).

## Appendix 9 - Candidate novel microRNA expression across a human tissue panel

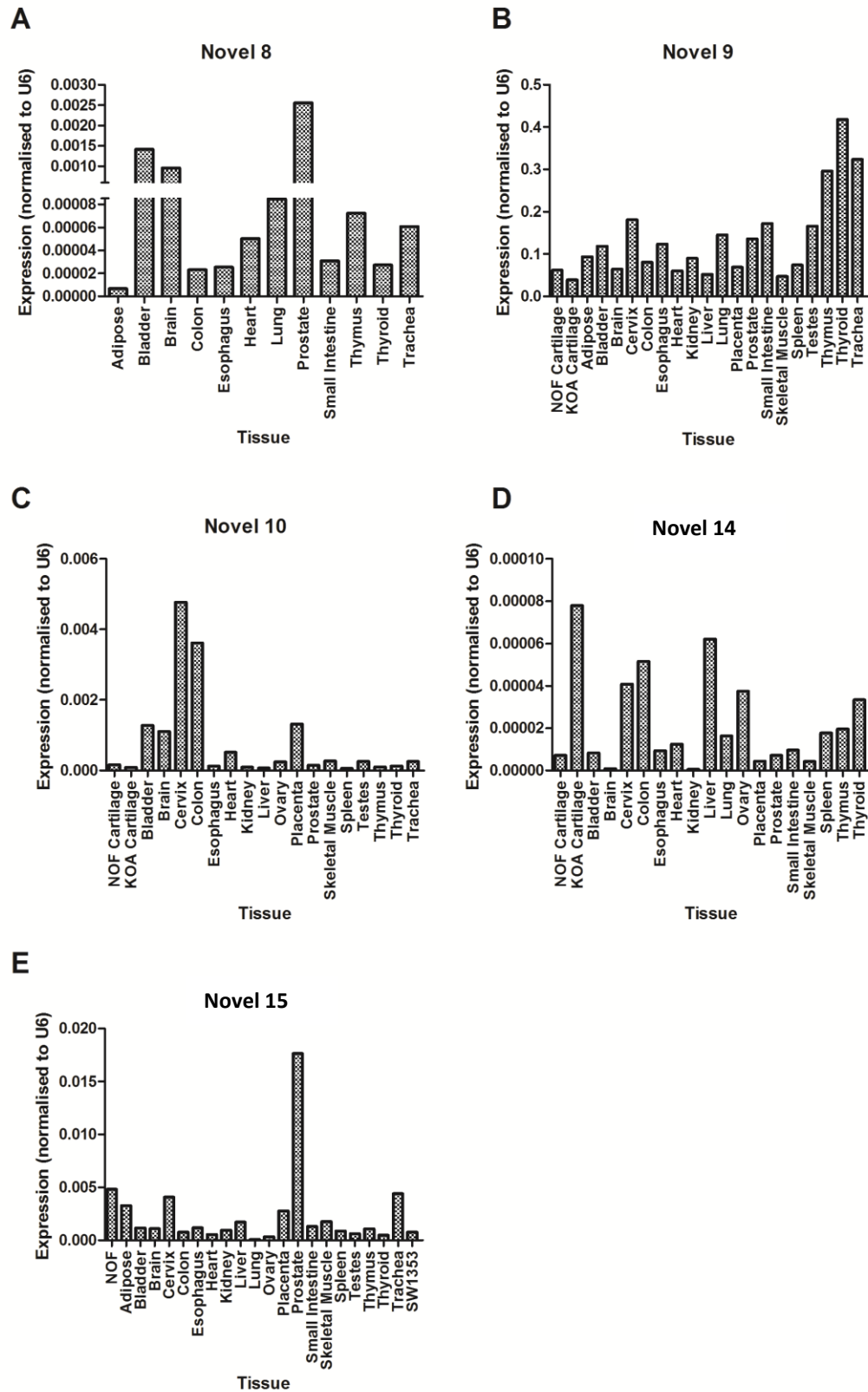


**Figure A9.1** Candidate novel microRNA expression across a human tissue panel. Expression levels of candidate microRNAs (A) novel 1, (B) novel 3, (C) novel 4, (D) novel 5 and (E) novel 6 were obtained by qRT-PCR from RNA isolated from 23 human

tissues including adipose, bladder, brain, cervix, colon, oesophagus, heart, kidney, liver, lung, ovary, placenta, prostate, small intestine, skeletal muscle, spleen, testes, thymus, thyroid, trachea, osteoarthritic knee cartilage tissue (KOA), and neck of femur fracture cartilage (NOF). Data were normalised to *U6* RNA expression. RNA of cartilage samples was from 3 patients. RNA from all remaining tissues was from a pool of 3 patients with n= 3 technical replicates.



## Appendix 9 - Candidate novel microRNA expression across a human tissue panel



**Figure A9.2 - Candidate novel microRNA expression across a human tissue panel.** Expression levels of candidate microRNAs (A) novel 8, (B) novel 9, (C) novel 10, (D)

novel 14 and (E) novel 15 were obtained by qRT-PCR from RNA isolated from 23 human tissues including adipose, bladder, brain, cervix, colon, oesophagus, heart, kidney, liver, lung, ovary, placenta, prostate, small intestine, skeletal muscle, spleen, testes, thymus, thyroid, trachea, osteoarthritic knee cartilage tissue (KOA), and neck of femur fracture cartilage (NOF). Candidate novels 12, 13 and 18 were undetectable. Data were normalised to *U6* RNA expression. RNA of cartilage samples was from 3 patients. RNA from all remaining tissues was from a pool of 3 patients with n= 3 technical replicates.

## Appendix 10 – Computational target search of miR-6509-5p

Target sites			Target information		
6mer	All 7mers	8mer	Ensemble gene ID	Gene	Gene name
6	8	3	ENSG00000113441	LNPEP	leucyl/cystinyl aminopeptidase
5	6	2	ENSG00000133703	KRAS	Kirsten rat sarcoma viral oncogene homolog
4	6	2	ENSG00000082258	CCNT2	cyclin T2
3	5	2	ENSG00000145012	LPP	LIM domain containing preferred translocation partner in lipoma
3	5	2	ENSG00000146373	RNF217	ring finger protein 217
3	5	2	ENSG00000198265	HELZ	helicase with zinc finger
4	4	2	ENSG00000196632	WNK3	WNK lysine deficient protein kinase 3
3	4	2	ENSG00000081014	AP4E1	adaptor-related protein complex 4, epsilon 1 subunit
3	4	2	ENSG00000120948	TARDBP	TAR DNA binding protein
3	4	2	ENSG00000163545	NUAK2	NUAK family, SNF1-like kinase, 2
2	4	2	ENSG00000118965	WDR35	WD repeat domain 35
2	4	2	ENSG00000131626	PPFIA1	protein tyrosine phosphatase, receptor type, f polypeptide, interacting protein, alpha 1
2	4	2	ENSG00000135932	CAB39	calcium binding protein 39
2	4	2	ENSG00000154114	TBCEL	tubulin folding cofactor E-like
2	4	2	ENSG00000170234	PWWP2 A	PWWP domain containing 2A
2	4	2	ENSG00000241644	INMT	indolethylamine N-methyltransferase
6	5	1	ENSG00000163428	LRRCS58	leucine rich repeat containing 58
5	5	1	ENSG00000117500	TMED5	transmembrane emp24 protein transport domain containing 5
5	5	1	ENSG00000149212	SESN3	sestrin 3
5	5	1	ENSG00000169905	TOR1AIP2	torsin A interacting protein 2
4	5	1	ENSG00000158528	PPP1R9 A	protein phosphatase 1, regulatory subunit 9A
4	5	1	ENSG00000172469	MANEA	mannosidase, endo-alpha
5	4	1	ENSG00000046889	PREX2	phosphatidylinositol-3,4,5-trisphosphate-dependent Rac exchange factor 2
5	4	1	ENSG00000071537	SEL1L	sel-1 suppressor of lin-12-like
5	4	1	ENSG00000204406	MBD5	methyl-CpG binding domain protein 5
5	4	1	ENSG00000205189	ZBTB10	zinc finger and BTB domain containing 10
4	4	1	ENSG00000060982	BCAT1	branched chain amino-acid transaminase 1
4	4	1	ENSG00000070367	EXOC5	exocyst complex component 5
4	4	1	ENSG00000149187	CELF1	CUGBP, Elav-like family member 1
4	4	1	ENSG00000149948	HMGA2	high mobility group AT-hook 2
3	4	1	ENSG00000097046	CDC7	cell division cycle
3	4	1	ENSG00000122042	UBL3	ubiquitin-like 3
3	4	1	ENSG00000133103	COG6	component of oligomeric golgi complex 6
3	4	1	ENSG00000143322	ABL2	c-abl oncogene 2, non-receptor tyrosine kinase
3	4	1	ENSG00000154429	CCSAP	centriole, cilia and spindle-associated protein

## Appendix 10 – Computational target search of miR-6509-5p

Target sites			Target information		
6mer	All 7mers	8mer	Ensemble gene ID	Gene	Gene name
3	4	1	ENSG00000164022	AIMP1	aminoacyl tRNA synthetase complex-interacting multifunctional protein 1
3	4	1	ENSG00000198010	DLGAP2	discs, large (Drosophila) homolog-associated protein 2
6	3	1	ENSG00000055609	MLL3	lysine (K)-specific methyltransferase 2C
5	3	1	ENSG00000172765	TMCC1	transmembrane and coiled-coil domain family 1
4	3	1	ENSG00000052723	SIKE1	suppressor of IKBKE 1
4	3	1	ENSG00000082175	PGR	progesterone receptor
4	3	1	ENSG00000101639	CEP192	centrosomal protein 192kDa
4	3	1	ENSG00000102271	KLHL4	kelch-like family member 4
4	3	1	ENSG00000135913	USP37	ubiquitin specific peptidase 37
4	3	1	ENSG00000151338	MIPOL1	mirror-image polydactyly 1
3	3	1	ENSG00000072364	AFF4	AF4/FMR2 family, member 4
3	3	1	ENSG00000106692	FKTN	fukutin
3	3	1	ENSG00000109158	GABRA4	gamma-aminobutyric acid (GABA) A receptor, alpha 4
3	3	1	ENSG00000112874	NUDT12	nudix (nucleoside diphosphate linked moiety X)-type motif 12
3	3	1	ENSG00000115137	DNAJC2 7	DnaJ (Hsp40) homolog, subfamily C, member 27
3	3	1	ENSG00000119729	RHOQ	ras homolog family member Q
3	3	1	ENSG00000120137	PANK3	pantothenate kinase 3
3	3	1	ENSG00000124783	SSR1	signal sequence receptor, alpha
3	3	1	ENSG00000128989	ARPP19	cAMP-regulated phosphoprotein, 19kDa
3	3	1	ENSG00000133104	SPG20	spastic paraplegia 20 (Troyer syndrome)
3	3	1	ENSG00000134352	IL6ST	interleukin 6 signal transducer (gp130, oncostatin M receptor)
3	3	1	ENSG00000139173	TMEM1 17	transmembrane protein 117
3	3	1	ENSG00000145335	SNCA	synuclein, alpha (non A4 component of amyloid precursor)
3	3	1	ENSG00000163904	SEN2	SUMO1/sentrin/SMT3 specific peptidase 2
3	3	1	ENSG00000168152	THAP9	THAP domain containing 9
3	3	1	ENSG00000175414	ARL10	ADP-ribosylation factor-like 10
3	3	1	ENSG00000178295	GEN1	GEN1 Holliday junction 5' flap endonuclease
3	3	1	ENSG00000185518	SV2B	synaptic vesicle glycoprotein 2B
3	3	1	ENSG00000185760	KCNQ5	potassium voltage-gated channel, KQT-like subfamily, member 5
3	3	1	ENSG00000213281	NRAS	neuroblastoma RAS viral (v-ras)
3	3	1	ENSG00000257127	CLU1	chronic lymphocytic leukemia up-regulated 1
2	3	1	ENSG00000003989	SLC7A2	solute carrier family 7, member 2
2	3	1	ENSG00000006576	PHTF2	putative homeodomain transcription factor 2
2	3	1	ENSG00000079387	SEN1	SUMO1/sentrin specific peptidase 1
2	3	1	ENSG00000084110	HAL	histidine ammonia-lyase

## Appendix 10 – Computational target search of miR-6509-5p

Target sites			Target information		
6mer	All 7mers	8mer	Ensemble gene ID	Gene	Gene name
2	3	1	ENSG00000089916	GPATCH2L	G patch domain containing 2-like
2	3	1	ENSG00000101938	CHRDL1	chordin-like 1
2	3	1	ENSG00000110987	BCL7A	B-cell CLL/lymphoma 7A
2	3	1	ENSG00000112246	SIM1	single-minded family bHLH transcription factor 1
2	3	1	ENSG00000113319	RASGRF2	Ras protein-specific guanine nucleotide-releasing factor 2
2	3	1	ENSG00000120217	CD274	CD274 molecule
2	3	1	ENSG00000120963	ZNF706	zinc finger protein 706
2	3	1	ENSG00000123728	RAP2C	RAP2C, member of RAS oncogene family
2	3	1	ENSG00000133704	IPO8	importin 8
2	3	1	ENSG00000145545	SRD5A1	steroid-5-alpha-reductase, alpha polypeptide 1
2	3	1	ENSG00000145833	DDX46	DEAD (Asp-Glu-Ala-Asp) box polypeptide 46
2	3	1	ENSG00000146776	ATXN7L1	ataxin 7-like 1
2	3	1	ENSG00000150637	CD226	CD226 molecule
2	3	1	ENSG00000152495	CAMK4	calcium/calmodulin-dependent protein kinase IV
2	3	1	ENSG00000157106	SMG1	SMG1 phosphatidylinositol 3-kinase-related kinase
2	3	1	ENSG00000162981	FAM84A	family with sequence similarity 84, member A
2	3	1	ENSG00000162994	CLHC1	clathrin heavy chain linker domain containing 1
2	3	1	ENSG00000163492	CCDC141	coiled-coil domain containing 141
2	3	1	ENSG00000164284	GRPEL2	GrpE-like 2, mitochondrial
2	3	1	ENSG00000165192	ASB11	ankyrin repeat and SOCS box containing 11
2	3	1	ENSG00000171466	ZNF562	zinc finger protein 562
2	3	1	ENSG00000172575	RASGRP1	RAS guanyl releasing protein 1
2	3	1	ENSG00000186591	UBE2H	ubiquitin-conjugating enzyme E2H
2	3	1	ENSG00000186834	HEXIM1	hexamethylene bis-acetamide inducible 1
2	3	1	ENSG00000188419	CHM	choroideremia (Rab escort protein 1)
2	3	1	ENSG00000188906	LRRK2	leucine-rich repeat kinase 2
2	3	1	ENSG00000198814	GK	glycerol kinase
2	3	1	ENSG00000205835	GMNC	geminin coiled-coil domain containing
2	3	1	ENSG00000258588	TRIM6-TRIM34	TRIM6-TRIM34 readthrough
2	3	1	ENSG00000258659	TRIM34	tripartite motif containing 34

**Table A10.1 - Computational gene target search of miR-6509-5p using RStudio.**

All 4 target sites; 6mer, 7mer, 7merA1 and 8mer were identified within the 3'UTR

of all genes within the human genome as described on Ensembl, downloaded using Biomart. Target search was carried out with the R programming Biostrings package, and the top 100 genes of 4655 with target sites (ranked by size and number of target sites) is shown.

## Appendix 11 – Computational target search of miR-664b-3p

Target sites			Target information		
6mer	All 7mers	8mer	Ensemble gene ID	Gene	Gene name
6	8	4	ENSG00000164308	ERAP2	endoplasmic reticulum aminopeptidase 2
15	14	3	ENSG00000155966	AFF2	AF4/FMR2 family, member 2
10	10	3	ENSG00000134982	APC	adenomatous polyposis coli
12	9	3	ENSG00000055609	MLL3	membrane bound O-acyltransferase domain containing 2
7	9	3	ENSG00000138138	ATAD1	ATPase family, AAA domain containing 1
7	9	3	ENSG00000144824	PHLDB2	pleckstrin homology-like domain, family B, member 2
3	6	3	ENSG00000196166	C8orf86	chromosome 8 open reading frame 86
3	6	3	ENSG00000040199	PHLPP2	PH domain and leucine rich repeat protein phosphatase 2
15	12	2	ENSG00000168807	SNTB2	syntrophin, beta 2 (dystrophin-associated protein A1, 59kDa, basic component 2)
13	12	2	ENSG00000104290	FZD3	frizzled family receptor 3
10	10	2	ENSG00000166450	PRTG	protogenin
10	9	2	ENSG00000172795	DCP2	decapping mRNA 2
8	9	2	ENSG00000133739	LRRCC1	lysine (K)-specific methyltransferase 2C
12	8	2	ENSG00000152495	CAMK4	calcium/calmodulin-dependent protein kinase IV
11	8	2	ENSG00000010244	ZNF207	zinc finger protein 207
10	8	2	ENSG00000169213	RAB3B	RAB3B, member RAS oncogene family
9	8	2	ENSG00000154429	CCSAP	centriole, cilia and spindle-associated protein
8	8	2	ENSG00000036530	CYP46A1	cytochrome P450, family 46, subfamily A, polypeptide 1
8	8	2	ENSG00000143797	MBOAT2	leucine rich repeat and coiled-coil centrosomal protein 1
8	8	2	ENSG00000147862	NFIB	nuclear factor I/B [Source:HGNC
8	8	2	ENSG00000051382	PIK3CB	phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit beta
6	8	2	ENSG00000196865	NHLRC2	NHL repeat containing 2
14	7	2	ENSG00000143190	POU2F1	POU class 2 homeobox 1
10	7	2	ENSG00000082438	COBLL1	cordón-bleu WH2 repeat protein-like 1
7	7	2	ENSG00000137145	DENND4C	DENN/MADD domain containing 4C
5	7	2	ENSG00000157538	DSCR3	Down syndrome critical region gene 3
5	7	2	ENSG00000185231	MC2R	melanocortin 2 receptor (adrenocorticotrophic hormone)
9	6	2	ENSG00000198087	CD2AP	CD2-associated protein
8	6	2	ENSG00000143079	CTTNBP2NL	CTTNBP2 N-terminal like
7	6	2	ENSG00000248905	FMN1	formin 1
6	6	2	ENSG00000144959	NCEH1	neutral cholesterol ester hydrolase 1
5	6	2	ENSG00000102531	FNDC3A	fibronectin type III domain containing 3A
5	6	2	ENSG00000145626	UGT3A1	UDP glycosyltransferase 3 family, polypeptide A1

## Appendix 11 – Computational target search of miR-664b-3p

Target sites			Target information		
6mer	All 7mers	8mer	Ensemble gene ID	Gene	Gene name
5	6	2	ENSG00000147488	ST18	suppression of tumorigenicity 18 (breast carcinoma) (zinc finger protein)
5	6	2	ENSG00000177707	PVRL3	poliovirus receptor-related 3
5	6	2	ENSG00000182168	UNC5C	unc-5 homolog C
4	6	2	ENSG00000116985	BMP8B	bone morphogenetic protein 8b
4	6	2	ENSG00000118689	FOXO3	forkhead box O3
8	5	2	ENSG00000124209	RAB22A	RAB22A, member RAS oncogene family
7	5	2	ENSG00000064933	PMS1	PMS1 postmeiotic segregation increased 1
7	5	2	ENSG00000112701	SENP6	SUMO1/sentrin specific peptidase 6
7	5	2	ENSG00000161048	NAPEPLD	N-acyl phosphatidylethanolamine phospholipase D
7	5	2	ENSG00000163762	TM4SF18	transmembrane 4 L six family member 18
7	5	2	ENSG00000170522	ELOVL6	ELOVL fatty acid elongase 6
6	5	2	ENSG00000064393	HIPK2	homeodomain interacting protein kinase 2
6	5	2	ENSG00000113273	ARSB	arylsulfatase B
5	5	2	ENSG00000047932	GOPC	golgi-associated PDZ and coiled-coil motif containing
5	5	2	ENSG00000073803	MAP3K13	mitogen-activated protein kinase kinase kinase 13
5	5	2	ENSG00000170160	CCDC144A	coiled-coil domain containing 144A
5	5	2	ENSG00000266302	CCDC144A	Coiled-coil domain-containing protein 144A
4	5	2	ENSG00000122145	TBX22	T-box 22
4	5	2	ENSG00000132849	INADL	InaD-like
4	5	2	ENSG00000145349	CAMK2D	calcium/calmodulin-dependent protein kinase II delta
4	5	2	ENSG00000164307	ERAP1	endoplasmic reticulum aminopeptidase 1
4	5	2	ENSG00000171435	KSR2	kinase suppressor of ras 2
4	5	2	ENSG00000187699	C2orf88	chromosome 2 open reading frame 88
4	5	2	ENSG00000196715	VKORC1L1	vitamin K epoxide reductase complex, subunit 1-like 1
4	5	2	ENSG00000204186	ZDBF2	zinc finger, DBF-type containing 2
3	5	2	ENSG00000054796	SPO11	SPO11 meiotic protein covalently bound to DSB
3	5	2	ENSG00000054965	FAM168A	family with sequence similarity 168, member A
3	5	2	ENSG00000065243	PKN2	protein kinase N2
3	5	2	ENSG00000070190	DAPP1	dual adaptor of phosphotyrosine and 3-phosphoinositides
3	5	2	ENSG00000084093	REST	RE1-silencing transcription factor
3	5	2	ENSG00000112276	BVES	blood vessel epicardial substance
3	5	2	ENSG00000120868	APAF1	apoptotic peptidase activating factor 1
3	5	2	ENSG00000125827	TMX4	thioredoxin-related transmembrane protein 4



## Appendix 11 – Computational target search of miR-664b-3p

Target sites			Target information		
6mer	All 7mers	8mer	Ensemble gene ID	Gene	Gene name
3	5	2	ENSG00000131042	LILRB2	leukocyte immunoglobulin-like receptor, subfamily B, member 2
3	5	2	ENSG00000145757	SPATA9	spermatogenesis associated 9
3	5	2	ENSG00000149260	CAPN5	calpain 5
3	5	2	ENSG00000196935	SRGAP1	SLIT-ROBO Rho GTPase activating protein 1
3	5	2	ENSG00000234631	AL162407.1	CDNA FLJ20147 fis, Uncharacterized protein
8	4	2	ENSG00000134852	CLOCK	clock circadian regulator
5	4	2	ENSG00000011258	MBTD1	mbt domain containing 1
4	4	2	ENSG00000009694	TENM1	teneurin transmembrane protein 1
4	4	2	ENSG00000082153	BZW1	basic leucine zipper and W2 domains 1
4	4	2	ENSG00000119729	RHOQ	ras homolog family member Q
4	4	2	ENSG00000134376	CRB1	crumbs homolog 1
4	4	2	ENSG00000149948	HMGA2	high mobility group AT-hook 2
4	4	2	ENSG00000178217	SH2D4B	SH2 domain containing 4B
3	4	2	ENSG00000038945	MSR1	macrophage scavenger receptor 1
3	4	2	ENSG00000127993	RBM48	RNA binding motif protein 48
3	4	2	ENSG00000134910	STT3A	STT3A, subunit of the oligosaccharyltransferase complex
3	4	2	ENSG00000144597	EAF1	ELL associated factor 1
3	4	2	ENSG00000145740	SLC30A5	solute carrier family 30 (zinc transporter), member 5
3	4	2	ENSG00000186417	GLDN	gliomedin
3	4	2	ENSG00000196932	TMEM26	transmembrane protein 26
3	4	2	ENSG00000197170	PSMD12	proteasome (prosome, macropain) 26S subunit, non-ATPase, 12
2	4	2	ENSG00000100027	YPEL1	yippee-like 1
2	4	2	ENSG00000103550	C16orf88	lysine-rich nucleolar protein 1
2	4	2	ENSG00000109684	CLNK	cytokine-dependent hematopoietic cell linker
2	4	2	ENSG00000116678	LEPR	leptin receptor
2	4	2	ENSG00000120686	UFM1	ubiquitin-fold modifier 1
2	4	2	ENSG00000147224	PRPS1	phosphoribosyl pyrophosphate synthetase 1
2	4	2	ENSG00000152457	DCLRE1C	DNA cross-link repair 1C
2	4	2	ENSG00000152578	GRIA4	glutamate receptor, ionotropic, AMPA 4
2	4	2	ENSG00000162511	LAPTM5	lysosomal protein transmembrane 5
2	4	2	ENSG00000169504	CLIC4	chloride intracellular channel 4
2	4	2	ENSG00000170921	TANC2	tetratricopeptide repeat, ankyrin repeat and coiled-coil containing 2
2	4	2	ENSG00000182712	CMC4	C-x(9)-C motif containing 4
2	4	2	ENSG00000214827	MTCP1	mature T-cell proliferation 1

**Table A11.1 - Computational gene target search of miR-664b-3p using RStudios.**

All 4 target sites; 6mer, 7merm8, 7merA1 and 8mer were identified within the 3'UTR of all genes within the human genome as described on Ensembl, downloaded using Biomart. Target search was carried out with the R programming Biostrings package, and the top 100 genes of 9690 with target sites (ranked by size and number of target sites) is shown.

## Appendix 12 – Computational target search of miR-3085-3p

Target sites			Target information		
6mer	All 7mers	8mer	Ensemble gene ID	Gene	Gene name
8	10	4	ENSG00000042832	TG	thyroglobulin
12	11	3	ENSG00000170921	TANC2	tetratricopeptide repeat, ankyrin repeat and coiled-coil containing 2
11	11	3	ENSG00000136944	LMX1B	LIM homeobox transcription factor 1, beta
9	9	3	ENSG00000127554	GFER	growth factor, augments of liver regeneration
6	8	3	ENSG00000072135	PTPN18	protein tyrosine phosphatase, non-receptor type 18
6	7	3	ENSG00000156639	ZFAND3	zinc finger, AN1-type domain 3
5	6	3	ENSG00000110888	CAPRN2	caprin family member 2
3	6	3	ENSG00000073331	ALPK1	alpha-kinase 1
10	8	2	ENSG00000130052	STARD8	StAR-related lipid transfer (START) domain containing 8
9	8	2	ENSG00000114735	HEMK1	HemK methyltransferase family member 1
8	8	2	ENSG00000164742	ADCY1	adenylate cyclase 1 (brain)
6	8	2	ENSG00000100106	TRIOBP	TRIO and F-actin binding protein
11	7	2	ENSG00000196092	PAX5	paired box 5
9	7	2	ENSG00000136279	DBNL	drebrin-like
9	7	2	ENSG00000174437	ATP2A2	ATPase, Ca <sup>++</sup> transporting, cardiac muscle, slow twitch 2
8	7	2	ENSG00000072501	SMC1A	structural maintenance of chromosomes 1A
7	7	2	ENSG00000090905	TNRC6A	trinucleotide repeat containing 6A
7	7	2	ENSG00000142765	SYTL1	synaptotagmin-like 1
7	7	2	ENSG00000169118	CSNK1G1	casein kinase 1, gamma 1
6	7	2	ENSG00000119720	NRDE2	NRDE-2, necessary for RNA interference, domain containing
6	7	2	ENSG00000123643	SLC36A1	solute carrier family 36 (proton/amino acid symporter), member 1
6	7	2	ENSG00000258388	EGFL8	PPT2-EGFL8 readthrough
8	6	2	ENSG00000092847	EIF2C1	argonaute RISC catalytic component 1
6	6	2	ENSG00000078900	TP73	tumor protein p73
6	6	2	ENSG00000213901	SLC23A3	solute carrier family 23, member 3
5	6	2	ENSG00000141736	ERBB2	v-erb-b2 avian erythroblastic leukemia viral oncogene homolog 2
5	6	2	ENSG00000160218	TRAPPC10	trafficking protein particle complex 10
5	6	2	ENSG00000185838	GNB1L	guanine nucleotide binding protein (G protein), beta polypeptide 1-like
5	6	2	ENSG00000212719	C17orf51	chromosome 17 open reading frame 51
4	6	2	ENSG00000134160	TRPM1	transient receptor potential cation channel, subfamily M, member 1
4	6	2	ENSG00000136051	KIAA1033	KIAA1033
4	6	2	ENSG00000138092	CENPO	centromere protein O
4	6	2	ENSG00000143217	PVRL4	poliovirus receptor-related 4
4	6	2	ENSG00000157570	TSPAN18	tetraspanin 18

## Appendix 12 – Computational target search of miR-3085-3p

Target sites			Target information		
6mer	All 7mers	8mer	Ensemble gene ID	Gene	Gene name
4	6	2	ENSG00000204220	PFDN6	prefoldin subunit 6
7	5	2	ENSG00000153404	PLEKHG4B	pleckstrin homology domain containing, family G member 4B
6	5	2	ENSG00000137070	IL11RA	interleukin 11 receptor, alpha
6	5	2	ENSG00000164061	BSN	bassoon presynaptic cytomatrix protein
6	5	2	ENSG00000169057	MECP2	methyl CpG binding protein 2 (Rett syndrome)
5	5	2	ENSG00000102786	INTS6	integrator complex subunit 6
5	5	2	ENSG00000111331	OAS3	2'-5'-oligoadenylate synthetase 3, 100kDa
5	5	2	ENSG00000121671	CRY2	cryptochrome 2 (photolyase-like)
5	5	2	ENSG00000134516	DOCK2	dedicator of cytokinesis 2
4	5	2	ENSG00000003400	CASP10	caspase 10, apoptosis-related cysteine peptidase
4	5	2	ENSG00000137135	ARHGEF39	Rho guanine nucleotide exchange factor (GEF) 39
4	5	2	ENSG00000163820	FYCO1	FYVE and coiled-coil domain containing 1
3	5	2	ENSG00000081059	TCF7	transcription factor 7
3	5	2	ENSG00000113494	PRLR	prolactin receptor
3	5	2	ENSG00000154975	CA10	carbonic anhydrase X
3	5	2	ENSG00000155229	MMS19	MMS19 nucleotide excision repair homolog ( <i>S. cerevisiae</i> )
3	5	2	ENSG00000166912	MTMR10	myotubularin related protein 10
3	5	2	ENSG00000173210	ABLIM3	actin binding LIM protein family, member 3
3	5	2	ENSG00000179409	GEMIN4	gem (nuclear organelle) associated protein 4
3	5	2	ENSG00000225697	SLC26A6	solute carrier family 26 (anion exchanger), member 6
5	4	2	ENSG00000106246	PTCD1	pentatricopeptide repeat domain 1
5	4	2	ENSG00000111752	PHC1	polyhomeotic homolog 1
5	4	2	ENSG00000120645	IQSEC3	IQ motif and Sec7 domain 3
5	4	2	ENSG00000152104	PTPN14	protein tyrosine phosphatase, non-receptor type 14
5	4	2	ENSG00000196209	SIRPB2	signal-regulatory protein beta 2
4	4	2	ENSG00000158786	PLA2G2F	phospholipase A2, group IIF
4	4	2	ENSG00000175662	TOM1L2	target of myb1-like 2 (chicken)
4	4	2	ENSG00000185163	DDX51	DEAD (Asp-Glu-Ala-Asp) box polypeptide 51
3	4	2	ENSG00000079387	SENP1	SUMO1/sentrin specific peptidase 1
3	4	2	ENSG00000126822	PLEKHG3	pleckstrin homology domain containing, family G member 3
3	4	2	ENSG00000144583	MARCH4	membrane-associated ring finger (C3HC4) 4, E3 ubiquitin protein ligase
3	4	2	ENSG00000161904	LEMD2	LEM domain containing 2
3	4	2	ENSG00000162510	MATN1	matrilin 1, cartilage matrix protein
3	4	2	ENSG00000182040	USH1G	Usher syndrome 1G (autosomal recessive)

## Appendix 12 – Computational target search of miR-3085-3p

Target sites			Target information		
6mer	All 7mers	8mer	Ensemble gene ID	Gene	Gene name
3	4	2	ENSG00000196922	ZNF252P	zinc finger protein 252, pseudogene
2	4	2	ENSG00000023902	PLEKHO1	pleckstrin homology domain containing, family O member 1
2	4	2	ENSG00000053108	FSTL4	folliculin-like 4
2	4	2	ENSG00000054392	HHAT	hedgehog acyltransferase
2	4	2	ENSG00000073417	PDE8A	phosphodiesterase 8A
2	4	2	ENSG00000103222	ABCC1	ATP-binding cassette, sub-family C (CFTR/MRP), member 1
2	4	2	ENSG00000106609	TMEM248	transmembrane protein 248
2	4	2	ENSG00000107331	ABCA2	ATP-binding cassette, sub-family A (ABC1), member 2
2	4	2	ENSG00000115594	IL1R1	interleukin 1 receptor, type I
2	4	2	ENSG00000131373	HACL1	2-hydroxyacyl-CoA lyase 1
2	4	2	ENSG00000136878	USP20	ubiquitin specific peptidase 20
2	4	2	ENSG00000141141	DDX52	DEAD (Asp-Glu-Ala-Asp) box polypeptide 52
2	4	2	ENSG00000157227	MMP14	matrix metalloproteinase 14 (membrane-inserted)
2	4	2	ENSG00000173698	GPR64	G protein-coupled receptor 64
2	4	2	ENSG00000183837	PNMA3	paraneoplastic Ma antigen 3
2	4	2	ENSG00000187987	ZSCAN23	zinc finger and SCAN domain containing 23
2	4	2	ENSG00000214022	REPIN1	replication initiator 1
2	4	2	ENSG00000254806	SYS1-DBNDD2	SYS1-DBNDD2 readthrough
15	15	1	ENSG00000072840	EVC	Ellis van Creveld syndrome
11	10	1	ENSG00000149187	CELF1	CUGBP, Elav-like family member 1
14	9	1	ENSG00000167291	TBC1D16	TBC1 domain family, member 16
11	8	1	ENSG00000149639	SOGA1	suppressor of glucose, autophagy associated 1
9	8	1	ENSG00000146067	FAM193B	family with sequence similarity 193, member B
10	7	1	ENSG00000115461	IGFBP5	insulin-like growth factor binding protein 5
7	7	1	ENSG00000122863	CHST3	carbohydrate (chondroitin 6) sulfotransferase 3
7	7	1	ENSG00000153885	KCTD15	potassium channel tetramerization domain containing 15
9	6	1	ENSG00000164627	KIF6	kinesin family member 6
9	6	1	ENSG00000170006	TMEM154	transmembrane protein 154
9	6	1	ENSG00000181873	IBA57	IBA57, iron-sulfur cluster assembly homolog ( <i>S. cerevisiae</i> )
9	6	1	ENSG00000220256	AC093802.1	Uncharacterized protein
8	6	1	ENSG00000133454	MYO18B	Myosin XVIIIIB
8	6	1	ENSG00000131389	SLC6A6	Solute carrier family 6

**Table A12.1 - Computational gene target search of microRNA miR-3085-3p using RStudio.** All 4 target sites; 6mer, 7merm8, 7merA1 and 8mer were identified within the 3'UTR of all genes within the human genome as described on ensembl, downloaded using Biomart. Target search was carried out with the R programming Biostrings package, and the top 100 genes of 8930 genes with target sites (ranked by size and number of target sites) is shown.

## Bibliography

Aagaard, L., and Rossi, J.J. (2007). RNAi therapeutics: principles, prospects and challenges. *Adv Drug Deliv Rev* 59, 75-86.

Abouheif, M.M., Nakasa, T., Shibuya, H., Niimoto, T., Kongcharoensombat, W., and Ochi, M. (2010). Silencing microRNA-34a inhibits chondrocyte apoptosis in a rat osteoarthritis model in vitro. *Rheumatology (Oxford)* 49, 2054-2060.

Aigner, T., Soeder, S., and Haag, J. (2006). IL-1beta and BMPs--interactive players of cartilage matrix degradation and regeneration. *Eur Cell Mater* 12, 49-56; discussion 56.

Aigner, T., Vornehm, S.I., Zeiler, G., Dudhia, J., von der Mark, K., and Bayliss, M.T. (1997). Suppression of cartilage matrix gene expression in upper zone chondrocytes of osteoarthritic cartilage. *Arthritis Rheum* 40, 562-569.

Aigner, T., Zien, A., Gehrsitz, A., Gebhard, P.M., and McKenna, L. (2001). Anabolic and catabolic gene expression pattern analysis in normal versus osteoarthritic cartilage using complementary DNA-array technology. *Arthritis Rheum* 44, 2777-2789.

Akbari Moqadam, F., Pieters, R., and den Boer, M.L. (2013). The hunting of targets: challenge in miRNA research. *Leukemia* 27, 16-23.

Akhtar, N., Rasheed, Z., Ramamurthy, S., Anbazhagan, A.N., Voss, F.R., and Haqqi, T.M. (2010). MicroRNA-27b regulates the expression of matrix metalloproteinase 13 in human osteoarthritis chondrocytes. *Arthritis Rheum* 62, 1361-1371.

Alaaeddine, N., Di Battista, J.A., Pelletier, J.P., Kiansa, K., Cloutier, J.M., and Martel-Pelletier, J. (1999). Differential effects of IL-8, LIF (pro-inflammatory) and IL-11 (anti-inflammatory) on TNF-alpha-induced PGE(2) release and on signalling pathways in human OA synovial fibroblasts. *Cytokine* 11, 1020-1030.

Almonte-Becerril, M., Navarro-Garcia, F., Gonzalez-Robles, A., Vega-Lopez, M.A., Lavalle, C., and Kouri, J.B. (2010). Cell death of chondrocytes is a combination between apoptosis and autophagy during the pathogenesis of Osteoarthritis within an experimental model. *Apoptosis* 15, 631-638.

Altuvia, Y., Landgraf, P., Lithwick, G., Elefant, N., Pfeffer, S., Aravin, A., Brownstein, M.J., Tuschl, T., and Margalit, H. (2005). Clustering and conservation patterns of human microRNAs. *Nucleic Acids Res* 33, 2697-2706.

Ambros, V. (2004). The functions of animal microRNAs. *Nature* 431, 350-355.

Amour, A., Slocombe, P.M., Webster, A., Butler, M., Knight, C.G., Smith, B.J., Stephens, P.E., Shelley, C., Hutton, M., Knauper, V., *et al.* (1998). TNF-alpha converting enzyme (TACE) is inhibited by TIMP-3. *FEBS Lett* 435, 39-44.

Apte, S.S., Olsen, B.R., and Murphy, G. (1996). The gene structure of tissue inhibitor of metalloproteinases (TIMP)-3 and its inhibitory activities define the distinct TIMP gene family. *J Biol Chem* 271, 2874.

- Araldi, E., and Schipani, E. (2010). MicroRNA-140 and the silencing of osteoarthritis. *Genes Dev* 24, 1075-1080.
- Arroll, B., and Goodyear-Smith, F. (2004). Corticosteroid injections for osteoarthritis of the knee: meta-analysis. *BMJ* 328, 869.
- Arroyo, J.D., Chevillet, J.R., Kroh, E.M., Ruf, I.K., Pritchard, C.C., Gibson, D.F., Mitchell, P.S., Bennett, C.F., Pogosova-Agadjanyan, E.L., Stirewalt, D.L., *et al.* (2011). Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. *Proc Natl Acad Sci U S A* 108, 5003-5008.
- Arthritis-Care (2012). OANation 2012 report.
- Aspden, R.M. (2011). Obesity punches above its weight in osteoarthritis. *Nat Rev Rheumatol* 7, 65-68.
- Attur, M., Samuels, J., Krasnokutsky, S., and Abramson, S.B. (2010). Targeting the synovial tissue for treating osteoarthritis (OA): where is the evidence? *Best Pract Res Clin Rheumatol* 24, 71-79.
- Babiarz, J.E., Ruby, J.G., Wang, Y., Bartel, D.P., and Blelloch, R. (2008). Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes Dev* 22, 2773-2785.
- Bail, S., Swerdel, M., Liu, H., Jiao, X., Goff, L.A., Hart, R.P., and Kiledjian, M. (2010). Differential regulation of microRNA stability. *RNA* 16, 1032-1039.
- Baker, M. (2010). MicroRNA profiling: separating signal from noise. *Nat Methods* 7, 687-692.
- Barad, O., Meiri, E., Avniel, A., Aharonov, R., Barzilai, A., Bentwich, I., Einav, U., Gilad, S., Hurban, P., Karov, Y., *et al.* (2004). MicroRNA expression detected by oligonucleotide microarrays: system establishment and expression profiling in human tissues. *Genome Res* 14, 2486-2494.
- Baragi, V.M., Becher, G., Bendele, A.M., Biesinger, R., Bluhm, H., Boer, J., Deng, H., Dodd, R., Essers, M., Feuerstein, T., *et al.* (2009). A new class of potent matrix metalloproteinase 13 inhibitors for potential treatment of osteoarthritis: Evidence of histologic and clinical efficacy without musculoskeletal toxicity in rat models. *Arthritis Rheum* 60, 2008-2018.
- Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell* 136, 215-233.
- Bartels, C.L., and Tsongalis, G.J. (2009). MicroRNAs: novel biomarkers for human cancer. *Clin Chem* 55, 623-631.
- Barter, M.J., Woods, S., and Young, D. A. (2012). microRNA in Chondrogenesis, Cartilage and Osteoarthritis. *Current Rheumatology Reviews* 8, 89-97.



- Bashirullah, A., Pasquinelli, A.E., Kiger, A.A., Perrimon, N., Ruvkun, G., and Thummel, C.S. (2003). Coordinate regulation of small temporal RNAs at the onset of *Drosophila* metamorphosis. *Dev Biol* 259, 1-8.
- Bau, B., Gebhard, P.M., Haag, J., Knorr, T., Bartnik, E., and Aigner, T. (2002). Relative messenger RNA expression profiling of collagenases and aggrecanases in human articular chondrocytes in vivo and in vitro. *Arthritis Rheum* 46, 2648-2657.
- Becker, H.M., Klier, M., Schuler, C., McKenna, R., and Deitmer, J.W. (2011). Intramolecular proton shuttle supports not only catalytic but also noncatalytic function of carbonic anhydrase II. *Proc Natl Acad Sci U S A* 108, 3071-3076.
- Beitzinger, M., Peters, L., Zhu, J.Y., Kremmer, E., and Meister, G. (2007). Identification of human microRNA targets from isolated argonaute protein complexes. *RNA Biol* 4, 76-84.
- Bejerano, G., Pheasant, M., Makunin, I., Stephen, S., Kent, W.J., Mattick, J.S., and Haussler, D. (2004). Ultraconserved elements in the human genome. *Science* 304, 1321-1325.
- Bellido, M., Lugo, L., Roman-Blas, J.A., Castaneda, S., Caeiro, J.R., Dapia, S., Calvo, E., Largo, R., and Herrero-Beaumont, G. (2010). Subchondral bone microstructural damage by increased remodelling aggravates experimental osteoarthritis preceded by osteoporosis. *Arthritis Res Ther* 12, R152.
- Benetti, R., Gonzalo, S., Jaco, I., Munoz, P., Gonzalez, S., Schoeftner, S., Murchison, E., Andl, T., Chen, T., Klatt, P., *et al.* (2008). A mammalian microRNA cluster controls DNA methylation and telomere recombination via Rbl2-dependent regulation of DNA methyltransferases. *Nat Struct Mol Biol* 15, 998.
- Benito, M.J., Veale, D.J., FitzGerald, O., van den Berg, W.B., and Bresnihan, B. (2005). Synovial tissue inflammation in early and late osteoarthritis. *Ann Rheum Dis* 64, 1263-1267.
- Berchtold, S., Manncke, B., Klenk, J., Geisel, J., Autenrieth, I.B., and Bohn, E. (2008). Forced IFIT-2 expression represses LPS induced TNF-alpha expression at posttranscriptional levels. *BMC Immunol* 9, 75.
- Berenbaum, F. (2012). Diabetes-induced osteoarthritis: from a new paradigm to a new phenotype. *Postgrad Med J* 88, 240-242.
- 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., and Hannon, G.J. (2003). Dicer is essential for mouse development. *Nat Genet* 35, 215-217.
- Betancur, J.G., and Tomari, Y. (2012). Dicer is dispensable for asymmetric RISC loading in mammals. *RNA* 18, 24-30.
- Betel, D., Wilson, M., Gabow, A., Marks, D.S., and Sander, C. (2008). The microRNA.org resource: targets and expression. *Nucleic Acids Res* 36, D149-153.

Beyer, C., Zampetaki, A., Lin, N.Y., Kleyer, A., Perricone, C., Iagnocco, A., Distler, A., Langley, S.R., Gelse, K., Sesselmann, S., *et al.* (2014). Signature of circulating microRNAs in osteoarthritis. *Ann Rheum Dis*.

Bhosale, A.M., and Richardson, J.B. (2008). Articular cartilage: structure, injuries and review of management. *Br Med Bull* 87, 77-95.

Bierma-Zeinstra, S.M., and Verhagen, A.P. (2011). Osteoarthritis subpopulations and implications for clinical trial design. *Arthritis Res Ther* 13, 213.

Billingham, R.C., Dahlberg, L., Ionescu, M., Reiner, A., Bourne, R., Rorabeck, C., Mitchell, P., Hambor, J., Diekmann, O., Tschesche, H., *et al.* (1997). Enhanced cleavage of type II collagen by collagenases in osteoarthritic articular cartilage. *J Clin Invest* 99, 1534-1545.

Blanco, F.J., Guitian, R., Vazquez-Martul, E., de Toro, F.J., and Galdo, F. (1998). Osteoarthritis chondrocytes die by apoptosis. A possible pathway for osteoarthritis pathology. *Arthritis Rheum* 41, 284-289.

Bluteau, G., Conrozier, T., Mathieu, P., Vignon, E., Herbage, D., and Mallein-Gerin, F. (2001). Matrix metalloproteinase-1, -3, -13 and aggrecanase-1 and -2 are differentially expressed in experimental osteoarthritis. *Biochim Biophys Acta* 1526, 147-158.

Bobick, B.E., and Kulyk, W.M. (2008). Regulation of cartilage formation and maturation by mitogen-activated protein kinase signaling. *Birth Defects Res C Embryo Today* 84, 131-154.

Bohnsack, M.T., Czaplinski, K., and Gorlich, D. (2004). Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA* 10, 185-191.

Boot-Handford, R.P., Tuckwell, D.S., Plumb, D.A., Rock, C.F., and Poulson, R. (2003). A novel and highly conserved collagen (pro(α)1(XXVII)) with a unique expression pattern and unusual molecular characteristics establishes a new clade within the vertebrate fibrillar collagen family. *J Biol Chem* 278, 31067-31077.

Borchert, G.M., Lanier, W., and Davidson, B.L. (2006). RNA polymerase III transcribes human microRNAs. *Nat Struct Mol Biol* 13, 1097-1101.

Brennecke, J., Stark, A., Russell, R.B., and Cohen, S.M. (2005). Principles of microRNA-target recognition. *PLoS Biol* 3, e85.

Brew, C.J., Clegg, P.D., Boot-Handford, R.P., Andrew, J.G., and Hardingham, T. (2010). Gene expression in human chondrocytes in late osteoarthritis is changed in both fibrillated and intact cartilage without evidence of generalised chondrocyte hypertrophy. *Ann Rheum Dis* 69, 234-240.

Brew, K., and Nagase, H. (2010). The tissue inhibitors of metalloproteinases (TIMPs): an ancient family with structural and functional diversity. *Biochim Biophys Acta* 1803, 55-71.

- Cai, X., Hagedorn, C.H., and Cullen, B.R. (2004). Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA* 10, 1957-1966.
- Caiment, F., Charlier, C., Hadfield, T., Cockett, N., Georges, M., and Baurain, D. (2010). Assessing the effect of the CLPG mutation on the microRNA catalog of skeletal muscle using high-throughput sequencing. *Genome Res* 20, 1651-1662.
- Carter, D.R., and Beaupre, G. S., ed. (2007). *Skeletal Function and Form: Mechanobiology of Skeletal Development, Ageing, and Regeneration* (Cambridge University Press).
- Carthew, R.W., and Sontheimer, E.J. (2009). Origins and Mechanisms of miRNAs and siRNAs. *Cell* 136, 642-655.
- Cawston, T.E., Curry, V.A., Summers, C.A., Clark, I.M., Riley, G.P., Life, P.F., Spaul, J.R., Goldring, M.B., Koshy, P.J., Rowan, A.D., *et al.* (1998). The role of oncostatin M in animal and human connective tissue collagen turnover and its localization within the rheumatoid joint. *Arthritis Rheum* 41, 1760-1771.
- Cheloufi, S., Dos Santos, C.O., Chong, M.M., and Hannon, G.J. (2010). A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature* 465, 584-589.
- Chen, F.H., Thomas, A.O., Hecht, J.T., Goldring, M.B., and Lawler, J. (2005). Cartilage oligomeric matrix protein/thrombospondin 5 supports chondrocyte attachment through interaction with integrins. *J Biol Chem* 280, 32655-32661.
- Chevalier, X. (1993). Fibronectin, cartilage, and osteoarthritis. *Semin Arthritis Rheum* 22, 307-318.
- Chiang, H.R., Schoenfeld, L.W., Ruby, J.G., Auyeung, V.C., Spies, N., Baek, D., Johnston, W.K., Russ, C., Luo, S., Babiarz, J.E., *et al.* (2010). Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. *Genes Dev* 24, 992-1009.
- Chong, K.W., Chanalaris, A., Burleigh, A., Jin, H., Watt, F.E., Saklatvala, J., and Vincent, T.L. (2013). Fibroblast growth factor 2 drives changes in gene expression following injury to murine cartilage in vitro and in vivo. *Arthritis Rheum* 65, 2346-2355.
- Ciaudo, C., Servant, N., Cognat, V., Sarazin, A., Kieffer, E., Viville, S., Colot, V., Barillot, E., Heard, E., and Voinnet, O. (2009). Highly dynamic and sex-specific expression of microRNAs during early ES cell differentiation. *PLoS Genet* 5, e1000620.
- Collins, A., McGahon, M.K., Yarham, J.M., Daly, A., Guduric-Fuchs, J., and Simpson, D.A. (2013). Next-generation sequencing analysis of rat heart left ventricle: novel microRNAs, novel isomiRs and transmural distribution of microRNA expression. *The FASEB journal*.

Cordes, K.R., Sheehy, N.T., White, M.P., Berry, E.C., Morton, S.U., Muth, A.N., Lee, T.H., Miano, J.M., Ivey, K.N., and Srivastava, D. (2009). miR-145 and miR-143 regulate smooth muscle cell fate and plasticity. *Nature* 460, 705-710.

Cummins, J.M., He, Y., Leary, R.J., Pagliarini, R., Diaz, L.A., Jr., Sjoblom, T., Barad, O., Bentwich, Z., Szafranska, A.E., Labourier, E., *et al.* (2006). The colorectal microRNAome. *Proc Natl Acad Sci U S A* 103, 3687-3692.

Da, R.R., Qin, Y., Baeten, D., and Zhang, Y. (2007). B cell clonal expansion and somatic hypermutation of Ig variable heavy chain genes in the synovial membrane of patients with osteoarthritis. *J Immunol* 178, 557-565.

Da Sacco, L., and Masotti, A. (2012). Recent Insights and Novel Bioinformatics Tools to Understand the Role of MicroRNAs Binding to 5' Untranslated Region. *Int J Mol Sci* 14, 480-495.

Dahlberg, L., Billingham, R.C., Manner, P., Nelson, F., Webb, G., Ionescu, M., Reiner, A., Tanzer, M., Zukor, D., Chen, J., *et al.* (2000). Selective enhancement of collagenase-mediated cleavage of resident type II collagen in cultured osteoarthritic cartilage and arrest with a synthetic inhibitor that spares collagenase 1 (matrix metalloproteinase 1). *Arthritis Rheum* 43, 673-682.

Davis, M.A. (1988). Epidemiology of osteoarthritis. *Clin Geriatr Med* 4, 241-255.

de Jong, H., Berlo, S.E., Hombrink, P., Otten, H.G., van Eden, W., Lafeber, F.P., Heurkens, A.H., Bijlsma, J.W., Glant, T.T., and Prakken, B.J. (2010). Cartilage proteoglycan aggrecan epitopes induce proinflammatory autoreactive T-cell responses in rheumatoid arthritis and osteoarthritis. *Ann Rheum Dis* 69, 255-262.

Diaz-Prado, S., Cicione, C., Muinos-Lopez, E., Hermida-Gomez, T., Oreiro, N., Fernandez-Lopez, C., and Blanco, F.J. (2012). Characterization of microRNA expression profiles in normal and osteoarthritic human chondrocytes. *BMC Musculoskelet Disord* 13, 144.

Dieppe, P.A., and Lohmander, L.S. (2005). Pathogenesis and management of pain in osteoarthritis. *Lancet* 365, 965-973.

Distel, E., Cadoudal, T., Durant, S., Poignard, A., Chevalier, X., and Benelli, C. (2009). The infrapatellar fat pad in knee osteoarthritis: an important source of interleukin-6 and its soluble receptor. *Arthritis Rheum* 60, 3374-3377.

Drabek, K., van de Peppel, J., Eijken, M., and van Leeuwen, J.P. (2011). GPM6B regulates osteoblast function and induction of mineralization by controlling cytoskeleton and matrix vesicle release. *J Bone Miner Res* 26, 2045-2051.

Dreier, R., Opolka, A., Grifka, J., Bruckner, P., and Grassel, S. (2008). Collagen IX-deficiency seriously compromises growth cartilage development in mice. *Matrix Biol* 27, 319-329.

Driban, J.B., Sitler, M.R., Barbe, M.F., and Balasubramanian, E. (2010). Is osteoarthritis a heterogeneous disease that can be stratified into subsets? *Clin Rheumatol* 29, 123-131.

- Dudek, K.A., Lafont, J.E., Martinez-Sanchez, A., and Murphy, C.L. (2010). Type II collagen expression is regulated by tissue-specific miR-675 in human articular chondrocytes. *J Biol Chem* 285, 24381-24387.
- Dunn, W., DuRaine, G., and Reddi, A.H. (2009). Profiling microRNA expression in bovine articular cartilage and implications for mechanotransduction. *Arthritis Rheum* 60, 2333-2339.
- Durigova, M., Nagase, H., Mort, J.S., and Roughley, P.J. (2011). MMPs are less efficient than ADAMTS5 in cleaving aggrecan core protein. *Matrix Biol* 30, 145-153.
- Duursma, A.M., Kedde, M., Schrier, M., le Sage, C., and Agami, R. (2008). miR-148 targets human DNMT3b protein coding region. *RNA* 14, 872-877.
- Eltawil, N.M., De Bari, C., Achan, P., Pitzalis, C., and Dell'accio, F. (2009). A novel in vivo murine model of cartilage regeneration. Age and strain-dependent outcome after joint surface injury. *Osteoarthritis Cartilage* 17, 695-704.
- Ender, C., Krek, A., Friedlander, M.R., Beitzinger, M., Weinmann, L., Chen, W., Pfeffer, S., Rajewsky, N., and Meister, G. (2008). A human snoRNA with microRNA-like functions. *Mol Cell* 32, 519-528.
- Englund, M., Guermazi, A., Roemer, F.W., Yang, M., Zhang, Y., Nevitt, M.C., Lynch, J.A., Lewis, C.E., Torner, J., and Felson, D.T. (2010). Meniscal pathology on MRI increases the risk for both incident and enlarging subchondral bone marrow lesions of the knee: the MOST Study. *Ann Rheum Dis* 69, 1796-1802.
- Enomoto, H., Enomoto-Iwamoto, M., Iwamoto, M., Nomura, S., Himeno, M., Kitamura, Y., Kishimoto, T., and Komori, T. (2000). Cbfa1 is a positive regulatory factor in chondrocyte maturation. *J Biol Chem* 275, 8695-8702.
- Eyre, D.R., Weis, M.A., and Wu, J.J. (2006). Articular cartilage collagen: an irreplaceable framework? *Eur Cell Mater* 12, 57-63.
- Fassler, R., Schnegelsberg, P.N., Dausman, J., Shinya, T., Muragaki, Y., McCarthy, M.T., Olsen, B.R., and Jaenisch, R. (1994). Mice lacking alpha 1 (IX) collagen develop noninflammatory degenerative joint disease. *Proc Natl Acad Sci U S A* 91, 5070-5074.
- Fedak, P.W., Smookler, D.S., Kassiri, Z., Ohno, N., Leco, K.J., Verma, S., Mickle, D.A., Watson, K.L., Hojilla, C.V., Cruz, W., *et al.* (2004). TIMP-3 deficiency leads to dilated cardiomyopathy. *Circulation* 110, 2401-2409.
- Fehniger, T.A., Wylie, T., Germino, E., Leong, J.W., Magrini, V.J., Koul, S., Keppel, C.R., Schneider, S.E., Koboldt, D.C., Sullivan, R.P., *et al.* (2010). Next-generation sequencing identifies the natural killer cell microRNA transcriptome. *Genome Res* 20, 1590-1604.
- Fehr, J.E., Trotter, G.W., Oxford, J.T., and Hart, D.A. (2000). Comparison of Northern blot hybridization and a reverse transcriptase-polymerase chain reaction technique for measurement of mRNA expression of metalloproteinases and matrix components in articular cartilage and synovial membrane from horses with osteoarthritis. *Am J Vet Res* 61, 900-905.

- Fensterl, V., and Sen, G.C. (2011). The ISG56/IFIT1 gene family. *J Interferon Cytokine Res* 31, 71-78.
- Fernandes, J.C., Martel-Pelletier, J., and Pelletier, J.P. (2002). The role of cytokines in osteoarthritis pathophysiology. *Biorheology* 39, 237-246.
- Fichtlscherer, S., De Rosa, S., Fox, H., Schwietz, T., Fischer, A., Liebetrau, C., Weber, M., Hamm, C.W., Roxe, T., Muller-Ardogan, M., *et al.* (2010). Circulating microRNAs in patients with coronary artery disease. *Circ Res* 107, 677-684.
- Fischer, N., Konevega, A.L., Wintermeyer, W., Rodnina, M.V., and Stark, H. (2010). Ribosome dynamics and tRNA movement by time-resolved electron cryomicroscopy. *Nature* 466, 329-333.
- Flicek, P., Amode, M.R., Barrell, D., Beal, K., Billis, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fitzgerald, S., *et al.* (2014). Ensembl 2014. *Nucleic Acids Res* 42, D749-755.
- Fosang, A.J., and Beier, F. (2011). Emerging Frontiers in cartilage and chondrocyte biology. *Best Pract Res Clin Rheumatol* 25, 751-766.
- Fosang, A.J., Last, K., Stanton, H., Golub, S.B., Little, C.B., Brown, L., and Jackson, D.C. (2010). Neopeptide antibodies against MMP-cleaved and aggrecanase-cleaved aggrecan. *Methods Mol Biol* 622, 312-347.
- Foster, D.N., and Foster, L.K. (1997). Immortalized cell lines for virus growth (US).
- Franklin, J., Englund, M., Ingvarsson, T., and Lohmander, S. (2010). The association between hip fracture and hip osteoarthritis: a case-control study. *BMC Musculoskelet Disord* 11, 274.
- Freemont, A.J., and Hoyland, J. (2006). Lineage plasticity and cell biology of fibrocartilage and hyaline cartilage: its significance in cartilage repair and replacement. *Eur J Radiol* 57, 32-36.
- Fregly, B.J. (2012). Gait modification to treat knee osteoarthritis. *HSS J* 8, 45-48.
- French, M.M.a.A., K.A. (2003). Differentiation Factors and Articular Cartilage Regeneration. *Topics in Tissue Engineering Chapter 2*, 2-19.
- Friedlander, M.R., Mackowiak, S.D., Li, N., Chen, W., and Rajewsky, N. (2012). miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res* 40, 37-52.
- Friedman, R.C., Farh, K.K., Burge, C.B., and Bartel, D.P. (2009). Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* 19, 92-105.
- Fu, H., Tie, Y., Xu, C., Zhang, Z., Zhu, J., Shi, Y., Jiang, H., Sun, Z., and Zheng, X. (2005). Identification of human fetal liver miRNAs by a novel method. *FEBS Lett* 579, 3849-3854.

- Gantier, M.P., McCoy, C.E., Rusinova, I., Saulep, D., Wang, D., Xu, D., Irving, A.T., Behlke, M.A., Hertzog, P.J., Mackay, F., *et al.* (2011). Analysis of microRNA turnover in mammalian cells following Dicer1 ablation. *Nucleic Acids Res* 39, 5692-5703.
- Gao, S., von der Malsburg, A., Dick, A., Faelber, K., Schroder, G.F., Haller, O., Kochs, G., and Daumke, O. (2011). Structure of myxovirus resistance protein a reveals intra- and intermolecular domain interactions required for the antiviral function. *Immunity* 35, 514-525.
- Gebauer, M., Saas, J., Sohler, F., Haag, J., Soder, S., Pieper, M., Bartnik, E., Beninga, J., Zimmer, R., and Aigner, T. (2005). Comparison of the chondrosarcoma cell line SW1353 with primary human adult articular chondrocytes with regard to their gene expression profile and reactivity to IL-1beta. *Osteoarthritis Cartilage* 13, 697-708.
- Gendron, C., Kashiwagi, M., Lim, N.H., Enghild, J.J., Thogersen, I.B., Hughes, C., Caterson, B., and Nagase, H. (2007). Proteolytic activities of human ADAMTS-5: comparative studies with ADAMTS-4. *J Biol Chem* 282, 18294-18306.
- Girard, A., Sachidanandam, R., Hannon, G.J., and Carmell, M.A. (2006). A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature* 442, 199-202.
- Girkontaite, I., Frischholz, S., Lammi, P., Wagner, K., Swoboda, B., Aigner, T., and Von der Mark, K. (1996). Immunolocalization of type X collagen in normal fetal and adult osteoarthritic cartilage with monoclonal antibodies. *Matrix Biol* 15, 231-238.
- Git, A., Dvinge, H., Salmon-Divon, M., Osborne, M., Kutter, C., Hadfield, J., Bertone, P., and Caldas, C. (2010). Systematic comparison of microarray profiling, real-time PCR, and next-generation sequencing technologies for measuring differential microRNA expression. *RNA* 16, 991-1006.
- Glasson, S.S., Askew, R., Sheppard, B., Carito, B.A., Blanchet, T., Ma, H.L., Flannery, C.R., Kanki, K., Wang, E., Peluso, D., *et al.* (2004). Characterization of and osteoarthritis susceptibility in ADAMTS-4-knockout mice. *Arthritis Rheum* 50, 2547-2558.
- Goldring, M.B. (2000). The role of the chondrocyte in osteoarthritis. *Arthritis Rheum* 43, 1916-1926.
- Goldring, M.B. (2012). Chondrogenesis, chondrocyte differentiation, and articular cartilage metabolism in health and osteoarthritis. *Ther Adv Musculoskelet Dis* 4, 269-285.
- Goldring, M.B., and Marcu, K.B. (2009). Cartilage homeostasis in health and rheumatic diseases. *Arthritis Res Ther* 11, 224.
- Goldring, M.B., Otero, M., Plumb, D.A., Dragomir, C., Favero, M., El Hachem, K., Hashimoto, K., Roach, H.I., Olivotto, E., Borzi, R.M., *et al.* (2011). Roles of inflammatory and anabolic cytokines in cartilage metabolism: signals and multiple effectors converge upon MMP-13 regulation in osteoarthritis. *Eur Cell Mater* 21, 202-220.

- Goldring, M.B., Tsuchimochi, K., and Ijiri, K. (2006). The control of chondrogenesis. *J Cell Biochem* 97, 33-44.
- Goldring, S.R., and Goldring, M.B. (2006). Clinical aspects, pathology and pathophysiology of osteoarthritis. *J Musculoskelet Neuronal Interact* 6, 376-378.
- Gregory, R.I., Chendrimada, T.P., Cooch, N., and Shiekhattar, R. (2005). Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. *Cell* 123, 631-640.
- Gregory, R.I., Yan, K.P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., and Shiekhattar, R. (2004). The Microprocessor complex mediates the genesis of microRNAs. *Nature* 432, 235-240.
- Grgurevic, L., Macek, B., Durdevic, D., and Vukicevic, S. (2007). Detection of bone and cartilage-related proteins in plasma of patients with a bone fracture using liquid chromatography-mass spectrometry. *Int Orthop* 31, 743-751.
- Griffiths-Jones, S. (2004). The microRNA Registry. *Nucleic Acids Res* 32, D109-111.
- Griffiths-Jones, S., Grocock, R.J., van Dongen, S., Bateman, A., and Enright, A.J. (2006). miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res* 34, D140-144.
- Griffiths-Jones, S., Saini, H.K., van Dongen, S., and Enright, A.J. (2008). miRBase: tools for microRNA genomics. *Nucleic Acids Res* 36, D154-158.
- Grimshaw, M.J., and Mason, R.M. (2000). Bovine articular chondrocyte function in vitro depends upon oxygen tension. *Osteoarthritis Cartilage* 8, 386-392.
- Guilak, F., Ratcliffe, A., Lane, N., Rosenwasser, M.P., and Mow, V.C. (1994). Mechanical and biochemical changes in the superficial zone of articular cartilage in canine experimental osteoarthritis. *J Orthop Res* 12, 474-484.
- Guo, H., Ingolia, N.T., Weissman, J.S., and Bartel, D.P. (2010). Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* 466, 835-840.
- Habbe, N., Koorstra, J.B., Mendell, J.T., Offerhaus, G.J., Ryu, J.K., Feldmann, G., Mullendore, M.E., Goggins, M.G., Hong, S.M., and Maitra, A. (2009). MicroRNA miR-155 is a biomarker of early pancreatic neoplasia. *Cancer Biol Ther* 8, 340-346.
- Hafner, M., Landgraf, P., Ludwig, J., Rice, A., Ojo, T., Lin, C., Holoch, D., Lim, C., and Tuschl, T. (2008). Identification of microRNAs and other small regulatory RNAs using cDNA library sequencing. *Methods* 44, 3-12.
- Halasz, K., Kassner, A., Morgelin, M., and Heinegard, D. (2007). COMP acts as a catalyst in collagen fibrillogenesis. *J Biol Chem* 282, 31166-31173.
- Haller, O., and Kochs, G. (2011). Human MxA protein: an interferon-induced dynamin-like GTPase with broad antiviral activity. *J Interferon Cytokine Res* 31, 79-87.



Hammond, S.M., Bernstein, E., Beach, D., and Hannon, G.J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* *404*, 293-296.

Hansen, K.F., Sakamoto, K., and Obrietan, K. (2011). MicroRNAs: a potential interface between the circadian clock and human health. *Genome Med* *3*, 10.

Harfe, B.D., McManus, M.T., Mansfield, J.H., Hornstein, E., and Tabin, C.J. (2005). The RNaseIII enzyme Dicer is required for morphogenesis but not patterning of the vertebrate limb. *Proc Natl Acad Sci U S A* *102*, 10898-10903.

Hartmann, C. (2009). Transcriptional networks controlling skeletal development. *Curr Opin Genet Dev* *19*, 437-443.

Haussecker, D., Huang, Y., Lau, A., Parameswaran, P., Fire, A.Z., and Kay, M.A. (2010). Human tRNA-derived small RNAs in the global regulation of RNA silencing. *RNA* *16*, 673-695.

Hayashi, S., Murakami, Y., and Matsufuji, S. (1996). Ornithine decarboxylase antizyme: a novel type of regulatory protein. *Trends Biochem Sci* *21*, 27-30.

Hedges, S.B., Dudley, J., and Kumar, S. (2006). TimeTree: a public knowledge-base of divergence times among organisms. *Bioinformatics* *22*, 2971-2972.

Heinegard, D., and Saxne, T. (2011). The role of the cartilage matrix in osteoarthritis. *Nat Rev Rheumatol* *7*, 50-56.

Hendrickson, D.G., Hogan, D.J., McCullough, H.L., Myers, J.W., Herschlag, D., Ferrell, J.E., and Brown, P.O. (2009). Concordant regulation of translation and mRNA abundance for hundreds of targets of a human microRNA. *PLoS Biol* *7*, e1000238.

Henke, J.I., Goergen, D., Zheng, J., Song, Y., Schuttler, C.G., Fehr, C., Junemann, C., and Niepmann, M. (2008). microRNA-122 stimulates translation of hepatitis C virus RNA. *EMBO J* *27*, 3300-3310.

Herrero-Beaumont, G., Roman-Blas, J.A., Castaneda, S., and Jimenez, S.A. (2009). Primary osteoarthritis no longer primary: three subsets with distinct etiological, clinical, and therapeutic characteristics. *Semin Arthritis Rheum* *39*, 71-80.

Hibio, N., Hino, K., Shimizu, E., Nagata, Y., and Ui-Tei, K. (2012). Stability of miRNA 5'terminal and seed regions is correlated with experimentally observed miRNA-mediated silencing efficacy. *Sci Rep* *2*, 996.

Higuchi, Y., Kawakami, S., and Hashida, M. (2010). Strategies for in vivo delivery of siRNAs: recent progress. *BioDrugs* *24*, 195-205.

Hino, S., Michiue, T., Asashima, M., and Kikuchi, A. (2003). Casein kinase I epsilon enhances the binding of Dvl-1 to Frat-1 and is essential for Wnt-3a-induced accumulation of beta-catenin. *J Biol Chem* *278*, 14066-14073.

Hoeven, T.A., de Vries, P.S., Dehghan, A., Franco, O.H., Hofman, A., Bindels, P.J., Bierma-Zeinstra, S.M., and Van Meurs, J.B. (2014). Is type 2 diabetes related to

osteoarthritis? a mendelian randomisation study that investigates glycaemic traits and osteoarthritis. *Osteoarthritis and Cartilage* 22, S229.

Hon, L.S., and Zhang, Z. (2007). The roles of binding site arrangement and combinatorial targeting in microRNA repression of gene expression. *Genome Biol* 8, R166.

Hong, E., and Reddi, A.H. (2013). Dedifferentiation and redifferentiation of articular chondrocytes from surface and middle zones: changes in microRNAs-221/-222, -140, and -143/145 expression. *Tissue Eng Part A* 19, 1015-1022.

Hornung, V., Guenther-Biller, M., Bourquin, C., Ablasser, A., Schlee, M., Uematsu, S., Noronha, A., Manoharan, M., Akira, S., de Fougerolles, A., *et al.* (2005). Sequence-specific potent induction of IFN- $\alpha$  by short interfering RNA in plasmacytoid dendritic cells through TLR7. *Nat Med* 11, 263-270.

Hotamisligil, G.S. (2006). Inflammation and metabolic disorders. *Nature* 444, 860-867.

Houbaviy, H.B., Murray, M.F., and Sharp, P.A. (2003). Embryonic stem cell-specific MicroRNAs. *Dev Cell* 5, 351-358.

Howell, P.M., Jr., Li, X., Riker, A.I., and Xi, Y. (2010). MicroRNA in Melanoma. *Ochsner J* 10, 83-92.

Hu, H.Y., Yan, Z., Xu, Y., Hu, H., Menzel, C., Zhou, Y.H., Chen, W., and Khaitovich, P. (2009). Sequence features associated with microRNA strand selection in humans and flies. *BMC Genomics* 10, 413.

Hu, K., Radhakrishnan, P., Patel, R.V., and Mao, J.J. (2001). Regional structural and viscoelastic properties of fibrocartilage upon dynamic nanoindentation of the articular condyle. *J Struct Biol* 136, 46-52.

Ikeda, T., Kamekura, S., Mabuchi, A., Kou, I., Seki, S., Takato, T., Nakamura, K., Kawaguchi, H., Ikegawa, S., and Chung, U.I. (2004). The combination of SOX5, SOX6, and SOX9 (the SOX trio) provides signals sufficient for induction of permanent cartilage. *Arthritis Rheum* 50, 3561-3573.

Iliopoulos, D., Malizos, K.N., Oikonomou, P., and Tsezou, A. (2008). Integrative microRNA and proteomic approaches identify novel osteoarthritis genes and their collaborative metabolic and inflammatory networks. *PLoS One* 3, e3740.

Jakkula, E., Melkonieni, M., Kiviranta, I., Lohiniva, J., Raina, S.S., Perala, M., Warman, M.L., Ahonen, K., Kroger, H., Goring, H.H., *et al.* (2005). The role of sequence variations within the genes encoding collagen II, IX and XI in non-syndromic, early-onset osteoarthritis. *Osteoarthritis Cartilage* 13, 497-507.

Jangra, R.K., Yi, M., and Lemon, S.M. (2010). Regulation of hepatitis C virus translation and infectious virus production by the microRNA miR-122. *J Virol* 84, 6615-6625.

Jayaprakash, A.D., Jabado, O., Brown, B.D., and Sachidanandam, R. (2011). Identification and remediation of biases in the activity of RNA ligases in small-RNA deep sequencing. *Nucleic Acids Res* 39, e141.

- Jones, S.E., and Jomary, C. (2002). Secreted Frizzled-related proteins: searching for relationships and patterns. *Bioessays* 24, 811-820.
- Jones, S.W., Watkins, G., Le Good, N., Roberts, S., Murphy, C.L., Brockbank, S.M., Needham, M.R., Read, S.J., and Newham, P. (2009). The identification of differentially expressed microRNA in osteoarthritic tissue that modulate the production of TNF-alpha and MMP13. *Osteoarthritis Cartilage* 17, 464-472.
- Jopling, C.L., Yi, M., Lancaster, A.M., Lemon, S.M., and Sarnow, P. (2005). Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. *Science* 309, 1577-1581.
- Joyce, C.E., Zhou, X., Xia, J., Ryan, C., Thrash, B., Menter, A., Zhang, W., and Bowcock, A.M. (2011). Deep sequencing of small RNAs from human skin reveals major alterations in the psoriasis miRNAome. *Hum Mol Genet* 20, 4025-4040.
- Kaneko, S., Satoh, T., Chiba, J., Ju, C., Inoue, K., and Kagawa, J. (2000). Interleukin-6 and interleukin-8 levels in serum and synovial fluid of patients with osteoarthritis. *Cytokines Cell Mol Ther* 6, 71-79.
- Kao, X.B., Gao, Y., Chen, J.H., Chen, Q., Wang, Z.L., and Wang, Z. (2013). [Role of JNK signaling pathway in chondrocyte apoptosis induced by nitric oxide]. *Zhonghua Lao Dong Wei Sheng Zhi Ye Bing Za Zhi* 31, 271-275.
- Kapinas, K., and Delany, A.M. (2011). MicroRNA biogenesis and regulation of bone remodeling. *Arthritis Res Ther* 13, 220.
- Kapoor, M., Martel-Pelletier, J., Lajeunesse, D., Pelletier, J.P., and Fahmi, H. (2011). Role of proinflammatory cytokines in the pathophysiology of osteoarthritis. *Nat Rev Rheumatol* 7, 33-42.
- Karlsen, T.A., Shahdadfar, A., and Brinckmann, J.E. (2011). Human primary articular chondrocytes, chondroblasts-like cells, and dedifferentiated chondrocytes: differences in gene, microRNA, and protein expression and phenotype. *Tissue Eng Part C Methods* 17, 219-227.
- Kashiwagi, M., Tortorella, M., Nagase, H., and Brew, K. (2001). TIMP-3 is a potent inhibitor of aggrecanase 1 (ADAM-TS4) and aggrecanase 2 (ADAM-TS5). *J Biol Chem* 276, 12501-12504.
- Kasprzyk, A. (2011). BioMart: driving a paradigm change in biological data management. *Database (Oxford)* 2011, bar049.
- Kato, M., de Lencastre, A., Pincus, Z., and Slack, F.J. (2009). Dynamic expression of small non-coding RNAs, including novel microRNAs and piRNAs/21U-RNAs, during *Caenorhabditis elegans* development. *Genome Biol* 10, R54.
- Kawamata, T., Yoda, M., and Tomari, Y. (2011). Multilayer checkpoints for microRNA authenticity during RISC assembly. *EMBO Rep* 12, 944-949.

- Kelleher, D.J., Karaoglu, D., Mandon, E.C., and Gilmore, R. (2003). Oligosaccharyltransferase isoforms that contain different catalytic STT3 subunits have distinct enzymatic properties. *Mol Cell* *12*, 101-111.
- Kempson, G.E. (1982). Relationship between the tensile properties of articular cartilage from the human knee and age. *Ann Rheum Dis* *41*, 508-511.
- Kerkhof, H.J., Meulenbelt, I., Akune, T., Arden, N.K., Aromaa, A., Bierma-Zeinstra, S.M., Carr, A., Cooper, C., Dai, J., Doherty, M., *et al.* (2011). Recommendations for standardization and phenotype definitions in genetic studies of osteoarthritis: the TREAT-OA consortium. *Osteoarthritis Cartilage* *19*, 254-264.
- Kertesz, M., Iovino, N., Unnerstall, U., Gaul, U., and Segal, E. (2007). The role of site accessibility in microRNA target recognition. *Nat Genet* *39*, 1278-1284.
- Kevorkian, L., Young, D.A., Darrah, C., Donell, S.T., Shepstone, L., Porter, S., Brockbank, S.M., Edwards, D.R., Parker, A.E., and Clark, I.M. (2004). Expression profiling of metalloproteinases and their inhibitors in cartilage. *Arthritis Rheum* *50*, 131-141.
- Kim, V.N. (2005). MicroRNA biogenesis: coordinated cropping and dicing. *Nat Rev Mol Cell Biol* *6*, 376-385.
- Kim, V.N., Han, J., and Siomi, M.C. (2009). Biogenesis of small RNAs in animals. *Nat Rev Mol Cell Biol* *10*, 126-139.
- Kim, Y.K., and Kim, V.N. (2007). Processing of intronic microRNAs. *EMBO J* *26*, 775-783.
- Kiss, A.M., Jady, B.E., Bertrand, E., and Kiss, T. (2004). Human box H/ACA pseudouridylation guide RNA machinery. *Mol Cell Biol* *24*, 5797-5807.
- Kiss, T. (2002). Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. *Cell* *109*, 145-148.
- Kobayashi, M., Squires, G.R., Mousa, A., Tanzer, M., Zukor, D.J., Antoniou, J., Feige, U., and Poole, A.R. (2005). Role of interleukin-1 and tumor necrosis factor alpha in matrix degradation of human osteoarthritic cartilage. *Arthritis Rheum* *52*, 128-135.
- Kobayashi, T., Lu, J., Cobb, B.S., Rodda, S.J., McMahon, A.P., Schipani, E., Merckenschlager, M., and Kronenberg, H.M. (2008). Dicer-dependent pathways regulate chondrocyte proliferation and differentiation. *Proc Natl Acad Sci U S A* *105*, 1949-1954.
- Kobe, B., and Deisenhofer, J. (1994). The leucine-rich repeat: a versatile binding motif. *Trends Biochem Sci* *19*, 415-421.
- Komori, T., Yagi, H., Nomura, S., Yamaguchi, A., Sasaki, K., Deguchi, K., Shimizu, Y., Bronson, R.T., Gao, Y.H., Inada, M., *et al.* (1997). Targeted disruption of *Cbfa1* results in a complete lack of bone formation owing to maturational arrest of osteoblasts. *Cell* *89*, 755-764.

- Kozomara, A., and Griffiths-Jones, S. (2011). miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res* 39, D152-157.
- Krek, A., Grun, D., Poy, M.N., Wolf, R., Rosenberg, L., Epstein, E.J., MacMenamin, P., da Piedade, I., Gunsalus, K.C., Stoffel, M., *et al.* (2005). Combinatorial microRNA target predictions. *Nat Genet* 37, 495-500.
- Kruger, J., and Rehmsmeier, M. (2006). RNAhybrid: microRNA target prediction easy, fast and flexible. *Nucleic Acids Res* 34, W451-454.
- Kuhn, D.E., Martin, M.M., Feldman, D.S., Terry, A.V., Jr., Nuovo, G.J., and Elton, T.S. (2008). Experimental validation of miRNA targets. *Methods* 44, 47-54.
- Kuivaniemi, H., Tromp, G., and Prockop, D.J. (1997). Mutations in fibrillar collagens (types I, II, III, and XI), fibril-associated collagen (type IX), and network-forming collagen (type X) cause a spectrum of diseases of bone, cartilage, and blood vessels. *Hum Mutat* 9, 300-315.
- Kumar, S., and Hedges, S.B. (2011). TimeTree2: species divergence times on the iPhone. *Bioinformatics* 27, 2023-2024.
- Kvist, A.J., Johnson, A.E., Morgelin, M., Gustafsson, E., Bengtsson, E., Lindblom, K., Aszodi, A., Fassler, R., Sasaki, T., Timpl, R., *et al.* (2006). Chondroitin sulfate perlecan enhances collagen fibril formation. Implications for perlecan chondrodysplasias. *J Biol Chem* 281, 33127-33139.
- Lalioti, M.D., Mirotsoy, M., Buresi, C., Peitsch, M.C., Rossier, C., Ouazzani, R., Baldy-Moulinier, M., Bottani, A., Malafosse, A., and Antonarakis, S.E. (1997). Identification of mutations in cystatin B, the gene responsible for the Unverricht-Lundborg type of progressive myoclonus epilepsy (EPM1). *Am J Hum Genet* 60, 342-351.
- Landgraf, P., Rusu, M., Sheridan, R., Sewer, A., Iovino, N., Aravin, A., Pfeffer, S., Rice, A., Kamphorst, A.O., Landthaler, M., *et al.* (2007). A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* 129, 1401-1414.
- Lark, M.W., Bayne, E.K., Flanagan, J., Harper, C.F., Hoerrner, L.A., Hutchinson, N.I., Singer, II, Donatelli, S.A., Weidner, J.R., Williams, H.R., *et al.* (1997). Aggrecan degradation in human cartilage. Evidence for both matrix metalloproteinase and aggrecanase activity in normal, osteoarthritic, and rheumatoid joints. *J Clin Invest* 100, 93-106.
- Leco, K.J., Waterhouse, P., Sanchez, O.H., Gowing, K.L., Poole, A.R., Wakeham, A., Mak, T.W., and Khokha, R. (2001). Spontaneous air space enlargement in the lungs of mice lacking tissue inhibitor of metalloproteinases-3 (TIMP-3). *J Clin Invest* 108, 817-829.
- Lee, I., Ajay, S.S., Yook, J.I., Kim, H.S., Hong, S.H., Kim, N.H., Dhanasekaran, S.M., Chinnaiyan, A.M., and Athey, B.D. (2009). New class of microRNA targets containing simultaneous 5'-UTR and 3'-UTR interaction sites. *Genome Res* 19, 1175-1183.
- Lee, R.B., and Urban, J.P. (1997). Evidence for a negative Pasteur effect in articular cartilage. *Biochem J* 321 ( Pt 1), 95-102.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S., *et al.* (2003). The nuclear RNase III Drosha initiates microRNA processing. *Nature* 425, 415-419.

Leighton, M.P., Nundlall, S., Starborg, T., Meadows, R.S., Suleman, F., Knowles, L., Wagener, R., Thornton, D.J., Kadler, K.E., Boot-Handford, R.P., *et al.* (2007). Decreased chondrocyte proliferation and dysregulated apoptosis in the cartilage growth plate are key features of a murine model of epiphyseal dysplasia caused by a *matn3* mutation. *Hum Mol Genet* 16, 1728-1741.

Lengner, C.J., Hassan, M.Q., Serra, R.W., Lepper, C., van Wijnen, A.J., Stein, J.L., Lian, J.B., and Stein, G.S. (2005). Nkx3.2-mediated repression of Runx2 promotes chondrogenic differentiation. *J Biol Chem* 280, 15872-15879.

Lestrade, L., and Weber, M.J. (2006). snoRNA-LBME-db, a comprehensive database of human H/ACA and C/D box snoRNAs. *Nucleic Acids Res* 34, D158-162.

Leung, A.K., and Sharp, P.A. (2010). MicroRNA functions in stress responses. *Mol Cell* 40, 205-215.

Li, J., Huang, J., Dai, L., Yu, D., Chen, Q., Zhang, X., and Dai, K. (2012a). miR-146a, an IL-1beta responsive miRNA, induces vascular endothelial growth factor and chondrocyte apoptosis by targeting Smad4. *Arthritis Res Ther* 14, R75.

Li, X., Gibson, G., Kim, J.S., Kroin, J., Xu, S., van Wijnen, A.J., and Im, H.J. (2011). MicroRNA-146a is linked to pain-related pathophysiology of osteoarthritis. *Gene* 480, 34-41.

Li, Y., Qiu, C., Tu, J., Geng, B., Yang, J., Jiang, T., and Cui, Q. (2014). HMDD v2.0: a database for experimentally supported human microRNA and disease associations. *Nucleic Acids Res* 42, D1070-1074.

Li, Y., Wang, H.Y., Wan, F.C., Liu, F.J., Liu, J., Zhang, N., Jin, S.H., and Li, J.Y. (2012b). Deep sequencing analysis of small non-coding RNAs reveals the diversity of microRNAs and piRNAs in the human epididymis. *Gene* 497, 330-335.

Liang, Z.J., Zhuang, H., Wang, G.X., Li, Z., Zhang, H.T., Yu, T.Q., and Zhang, B.D. (2012). MiRNA-140 is a negative feedback regulator of MMP-13 in IL-1beta-stimulated human articular chondrocyte C28/I2 cells. *Inflamm Res* 61, 503-509.

Lin, J., Zhang, W., Jones, A., and Doherty, M. (2004). Efficacy of topical non-steroidal anti-inflammatory drugs in the treatment of osteoarthritis: meta-analysis of randomised controlled trials. *BMJ* 329, 324.

Lin, L., Shen, Q., Zhang, C., Chen, L., and Yu, C. (2011). Assessment of the profiling microRNA expression of differentiated and dedifferentiated human adult articular chondrocytes. *J Orthop Res* 29, 1578-1584.

Linsen, S.E., de Wit, E., de Bruijn, E., and Cuppen, E. (2010). Small RNA expression and strain specificity in the rat. *BMC Genomics* 11, 249.

- Little, C.B., Barai, A., Burkhardt, D., Smith, S.M., Fosang, A.J., Werb, Z., Shah, M., and Thompson, E.W. (2009). Matrix metalloproteinase 13-deficient mice are resistant to osteoarthritic cartilage erosion but not chondrocyte hypertrophy or osteophyte development. *Arthritis Rheum* *60*, 3723-3733.
- Liu, C.G., Calin, G.A., Meloon, B., Gamliel, N., Sevignani, C., Ferracin, M., Dumitru, C.D., Shimizu, M., Zupo, S., Dono, M., *et al.* (2004). An oligonucleotide microchip for genome-wide microRNA profiling in human and mouse tissues. *Proc Natl Acad Sci U S A* *101*, 9740-9744.
- Liu, H.H., Tian, X., Li, Y.J., Wu, C.A., and Zheng, C.C. (2008). Microarray-based analysis of stress-regulated microRNAs in *Arabidopsis thaliana*. *RNA* *14*, 836-843.
- Liu, Q., Wang, F., and Axtell, M.J. (2014). Analysis of complementarity requirements for plant MicroRNA targeting using a *Nicotiana benthamiana* quantitative transient assay. *Plant Cell* *26*, 741-753.
- Liu, Z., Xu, J., Colvin, J.S., and Ornitz, D.M. (2002). Coordination of chondrogenesis and osteogenesis by fibroblast growth factor 18. *Genes Dev* *16*, 859-869.
- Lorenzo, P., Bayliss, M.T., and Heinegard, D. (1998). A novel cartilage protein (CILP) present in the mid-zone of human articular cartilage increases with age. *J Biol Chem* *273*, 23463-23468.
- Loughlin, J., Dowling, B., Mustafa, Z., and Chapman, K. (2002). Association of the interleukin-1 gene cluster on chromosome 2q13 with knee osteoarthritis. *Arthritis Rheum* *46*, 1519-1527.
- Lubberts, E., Joosten, L.A., van de Loo, F.A., van den Gersselaar, L.A., and van den Berg, W.B. (2000). Reduction of interleukin-17-induced inhibition of chondrocyte proteoglycan synthesis in intact murine articular cartilage by interleukin-4. *Arthritis Rheum* *43*, 1300-1306.
- Lumeng, C., Phelps, S., Crawford, G.E., Walden, P.D., Barald, K., and Chamberlain, J.S. (1999). Interactions between beta 2-syntrophin and a family of microtubule-associated serine/threonine kinases. *Nat Neurosci* *2*, 611-617.
- Luo, G., D'Souza, R., Hogue, D., and Karsenty, G. (1995). The matrix Gla protein gene is a marker of the chondrogenesis cell lineage during mouse development. *J Bone Miner Res* *10*, 325-334.
- Ma, M.T., He, M., Wang, Y., Jiao, X.Y., Zhao, L., Bai, X.F., Yu, Z.J., Wu, H.Z., Sun, M.L., Song, Z.G., *et al.* (2013). MiR-487a resensitizes mitoxantrone (MX)-resistant breast cancer cells (MCF-7/MX) to MX by targeting breast cancer resistance protein (BCRP/ABCG2). *Cancer Lett* *339*, 107-115.
- Mackie, E.J., Tatarczuch, L., and Mirams, M. (2011). The skeleton: a multi-functional complex organ: the growth plate chondrocyte and endochondral ossification. *J Endocrinol* *211*, 109-121.

- Mahmoodi, M., Sahebjam, S., Smookler, D., Khokha, R., and Mort, J.S. (2005). Lack of tissue inhibitor of metalloproteinases-3 results in an enhanced inflammatory response in antigen-induced arthritis. *Am J Pathol* 166, 1733-1740.
- Malemud, C.J. (2010). Anticytokine therapy for osteoarthritis: evidence to date. *Drugs Aging* 27, 95-115.
- Manavella, P.A., Koenig, D., and Weigel, D. (2012). Plant secondary siRNA production determined by microRNA-duplex structure. *Proc Natl Acad Sci U S A* 109, 2461-2466.
- Marques, J.T., and Williams, B.R. (2005). Activation of the mammalian immune system by siRNAs. *Nat Biotechnol* 23, 1399-1405.
- Martel-Pelletier, J., Boileau, C., Pelletier, J.P., and Roughley, P.J. (2008). Cartilage in normal and osteoarthritis conditions. *Best Pract Res Clin Rheumatol* 22, 351-384.
- Martel-Pelletier, J., Mineau, F., Jovanovic, D., Di Battista, J.A., and Pelletier, J.P. (1999). Mitogen-activated protein kinase and nuclear factor kappaB together regulate interleukin-17-induced nitric oxide production in human osteoarthritic chondrocytes: possible role of transactivating factor mitogen-activated protein kinase-activated protein kinase (MAPKAPK). *Arthritis Rheum* 42, 2399-2409.
- Martin, M.S., Van Sell, S., and Danter, J. (2012). Glucosamine and chondroitin: an appropriate adjunct treatment of symptomatic osteoarthritis of the knee. *Orthop Nurs* 31, 160-166.
- Martinez-Sanchez, A., Dudek, K.A., and Murphy, C.L. (2012). Regulation of human chondrocyte function through direct inhibition of cartilage master regulator SOX9 by microRNA-145 (miRNA-145). *J Biol Chem* 287, 916-924.
- Matera, A.G., Terns, R.M., and Terns, M.P. (2007). Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. *Nat Rev Mol Cell Biol* 8, 209-220.
- Matsuda, A., Ogawa, M., Yanai, H., Naka, D., Goto, A., Ao, T., Tanno, Y., Takeda, K., Watanabe, Y., Honda, K., *et al.* (2011). Generation of mice deficient in RNA-binding motif protein 3 (RBM3) and characterization of its role in innate immune responses and cell growth. *Biochem Biophys Res Commun* 411, 7-13.
- McAlinden, A., Varghese, N., Wirthlin, L., and Chang, L.W. (2013). Differentially Expressed MicroRNAs in Chondrocytes from Distinct Regions of Developing Human Cartilage. *PLoS One* 8, e75012.
- Medzhitov, R., Preston-Hurlburt, P., Kopp, E., Stadlen, A., Chen, C., Ghosh, S., and Janeway, C.A., Jr. (1998). MyD88 is an adaptor protein in the hToll/IL-1 receptor family signaling pathways. *Mol Cell* 2, 253-258.
- Meister, G., and Tuschl, T. (2004). Mechanisms of gene silencing by double-stranded RNA. *Nature* 431, 343-349.
- Mendias, C.L., Gumucio, J.P., and Lynch, E.B. (2012). Mechanical loading and TGF-beta change the expression of multiple miRNAs in tendon fibroblasts. *J Appl Physiol* (1985) 113, 56-62.



- Mendler, M., Eich-Bender, S.G., Vaughan, L., Winterhalter, K.H., and Bruckner, P. (1989). Cartilage contains mixed fibrils of collagen types II, IX, and XI. *J Cell Biol* 108, 191-197.
- Mestdagh, P., Lefever, S., Pattyn, F., Ridzon, D., Fredlund, E., Fieuw, A., Ongenaert, M., Vermeulen, J., De Paepe, A., Wong, L., *et al.* (2011). The microRNA body map: dissecting microRNA function through integrative genomics. *Nucleic Acids Res* 39, e136.
- Metsaranta, M., Garofalo, S., Decker, G., Rintala, M., de Crombrughe, B., and Vuorio, E. (1992). Chondrodysplasia in transgenic mice harboring a 15-amino acid deletion in the triple helical domain of pro alpha 1(II) collagen chain. *J Cell Biol* 118, 203-212.
- Mignery, G.A., and Sudhof, T.C. (1990). The ligand binding site and transduction mechanism in the inositol-1,4,5-triphosphate receptor. *EMBO J* 9, 3893-3898.
- Millward-Sadler, S.J., Wright, M.O., Lee, H., Caldwell, H., Nuki, G., and Salter, D.M. (2000). Altered electrophysiological responses to mechanical stimulation and abnormal signalling through alpha5beta1 integrin in chondrocytes from osteoarthritic cartilage. *Osteoarthritis Cartilage* 8, 272-278.
- Minina, E., Kreschel, C., Naski, M.C., Ornitz, D.M., and Vortkamp, A. (2002). Interaction of FGF, Ihh/Pthlh, and BMP signaling integrates chondrocyte proliferation and hypertrophic differentiation. *Dev Cell* 3, 439-449.
- Mitra, R., and Bandyopadhyay, S. (2011). MultiMiTar: a novel multi objective optimization based miRNA-target prediction method. *PLoS One* 6, e24583.
- Miyaki, S., Nakasa, T., Otsuki, S., Grogan, S.P., Higashiyama, R., Inoue, A., Kato, Y., Sato, T., Lotz, M.K., and Asahara, H. (2009). MicroRNA-140 is expressed in differentiated human articular chondrocytes and modulates interleukin-1 responses. *Arthritis Rheum* 60, 2723-2730.
- Miyaki, S., Sato, T., Inoue, A., Otsuki, S., Ito, Y., Yokoyama, S., Kato, Y., Takemoto, F., Nakasa, T., Yamashita, S., *et al.* (2010). MicroRNA-140 plays dual roles in both cartilage development and homeostasis. *Genes Dev* 24, 1173-1185.
- Miyamoto, Y., Shi, D., Nakajima, M., Ozaki, K., Sudo, A., Kotani, A., Uchida, A., Tanaka, T., Fukui, N., Tsunoda, T., *et al.* (2008). Common variants in DVWA on chromosome 3p24.3 are associated with susceptibility to knee osteoarthritis. *Nat Genet* 40, 994-998.
- Mobasher, A., Richardson, S., Mobasher, R., Shakibaei, M., and Hoyland, J.A. (2005). Hypoxia inducible factor-1 and facilitative glucose transporters GLUT1 and GLUT3: putative molecular components of the oxygen and glucose sensing apparatus in articular chondrocytes. *Histol Histopathol* 20, 1327-1338.
- Moren, A., Ichijo, H., and Miyazono, K. (1992). Molecular cloning and characterization of the human and porcine transforming growth factor-beta type III receptors. *Biochem Biophys Res Commun* 189, 356-362.

- Moretti, F., Thermann, R., and Hentze, M.W. (2010). Mechanism of translational regulation by miR-2 from sites in the 5' untranslated region or the open reading frame. *RNA* 16, 2493-2502.
- Morin, R.D., O'Connor, M.D., Griffith, M., Kuchenbauer, F., Delaney, A., Prabhu, A.L., Zhao, Y., McDonald, H., Zeng, T., Hirst, M., *et al.* (2008). Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Res* 18, 610-621.
- Mourelatos, Z., Dostie, J., Paushkin, S., Sharma, A., Charroux, B., Abel, L., Rappsilber, J., Mann, M., and Dreyfuss, G. (2002). miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. *Genes Dev* 16, 720-728.
- Moxon, S., Schwach, F., Dalmay, T., Maclean, D., Studholme, D.J., and Moulton, V. (2008). A toolkit for analysing large-scale plant small RNA datasets. *Bioinformatics* 24, 2252-2253.
- Mu, P., Nagahara, S., Makita, N., Tarumi, Y., Kadomatsu, K., and Takei, Y. (2009). Systemic delivery of siRNA specific to tumor mediated by atelocollagen: combined therapy using siRNA targeting Bcl-xL and cisplatin against prostate cancer. *Int J Cancer* 125, 2978-2990.
- Muller-Glauser, W., Humbel, B., Glatt, M., Strauli, P., Winterhalter, K.H., and Bruckner, P. (1986). On the role of type IX collagen in the extracellular matrix of cartilage: type IX collagen is localized to intersections of collagen fibrils. *J Cell Biol* 102, 1931-1939.
- Murdoch, A.D., Grady, L.M., Ablett, M.P., Katopodi, T., Meadows, R.S., and Hardingham, T.E. (2007). Chondrogenic differentiation of human bone marrow stem cells in transwell cultures: generation of scaffold-free cartilage. *Stem Cells* 25, 2786-2796.
- Nagae, M., Re, S., Mihara, E., Nogi, T., Sugita, Y., and Takagi, J. (2012). Crystal structure of alpha5beta1 integrin ectodomain: atomic details of the fibronectin receptor. *J Cell Biol* 197, 131-140.
- Nagase, H., and Kashiwagi, M. (2003). Aggrecanases and cartilage matrix degradation. *Arthritis Res Ther* 5, 94-103.
- Nagata, Y., Nakasa, T., Mochizuki, Y., Ishikawa, M., Miyaki, S., Shibuya, H., Yamasaki, K., Adachi, N., Asahara, H., and Ochi, M. (2009). Induction of apoptosis in the synovium of mice with autoantibody-mediated arthritis by the intraarticular injection of double-stranded MicroRNA-15a. *Arthritis Rheum* 60, 2677-2683.
- Nakamura, Y., Inloes, J.B., Katagiri, T., and Kobayashi, T. (2011). Chondrocyte-specific microRNA-140 regulates endochondral bone development and targets Dnpep to modulate bone morphogenetic protein signaling. *Mol Cell Biol* 31, 3019-3028.
- Nandakumar, J., and Shuman, S. (2004). How an RNA ligase discriminates RNA versus DNA damage. *Mol Cell* 16, 211-221.

Nguyen, B.V., Wang, Q.G., Kuiper, N.J., El Haj, A.J., Thomas, C.R., and Zhang, Z. (2010). Biomechanical properties of single chondrocytes and chondrons determined by micromanipulation and finite-element modelling. *J R Soc Interface* 7, 1723-1733.

NHS (accessed 03/04/14). NHS Choices: Osteoarthritis treatment. <http://www.nhs.uk/Conditions/Osteoarthritis/Pages/treatment.aspx>.

Nicolas, F.E., Hall, A.E., Csorba, T., Turnbull, C., and Dalmay, T. (2012). Biogenesis of Y RNA-derived small RNAs is independent of the microRNA pathway. *FEBS Lett* 586, 1226-1230.

Nicolas, F.E., Pais, H., Schwach, F., Lindow, M., Kauppinen, S., Moulton, V., and Dalmay, T. (2008). Experimental identification of microRNA-140 targets by silencing and overexpressing miR-140. *RNA* 14, 2513-2520.

Nishioka, K. (2004). Autoimmune response in cartilage-delivered peptides in a patient with osteoarthritis. *Arthritis Res Ther* 6, 6-7.

Nuttall, R.K., Pennington, C.J., Taplin, J., Wheal, A., Yong, V.W., Forsyth, P.A., and Edwards, D.R. (2003). Elevated membrane-type matrix metalloproteinases in gliomas revealed by profiling proteases and inhibitors in human cancer cells. *Mol Cancer Res* 1, 333-345.

Okuhara, A., Nakasa, T., Shibuya, H., Niimoto, T., Adachi, N., Deie, M., and Ochi, M. (2012). Changes in microRNA expression in peripheral mononuclear cells according to the progression of osteoarthritis. *Mod Rheumatol* 22, 446-457.

Oldershaw, R.A., Baxter, M.A., Lowe, E.T., Bates, N., Grady, L.M., Soncin, F., Brison, D.R., Hardingham, T.E., and Kimber, S.J. (2010). Directed differentiation of human embryonic stem cells toward chondrocytes. *Nat Biotechnol* 28, 1187-1194.

Olejniczak, M., Polak, K., Galka-Marciniak, P., and Krzyzosiak, W.J. (2011). Recent advances in understanding of the immunological off-target effects of siRNA. *Curr Gene Ther* 11, 532-543.

Olsen, B.R., Reginato, A.M., and Wang, W. (2000). Bone development. *Annu Rev Cell Dev Biol* 16, 191-220.

Ono, M., Yamada, K., Avolio, F., Scott, M.S., van Koningsbruggen, S., Barton, G.J., and Lamond, A.I. (2010). Analysis of human small nucleolar RNAs (snoRNA) and the development of snoRNA modulator of gene expression vectors. *Mol Biol Cell* 21, 1569-1584.

Orom, U.A., Nielsen, F.C., and Lund, A.H. (2008). MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. *Mol Cell* 30, 460-471.

Otte, P. (1991). Basic cell metabolism of articular cartilage. Manometric studies. *Z Rheumatol* 50, 304-312.

Ozsolak, F., Poling, L.L., Wang, Z., Liu, H., Liu, X.S., Roeder, R.G., Zhang, X., Song, J.S., and Fisher, D.E. (2008). Chromatin structure analyses identify miRNA promoters. *Genes Dev* 22, 3172-3183.

- Pages, H., Aboyoun, P., Gentleman, R., and DebRoy, S. Biostrings: String objects representing biological sequences, and matching algorithms. R package version 2.28.0.
- Pages, H., Aboyoun, P., Gentleman, R., and DebRoy, S. (2012). Biostrings: String objects representing biological sequences, and matching algorithms. R package version 2.28.0.
- Pais, H., Nicolas, F.E., Soond, S.M., Swingler, T.E., Clark, I.M., Chantry, A., Moulton, V., and Dalmay, T. (2010). Analyzing mRNA expression identifies Smad3 as a microRNA-140 target regulated only at protein level. *RNA* 16, 489-494.
- Pallares, I., Bonet, R., Garcia-Castellanos, R., Ventura, S., Aviles, F.X., Vendrell, J., and Gomis-Ruth, F.X. (2005). Structure of human carboxypeptidase A4 with its endogenous protein inhibitor, latexin. *Proc Natl Acad Sci U S A* 102, 3978-3983.
- Pando, R., Even-Zohar, N., Shtauf, B., Edry, L., Shomron, N., Phillip, M., and Gat-Yablonski, G. (2012). MicroRNAs in the growth plate are responsive to nutritional cues: association between miR-140 and SIRT1. *J Nutr Biochem* 23, 1474-1481.
- Park, C.Y., Jeker, L.T., Carver-Moore, K., Oh, A., Liu, H.J., Cameron, R., Richards, H., Li, Z., Adler, D., Yoshinaga, Y., *et al.* (2012). A resource for the conditional ablation of microRNAs in the mouse. *Cell Rep* 1, 385-391.
- Peters, L., and Meister, G. (2007). Argonaute proteins: mediators of RNA silencing. *Mol Cell* 26, 611-623.
- Pichlmair, A., Lassnig, C., Eberle, C.A., Gorna, M.W., Baumann, C.L., Burkard, T.R., Burckstummer, T., Stefanovic, A., Krieger, S., Bennett, K.L., *et al.* (2011). IFIT1 is an antiviral protein that recognizes 5'-triphosphate RNA. *Nat Immunol* 12, 624-630.
- Piecha, D., Weik, J., Kheil, H., Becher, G., Timmermann, A., Jaworski, A., Burger, M., and Hofmann, M.W. (2010). Novel selective MMP-13 inhibitors reduce collagen degradation in bovine articular and human osteoarthritis cartilage explants. *Inflamm Res* 59, 379-389.
- Pilote, J., Dupont-Versteegden, E.E., and Vanderklish, P.W. (2011). Widespread regulation of miRNA biogenesis at the Dicer step by the cold-inducible RNA-binding protein, RBM3. *PLoS One* 6, e28446.
- Plumb, D.A., Ferrara, L., Torbica, T., Knowles, L., Mironov, A., Jr., Kadler, K.E., Briggs, M.D., and Boot-Handford, R.P. (2011). Collagen XXVII organises the pericellular matrix in the growth plate. *PLoS One* 6, e29422.
- Poole, A.R., Guilak, F., and Abramson, S.B. (2007). Etiopathogenesis of osteoarthritis. *Osteoarthritis: Diagnosis and Medical/Surgical Management*, 27-49.
- Poole, A.R., Kobayashi, M., Yasuda, T., Lavery, S., Mwale, F., Kojima, T., Sakai, T., Wahl, C., El-Maadawy, S., Webb, G., *et al.* (2002). Type II collagen degradation and its regulation in articular cartilage in osteoarthritis. *Ann Rheum Dis* 61 Suppl 2, ii78-81.
- Poole, C.A. (1997). Articular cartilage chondrons: form, function and failure. *J Anat* 191 ( Pt 1), 1-13.

- Pratta, M.A., Yao, W., Decicco, C., Tortorella, M.D., Liu, R.Q., Copeland, R.A., Magolda, R., Newton, R.C., Trzaskos, J.M., and Arner, E.C. (2003). Aggrecan protects cartilage collagen from proteolytic cleavage. *J Biol Chem* 278, 45539-45545.
- Price, J.S., Waters, J.G., Darrah, C., Pennington, C., Edwards, D.R., Donell, S.T., and Clark, I.M. (2002). The role of chondrocyte senescence in osteoarthritis. *Aging Cell* 1, 57-65.
- R Core Team (2013). R: A language and environment for statistical computing. R Foundation for statistical computing, Vienna, Austria.
- Raia, R.M., Calin, G.A. (2011). Non-coding RNAs and cancer: microRNAs and beyond. *Journal of Nucleic Acids investigation* 2, 27-30.
- Ramos-Molina, B., Lambertos, A., Lopez-Contreras, A.J., and Penafiel, R. (2013). Mutational analysis of the antizyme-binding element reveals critical residues for the function of ornithine decarboxylase. *Biochim Biophys Acta* 1830, 5157-5165.
- Reginster, J.Y., Neuprez, A., Lecart, M.P., Sarlet, N., and Bruyere, O. (2012). Role of glucosamine in the treatment for osteoarthritis. *Rheumatol Int* 32, 2959-2967.
- Reigstad, L.J., Varhaug, J.E., and Lillehaug, J.R. (2005). Structural and functional specificities of PDGF-C and PDGF-D, the novel members of the platelet-derived growth factors family. *FEBS J* 272, 5723-5741.
- Reynolds, A., Anderson, E.M., Vermeulen, A., Fedorov, Y., Robinson, K., Leake, D., Karpilow, J., Marshall, W.S., and Khvorova, A. (2006). Induction of the interferon response by siRNA is cell type- and duplex length-dependent. *RNA* 12, 988-993.
- Rhee, D.K., Marcelino, J., Baker, M., Gong, Y., Smits, P., Lefebvre, V., Jay, G.D., Stewart, M., Wang, H., Warman, M.L., *et al.* (2005). The secreted glycoprotein lubricin protects cartilage surfaces and inhibits synovial cell overgrowth. *J Clin Invest* 115, 622-631.
- Robertson, C.M., Pennock, A.T., Harwood, F.L., Pomerleau, A.C., Allen, R.T., and Amiel, D. (2006). Characterization of pro-apoptotic and matrix-degradative gene expression following induction of osteoarthritis in mature and aged rabbits. *Osteoarthritis Cartilage* 14, 471-476.
- Rodrigo, I., Hill, R.E., Balling, R., Munsterberg, A., and Imai, K. (2003). Pax1 and Pax9 activate Bapx1 to induce chondrogenic differentiation in the sclerotome. *Development* 130, 473-482.
- Rollin, R., Marco, F., Jover, J.A., Garcia-Asenjo, J.A., Rodriguez, L., Lopez-Duran, L., and Fernandez-Gutierrez, B. (2008). Early lymphocyte activation in the synovial microenvironment in patients with osteoarthritis: comparison with rheumatoid arthritis patients and healthy controls. *Rheumatol Int* 28, 757-764.
- Rousseau, J., and Garnero, P. (2012). Biological markers in osteoarthritis. *Bone* 51, 265-277.

Rowan, A.D., Koshy, P.J., Shingleton, W.D., Degnan, B.A., Heath, J.K., Vernallis, A.B., Spaul, J.R., Life, P.F., Hudson, K., and Cawston, T.E. (2001). Synergistic effects of glycoprotein 130 binding cytokines in combination with interleukin-1 on cartilage collagen breakdown. *Arthritis Rheum* *44*, 1620-1632.

Ruegger, S., and Grosshans, H. (2012). MicroRNA turnover: when, how, and why. *Trends Biochem Sci* *37*, 436-446.

Sahebjam, S., Khokha, R., and Mort, J.S. (2007). Increased collagen and aggrecan degradation with age in the joints of Timp3(-/-) mice. *Arthritis Rheum* *56*, 905-909.

Sahni, M., Ambrosetti, D.C., Mansukhani, A., Gertner, R., Levy, D., and Basilico, C. (1999). FGF signaling inhibits chondrocyte proliferation and regulates bone development through the STAT-1 pathway. *Genes Dev* *13*, 1361-1366.

Sakkas, L.I., and Platsoucas, C.D. (2007). The role of T cells in the pathogenesis of osteoarthritis. *Arthritis Rheum* *56*, 409-424.

Sakurai, K., Amarzguioui, M., Kim, D.H., Alluin, J., Heale, B., Song, M.S., Gatignol, A., Behlke, M.A., and Rossi, J.J. (2011). A role for human Dicer in pre-RISC loading of siRNAs. *Nucleic Acids Res* *39*, 1510-1525.

Sandell, L.J., and Aigner, T. (2001). Articular cartilage and changes in arthritis. An introduction: cell biology of osteoarthritis. *Arthritis Res* *3*, 107-113.

Sandell, L.J., Nalin, A.M., and Reife, R.A. (1994). Alternative splice form of type II procollagen mRNA (IIA) is predominant in skeletal precursors and non-cartilaginous tissues during early mouse development. *Dev Dyn* *199*, 129-140.

Santibanez, J.F., Quintanilla, M., and Bernabeu, C. (2011). TGF-beta/TGF-beta receptor system and its role in physiological and pathological conditions. *Clin Sci (Lond)* *121*, 233-251.

Saugstad, J.A. (2010). MicroRNAs as effectors of brain function with roles in ischemia and injury, neuroprotection, and neurodegeneration. *J Cereb Blood Flow Metab* *30*, 1564-1576.

Sawitzke, A.D., Shi, H., Finco, M.F., Dunlop, D.D., Harris, C.L., Singer, N.G., Bradley, J.D., Silver, D., Jackson, C.G., Lane, N.E., *et al.* (2010). Clinical efficacy and safety of glucosamine, chondroitin sulphate, their combination, celecoxib or placebo taken to treat osteoarthritis of the knee: 2-year results from GAIT. *Ann Rheum Dis* *69*, 1459-1464.

Saxne, T., Lindell, M., Mansson, B., Petersson, I.F., and Heinegard, D. (2003). Inflammation is a feature of the disease process in early knee joint osteoarthritis. *Rheumatology (Oxford)* *42*, 903-904.

Scanzello, C.R., Plaas, A., and Crow, M.K. (2008). Innate immune system activation in osteoarthritis: is osteoarthritis a chronic wound? *Curr Opin Rheumatol* *20*, 565-572.

Scanzello, C.R., Umoh, E., Pessler, F., Diaz-Torne, C., Miles, T., Dicarolo, E., Potter, H.G., Mandl, L., Marx, R., Rodeo, S., *et al.* (2009). Local cytokine profiles in knee

osteoarthritis: elevated synovial fluid interleukin-15 differentiates early from end-stage disease. *Osteoarthritis Cartilage* 17, 1040-1048.

Schaffner, F., Ray, A.M., and Dontenwill, M. (2013). Integrin alpha5beta1, the Fibronectin Receptor, as a Pertinent Therapeutic Target in Solid Tumors. *Cancers (Basel)* 5, 27-47.

Schett, G., Kleyer, A., Perricone, C., Sahinbegovic, E., Iagnocco, A., Zwerina, J., Lorenzini, R., Aschenbrenner, F., Berenbaum, F., D'Agostino, M.A., *et al.* (2013). Diabetes is an independent predictor for severe osteoarthritis: results from a longitudinal cohort study. *Diabetes Care* 36, 403-409.

Schnall-Levin, M., Rissland, O.S., Johnston, W.K., Perrimon, N., Bartel, D.P., and Berger, B. (2011). Unusually effective microRNA targeting within repeat-rich coding regions of mammalian mRNAs. *Genome Res* 21, 1395-1403.

Schneider, G.F., and Dekker, C. (2012). DNA sequencing with nanopores. *Nat Biotechnol* 30, 326-328.

Schoggins, J.W., and Rice, C.M. (2011). Interferon-stimulated genes and their antiviral effector functions. *Curr Opin Virol* 1, 519-525.

Schrauder, M.G., Strick, R., Schulz-Wendtland, R., Strissel, P.L., Kahmann, L., Loehberg, C.R., Lux, M.P., Jud, S.M., Hartmann, A., Hein, A., *et al.* (2012). Circulating micro-RNAs as potential blood-based markers for early stage breast cancer detection. *PLoS One* 7, e29770.

Schultz, M., Jin, W., Waheed, A., Moed, B.R., Sly, W., and Zhang, Z. (2011). Expression profile of carbonic anhydrases in articular cartilage. *Histochem Cell Biol* 136, 145-151.

Schwarz, D.S., Hutvagner, G., Du, T., Xu, Z., Aronin, N., and Zamore, P.D. (2003). Asymmetry in the assembly of the RNAi enzyme complex. *Cell* 115, 199-208.

Scott, M.S., Avolio, F., Ono, M., Lamond, A.I., and Barton, G.J. (2009). Human miRNA precursors with box H/ACA snoRNA features. *PLoS Comput Biol* 5, e1000507.

Seitz, H., Royo, H., Bortolin, M.L., Lin, S.P., Ferguson-Smith, A.C., and Cavaille, J. (2004). A large imprinted microRNA gene cluster at the mouse *Dlk1-Gtl2* domain. *Genome Res* 14, 1741-1748.

Sellam, J., and Berenbaum, F. (2010). The role of synovitis in pathophysiology and clinical symptoms of osteoarthritis. *Nat Rev Rheumatol* 6, 625-635.

Sewer, A., Paul, N., Landgraf, P., Aravin, A., Pfeffer, S., Brownstein, M.J., Tuschl, T., van Nimwegen, E., and Zavolan, M. (2005). Identification of clustered microRNAs using an ab initio prediction method. *BMC Bioinformatics* 6, 267.

Shin, C. (2008). Cleavage of the star strand facilitates assembly of some microRNAs into Ago2-containing silencing complexes in mammals. *Mol Cells* 26, 308-313.

- Smits, P., Li, P., Mandel, J., Zhang, Z., Deng, J.M., Behringer, R.R., de Crombrughe, B., and Lefebvre, V. (2001). The transcription factors L-Sox5 and Sox6 are essential for cartilage formation. *Dev Cell* 1, 277-290.
- Sorefan, K., Pais, H., Hall, A.E., Kozomara, A., Griffiths-Jones, S., Moulton, V., and Dalmay, T. (2012). Reducing sequencing bias of small RNAs. *Silence* 3, 4.
- Sotoodehnejadnematlahi, F., and Burke, B. (2013). Structure, function and regulation of versican: the most abundant type of proteoglycan in the extracellular matrix. *Acta Med Iran* 51, 740-750.
- St-Jacques, B., Hammerschmidt, M., and McMahon, A.P. (1999). Indian hedgehog signaling regulates proliferation and differentiation of chondrocytes and is essential for bone formation. *Genes Dev* 13, 2072-2086.
- Stanton, H., Golub, S.B., Rogerson, F.M., Last, K., Little, C.B., and Fosang, A.J. (2011). Investigating ADAMTS-mediated aggrecanolysis in mouse cartilage. *Nat Protoc* 6, 388-404.
- 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., *et al.* (2005). ADAMTS5 is the major aggrecanase in mouse cartilage in vivo and in vitro. *Nature* 434, 648-652.
- Steck, E., Benz, K., Lorenz, H., Loew, M., Gress, T., and Richter, W. (2001). Chondrocyte expressed protein-68 (CEP-68), a novel human marker gene for cultured chondrocytes. *Biochem J* 353, 169-174.
- Steck, E., Boeuf, S., Gabler, J., Werth, N., Schnatzer, P., Diederichs, S., and Richter, W. (2012). Regulation of H19 and its encoded microRNA-675 in osteoarthritis and under anabolic and catabolic in vitro conditions. *J Mol Med (Berl)* 90, 1185-1195.
- Steck, E., Braun, J., Peltari, K., Kadel, S., Kalbacher, H., and Richter, W. (2007). Chondrocyte secreted CRTAC1: a glycosylated extracellular matrix molecule of human articular cartilage. *Matrix Biol* 26, 30-41.
- Steck, E., Maurer, A.N., Kalbacher, H., and Richter, W. (2006). Chondrocyte secreted CRTAC1: an extracellular matrix molecule of human articular cartilage facilitates cell adhesion. *European Cellsand Materials* 12, 9.
- Stennard, F.A., Holloway, A.F., Hamilton, J., and West, A.K. (1994). Characterisation of six additional human metallothionein genes. *Biochim Biophys Acta* 1218, 357-365.
- Stokes, D.G., Liu, G., Coimbra, I.B., Piera-Velazquez, S., Crawl, R.M., and Jimenez, S.A. (2002). Assessment of the gene expression profile of differentiated and dedifferentiated human fetal chondrocytes by microarray analysis. *Arthritis Rheum* 46, 404-419.
- Stokes, D.G., Liu, G., Dharmavaram, R., Hawkins, D., Piera-Velazquez, S., and Jimenez, S.A. (2001). Regulation of type-II collagen gene expression during human chondrocyte de-differentiation and recovery of chondrocyte-specific phenotype in culture involves Sry-type high-mobility-group box (SOX) transcription factors. *Biochem J* 360, 461-470.



- Sun, G., Wu, X., Wang, J., Li, H., Li, X., Gao, H., Rossi, J., and Yen, Y. (2011a). A bias-reducing strategy in profiling small RNAs using Solexa. *RNA* *17*, 2256-2262.
- Sun, J., Zhong, N., Li, Q., Min, Z., Zhao, W., Sun, Q., Tian, L., Yu, H., Shi, Q., Zhang, F., *et al.* (2011b). MicroRNAs of rat articular cartilage at different developmental stages identified by Solexa sequencing. *Osteoarthritis Cartilage* *19*, 1237-1245.
- Suomi, S., Taipaleenmaki, H., Seppanen, A., Ripatti, T., Vaananen, K., Hentunen, T., Saamanen, A.M., and Laitala-Leinonen, T. (2008). MicroRNAs regulate osteogenesis and chondrogenesis of mouse bone marrow stromal cells. *Gene Regul Syst Bio* *2*, 177-191.
- Swales, C., and Athanasou, N.A. (2010). The pathobiology of osteoarthritis. *Orthopaedics and Trauma* *26*, 39-404.
- Swingler, T.E., Waters, J.G., Davidson, R.K., Pennington, C.J., Puente, X.S., Darrah, C., Cooper, A., Donell, S.T., Guile, G.R., Wang, W., *et al.* (2009). Degradome expression profiling in human articular cartilage. *Arthritis Res Ther* *11*, R96.
- Swingler, T.E., Wheeler, G., Carmont, V., Elliott, H.R., Barter, M.J., Abu-Elmagd, M., Donell, S.T., Boot-Handford, R.P., Hajihosseini, M.K., Munsterberg, A., *et al.* (2012). The expression and function of microRNAs in chondrogenesis and osteoarthritis. *Arthritis Rheum* *64*, 1909-1919.
- Takada, Y., Ye, X., and Simon, S. (2007). The integrins. *Genome Biol* *8*, 215.
- Tanco, S., Zhang, X., Morano, C., Aviles, F.X., Lorenzo, J., and Fricker, L.D. (2010). Characterization of the substrate specificity of human carboxypeptidase A4 and implications for a role in extracellular peptide processing. *J Biol Chem* *285*, 18385-18396.
- Tang, G., Reinhart, B.J., Bartel, D.P., and Zamore, P.D. (2003). A biochemical framework for RNA silencing in plants. *Genes Dev* *17*, 49-63.
- Tardif, G., Hum, D., Pelletier, J.P., Duval, N., and 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 Musculoskelet Disord* *10*, 148.
- Tazi-Ahnini, R., di Giovine, F.S., McDonagh, A.J., Messenger, A.G., Amadou, C., Cox, A., Duff, G.W., and Cork, M.J. (2000). Structure and polymorphism of the human gene for the interferon-induced p78 protein (MX1): evidence of association with alopecia areata in the Down syndrome region. *Hum Genet* *106*, 639-645.
- Tchetina, E.V. (2011). Developmental mechanisms in articular cartilage degradation in osteoarthritis. *Arthritis* *2011*, 683970.
- Terenzi, F., Hui, D.J., Merrick, W.C., and Sen, G.C. (2006). Distinct induction patterns and functions of two closely related interferon-inducible human genes, ISG54 and ISG56. *J Biol Chem* *281*, 34064-34071.
- Thakur, M., Dawes, J.M., and McMahon, S.B. (2013). Genomics of pain in osteoarthritis. *Osteoarthritis Cartilage* *21*, 1374-1382.

- Thomson, D.W., Bracken, C.P., and Goodall, G.J. (2011). Experimental strategies for microRNA target identification. *Nucleic Acids Res* 39, 6845-6853.
- Tomari, Y., Matranga, C., Haley, B., Martinez, N., and Zamore, P.D. (2004). A protein sensor for siRNA asymmetry. *Science* 306, 1377-1380.
- Treiber, T., Treiber, N., and Meister, G. (2012). Regulation of microRNA biogenesis and function. *Thromb Haemost* 107, 605-610.
- Troeberg, L., and Nagase, H. (2012). Proteases involved in cartilage matrix degradation in osteoarthritis. *Biochim Biophys Acta* 1824, 133-145.
- Tuddenham, L., Wheeler, G., Ntounia-Fousara, S., Waters, J., Hajihosseini, M.K., Clark, I., and Dalmay, T. (2006). The cartilage specific microRNA-140 targets histone deacetylase 4 in mouse cells. *FEBS Lett* 580, 4214-4217.
- Uechi, T., Tanaka, T., and Kenmochi, N. (2001). A complete map of the human ribosomal protein genes: assignment of 80 genes to the cytogenetic map and implications for human disorders. *Genomics* 72, 223-230.
- Umate, P., and Tuteja, N. (2010). microRNA access to the target helicases from rice. *Plant Signal Behav* 5, 1171-1175.
- Umlauf, D., Frank, S., Pap, T., and Bertrand, J. (2010). Cartilage biology, pathology, and repair. *Cell Mol Life Sci* 67, 4197-4211.
- Vagin, V.V., Sigova, A., Li, C., Seitz, H., Gvozdev, V., and Zamore, P.D. (2006). A distinct small RNA pathway silences selfish genetic elements in the germline. *Science* 313, 320-324.
- van der Kraan, P.M., Blaney Davidson, E.N., and van den Berg, W.B. (2010). A role for age-related changes in TGFbeta signaling in aberrant chondrocyte differentiation and osteoarthritis. *Arthritis Res Ther* 12, 201.
- van Rooij, E. (2011). The art of microRNA research. *Circ Res* 108, 219-234.
- van Rooij, E., Sutherland, L.B., Qi, X., Richardson, J.A., Hill, J., and Olson, E.N. (2007). Control of stress-dependent cardiac growth and gene expression by a microRNA. *Science* 316, 575-579.
- van Rossum, D.B., Patterson, R.L., Cheung, K.H., Barrow, R.K., Syrovatkina, V., Gessell, G.S., Burkholder, S.G., Watkins, D.N., Foskett, J.K., and Snyder, S.H. (2006). DANGER, a novel regulatory protein of inositol 1,4,5-trisphosphate-receptor activity. *J Biol Chem* 281, 37111-37116.
- Vega, R.B., Matsuda, K., Oh, J., Barbosa, A.C., Yang, X., Meadows, E., McAnally, J., Pomajzl, C., Shelton, J.M., Richardson, J.A., *et al.* (2004). Histone deacetylase 4 controls chondrocyte hypertrophy during skeletogenesis. *Cell* 119, 555-566.
- Visse, R., and Nagase, H. (2003). Matrix metalloproteinases and tissue inhibitors of metalloproteinases: structure, function, and biochemistry. *Circ Res* 92, 827-839.

- Viswanathan, S.R., Daley, G.Q., and Gregory, R.I. (2008). Selective blockade of microRNA processing by Lin28. *Science* 320, 97-100.
- von der Mark, K., Kirsch, T., Nerlich, A., Kuss, A., Weseloh, G., Gluckert, K., and Stoss, H. (1992). Type X collagen synthesis in human osteoarthritic cartilage. Indication of chondrocyte hypertrophy. *Arthritis Rheum* 35, 806-811.
- Wager-Smith, K., and Kay, S.A. (2000). Circadian rhythm genetics: from flies to mice to humans. *Nat Genet* 26, 23-27.
- Walker, G.D., Fischer, M., Gannon, J., Thompson, R.C., Jr., and Oegema, T.R., Jr. (1995). Expression of type-X collagen in osteoarthritis. *J Orthop Res* 13, 4-12.
- Wang, P., Fu, T., Wang, X., and Zhu, W. (2010). [Primary, study of miRNA expression patterns in laryngeal carcinoma by microarray]. *Lin Chung Er Bi Yan Hou Tou Jing Wai Ke Za Zhi* 24, 535-538.
- Wang, Z., Kim, J.H., Higashino, K., Kim, S.S., Wang, S., Seki, S., Hutton, W.C., and Yoon, S.T. (2012). Cartilage intermediate layer protein (CILP) regulation in intervertebral discs. The effect of age, degeneration, and bone morphogenetic protein-2. *Spine (Phila Pa 1976)* 37, E203-208.
- Wang, Z., Weitzmann, M.N., Sangadala, S., Hutton, W.C., and Yoon, S.T. (2013). Link protein N-terminal peptide binds to bone morphogenetic protein (BMP) type II receptor and drives matrix protein expression in rabbit intervertebral disc cells. *J Biol Chem* 288, 28243-28253.
- Wardale, R.J., and Duance, V.C. (1994). Characterisation of articular and growth plate cartilage collagens in porcine osteochondrosis. *J Cell Sci* 107 ( Pt 1), 47-59.
- Watanabe, Y., and Kanai, A. (2011). Systems Biology Reveals MicroRNA-Mediated Gene Regulation. *Front Genet* 2, 29.
- Waterhouse, P.M., Wang, M.B., and Lough, T. (2001). Gene silencing as an adaptive defence against viruses. *Nature* 411, 834-842.
- Weber, B.H., Vogt, G., Pruett, R.C., Stohr, H., and Felbor, U. (1994). Mutations in the tissue inhibitor of metalloproteinases-3 (TIMP3) in patients with Sorsby's fundus dystrophy. *Nat Genet* 8, 352-356.
- Weiland, M., Gao, X.H., Zhou, L., and Mi, Q.S. (2012). Small RNAs have a large impact: Circulating microRNAs as biomarkers for human diseases. *RNA Biology* 9, 850 - 859.
- Westholm, J.O., and Lai, E.C. (2011). Mirtrons: microRNA biogenesis via splicing. *Biochimie* 93, 1897-1904.
- Wiberg, C., Klatt, A.R., Wagener, R., Paulsson, M., Bateman, J.F., Heinegard, D., and Morgelin, M. (2003). Complexes of matrilin-1 and biglycan or decorin connect collagen VI microfibrils to both collagen II and aggrecan. *J Biol Chem* 278, 37698-37704.

Wienholds, E., Kloosterman, W.P., Miska, E., Alvarez-Saavedra, E., Berezikov, E., de Bruijn, E., Horvitz, H.R., Kauppinen, S., and Plasterk, R.H. (2005). MicroRNA expression in zebrafish embryonic development. *Science* 309, 310-311.

Wienholds, E., Koudijs, M.J., van Eeden, F.J., Cuppen, E., and Plasterk, R.H. (2003). The microRNA-producing enzyme Dicer1 is essential for zebrafish development. *Nat Genet* 35, 217-218.

Wiesler, E.R., Shen, J., and Papadonikolakis, A. (2007). Injuries of the Wrist and Hand. *Practical Orthopaedic Sports Medicine & Arthroscopy*.

Wildi, L.M., Raynauld, J.P., Martel-Pelletier, J., Beaulieu, A., Bessette, L., Morin, F., Abram, F., Dorais, M., and Pelletier, J.P. (2011). Chondroitin sulphate reduces both cartilage volume loss and bone marrow lesions in knee osteoarthritis patients starting as early as 6 months after initiation of therapy: a randomised, double-blind, placebo-controlled pilot study using MRI. *Ann Rheum Dis* 70, 982-989.

Willmann, M.R., Mehalick, A.J., Packer, R.L., and Jenik, P.D. (2011). MicroRNAs regulate the timing of embryo maturation in Arabidopsis. *Plant Physiol* 155, 1871-1884.

Winter, J., Jung, S., Keller, S., Gregory, R.I., and Diederichs, S. (2009). Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nat Cell Biol* 11, 228-234.

Wu, C.W., Tchetina, E.V., Mwale, F., Hasty, K., Pidoux, I., Reiner, A., Chen, J., Van Wart, H.E., and Poole, A.R. (2002). Proteolysis involving matrix metalloproteinase 13 (collagenase-3) is required for chondrocyte differentiation that is associated with matrix mineralization. *J Bone Miner Res* 17, 639-651.

Wu, Q., Lu, Z., Li, H., Lu, J., Guo, L., and Ge, Q. (2011). Next-generation sequencing of microRNAs for breast cancer detection. *J Biomed Biotechnol* 2011, 597145.

Xu, C.Z., Xie, J., Jin, B., Chen, X.W., Sun, Z.F., Wang, B.X., and Dong, P. (2013). Gene and microRNA expression reveals sensitivity to paclitaxel in laryngeal cancer cell line. *Int J Clin Exp Pathol* 6, 1351-1361.

Yamasaki, K., Nakasa, T., Miyaki, S., Ishikawa, M., Deie, M., Adachi, N., Yasunaga, Y., Asahara, H., and Ochi, M. (2009). Expression of MicroRNA-146a in osteoarthritis cartilage. *Arthritis Rheum* 60, 1035-1041.

Yan, C., Wang, Y., Shen, X.Y., Yang, G., Jian, J., Wang, H.S., Chen, G.Q., and Wu, Q. (2011). MicroRNA regulation associated chondrogenesis of mouse MSCs grown on polyhydroxyalkanoates. *Biomaterials* 32, 6435-6444.

Yang, B., Guo, H., Zhang, Y., Chen, L., Ying, D., and Dong, S. (2011a). MicroRNA-145 regulates chondrogenic differentiation of mesenchymal stem cells by targeting Sox9. *PLoS One* 6, e21679.

Yang, J., Qin, S., Yi, C., Ma, G., Zhu, H., Zhou, W., Xiong, Y., Zhu, X., Wang, Y., He, L., *et al.* (2011b). MiR-140 is co-expressed with Wwp2-C transcript and activated by Sox9 to target Sp1 in maintaining the chondrocyte proliferation. *FEBS Lett* 585, 2992-2997.

- Yang, X., Chen, L., Xu, X., Li, C., Huang, C., and Deng, C.X. (2001). TGF-beta/Smad3 signals repress chondrocyte hypertrophic differentiation and are required for maintaining articular cartilage. *J Cell Biol* 153, 35-46.
- Yi, R., Qin, Y., Macara, I.G., and Cullen, B.R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev* 17, 3011-3016.
- Yin, S., Ho, C.K., and Shuman, S. (2003). Structure-function analysis of T4 RNA ligase 2. *J Biol Chem* 278, 17601-17608.
- Zeggini, E., Panoutsopoulou, K., Southam, L., Rayner, N.W., Day-Williams, A.G., Lopes, M.C., Boraska, V., Esko, T., Evangelou, E., Hoffman, A., *et al.* (2012). Identification of new susceptibility loci for osteoarthritis (arcOGEN): a genome-wide association study. *Lancet* 380, 815-823.
- Zhang, B., Pan, X., Cannon, C.H., Cobb, G.P., and Anderson, T.A. (2006a). Conservation and divergence of plant microRNA genes. *Plant J* 46, 243-259.
- Zhang, H., Kolb, F.A., Brondani, V., Billy, E., and Filipowicz, W. (2002). Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. *EMBO J* 21, 5875-5885.
- Zhang, M., Liu, L., Xiao, T., and Guo, W. (2012a). [Detection of the expression level of miR-140 using realtime fluorescent quantitative PCR in knee synovial fluid of osteoarthritis patients]. *Zhong Nan Da Xue Xue Bao Yi Xue Ban* 37, 1210-1214.
- Zhang, R., Murakami, S., Coustry, F., Wang, Y., and de Crombrughe, B. (2006b). Constitutive activation of MKK6 in chondrocytes of transgenic mice inhibits proliferation and delays endochondral bone formation. *Proc Natl Acad Sci U S A* 103, 365-370.
- Zhang, Z., Kang, Y., Zhang, H., Duan, X., Liu, J., Li, X., and Liao, W. (2012b). Expression of microRNAs during chondrogenesis of human adipose-derived stem cells. *Osteoarthritis Cartilage* 20, 1638-1646.
- Zhang, Z., Qin, Y.W., Brewer, G., and Jing, Q. (2012c). MicroRNA degradation and turnover: regulating the regulators. *Wiley Interdiscip Rev RNA* 3, 593-600.
- Zhou, S., Cui, Z., and Urban, J.P. (2004). Factors influencing the oxygen concentration gradient from the synovial surface of articular cartilage to the cartilage-bone interface: a modeling study. *Arthritis Rheum* 50, 3915-3924.
- Zhuo, Q., Yang, W., Chen, J., and Wang, Y. (2012). Metabolic syndrome meets osteoarthritis. *Nat Rev Rheumatol* 8, 729-737.
- Zinovyev, A., Morozova, N., Nonne, N., Barillot, E., Harel-Bellan, A., and Gorban, A.N. (2010). Dynamical modeling of microRNA action on the protein translation process. *BMC Syst Biol* 4, 13.