THE ROLE OF OSTEOARTHRITIS REGULATED microRNAs

IN SKELETAL DEVELOPMENT PATHWAYS

Steven Woods

Thesis submitted for Doctor of Philosophy

Institute of Cellular Medicine

July 2013

Abstract

Cellular changes occur during osteoarthritis (OA), which lead to an alteration in phenotype of the resident cartilage cell, the chondrocyte, and subsequent destruction of the tissue. Many of these changes are unknown. I hypothesise these changes may be due in part to microRNAs (miRNAs), small non-coding RNAs that regulate the expression of a discrete repertoire of genes through base-specific interactions within the target genes 3'untranslated region. A previous screen in our laboratory has identified a number of miRNAs differentially expressed in OA cartilage.

The objectives of this study were; (1) to investigate the genetic association of the most extensively studied cartilage miRNA (miR-140), and its targets, with OA, and (2) to assess the function of other, differentially expressed, and less well studied miRNAs, in development signalling pathways, namely miR-125b and miR-324-5p, but focusing on miR-324-5p.

Here I identified; (1) SNPs within the miR-140 locus and a predicted, OA-associated, miR-140 target, that may affect function; and (2) miR-125b and miR-324-5p as regulators of Hedgehog (Hh) signalling, likely to play a role in skeletal development. miR-324-5p regulates Hh signalling in human and mouse, yet the mechanism appears unconserved. In humans, miR-324-5p targets SMO and GLI1. Using Stable Isotope Labelling with Amino acids in Cell culture (SILAC) mass spectrometry and whole-genome microarrays, I identified novel miR-324-5p targets, and validated Glypican1 (Gpc1) as a direct target of miR-324-5p and a regulator of Hh signalling in mouse. In addition to regulation of Hh, miR-324-5p regulates Wnt signalling, in which it forms a negative feedback loop.

Together, this body of work demonstrates how miRNAs, their targets and their function can be linked in their expression and association with OA.

Acknowledgements

Contents

Figures

xii

Tables

List of abbreviations

ADAMTS	A Disintegrin And Metalloproteinase with Thrombospondin Motifs
AGEs	Advanced glycation end products
APP	Amyloid precursor protein
APS	Ammonium persulphate
BMPs	Bone morphogenic proteins
BSA	Bovine serum albumin
CDS	Coding sequence
ceRNAs	Competing endogenous RNAs
Chr	Chromosome
COX2	Cyclooxygenase 2
dH2O	Distilled H2O
DF	Dharmafect transfection reagent
Dhh	Desert hedgehog
dlp	Dally-like protein
DMEM	Dulbecco's modified Eagle's medium
DMSO	Dimethyl sulphoxide
Dvl2	Dishevelled
ECM	Extracellular matrix
FBS	Foetal bovine serum
FGF	Fibroblast growth factor
FHD	Fugene HD transfection reagent
FZD	Frizzled
GAG	Glycosaminoglycans
GHCl	Guanidinium hydrochloride
GO	Gene ontlogy
Gpc1	Glypican1
GPI	Glycophosphatidylinositol
GSK3	Glycogen synthase kinase 3
GWAS	Genome wide association scans
HAC	Human articular chondrocytes
HCV	Hepatitis C virus
Hdac4	Histone deacetylase 4
Hh	Hedgehog
HP	Hairpin
HRP	Horseradish peroxidase
HS	Heparin sulphate
Hspg2	Perlecan
HSPGs	Heparin sulphate proteoglycans
IFT	Intra flagella transport
IHH	Indian hedgehog
LC-MS/MS	Liquid chromatography tandem mass spectrometry,
LD	Linkage disequilibrium
LEF	Lymphoid-binding factor

LNA	Locked nucleic acid
LRP	Low density lipoprotein
MEM	Minimum essential medium
miRNAs	microRNAs
MMPs	Matrix metalloproteinases
MPs	Metalloproteinases
MS	Mass spectrometry
MT	Mutant
NOF	Neck of femur
NSAIDS	Non steroidal anti-inflammatory drugs
OA	Osteoarthritis
OE	Overexpression
OR	Odds ratio
PBS	Phosphate buffered saline
PCP	Planar cell polarity pathway
PKA	Protein kinase A
Ptch1	Patched
PTHrP	Parathyroid hormone related peptide
Pur	Purmorphamine
RA	Retinoic acid
RFLP	Restriction Fragment Length Polymorphism
SAG	Smoothened agonist
SC	Sodium chlorate
Sdc1-4	Syndecans
Shh	Sonic hedgehog
SILAC	Stable Isotope Labelling with Amino acids in Cell culture
siRNA	Small interfering RNA
Smo	Smoothened
SNPs	Single nucleotide polymorphisms
TAE	Tris-acetate-ethylenediamine tetraacetic acid
TCF	T-cell factors
TEMED	N,N',N'-Tetramethylethylenediamine
TGE	TaqMan® Gene expression mix
TGF-β	Transforming growth factor β
TIMPs	Tissue inhibitors of metalloproteinases
TLRs	Toll-like receptors
UTR	Untranslated region
WNTs	Wingless-type MMTV integration site family members
WT	Wild-type

Chapter 1 Introduction

1.1   Osteoarthritis

Arthritis is a disease of the joint which involves reduced joint function and pain. Osteoarthritis (OA) is the most common form of arthritis, with the majority of over 65 year olds showing radiographic evidence of the disease [1]. OA can be divided into two main groups; primary OA, which has no known cause and mainly effects older people, and secondary OA, which has a known cause such as trauma or developmental defects, and tends to affect younger people than primary OA [2]. OA can affect joints such as the hip, knee, spine and hands, with each having slightly different disease characteristics, and as such can be considered as separate diseases. There are a number of risk factors for OA development such as age, obesity, mechanical injury and genetics [3].

OA involves joint space narrowing, subchondral bone thickening, subchondral cyst formation, osteophyte formation, cartilage calcification, vascular invasion and changes in ligaments, muscles, nerves, meniscus and synovium, but is epitomized by the loss of articular cartilage [1, 4]. OA is a major cause of pain and disability and can lead to social isolation, depression [5], and can even increase the risk of mortality [6]. The number of people with OA is set to rise due to obesity and the aging population, increasing the burden on society which currently requires 1-2.5% of gross national product [7].

There are no OA disease modifying drugs. Currently management for OA involves exercise, lifestyle changes, non steroidal anti-inflammatory drugs (NSAIDS) and analgesics. NSAIDS which selectively target cyclooxygenase 2 (COX2), an enzyme responsible for mediating inflammation and pain, have been used in the treatment of OA and were effective in reducing pain and inflammation, but do not target the cause or progression of the disease [8]. In 2004, one of these drugs was withdrawn from use due to increased risk of heart attack and stroke [9]. In addition to NSAIDs, injections of corticosteroids, MMP inhibitors [10], IL-1β inhibitors [11] and glucosamine and chondroitin sulfate therapeutic agents [12] have been used treat OA but with mixed results [13]. More recently, autologous implants (where mesenchymal cells are extracted from bone marrow and differentiated *in vitro* are

implanted back into the patient) have been used to treat secondary OA, but again this approach has had mixed results [14]. As there are currently no OA disease modifying drugs it means joint replacement surgery and pain management are often the only option for patients suffering from OA [15].

1.2 The joint structure

There are a number of tissues which make up the joint including articular cartilage, synovial fluid, tendon, ligament and subchondral bone. In the knee there is also fat pad, menisci and the patella. Surrounding the joint is the fibrous capsule and synovial membrane. All these tissues are essential for the pain free and frictionless articulation of joints. The tissues are all linked, with damage in one adversely affecting the others, all having the potential to contribute to the pain and joint destruction seen in OA.

1.2.1 Articular cartilage

Articular cartilage covers the ends of long bones, is load bearing, essential for smooth pain free joint movement [16], and maintained by a single cell type, the chondrocyte [17]. In healthy cartilage the matrix is constantly being turned over by the chondrocyte, in a balance of catabolic and anabolic activity. Articular cartilage is avascular, alymphatic and aneuronal, meaning the chondrocyte obtains its nutrients from the surrounding synovial fluid. The cartilage extracellular matrix (ECM) predominately consists of proteoglycan and collagen. Aggrecan is the most abundant proteoglycan in the cartilage, it consists of a protein core with three globular domains (G1, G2 and G3) and glycosaminoglycan (GAG) chains. G1 binds to hyaluronic acid (via link protein) and there are many GAG chains between G2 and G3, which consist of chondroitin sulphate and keratin sulphate [18]. The GAG on aggrecan is negatively charged which attracts water and is surrounded by a collagen network. The most abundant collagen in articular cartilage is type II collagen. The combination of the hydrostatic pressure created by aggrecan and structure of collagen means the cartilage is firm and resists compression.

The articular cartilage can be divided into four zones; the superficial zone, which has the most organised and highest collagen content, providing a smooth surface for articulation [19]; the middle zone, which is less organised; the deep zone, which has a high

proteoglycan content; and the calcified zone, which is a layer of calcified cartilage that lies directly on the subchondral bone.

1.2.2   Subchondral bone

The subchondral bone can be split into two zones, the subchondral bone plate which lies directly beneath the calcified zone of articular cartilage, and the trabecular bone which is beneath the subchondral bone plate. Both are primarily composed of collagen type I and mineralised with inorganic crystals of hydroxyapatite. The two main types of cell within the bone are osteoblasts which form bone, and osteoclasts which resorb bone.

1.2.3   Joint capsule, synovial membrane and synovial fluid

Surrounding the joint is the joint capsule. The joint capsule is lined by the synovial membrane, which is responsible for fighting infection and secreting synovial fluid. The synovial fluid provides the chondrocyte with nutrients and also contains lubricin to reduce joint friction, and mesenchymal progenitor cells which are prepared to mediate joint repair.

Other joint tissues include; the fat pad which acts as a cushion and is a source of adipokines and mesenchymal progenitor cells; the menisci, which are made from fibrous cartilage and maintain joint stability; ligaments, which are mainly composed of type I collagen and attach bone to bone to also stabilise the joint; and the tendons, which are also predominately composed of type I collagen and attach muscle to bone to allow for joint movement.

1.3   Joint maintenance

A number of factors can influence the activity of the chondrocyte including growth factors and mechanical loading, which can both potentiate each other in their activation of matrix degrading enzymes [20]. This means the response to these factors needs to be finely regulated in order maintain the anabolic and catabolic balance of the chondrocyte, and ultimately maintain healthy cartilage.

Metalloproteinases (MPs) are a subtype of protease [21]. Matrix metalloproteinases (MMPs) and A Disintegrin And Metalloproteinase with Thrombospondin Motifs 4 and 5 (ADAMTS-4 and 5) are both subtypes of MPs [21]. MPs are essential during skeletal

development and joint maintenance [16]. Different MPs can cleave different parts of the ECM, collagenases such as MMP1, 8 and 13 can cleave collagen while aggrecanases, such as ADAMTS-4 and 5 can degrade aggrecan. MMPs and ADAMTS are regulated by tissue inhibitors of metalloproteinases (TIMPs)[22].

1.4   OA risk factors

Age is the main risk factor for OA. From a biochemical point of view there is less aggrecan in old cartilage than young [4, 23], meaning the cartilage becomes softer and weaker, increasing susceptibility to injury. During aging there is also an accumulation of advanced glycation end products (AGEs) [4]. However not all old people get OA, indicating other factors must also play a role.

Other risk factors for OA include obesity, gender, mechanical injury, abnormal joint loading, nutrition, and genetics. Mechanical injury and abnormal joint loading can both follow one another. Likewise muscle strength, leg length inequality, abnormal joint morphology and misalignment of the joint can all lead to abnormal joint loading and increased chance of mechanical injury. Abnormal joint loading such as static compression has been found to cause the degredation of cartilage, and is thought to contribute to OA [24].

Obesity is a risk factor for OA [25, 26], due in part to increased loading. However obesity can also increase the incidence of OA in non loading joints such as the hand [26], suggesting other factors in obese individuals must also play a role in increasing the incidence of OA. Indeed, adipokines such as leptin produced by adipose tissue, can up regulate and activate MMPs [27].

Interestingly, females tend to have an increased risk and severity of OA [28], possibly due to differences in levels of hormones [29], or due to differences in cartilage thickness, with females tending to have thinner cartilage than males [30]. In addition, the Chinese more resistant to hand OA [31] and hip OA [32], although Chinese woman are more susceptible to knee OA [33], suggesting ethnicity/genetics are important for OA pathogenesis.

Twin studies have identified genetics as contributing to the susceptibility of OA [34]. More specifically genes which contribute to OA susceptibility such as *GDF5* were found during hypothesis-lead candidate gene studies [35]. Studies to find additional susceptibility genes for OA now use hypothesis-free approaches. Linkage studies compare the genomes of effected siblings and attempt to identify regions of DNA associated with OA, regions on chromosomes 2 [36], 4 [37], 6 [37], 11 [38], and 16 [37] have all been associated with OA.

More recently genome wide association scans (GWAS), have attempted to identify additional signals associated with OA. Because many regions of DNA are in linkage disequilibrium, the number of SNPs needed to be genotyped to scan the whole genome is considerably less than the total number of SNPs in the genome. OA GWAS studies have been performed in by UK [39], Japanese [40] and Dutch [41] groups. The UK study identified 8 susceptibility signals either within or close to *GLT8D1, GNL3, ASTN2, FILIP1/SENP6, KLHDC5/PTHLH, CHST11, TP63, FTO and SUPT3H/CDC5L* [39]. Understanding why some people are more susceptible to OA than others will increase our understanding of the molecular mechanisms involved in OA.

1.5 OA pathology

OA is a disease that can affect multiple components of the joint. However, it is not completely clear which joint component the initiating factor occurs in. In reality, it is likely to be a combination of factors from different joint tissues with a large intra-individual variability. Nevertheless, as the disease progresses the number of tissues involved in the disease increases until the joint no longer functions. The structural changes within the joint lead to an altered biochemical environment which further contributes to the loss of articular cartilage.

MMPs are essential during skeletal development and in joint maintenance, but their deregulation is recognised as a major contributor to OA. Mice lacking MMP13 are somewhat resistant to cartilage erosion [42] and mice lacking ADAMTS5 are also resistant to cartilage loss [43], indicating MMP13 and ADAMTS5 are involved in OA pathogenesis. MMP13 (and MMP1) have been found to be increased in OA, leading to the degradation of type II collagen in OA [16]. There is also an imbalance in the level of MMPs and TIMPs,

which leads to ECM destruction and release of matrix fragments. The matrix fragments have been suggested to cause an inflammatory response in the synovium, which then responds by secreting MMPs and cytokines, further contributing to a vicious cycle of the inflammation and ECM degradation. Chondrocytes respond to the increased levels of MMPs, cytokines and growth factors with a change to a more hypertrophic phenotype [44], a chondrocyte phenotype also seen during skeletal development (Section 1.6), suggesting an attempt to re-initiate developmental signalling pathways. Although the chondrocytes do lay down new matrix they are unable to replicate the ECM that was laid down during development. This altered ECM may further contribute to the disease. During OA the cartilage become rough, thins and eventually can be absent from some parts of the joint. There is also vascular invasion of the cartilage from the subchondral bone [45].

From a biochemical point of view, there is cleavage of type II collagen and loss of proteoglycan, initially at the surface, as the disease progresses the loss of proteoglycan continues and the organisation of the collagen becomes affected throughout the cartilage [45].

As for the chondrocytes themselves, in healthy cartilage they have low activity, in OA the chondrocytes become 'activated', causing proliferation, increased secretion of proinflammatory cytokines, increased secretion of catabolic factors, increased secretion of anabolic factors, changes in chondrocyte gene expression [46], and changes in metabolic activity leading to altered ECM remodelling, and eventual erosion and fibrillation of the cartilage. Interestingly, there is also increased chondrocyte cell death and senescence in OA [47]. Studies have shown developmental pathways such as Hh (Hedgehog) [48], wingless-type MMTV integration site family members (Wnt) [49] and TGFβ (transforming growth factor β) [50] are involved in OA and contribute to cartilage breakdown. At the same time, there is thickening of the subchondral plate, formation of boney out growths at the joint margin called osteophytes and formation of bone cysts due to the action of osteoclasts.

Many of these changes are now quite well understood and are often said to be reminiscent of skeletal development [51]. However the mechanisms leading to this altered

6

gene expression is less well understood. In brief, the articular cartilage in OA is likely to become damaged, the chondrocyte then responds by re-instating developmental pathways in an attempt to repair the cartilage, the chondrocyte is unsuccessful and this may contribute further to the disease.

1.6   Development

Pathways such as Hedgehog (Hh) and wingless-type MMTV integration site family members (Wnt) signalling play a role in development, and adult tissue homeostasis. Defects in Hh and Wnt signalling have been shown to cause developmental abnormalities, and are often altered in diseases such as OA. The reinitiation/alteration of these signalling pathways is likely to be a failed attempt of the chondrocyte to initiate repair mechanisms. During development varying amount of different types of morphogen elicit a wide range of cellular responses to produce the patterning of tissues.

1.6.1   Digit formation

The formation of each digit is dependent on an increasing concentration of Sonic Hedgehog (Shh), with the thumb being able to form independent of Shh and little finger being highly dependent on the presence of Shh [52]. The separation of the digits is also a tightly regulated process. It is thought bone morphogenic proteins (BMPs) induce apoptosis in the interdigit space, and Noggin prevents BMP induced apoptosis within the digits [53].

1.6.2   Endochondral ossification

The development of longitudinal bones (endochondral ossification) involves a number of different cell populations, and tight regulation of signalling pathways. During endochondral ossification pluripotent mesenchymal cells differentiate into chondrocytes and lay down a cartilage model that extends from the growth plate which will eventually be degraded and turned into bone by osteoblasts.

The process involves the signalling of many morphogens including TGF-β [54], BMPs, fibroblast growth factor (FGF), parathyroid hormone related peptide (PTHrP), WNTs [54] and Indian hedgehog (IHH), which lead to the activation of a number of transcription

factors such as Sox9 [55], Runx2 [56] and Gli1 [57]. Briefly MSCs proliferate, condense then undergo chondrogenic differentiation via up regulation of Sox9/5/6 activity [58] and make a cartilage model, they then undergo hypertrophy and terminally differentiate [59]. Vascularisation of the cartilage then occurs which brings in osteoclasts which lay down bone [59]. The chondrocytes at the ends of long bones are prevented from undergoing hypertrophy and completing this process, meaning cartilage remains [59]. These chondrocytes then function to maintain this functional cartilage layer, known as articular cartilage, which is essential for smooth pain free joint movement.

IHH and PTHrP form a negative feedback loop to regulate chondrocyte hypertrophy. IHH is secreted by prehypertrophic chondrocytes to maintain chondrocyte proliferation and induce hypertrophic chondrocyte differentiation. IHH also induces the expression of PTHrP from the periarticular chondrocytes. PTHrP inhibits hypertrophic differentiation, to maintain the cartilage at the ends of long bones [60]. BMPs also have a role in endochondral ossification [61]. Cbfa1/Runx2 [62] and Osterix [63] are essential for osteoblast formation.

1.6.3 Intramembranous ossification
Unlike endochondral ossification, intramembranous ossification does not involve cartilage. Mesenchymal cells differentiate into osteoblasts which secrete bone matrix, blood vessels then infiltrate the bone and leads to the formation of bone marrow.

1.7   Hedgehog signalling

The Hh pathway was first discovered in Drosophila. It gained the name 'hedgehog' because when the *Hh* gene was mutated the fly developed spikes on the anterior half of its body [64]. Since then three Hh morphogens; Sonic (Shh), Desert (Dhh) and Indian (Ihh) Hedgehog have been discovered in mammals and have been shown to play roles in the proliferation and differentiation of cells.

1.7.1 Hedgehog signal transduction

The Hh signalling pathway is largely conserved throughout species and has important roles in embryonic and cartilage development [65, 66]. The pathway involves a cell organelle

called the primary cilia. In mammals the three homologous of the Hh signalling molecule; Shh, Dhh, and Ihh [67], are all thought to signal through the same pathway. The Hh pathway consists of the receptor Patched (Ptch1), signal transducer Smoothened (Smo) and the Gli transcription factors. Ptch1 is the transmembrane receptor for the Hh ligand and is required for signal transduction, Smo is a seven transmembrane signal transducer which signals through the Gli transcription factors. In mammals there are three Gli transcription factors (Gli1-3), interestingly active Gli1 leads to the expression of Hh target genes including Gli1 itself [67]. Shh, Ihh and Dhh all have the ability to bind to the membrane receptor Ptch1, causing a loss of the inhibition on the Hh signal transducer SMO, allowing accumulation of the active Hh transcription factor Gli1.

Inactive state: In the absence of a stimulatory Hh ligand, Ptch1 is localised to the cilia and prevents Smo from activating the pathway by stopping it from entering the primary cilium [68]. These circumstances allow the phosphorylation and degradation of Gli proteins. Gli proteins are cleaved into small repressors, a process that also occurs in, and requires, the primary cilium [69]. Glycogen synthase kinase 3 (GSK3) and Protein kinase A (PKA) are both thought to play a role in the inhibition of Hh signalling [70](Figure 1.1).

Active state: A Hh ligand is able to overcome the inhibition by GSK3 and PKA by binding to the membrane receptor Ptch1, upon interacting with the Hh ligand, Ptch1 moves away from the cell surface, diminishing this inhibition and allowing activation of the pathway activator Smo [71], which then relocalises to the membrane of the primary cilia [68]. Ptch1 contains a sterol sensing domain and has been associated with vesicle trafficking. This allows Smo to be released from intracellular vesicles and enter the primary cilia (a process involving oxysterols)[68], the presence of Smo in the primary cilia prevents phosphorylation and degradation of Gli proteins, allowing the full length active form of Gli1 to remain and increase in concentration. The active form of Gli then moves to the nucleus where it initiates the transcription of Hh response genes, including *Gli1* [67], *Hhip* and *Ptch1*, in feedback loops to further activate the pathway and also prevent the spread of the Hh signal [72]. As *Gli1* is also a Hh response gene, it can be used as an indicator of active Hh signalling.

In addition to activation of Hh signal through ligand binding, it is also possible to activate Hh signalling using synthetic compounds that directly antagonize Smo, such as Purmorphamine (Pur), a purine derivative [73] and Smoothened agonist (SAG) [74].

1.7.2 Hedgehog in development and OA

Of all cytokines involved in chondrogenesis and osteogenesis, Ihh plays a particularly interesting role [75]. Ihh is first expressed in the cartilaginous condensation [60] and is produced in the prehypertrophic and hypertrophic zones, which forms a gradient that opposes that of PTHrP, to regulate hypertrophy. PTHrP is secreted from resting chondrocytes and functions to inhibit chondrocyte differentiation, meaning proliferation decreases as the chondrocytes move away from the growth plate. Ihh increases chondrocyte hypertrophy via the Wnt and BMP pathways [76]. Ihh causes the proliferation of columnar chondrocytes and prevents the differentiation of proliferating chondrocytes [77]. Ihh also signals to preosteoblasts in the adjacent pericardium [75] to aid Runx2 and Osterix in osteoblast differentiation [78]. The role of Ihh continues postnatally to maintain growth plate and trabecular bone [79].

The role of Ihh in skeletal development was illustrated by in mice, *Ihh-/-* mice have abnormal bone growth and all skeletal structures are reduced in size [80], while *Smo* knockout in cartilage decreases chondrocyte proliferation [81]. In humans, defects in Hh can cause polydactyl [82]. Ihh protein is more abundant in OA cartilage compared to normal [83]. More recently, Lin *et al.* found surgical induction of OA in mice initiates Hh signalling, and the inhibition of Hh signalling causes a reduction in OA severity, in surgically-induced OA-mice [48]. More specifically, Hh signalling was suggested to contribute to OA severity through Runx2 regulation of ADAMTS5 [48].

Figure 1.1 Simplified schematic of the Hh signalling pathway

A.   Hh ligand absent

B.   Hh ligand present

**Ubiquitination and Degradation
No pathway activation**

**Transcription of Hh
target genes (including
GLI1 and PTCH)**
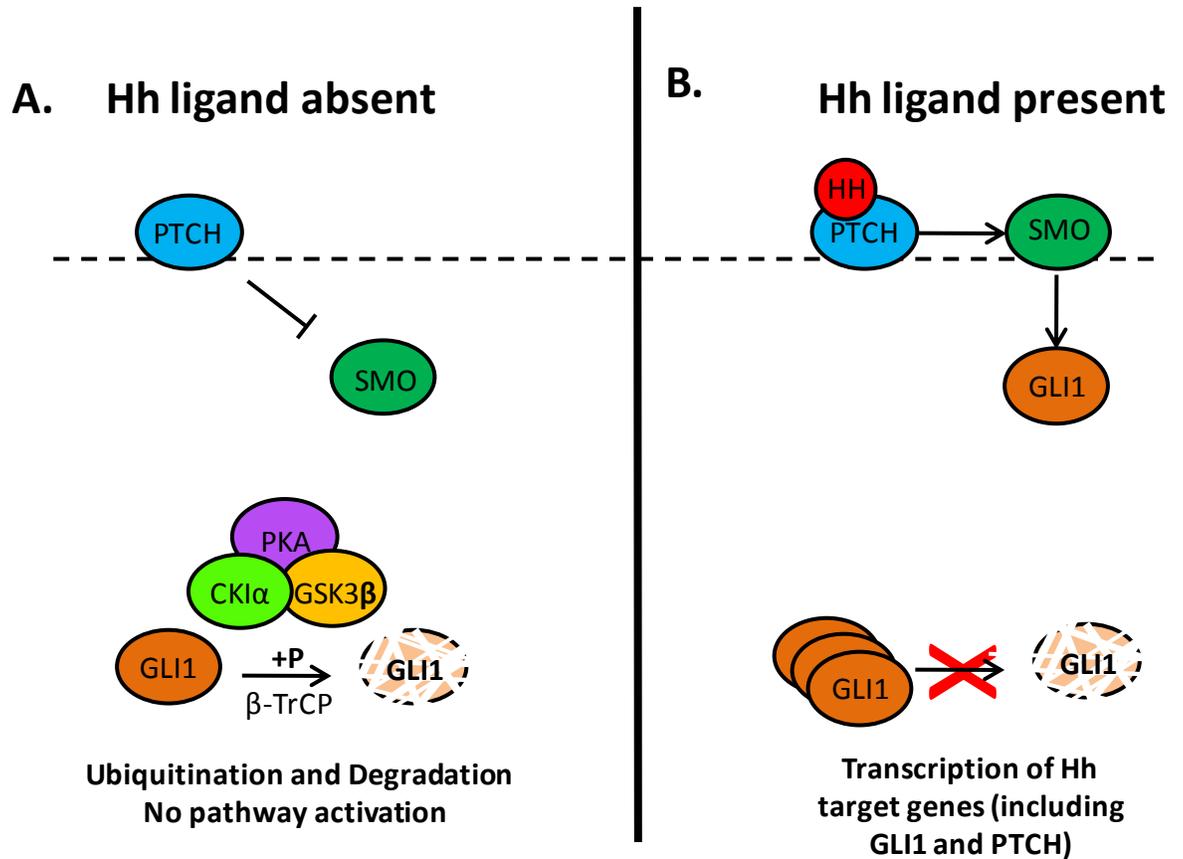
Figure 1.1 Simplified schematic of the Hh signalling pathway. (A) In the absence of Hh ligand, Ptch1 inhibits Smo by preventing it from entering the primary cilia (inhibitory arrow) meaning Smo cannot activate Gli1 (PKA/CKIα/GSK3β cause Gli1 degradation). (B) The binding of Hh to Ptch1 relieves its inhibitory effect on Smo, allowing Smo to enter the primary cilia and allowing the activation of the transcriptional activator Gli1 (and prevent Gli1 degradation by PKA/CKIα/GSK3β) and thus cause Hh target gene expression including that of Gli1 and Ptch1.

1.7.3 Primary cilia

The primary cilium is an important cell organelle for Hh signalling and is found on almost all growth arrested cells. Its importance in Hh signalling was first noticed in 2003, where a mouse mutant in Wimple (IFT172, required for primary cilia formation) was found to have defective Hh signalling [84]. Since then many of the Hh signalling components such as Gli [69], Smo [85] and Ptch1 [68] have been found to localise to the primary cilium.

The structure of the primary cilium relates to its function. The primary cilium extends into the extracellular space, allowing for the detection of environmental stimuli. They have a high surface area to volume ratio, allowing for a high number of membrane receptors to be expressed. At the base of the cilium lies the basal body. It is centriole derived, and organises the cilium microtubules [86]. The entry of proteins from the cell body to the primary cilium is tightly regulated, by a region at its base, called the transition zone, allowing for compartmentalization of signalling molecules [87]. The Hh signalling pathway utilises this compartmentalization. Primary cilia bending causes an increase in intracellular Ca^{2+}, an example of this is where urine passes over the cilia in the kidney cells [88], this effect is abolished by removal of the primary cilia [89]. The primary cilium has also been shown to play a role in sensing mechanical stimuli in the cartilage ECM surrounding the chondrocyte [90].

The primary cilium is constructed at the tip, away from the site of protein synthesis within the cell, they therefore need, and utilise, a process known as intra flagella transport (IFT) [91], to transport proteins from the basal body to the site of construction. As well as functioning in construction, IFT also regulates the signalling events that go on within the primary cilium. The IFT can move proteins at speeds of up to 1μm per second, meaning the primary cilium is efficient in cell signalling [92]. A study by McGlasham *et al.* in 2008 showed the length and number of primary cilia in cartilage are increased during OA [93], possibly due to IL1 [94]. This may lead to the altered Hh signalling and sensitivity to mechanical stimuli seen in OA.

There are a number of proteins associated with, and involved in primary cilia and cilliogenesis, such as BBS and IFT proteins. In humans, defects in BBS proteins have been

associated with obesity, loss of the sense of smell, mental retardation/learning disorders, diabetes, male infertility, kidney failure and sight loss. Mutations in IFT proteins have been associated with abnormal developmental phenotypes in mice, many of which are possibly related to altered Hh signalling pathway activity [69, 84, 95].

1.8 Wnt signalling

The Wnt family of ligands consists of 19 ligands, traditionally said to be either canonical or non canonical. Similar to Hh, the Wnt pathway also plays a role in bone development [96]. Wnt is actually regulated by Ihh during osteoblast formation [97]. In addition, Ihh is also regulated by canonical Wnt signallng [98]. The order of Hh and Wnt activation can determine the differentiation of osteoblasts and the differentiation of chondrocytes [99].

1.8.1 Wnt signal transduction

Canonical: In the absence of Wnt ligand, Axin, Gsk3, APC and CKI aid to degrade β-catenin [100]. In the presence of Wnt, Wnt binds a Frizzled receptor which leads to the recruitment of the LRP5/6 (low density lipoprotein) co-receptor, this then prevents Axin/Gsk3 phosphorylation of β-catenin leading to its stabilisation. β-catenin then accumulates in the nucleus and interacts with lymphoid-binding factor (LEF) and T-cell factors (TCF) to cause the expression of a number of Wnt target genes, including Runx2 [100] (Figure 1.2).

Non canonical: In addition to canonical Wnt signalling there are also a non-canonical Wnt signalling pathways such as the Wnt/calcium pathway and the planar cell polarity pathway (PCP) [101]. Unlike canonical Wnt signalling, these pathways do not require LRP and instead require ROR. The calcium signalling of the non canonical Wnt signalling can inhibit the canonical Wnt pathway [102].

1.8.2 Wnt in development and OA

The importance of canonical Wnt signalling in human development is shown by a naturally occurring loss of functional LRP5 causing decreased bone mass [103], the opposite is also true in that gain of function leads to increased bone mass [104]. Similar phenotypes are also observed in mice with abnormal Lrp5 [105]. Non canonical Wnt signalling requires ROR, therefore mutations in *Ror2* in mice is likely to be indicative of loss of non canonical

Wnt signalling. Human mutations in *ROR2* cause skeletal abnormalities such as brachadactyl type B, Robinow syndrome, spinal defects and a dysmorphic face [106, 107], with similar phenotypes observed in mice [108, 109]. A number of other Wnt signalling components have been shown to cause defects in bone development [96]. These studies show both the canonical and non-canonical Wnt signalling play a role in skeletogenesis.

Figure 1.2 Simplified schematic of canonical Wnt signalling
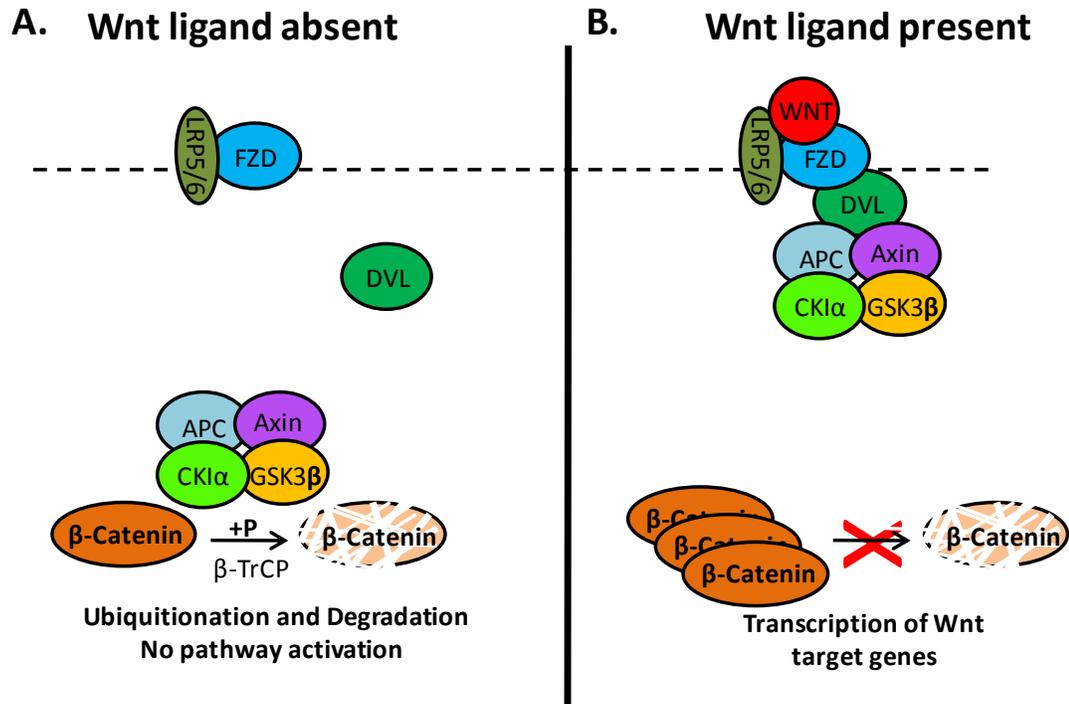
Figure 1.2 Canonical Wnt signalling. (A) In the absence of Wnt ligand APC/Axin/CKIα/GSK3β mediate the phosphorylation and degradation of β-catenin. (B) In the presence of Wnt ligand, Wnt ligand binds to Frizzled (FZD), and with the recruitment of LRP5/6, DVL prevents APC/Axin/CKIα/GSK3β from phosphorylating β-catenin meaning it is stabilised and translocates to the nucleus causing the activation of Wnt target genes.

1.9 microRNAs

I hypothesise miRNAs have a role in endrochondral osscification, chondrogenesis and OA, via regulation of the signalling pathways described. miRNAs are small (approximately 22nt), single stranded, non-coding RNAs that modulate gene expression through base specific interactions within the target genes 3'untranslated region (UTR) to block translation [110]. The first miRNA was described in 1993 by Lee *et al.* [111] and since then, hundreds have been identified (www.mirbase.org). Each miRNA has the ability to regulate the expression of a large number of genes, and each gene is postulated to be regulated by many miRNAs [110].

miRNAs were originally mooted as providing 'fine tuning' to the control of gene expression. Recently, particularly through the generation of knockout and transgenic mice for specific miRNAs, it has become clear that they can be major regulators in development and play a significant role in various pathologies and defining cell phenotype.

All known miRNAs are registered on miRBase (www.mirbase.org) [112]. miRBase is managed by The University of Manchester and contains information on miRNA stem loop sequence, deep sequencing data, genomic location, miRNA families, the mature sequence of both miRNAs encoded from the stem loop, any relevant references for each miRNA and links to other internet sites such as to the target prediction site PicTar (http://pictar.mdc-berlin.de/).

1.9.1 miRNA processing and mechanisms of action

miRNAs are either transcribed from within genes, usually intronic, or from outside of genes (intergenic) where they have their own promoter. Several miRNAs can be encoded from the same RNA transcript and are called polycistronic miRNAs. Initial miRNA transcripts (pri-miRNAs) are transcribed in a 5' to 3' direction by the RNA polymerase II transcription factor and vary in size but always contain a stem-loop.

Pri-miRNAs are then cleaved by the nuclear ribonuclease Drosha (which is a dsRNA-specific RNase-III-type endonucluease and is associated with the RNA binding protein; DiGeorge syndrome critical region gene 8; DGCR8) in the nucleus [113], leaving only the stem loop and a one or two nt overhang. This molecule is called a pre-miRNA which is

16

around 80 nucleotides in length. Following processing to short stem-loop structures, the pre-miRNA is exported from the nucleus usually by Exportin 5 [114]. The cytoplasmic dsRNA-specific RNase-III-type endonucluease Dicer removes the stem-loop forming a miRNA duplex of about 21nt in length consisting of two arms of the loop (often named -5p and -3p respectively). After Dicer cleavage a two nt 3' overhang remains (essential for the duplex to be incorporated into RISC). In a concerted process Dicer assigns miRNAs to one of a family of four Argonaute proteins [115].

Argonautes tightly bind to single stranded miRNAs in the core of the RISC complex, meaning only one of the miRNA strands is incorporated. The other strand is generally degraded, a choice based upon the intrinsic thermodynamic properties of the strands [116] and pri-miRNA sequence [117]. There are 4 different Ago proteins in humans, the reason for having more than one Ago protein is currently unclear, although each Ago protein may be sequence specific. Ago 1-4 can all be used for miRNAs [118]. Ago 2 also required for siRNA-triggered mRNA degradation [118].

The RISC complex then uses the miRNA as a guide to complementary base pair with the miRNA targets. The RISC-miRNA complex usually base pairs via the seed region (nucleotides 2-7) of the miRNA to target sites usually on the 3'UTR of target genes [119]. Both -5p and -3p miRNAs of the stem loop have to potential to be incorporated into the RISC complex and target their own discrete repertoire of genes. The miRNA:RISC:Target interaction then leads to the reduced translation and mRNA degradation of the target. [116, 120-122]. miRNAs themselves can be differentially expressed and can play a major role in defining cell phenotype and disease severity [110].

Figure 1.3 miRNA processing pathway
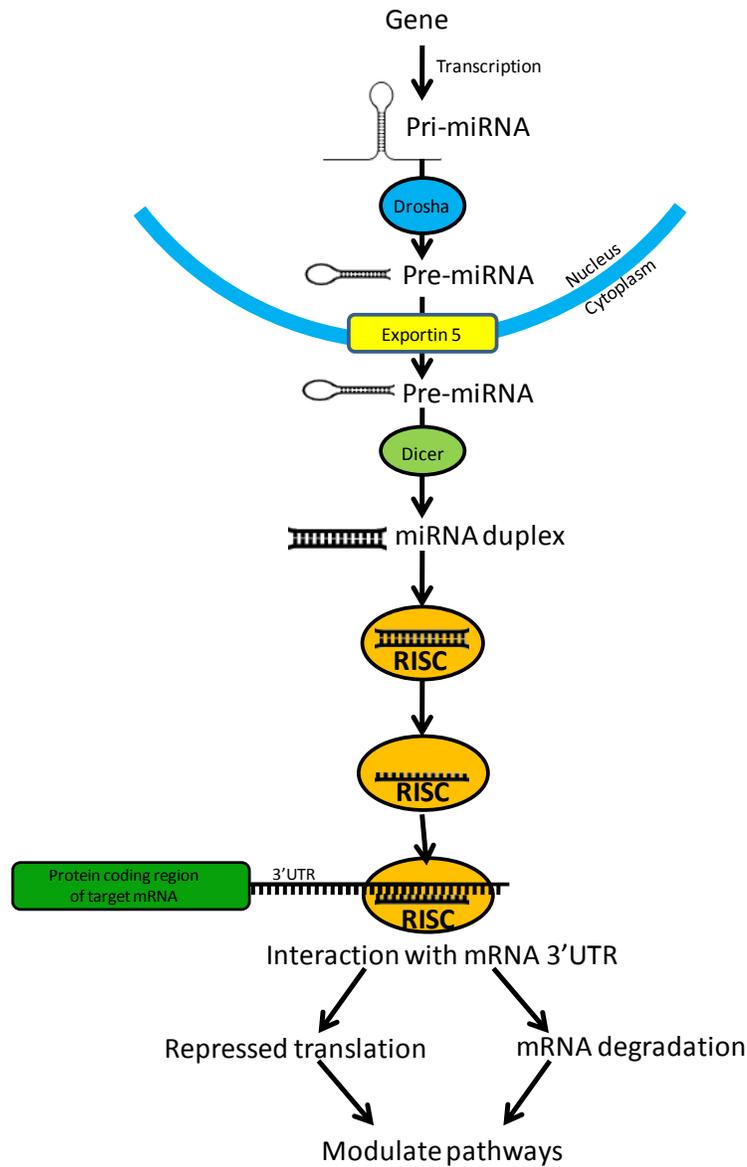
Figure 1.3 A schematic showing the main pathways involved in miRNA formation and function. miRNAs are transcribed by RNA polymerase II transcription into Pri-miRNAs, which are then cleaved by Drosha [113], leaving only the stem loop, called a pre-miRNA, which are exported from the nucleus usually by Exportin 5. Dicer then removes the stem-loop forming a miRNA duplex. One of the miRNA strands is incorporated into RISC. The other strand is generally degraded. The RISC complex then uses the miRNA as a guide to complementary base pair with the miRNA targets. The RISC-miRNA complex usually base pairs via the seed region (nucleotides 2-7) of the miRNA to target sites usually on the 3'UTR of target genes [119]. Both -5p and -3p miRNAs of the stem loop have to potential to be incorporated into the RISC complex and target their own discrete repertoire of genes and pathways.

1.9.2 miRNA regulation

As well as regulating the expression of various genes, miRNAs themselves are subject to various regulatory mechanisms, similar to normal protein coding genes, miRNA expression can be controlled by their (or their host genes) promoter. In addition, miRNAs participate in a large regulatory network where they are both regulators and regulated by other mRNAs and even miRNAs [123]. It has been suggested many miRNA targets are actually 'pseudo targets' whose function is to bind the miRNA and prevent its action on other 'real' miRNA targets [124]. Other types of RNA including coding mRNAs, pseudogenes and long non-coding RNAs can also act as competing endogenous RNAs (ceRNAs) to regulate the levels of miRNA available, causing altered expression of 'real' targets [125].

1.9.3 miRNA target prediction

miRNAs target many gene transcripts to negatively regulate them, potentially meaning each miRNA is involved in a number of cellular processes. Following the identification of miRNAs, work to identify miRNA targets began and it was quickly realised miRNA target prediction was complex. miRNAs do not need to be fully complementary to their target. In 2003 Stark *et al.* realized the importance of the seed region (nucleotides 2-8) in Drosophila [126]. Most miRNAs need to be fully complementary to their targets within the seed region, although this is not always the case [127]. There are a number of online target prediction programs for miRNAs, each has a different prediction algorithm and therefore each predicts a different repertoire of targets.

TargetScan

Targetscan (www.targetscan.org) is an online miRNA prediction software and was the first program to exploit the seed in its algorithm to predict vertebrate miRNA targets [128]. Originally Targetscan searched for 7mer seed regions and then took into consideration the rest of the miRNA-mRNA interaction [128]. Targetscan now searches for 8mer and 7mer seed matches, taking into account flanking by Adenosine, species conservation, and context, it also allows mismatches within the seed as long as there is compensatory 3' pairing. The most recent release (6.0) takes into account target site abundance and includes all 3'UTRs from RefSeq [119, 129, 130].

PicTar

PicTar (http://pictar.mdc-berlin.de/), searches for miRNA targets within conserved regions of 3'UTRs. It also takes into account the expression of the miRNA and the potential target. It then calculates the free energy of the potential miRNA-Target duplexes in each species and gives them a combined score for all species. miRNAs with multiple alignments are also favoured. PicTar tends to predict fewer targets but is more specific than other target prediction algorithms [131].

DIANA-microT

DIANA-microT (http://diana.cslab.ece.ntua.gr/microT/), uses a 9 nucleotide shifting window that moves along the 3'UTR. At each position it calculates the free energy and considers the number of consecutive complementary bases. For potential miRNA:Target interactions it then also considers species conservation. It also gives a signal:noise ratio (chosen miRNA binding to 3'UTR:random miRNA sequence binding to 3'UTR) [132].

miRanda

The miRanda algorithm searches for complementary miRNA binding sites in 3'UTRs. Each potential site is then scored based on binding energy, position within the 3'UTR and species conservation, the score is penalized by mismatches and gaps in binding [133]. miRGen (http://www.diana.pcbi.upenn.edu/cgi-bin/miRGen/v3/Targets) and MicroCosm (http://www.ebi.ac.uk/enright-srv/microcosm/htdocs/targets/v5/) both use the miRanda algorithm.

miRDB

miRDB (http://mirdb.org/miRDB/l) was created by using SVM (support vector learning machine) analysis of thousands of targets impacted by miRNAs. The features of these targets are then combined with already well known miRNA target features to create the algorithm (MirTarget2).

With ongoing experiments to validate miRNA targets additional information can be fed into the algorithms to improve the accuracy in which they predicted targets, with the eventually goal being to create a perfect algorithm. It remains to be seen what will come first, a perfect algorithm or validation of every target of every miRNA. Due to the sheer

number of miRNAs, miRNAs targets and the complexity of their interactions both at this stage seem unlikely. There is however an increasing number of validated miRNA targets.

miRTarBase (validated targets)

miRTarBase (http://miRTarBase.mbc.nctu.edu.tw/) is a database of 'validated' miRNA targets [134]. It also provides information on the validation technique (microarray, luciferase assay, western blot, proteomics, or real-time RT-PCR) as not all 'validated' targets on the database are validated via luciferase. It relies on manual surveying of the literature meaning a number of validated miRNA targets are likely to be missed.

1.10   miRNAs in development

The overall importance of miRNAs in development was demonstrated by Berstein *et al.* in 2003, where the global deletion of the miRNA processing enzyme Dicer led to embryonic lethality in mice [135]. This demonstrated miRNAs are essential for mouse survival, but does not indicate whether or not miRNAs are involved in skeletogenesis. In addition conditional knockout of Dicer in *Prx1* (a gene expressed in the developing skeleton) expressing cells also caused skeletal defects [136]. Kobayashi *et al.* used a 'collagen II-Cre floxed Dicer' mouse to produce a cartilage restricted, Dicer null mouse. The mice were viable, but lack virtually all cartilage miRNAs and display severe developmental defects [137]. More specifically there is a decrease in chondrocyte proliferation and a faster onset of hypertrophy [137]. Knockout of Dicer in the skeleton shows miRNAs have a role in skeletogenesis, bur does no indicate which.

The most extensively studied miRNA in cartilage is miR-140, originally identified as being cartilage specific in Danio rerio [138]. Tuddenham *et al.* later showed miR-140 was also cartilage specific in the developing mouse [139]. Following differentiation of mesenchymal cells to cartilage the expression of miR-140 is increased [140]. This is because miR-140 is under the control of cartilage specific transcription factor Sox9 [141]. Due to the restricted expression of miR-140 to cartilage, it was likely it played a role in development. Miyaki *et al.* demonstrated this by creating a miR-140 null mouse [142]. miR-140 mice have shorter long bones and craniofacial deformities [142]. The exact mechanisms by which miR-140 mediates its effects on skeletal development are not yet fully understood. miR-140 is

however known to directly target HDAC4 [143], CXCL12 [144], SMAD3 [145], DNPEP [146], VEGF [147] and SP1 [141], many of which have been shown to play a role in cartilage and OA. Taken together this suggests miR-140 plays a major role in skeletal development.

The phenotype of the miR-140 null mouse [142] is less severe than the Col2-Cre:Dicer$^{fl/fl}$ mouse [137] suggesting miR-140 is not the only miRNA involved in skeletal development. In addition to miR-140, other miRNAs have been shown to regulate the migration and differentiation of the chondrocyte as well as directly regulating the matrix components secreted by the chondrocyte. Both miR-488 [148] and miR-34a [149] have been shown to regulate chondrocyte/mesenchymal cell migration. Following migration and condensation the chondrocyte transcription factor SOX9 is upregulated leading to type II collagen and aggrecan expression. miR-145 directly targets SOX9 expression decreasing chondrocyte differentiation (decreased Col2a1 and aggrecan) and increase hypertrophic differentiation (increased MMP13) [150]. miR-365 is a mechanosensitive miRNA and can induce chondrocyte differentiation by targeting HDAC4 [151], similar to miR-140. A number of miRNAs have been shown to regulate secreted matrix constituents including miR-675 which is regulated by miR-145 and targets Col2a1 expression [152], miR-29a and miR-29b also regulate Col2a1 [153] and miR-1 regulates aggrecan expression [154]. In addition to those already mentioned, miR-455 is a cartilage specific miRNA which is expressed from the Col27a1 gene, regulates TGFβ signalling and may play a role in digit formation [155].

1.11   miRNAs in OA

The evident role of miRNAs in skeletal development and the role of developmental pathways in OA suggest miRNAs are involved in OA pathogenesis. Profiling miRNAs in healthy and OA cartilage attempted to determine the relative importance of each miRNA in OA [156]. Iliopoutos *et al.* showed miR-140 expression was decreased in OA, further suggesting a role of miR-140 in healthy cartilage maintenance [156]. They also found altered expression of miR-22 which targets PPARα and BMP7 leading to altered MMP13 and aggrecan expression. In another screen of miRNAS in OA, Jones *et al.* found altered expression of miR-9, miR-98 and miR-146a which are involved in the regulation of IL-1 [157]. IL-1 is a cytokine know to have a role in OA, in addition to the miRNAs mentioned

22

above there are a number of other miRNAs regulated by or involved in the regulation of IL-1 such as miR-34a [158], miR-199a-3p and miR-27b, the latter directly regulates MMP13 [159]. miR-146a, miR-155, miR-181a, miR-223 are all increased in blood of OA patients. Their function is unknown but they may serve an important role in the future as biomarkers, enabling for tailored therapies.

1.12   miRNA regulation of developmental pathways

Many of the miRNAs which regulate skeletal development or are involved in OA work through developmental signalling pathways. In addition to those known to have a role in OA and development there are likely to be a number of other miRNAs which have been shown to regulate these pathways and may have a yet to be discovered role in skeletal development and OA.

There is limited direct evidence for miRNAs regulating or being regulated by Wnt signalling to control skeletal development or OA. There are however a number of miRNAs which regulate Wnt and, due to Wnt involvement in development and OA, it would be surprising if these miRNAs did not have a role in either skeletal development or OA. miR-29b, miR-101 and miR-124 have all been shown to increase canonical Wnt signalling by reducing the expression of GSK3β [160].

The role of miRNA regulation of Hh signalling in development has also been demonstrated. In a study which knocked out all miRNA in the developing eye, there was altered Notch and Hh signalling [161] and is a study which knocked out all epidermal miRNAs there was loss of Shh [162]. Conditional knockout of Dicer in Shh expressing cells shows miRNAs play a role in preventing cell death in the Shh dependent digit formation in mice, although they do not appear to play a role in digit patterning in Shh expressing cells [136]. In a separate study, which also used conditional knockout of Dicer in Shh expressing cells, miRNAs were shown to play a role in lung development and patterning [163].

In addition to the studies that show miRNAs in general have a role in Hh regulation of development, a number of studies have shown specific miRNAs are important in the Hh regulation of development. Deletion of the miR-199a~214 cluster leads to skeletal

abnormalities in mice [164] which is likely to involve Hh, as miR-214 may form a feedback loop with Hh signalling. Twist1 a Hh regulator, induces the expression of both miR-199a and miR-214, and miR-214 has been found to target SuFu, another Hh regulator in *Danio renio* (zebrafish) [165]. miR-199a is a regulator of Smad/TGFβ signalling by targeting Smad1 [166]. Hh signalling can induce the expression of the miR-17~92 cluster which are a group of miRNAs involved in skeletal development and are also thought to act through TGFβ signalling [167].

A screen has shown a number of miRNAs have been found to be differentially expressed by altered Hh signalling in zebrafish [168]. miR-29 expression is repressed by Hh in human cholangiocarcinoma cells [169] and miR-206 is repressed by Shh in lung explants [170]. miR-25 is Hh responsive and regulates TNF signalling [171].

As well as being regulated by Hh, a number of miRNAs have been show to regulate multiple components of the Hh signalling pathway in Drosophila [172], Danio renio [165] and mammals. In humans the miR-302~367 cluster can indirectly regulate Shh signalling through CXCR4 repression [173]. miR-196 is a regulator of Hoxb8 and Shh in limb development [174]. miR-365 is a mechanosensitive miRNA in cartilage which increases the expression of Ihh by targeting HDAC4 [151]. Ferretti *et al.* showed miR-326, miR-125b and miR-324-5p can modulate the Hh pathway in cancer cells [127]. More specifically miR-326, miR-125b and miR-324-5p were predicted to suppress the pathway activator SMO, and miR-324-5p was also predicted to target the downstream transcription factor Gli1 in humans [127].

1.13   miRNA therapeutics

miRNAs are showing excellent potential as disease biomarkers to improve disease diagnosis. There is also research going on into an anti-miRNA therapy for the treatment of Hepatitis C virus (HCV). HCV RNA forms an interaction with miR-122, which is essential for virus reproduction [175]. The use of a locked nucleic acid (LNA) to inactivate miR-122 has proven useful in the treatment of HCV infected Chimps [176, 177] and is paving the way for future miRNA therapeutics.

1.14 Aims and hypothesis

Research has begun to investigate the effect of specific miRNAs in cartilage, but further work is required. With ongoing research, the role of each miRNA in cartilage is becoming clearer. To summarise, both miRNAs and developmental signalling pathways are important in cartilage development and OA. This project will focus on how miRNAs may act upon these signalling pathways to attenuate cartilage development and OA. This research will improve our understanding of OA biology, and may lead to anti/pro-miRNA therapy for OA. I hypothesise miRNAs have a role in endrochondral osscification/chondrogenesis and OA, via regulation of the signalling pathways described.

Chapter 2 Materials and methods

2.1   Materials

2.1.1   Antibodies

Anti-Human/mouse Gli1 (V812) rabbit polyclonal was purchased from Cell Signalling Technologies, New England Biolabs (NEB), Hitchin, UK. Anti-Human/mouse Smo (ab72130) was purchased from Abcam, Cambridge, UK.Anti-human/mouse Gpc1 (16700-1-AP) rabbit polyclonal was purchased from Protein Tech Group (Manchester, UK). Mouse monoclonal anti-acetylated tubulin (T-6793) was purchased from Sigma-Aldrich (Poole, UK). GAPDH was purchased from Chemicon International (Thermo Fisher, Uk). Alexa Fluor® 488 goat anti-mouse IgG, Alexa Fluor® 594 goat anti-mouse IgG, Alexa Fluor® 594 goat anti-rabbit IgG Alexa and Alexa Fluor® 488 goat anti-rabbit IgG antibodies were purchased from Life technologies (Paisley, UK).

2.1.2   Cell lines

2.1.2.1C3H10T1/2

C3H10T1/2 cells are mouse pluripotent mesenchymal cells, and are capable of differentiation into fat, cartilage and bone under the appropriate stimuli. The C3H10T1/2 cell line was isolated by Reznikoff *et al.* in 1972 from a line of C3H mouse embryo cells [178]. Full details of this cell line can be found at the American Type Culture Collection (ATCC) company website (www.atcc.org). Cells were a gift from Professor Ian Clark, University of East Anglia, Norwich, UK. The cell line was cultured in minimum essential medium (MEM) culture medium by the method outlined below. Stimulation of C3H10T1/2 cells with Ihh can activate the Hh signalling cascade [67]. Ihh and Bmp2 induces osteoblastogenesis as measured by alkaline phosphatase activity [179], and seeding in micromass followed by Bmp2 stimulation can induce chondrogenic differentiation as measured by alcian blue [180].

2.1.2.2SW1353

The SW1353 cell line was initiated from a primary grade II chondrosarcoma in the right humerus of a 72 year old Caucasian female in 1977. Cells were purchased from ATCC. The

cell line was cultured in Dulbecco's modified Eagle's medium (DMEM) culture medium by the method outlined below.

2.1.3 Cell culture reagents

DMEM medium was obtained from Invitrogen (Life technologies). Foetal bovine serum (FBS), MEM medium, penicillin-streptomycin solution (10000 U/ml and 10 mg/ml respectively), L-glutamine solution (200 mM), trypsin-EDTA solution (0.5g porcine trypsin and 0.2g EDTA per L), non-essential amino acid solution, Dimethyl sulphoxide (DMSO), Ampicillin sodium salt was obtained from Sigma-Aldrich Ltd. Phosphate buffered saline (PBS) was purchased from Lonza (Wokingham, UK). Tryptone, yeast extract and bacto-agar were purchased from Difco Laboratories (Detroit, MI, USA).

2.1.5 Transfection reagents

For small RNA transfections, cells were transfected with DharmaFECT® transfection reagents, siGENOME®/SMARTpool® small interfering RNA (siRNA) were purchased from Dharmacon (Cramlington, UK). For plasmid transfections cells were transfected with Fugene HD transfection reagent was purchased from Promega (Southampton, UK).

2.1.6 Cytokines, growth factors and other stimuli

Recombinant human Ihh, Recombinant Wnt3a, Recombinant BMP-2 were purchased from R&D systems (Abingdon, UK). Pur and SAG were purchased from (Calbiochem). All cytokines and agonists were stored at -80°C in buffer recommended by the supplier.

2.1.7 Immunoblotting reagents

Bovine serum albumin (BSA) (desiccate), Bradford reagent, N,N',N'-Tetramethylethylenediamine (TEMED), polyoxyethylene sorbitan monolaurate (Tween-20) and Kodak high-speed X-ray film were purchased from Sigma-Aldrich (Poole, UK). Ammonium persulphate (APS) was purchased from BDH Chemicals (Poole, UK). A 37.5:1 mix of acrylamide/bis-acrylamide was purchased from Amresco (Solon, OH, USA). Immobilon-P polyvinylidene difluoride (PVDF) 0.45 µM membrane was purchased from Millipore (Watford, UK). Enhanced chemiluminescence (ECL), ECL-plus and ECL-advanced Western blot detection reagents were purchased from Amersham Biosciences (Little Chalfont, UK). PageRuler™ pre-stained protein ladder, GeneRuler was purchased from

Fermentas Life Sciences (York, UK). Marvel non-fat dry milk powder was obtained from Premier Foods (St. Albans, UK).

2.1.6 Molecular biology reagents

TaqGold was purchased from Life technologies. GeneRuler™ 100kb and 1kb DNA ladders, *Eco*RI, BSR1 and *Hind*III were purchased from Fermentas Life Sciences (York, UK). Deoxyribonucleotide triphosphate (dNTP) was purchased from Bioline (London, UK). pMIR-Report vector was purchased from Ambion (Southampton, UK). TITANIUM™ Taq DNA Polymerase was purchased from TaKaRa Biomedicals (Wokingham, UK). Phire Hot Start II polymerase was purchased from Thermo Scientific. Agarose (electrophoresis grade) was obtained from Life Technologies. RNase- and DNase-free H_2O obtained from Sigma-Aldrich (Poole, UK) Real-time RT-PCR primers and probes were purchased from Sigma-Aldrich (Poole, UK). Probes library probes were purchased from Roche (Roche Diagnostics, Burgess Hill, UK). Moloney Murine Leukaemia Virus (M-MLV) was purchased from Life Technologies. TaqMan® Universal PCR Master Mix (2X) was purchased from Life Technologies (Foster City, CA, USA). VECTASHIELD mounting medium with DAPI was purchased from Vector Labs. β-mercaptoethanol, Triton X-100 and Gelatin from cold water fish skin were obtained from Sigma-Aldrich (Poole, UK).

2.1.7 Commercially available kits

The Cells-to-cDNA™ II Kit was purchased from Ambion (Ambion (Europe) Ltd., Huntingdon, UK). QIAquick® GelExtraction Kit and QIAquick DNA extraction kit were purchased from Qiagen (Crawley, UK). E.Z.N.A Tissue DNA extraction kit was purchased from Omega Bio-Tek (Norcross, GA, USA). BigDye Terminator v3.1 sequencing kit was purchased from Life Technologies.

All other standard laboratory chemicals and reagents, unless otherwise indicated, were commercially available from Sigma-Aldrich Ltd, Fisher Scientific or BDH Chemicals.

2.2   Methods

2.2.1   SNP genotyping assays

2.2.1.1Agarose Gel Electrophoresis

Agarose gels are prepared in the range of 1 – 2% (w/v) by dissolving the required amount of agarose in 1x Tris-acetate-ethylenediamine tetraacetic acid (TAE) containing 0.04 M Tris (pH 8), 5.7% (v/v) glacial acetic acid and 0.001M EDTA, through boiling. Ethidium bromide was added to cooled agarose at a final concentration of 0.2 µg/ml. Gels were poured, allowed to set and the required amount of DNA in loading buffer (0.125 M (pH 6.8), 2% (w/v) SDS, 10% (v/v) glycerol and 0.001% (w/v) bromophenol blue) loaded. Bands were separated at 100V for approximately 40 min in TAE and visualised on a ChemiGenius II BioImager (Syngene, Cambridge, UK).

2.2.1.2 Restriction Fragment Length Polymorphism (RFLP)

PCR was carried out on 50ng genomic DNA extracted from the blood of OA cases and controls. PCR was performed using TaqGold, a mix was prepared consisting of the following 50ng DNA, 7.5pM forward and reverse miR-140 RFLP primers (Table 2.1) (0.075µl of 100nM stock), 1.5µl 10XPCR buffer, 0.2mM dNTPs, 0.08 units Taq Gold, 2mM Mg++ (1.2µl) and H_2O to total 15µl. Thermo cycling conditions were as follows 95^oC for 5min followed by 30 cycles of 95^oC for 15s and 65^oC for 30s and a final step of 72^oC for 7min.

PCR products were then digested using *Bsr*I restriction enzyme. Digested PCR products were then separated by size through agarose Gel Electrophoresis stained with ethidium bromide and visualised using UV light (see above).

2.2.1.3Sequencing the miR-140 region

Genomic DNA of OA cases and controls was sequenced. PCR of the miR-140 region was performed as described for RFLP using miR-140 sequencing primers, both forward and reverse primers contained M13 primer binding sites (miR-140seq F and miR-140 seq R-Table 2.1). Following PCR, 10µl the reaction was cleaned up. 2.4U Shrimp Alkaline

Phosphatase (SAP) (2.4µl of 1U/µl), Exo1 and 1.4µl of 10x buffer were added. Mineral oil was then overlaid. Samples were then heated to 37°C for 15min and then 80°C for 15min.

Sequencing was performed on cleaned up PCR products generated from genomic DNA. Sequencing reactions consisting of the following; 6µl cleaned up PCR product, 0.5µl Big dye premix v3.1, 2µl 5x Big dye buffer, 1.2µl H$_2$O, 0.3µM primer were then subject to the following thermal conditions 25 cycles of 96°C for 10s, 50°C for 5s, 60°C for 4min then cooled to 4C. The reaction was then diluted with 10µl H$_2$O, mixed and briefly centrifuged. 20µl of diluted sequencing reaction was then transferred to fresh 3100 sequencing plate containing 60µl 100% ethanol and 0.1µl pellet paint, mixed and briefly centrifuged before being incubated at room temperature for 1hr. The solution was centrifuged (1600g for 45min) and supernatant discarded. The plate was inverted and briefly centrifuged on tissue paper to remove residual supernatant. 70µl 70% (v/v) ethanol was added was added to the pellet, vortexed and centrifuged (1600g for 15min). The supernatant was discarded and the plate inverted and briefly centrifuged on tissue paper to remove residual supernatant. 10µl MiliQ ultrapure H$_2$O was then added to each sample, mixed and briefly centrifuged. 10µl HiDi formamide (Invitrogen) was added t o each sample, mixed and briefly centrifuged. Samples were sequenced on the ABI 3100 Genetic Analyzer, and analysed using SeqScape software (Applied Biosystems, Warrington, UK).

2.2.1.4 ABI SNP Genotyping Assay

The applied biosystems SNP Genotyping Assay have a PCR primer pair and work in a similar way to the TaqMan assay, but have two allele-specific TaqMan® MGB (minor groove binding) probes containing distinct fluorescent dyes (VIC and FAM). Amplifications were preformed in 5µl final volumes in optical plates, following the manufactured instructions. Amplifications were preformed in a thermal cycler (Bio-Rad), the plate was then transferred to the Prism 7900HT sequence detection system (ABI) to read fluorescence and call SNPs.

2.2.2   3'Untranslated region and Gpc1 overexpression cloning

Cloning was performed using the infusion cloning system (Clontech). 3'UTR luciferase constructs were cloned from 3'UTRs of potential targets and pMIR-Report vector (Ambion, UK), which contains the luciferase reporter gene. Cloning was performed using the infusion cloning system. Transcriptional activation of the promoter generates luciferase enzyme from the reporter and results in the production of light when substrate is added. The intensity of the light produced is proportional to the amount of enzyme produced and therefore any effect of miRNAs on the 3'UTRs will be observed. Mouse Gpc1 overexpression construct was cloned into HaloTag® Fusion (N-Terminal) expression vector purchased from promega.

2.2.2.1 Infusion primer design

3'UTR sequences were obtained from NCBI. Primers were designed using the Primer3 software with the addition of vector specific 15-16nt infusion adaptors to the 5' end of both the forward and reverse primers, for example 5'-AAAGCTGCGCACTAGT-3' and 5'-ATCCTTTATTAAGCTT-3' for pMIR-Report. Primers for infusion are shown in Table 2.1.

2.2.2.2 PCR Amplification and purification

The PCR was carried out using 4µl 5x Phire buffer, 0.2mM dNTPs, 50ng gDNA, 0.4µl Phire DNA polymerase (Thermo Scientific), 500nM forward and reverse primers and H_2O in a 20µl reaction mix according to the manufacturer's instructions. Each primer pair PCR reaction was performed in quadruplicate, with each sample having a different annealing temperature (55^OC, ~56.5^OC, ~58^OC and 60^OC). PCR cycling were: 95°C for 1 min, 35 cycles of 95°C for 30 sec, gradient of 55-65 for 30 sec and 68°C for 2 min and then 68°C for 5 min. PCR products were visualised by gel electrophoresis as described. PCR reactions with single product of the correct size for each primer pair were combined and purified using Nucleospin extract II kit according to the manufactures protocol.

2.2.2.3 Preparation of plasmid

The pMIR-Report vector was digested with 10U *Hind*III and 10U *Spe*I in 50µl, in a reaction containing 5µl plasmid, 5µl 10x buffer M, (Roche), H_2O to 50µl. The reaction was incubated at 37^oC overnight before the digestion was then heat killed. The digested

plasmid was then run on a 1%TAE agarose gel as described above to verify correct digestion, and to perform gel extraction of the plasmid.

Digested plasmid was extracted from the gel and purified using the QIAquick® Gel extraction Kit according to manufacturer's instructions. Briefly the required band was excised with as little gel as possible. Gel was dissolved in Buffer QG and the plasmid extracted and purified using QIAquick® spin columns prior to elution in 30μL dH$_2$O.

2.2.2.4 Infusion cloning reaction

PCR products of 3'UTRs containing infusion (Clontech) sequence were cloned into pMIR-Report vector using infusion cloning. 10μl infusion reactions containing the following were created; 2μl 5x infusion HD enzyme premix, 1.5μl linearised vector (50ng/μl), 1μl Purified PCR amplicon (stock concentrations ~50ng/μl) and 5.5μl water. Reaction was incubated at 50°C for 15min before being placed on ice.

2.2.2.5 Infusion transformation

2.5μl of the infusion reaction was added to 50μl Stellar Competent cells in 10ml falcons. Reactions were incubated on ice for 30 minutes prior to transformation by heat-shock at 42°C for 45 seconds followed by incubation on ice for 1-2min. 500μl of S.O.C. medium was then added to each vial and cultures were shaken (220 rpm) at 37°C for 1 hour. A 50-75μl aliquot of each transformation was spread onto a pre-warmed agar plate containing 100 μg/ml ampicillin (10.3g LB (easy mix), 7.5g Agar (bacto) and 500μl water was autoclaved then cooled before 100ug/ml Ampicilin was added). Plates were incubated at 37°C for 16 hours in order for colonies to develop. The next day individual colonies were selected and incubated in 3-5ml of LB supplemented with ampicillin (100μg/ml) (10.3g LB (easy mix), and 500μl water was autoclaved then cooled before 100ug/ml Ampicilin was added) at 37°C and 220rpm overnight in an orbital shaker. Standard aseptic technique was used throughout the cloning procedure.

2.2.2.6 Small scale preparation of plasmid DNA

A 1.5 ml aliquot of a bacterial culture was transferred to an Eppendorf and centrifuged for 1min at 13,000 x g. The supernatant was discarded and cells resuspended in 100μl Buffer P1 (Qiagen) by vortexing. 200μl Buffer P2 (sodium hydroxide to lyse cells) (Qiagen) was

added to cells and mixed thoroughly by inverting 6 times. A 150μl aliquot of Buffer P3 (Qiagen) was then added and mixed immediately by inverting 6 times. Lysates were centrifuged at 13,000g for 3 min and the supernatant removed to a new Eppendorf containing 1 ml 100% ethanol. Samples were vortexed for 10-15 sec to mix thoroughly and centrifuged for 10 min at 13000 xg. All ethanol was removed and pellets allowed to air-dry (~15min) before resuspension in 50μl dH$_2$O.

2.2.2.7 Restriction digests to check insertion

A 2μl aliquot of each miniprep plasmid DNA was restriction endonuclease digested using the restriction enzyme HindIII and SpeI. Plasmid DNA was combined with 5U HindIII and SpeI, 1.5μl 10x buffer M and 11μl dH$_2$O in a 15μl reaction and incubated at 37°C for 1h. 10μl of each digest was then run out on a 1% agarose gel to determine if it contained an insert of the correct size.

2.2.2.8 Sequencing to check insertion

10μl miniprep and 100μl 3.2pmol/μl primer was sent to life sciences for sequencing (http://www.lifesciences.sourcebioscience.com/genomic-services.aspx).   Plasmids containing full insert and no errors were selected to study.

2.2.2.9 Sub cloning and large scale preparation of plasmid DNA

To create transfection quality plasmid, small scale plasmid preparations were re-transformed and large scale preparation performed. 1μl Plasmid (mini prep) was added to 25μl chemically competent DH5α (Invitrogen, Paisley, UK) in an Eppendorf and incubated on ice for 30min, before being heat shocked at 42°C for 30sec and returned to ice for 2min. 500μl LB was added and cells incubate at 37C for 45-60min, 75-150μl was spread on a pre-warmed agar plate containing 100 μg/ml ampicillin and incubated overnight. The next day several colonies were selected and incubated in 5ml of LB supplemented with ampicillin (100μg/ml) at 37°C and 220rpm for 5h in an orbital shaker. The 5ml of grown culture was then transferred to 250ml of LB supplemented with ampicillin (100μg/ml) and grown at 37°C overnight in an orbital shaker at 220rpm. Plasmids were then extracted using Qiagen Maxiprep Kit (Qiagen) according to the manufacturers protocol.

2.2.2.10 Site directed mutagenesis

To validate miRNA binding to 3'UTRs miRNA binding sites in 3'UTR were mutated. Positions complementary to positions 2, 3, 4 and 5 of the miRNA seed were mutated to the complementary sequence to ensure miRNA binding is disrupted. Primers were designed so they are complementary to the 3'UTR for 15 bp either side of the region to be mutated (Table 2.1).

Mutagenesis was performed using the QuikChange II kit (Agilent, Berkshire, UK), according to the manufactures protocol. Briefly, a reaction mix totalling 50µl and contained 5µl of 10x reaction buffer 50ng dsDNA template 125ng forward and reverse mutant primers, 0.01mM dNTP mix, 2.5U Pfu Ultra HF DNA polymerase and H_2O. Mix was then subject to the following cycling conditions; 95C for 30s, followed by 18 cycles of 95C for 30s, 55C for 1min and 68C for 1min per kb of plasmid. The reaction was then placed on ice for 2min. 1µl of *Dpn*1 restriction endonuclease enzyme was then added to each reaction, pipetted up and down and pulsed then incubated at 37C for 1hour.

Following mutagenesis reaction the new mutated plasmid was transformed into XL1 Super Competent cells according to the manufactures protocol. Briefly 1µl of each sample was added to 50µl XL1 Super Competent cells in 15ml falcons, gently swirled to mix and incubated on ice for 30 min. The cells were then heat shocked at 42C for 45 sec and placed on ice for 2min. 500µl of N2X broth (preheated to 42C) was added to the sample before being incubated for 1h at 37 at 37C in the shaking incubator (220rpm) and spread on to Ampicillin containing agar. Mutated plasmids were grown up and sequenced as previously described.

2.2.3 General cell culture

C3H10T1/2 were maintained in MEM culture medium containing 2mM L-glutamine, 1% non-essential amino acids, 10% FCS, 100U/ml penicillin, and 100µg/ml streptomycin. SW1353 were maintained in DMEM culture medium, containing 2mM L-glutamine, 1% non-essential amino acids, 10% FCS, 100U/ml penicillin, 100µg/ml streptomycin. Cells were grown in vented T75cm^2 flasks, at 37ºC in 5% (v/v) CO_2/humidified air until approximately 90% confluent. Cells were detached with trypsin and split into the

appropriate culture vessel for experimentation (either 96 (for RNA), 24 (for alkaline phosphatase assay) or 12 (protein for western) well plates and 6cm dish (protein for mass spectrometry) or into further T75cm^2 flasks for continuation of the line every 2-3 days.

For long term storage in liquid nitrogen, cells were detached with trypsin, pelleted by centrifugation at 200 x g for 5 minutes and resuspended in cryo-freezing medium (90% FCS with 10% DMSO), before being frozen slowly (1°C/min) in a cell freezing container overnight at -80°C prior to transfer to liquid nitrogen store.

2.2.4 Cell stimulation

Prior to stimulation, cells were plated in appropriate culture vessel at such a density to either ensure 80-90% confluence after 24h growing time or ensure appropriate confluence for transfection after 24h and then grown for a further 24h (either with or without transfection) depending on experiment type. Serum containing media was then aspirated off, and cells were washed twice in PBS to remove traces of serum, and then left for 16-20 hours in serum-free culture medium. Cells were then stimulated with recombinant Indian Hedgehog (R&D systems), Purmorphamine (Calbiochem), Recombinant BMP-2 (R&D systems) or smoothened agonist SAG (Calbiochem) for the time and concentration indicated in each figure.

2.2.5 siRNA and miRNA transfection

Pre-designed standard si*GENOME SMART*pool siRNAs against Gli1, Gpc1, SMO and Ptch and siGENOME® Non-Targeting Pool #2 (siControl2) were purchased from Dharmacon (Cramlington, UK). The use of a *SMART*pool maximises the chance of successful interference by providing four highly functional duplexes that target different regions of the target gene, while potentially minimising off-target effects due to the reduced concentration of each siRNA in the pool.

C3H10T1/2 cells were cultured as above, and seeded in a 96, 48 or 12 well plate (density of 10,000 cells/cm^2) and grown for 24h to reach around 50% confluence. Using Dharmafect transfection reagent (Dharmacon) cells were transfected for 24h (unless stated otherwise) with miRNAs and/or siRNAs (all 100nM). Control cells were also treated with Dharmafect alone and or miRNA or siRNA non-targeting control, miCon2 and siCon2

respectively (Dharmacon). Transfection was performed according to the manufacturer's instructions with the desired siRNA concentration. Briefly, two tubes were prepared (tube 1 and tube 2). In preparation of tube 1 mi/siRNAs were diluted to 20µM in 1 x DharmaFECT buffer. The appropriate volume of mi/siRNA (0.5µl per well for 96 well plate to give 100nM final) was added to the appropriate volume of H_2O and MEM/DMEM media (without additional reagents) to dilute the mi/siRNA 20-fold and to 10 x the final concentration required and ensuring half the volume of tube 1 was media. Tube 2 was prepared by diluting the appropriate amount of DharmaFECT transfection reagent with media alone to dilute the DharmaFECT transfection reagent 50-fold and to 10 x the final concentration required. Tube 1 and 2 were left for 5 minutes at room temperature. The contents of each tube were then combined and left for 20 minutes to form mi/siRNA containing liposomes. Sufficient serum-containing medium was then added to the mix to complete the desired volume of transfection medium for the culture vessel. The culture medium was removed from the vessel by aspiration, and the appropriate volume of transfection medium was added to each well. After the 24h incubated at 37^oC in 5% (v/v) CO_2, transfected cells were washed with PBS and used in appropriate experiments. Appropriate scaling of volumes was used to transect different sized culture vessels.

2.2.6 3'UTR Luciferase-miRNA co-transfection

To assess the effect miRNAs have on 3'UTRs, 3'UTR luciferase constructs and the appropriate miRNA were co transfected into either SW1353s or C3H10T1/2 cells (for human and mouse 3'UTR constructs respectively). Cells were plated out in 48 well plates, After 24h cells were transfected with 3'UTR luciferase constructs using FUGENE® HD transfection reagent.

Briefly a mix totalling 15µl per well and consisting of 200ng plasmid, 0.9µl Fugene HD and 8.7µl NIL media per well was created and left to incubate for 15min at room temperature. At the same time Dharmafect transfection for miRNAs was also prepared (all 50nM). Briefly, two tubes were prepared (tube 1 and tube 2). In preparation of tube 1 miRNAs were diluted to 20µM in 1 x DharmaFECT buffer (Dharmacon, Lafayette, CO, USA). 0.75µl (per 48 well) miRNA (to give 50nM final) was added to 14.25µl (per48well) H_2O and 15µl (per 48 well) 10T1/2nil media. Tube 2 was prepared by diluting 0.6µl (per48well)

DharmaFECT transfection reagent with 29.4µl (per 48well) 10T1/2nil media to dilute the DharmaFECT transfection reagent. Tube 1 and 2 were left for 5 minutes at room temperature. The 30µl of each tube were then combined (60µl) and left for 20 minutes to form miRNA containing liposomes. The 60µl DharmaFECT-miRNA mix and 15µl Fugene HD-plasmid mix were then both added to a fresh tube containing 225µl serum containing media (total 300µl), then gently mixed. The original media was then removed and replaced with the 300µl 'DharmaFECT-miRNA-Fugene HD-plasmid mix' and incubated at 37ºC in 5% (v/v) CO_2 for 24 hours.

2.2.7 Luciferase assay

Upon completion of the experiment (24h post transfection) cells were washed in PBS before the addition of 75µl 1 x reporter lysis buffer (Promega, Southampton, UK). Plates were frozen containing the lysis buffer at -20ºC until ready for use. Following equilibration to room temperature, lysates were briefly mixed by gentle agitation. Then 10µl of cell lysate was transferred to a cross-talk free, white walled 96-well plates (PerkinElmer, Wellesley, MA, USA) was mixed with 50µl of the luciferase substrate firefly luciferin (Promega, Southampton, UK), reconstituted in luciferase assay buffer according to the manufacturer's instructions and read immediately using the Berthold MicroLumatPlus LB96V (Bundoora, Australia).

2.2.8 Alkaline phosphatase assay
C3H10T1/2 cells were cultured as previously described, and plated at 36,000 cells per well in 24 well plates. After 24 h cells were transfected as described. After 48 h the cells were stimulated as shown in serum containing medium for 5 days. Medium was refreshed after 3 days. After 5 days of stimulation the cells were assessed for alkaline phosphatase activity. Briefly, cells were washed with PBS, 50mM HEPES pH 8.3 and fixed with 4% (w/v) paraformaldehyde for 5-10 min. 500µl alkaline phosphatase substrate (p-nitrophenylphosphate (Alkaline Phosphate Yellow Liquid Substrate for ELISA'), Sigma Aldrich) was added. In the presence of active alkaline phosphatase, p-nitrophenol is produced whose levels can then be measured. After 15 min (or when the yellow had

developed), 100µl was removed and placed in a 96 well plate and absorbance read at 405nm, using a TECAN plate reader and Xflor 4 software.

2.2.9 Micro mass/ alcian blue assay

C3H10T1/2 cells were cultured and transfected as previously described and indicated in each figure. Cells were trypsinised, counted, centrifuged, washed and resuspend at a concentration of 10×10^6/ml and plated into drops of 10µl, each containing 100,000 cells. Drops of cells were left to adhere for 2 h before appropriate medium was added. Micro mass cultures were cultured for the time indicated in each figure. After cell culture Alcian Blue was used as a marker of chondrogensis and extracted using Guanidinium hydrochloride (GHCl) to quantify. Alcian blue solution was prepared by adding 1 g Alcian blue to 100ml 0.1M HCl. Solution was mixed extensively at 37^oC for 3days.

2.2.10 Western blotting

2.2.10.1 Whole protein extraction

Cells were plated into the appropriate culture vessel and after the desired culture conditions, medium removed and cells rinsed with ice-cold PBS. Lysis buffer (50mM Tris-HCl, pH 7.4, 10% glycerol (v/v), 1mM EDTA, 1mM EGTA, 1mM Na_3VO_4, 5mM NaF, 10mM β glycerol phosphate, 5mM $Na_4P_2O_7$, 1% Triton X-100 (v/v), 1µM microcystin-LF and 1 Complete protease inhibitor Mini Protease Inhibitor Cocktail Tablet from (Roche Diagnostics) in a final volume of 50 ml) was prepared by adding 1µl 2-mercaptoethanol per ml lysis buffer. The PBS was then removed and replaced with ice-cold Lysis buffer (75µl/well for 24-well plate, 75µl/well for 12-well plate, 100µl/well for 6-well plate and 150µl/6cm dish). The cells were scraped into the lysis buffer then transferred to a fresh Eppendorf on ice. Samples were incubated for 20 minutes on ice, followed by centrifugation at 13,000 x g at 4^oC for 3 minutes. The supernatant was then removed and stored immediately at -80^oC, prior to protein quantification.

2.2.10.2 Protein Quantification

To determine protein concentration Bradford reagent (BioRad) was used. A 2mg/ml stock BSA protein standard solution (Pierce & Warriner, Chester, UK) was diluted to 0.4mg/ml BSA in distilled H_2O (dH_2O), and a series of standards from 0-4mg/ml added to a flat-

bottomed 96-well plate. Cell lysates and equal quantities of cell lysis buffer to act as blanks, were added to the 96-well plate. Bradford assay reagent was then added to all the wells according to the manufacturer's instructions. The samples were then gently mixed, left to stand for 5 minutes and the absorbance was read at 595nm using a Tecan Sunrise microplate absorbance reader. Protein concentrations were calculated and where possible protein equalized, prior to SDS-PAGE analysis.

2.2.10.3 SDS-PAGE

Cell lysates were thawed on ice and Laemmli sample buffer (0.1 M Tris-HCl, pH 6.8, 0.35 M SDS, 20% (v/v) glycerol, 0.01% (w/v) bromophenol blue and 5% (v/v) β-mercaptoethanol) added at a 5:1 ratio. Lysates were heated to 100°C for 5 min, cooled on ice and Proteins were then separated upon size by SDS-PAGE electrophoresis on appropriate 7.5 or 10% SDS-polyacrylamide gels at 180V for approximately 1 hour with pre-stained ladder SMO671 (Thermo).

2.2.10.4 Immunoblotting

Proteins were then transferred to PVDF membrane by electroblotting for 1-1.5 hours at 1mA/cm^2 in transfer buffer (20 mM Tris-HCl, 0.6 M glycine and 20% (v/v) methanol). Membranes were covered in blocking buffer (TBS-T (150 mM NaCl, 10 mM Tris-HCl, pH 7.5 and 0.1% (v/v) Tween-20), 5% non-fat dry milk powder (w/v)) for 1 hour at room temperature. Membranes were subsequently washed 3 times in TBS-T prior to incubation with primary antibody (diluted accordingly in to the manufacturer's instructions (1/1000 unless otherwise stated)) solution overnight at 4°C with gentle agitation. Following incubation, membranes were washed 3 times for 5 minutes in TBS-T and incubated with a horseradish peroxidase (HRP)-conjugated secondary antibody for 1 hour at room temperature and washed a further three times in TBS-T for 5 minutes each as before. Secondary antibodies were diluted 1/2000 in a solution of 5% (w/v) milk powder in TBS-T. The membranes were washed 3 x 5 min in TBS-T before visualisation on high-speed Kodak X-ray film (Sigma-Aldrich) using Amersham ECL, ELC plus or ECL advanced (GE Healthcare) according to manufacturer's instructions. Membranes were stripped of antibody in stripping buffer (1.5% (w/v) glycine, 0.1% (w/v) SDS and 1% (v/v) Tween 20) overnight with gentle agitation at room temperature. The membranes were washed 3 times for 5

minutes each in TBS-T, reblocked in blocking buffer for 1 hour at room temperature and incubated with the desired primary antibody overnight as described above. GAPDH (Chemico International) or β-tubulin antibodies were used as a loading control.

2.2.11 Real-time RT-PCR quantification of miRNAs and protein coding RNAs

2.2.11.1 RNA extraction using Cells to cDNA II

The Ambion Cells-to-cDNA II lysis buffer (Huntingdon, UK) lyses cells and inactivates RNases in a single step yielding a cell lysate that is competent for reverse transcription. This extraction method was used for cell numbers <1 x 10^4 seeded in a 96-well plate. Cell lysis was performed according to the manufacturer's instructions. Briefly, after the desired period of growth factor/cytokine stimulation, cell monolayers were washed with ice-cold PBS and lysed directly by addition of 30µl Cells-to-cDNA™ II Cell Lysis Buffer. The 30µl lysis buffer was then transferred to an ice-cold 96-well PCR plate, which was incubated at 75°C for 15 min to inactivate RNases.

2.2.11.2 Reverse transcription for protein coding genes following Cells to cDNA II

Expression of Gli1, Smo, Ptch, Gpc1 and 18s. mRNA was performed as previously described [181]. An 8µl aliquot of each RNA sample was transferred to a new 96-well PCR plate for reverse transcription. RNA samples were combined with 0.625 mM dNTP and 0 .2 µg p(dN)$_6$ and incubated at 70°C for 5 min. Samples were cooled on ice followed by the addition of a reaction mix consisting of 10 mM DTT, 100 U M-MLV, 4µl 5X First-Strand Buffer (Invitrogen) and 1.5µl dH$_2$O added to each sample giving a final reaction volume of 20µl. The plate was then incubated at 37ºC for 50 min, followed by 70°C for 15 min. Samples were diluted by addition of 30µl dH$_2$O (Sigma-Aldrich, Poole, UK) for target gene quantification, a further 1:10 dilution in H$_2$O for quantification of 18S housekeeping gene, and stored at -20°C for further use. 5µl aliquots were used in PCR reactions.

2.2.11.3 Reverse transcription for miRNAs following Cells to cDNA II

MiRNA Reverse Transcription (Applied Biosystems) was performed according to the manufactures instructions, but in 15µl reactions. Briefly, 5µl total RNA sample was mixed with 0.15µl dNTPs (100mM), 1µl multiscribe (50U/µl), 1.5µl 10x RT buffer, 0.188µl RNase inhibitor, 4.162µl Nuclease free water and 3µl 5x miRNA specific RT Primer. The reaction is

then subject to the following thermal conditions, 16^oC for 30 min, 42^oC for 30 min, 85^oC for 5 min then cooled to 4^oC. 5µl of the above reaction was diluted 1/6 with 25µl H_2O and 4µl of this added to 10µl reaction real time reaction (total dilution of RT product into real-time reaction was 15 fold).

2.2.11.4 real-time RT-PCR of protein coding genes

Primers for real-time RT-PCR were designed with Universal Probe library (Roche Diagnostics, Burgess Hill, UK). All primers were designed to span an intron-exon boundary to prevent amplification of any contaminating gDNA.

PCR reactions for Roche probe library assays were prepared by combining 4µl diluted cDNA with 5µl of TaqMan® Gene Expression Master Mix (Applied Biosystems), 30nM of each primer and 15nM probe in a final volume of 10µl. Cycling conditions were: 50°C for 2 min, 95°C for 10 min and 40 cycles of; 95°C for 15 sec, 60°C for 1 min.

TaqMan gene expression assays are made of a ready mix of primers and probe required for real-time RT-PCR detection that is based on the TaqMan® Probe-based real-time RT-PCR system described above.

2.2.11.5 real-time RT-PCR of miRNAs

Real-time RT-PCR was carried out using Applied Biosystems assays for miRNAs. Individual miRNA RTs are performed separately along with a standard RNA RT for 18S. 4µl cDNA of the diluted miRNA RT reaction was added to each 10µl miRNA real-time RT-PCR reaction. Each reaction consists of the following; 5µl TaqMan® Gene expression mix (TGE), 0.5µl H_2O, 0.5µl Assay (primer and probe for the studied miRNA), 4µl cDNA (1 in 6 diluted RT product). The reaction was then subject to the following thermal conditions; 95C for 10 min followed by 40 cycles of 95^oC for 15s and 60^oC for 60s and read using the ABI PRISM 7900HT Sequence Detection System (Applied Biosystems).

2.2.12 Immunofluorescent cell imaging

Cells were seeded into chamber slides and subject to the appropriate experimental conditions. Afterwards cell monolayers were washed twice in PBS for 5 min and fixed in 4% (w/v) cold preparation of paraformaldehyde in PBS for 10 min at room temperature.

Cells were permebilised and blocked with PBG-Saponin solution (0.2% (v/v) fish skin gelatine, 0.5% (w/v) BSA and 0.5% (v/v) Triton-X-100 in PBS) for 10 min or PBG- Triton solution for 45 min at room temperature. Cells were incubated with primary antibody at a final concentration of 1 μg/ml for 45mins. Cells were then incubated with a secondary antibody (either; Alexa Fluor® 488 goat anti-mouse IgG, Alexa Fluor® 594 goat anti-mouse IgG, Alexa Fluor® 594 goat anti-rabbit IgG Alexa and Alexa Fluor® 488 goat anti-rabbit IgG) diluted 1:1000 (v/v). Cells were washed twice in between the antibody incubations with PBG-Saponin or PBG-Triton. Slides were washed twice in PBS and mounted in Vectashield with DAPI (Vactorlabs). Cells were visualised using confocal and fluorescent microscopy. Images were analysed using ImageJ analysis software (Wayne Rasband, NIH, USA). Individual cells were selected and the average quantity of red (acetylated alpha tubulin) and/or green (Smo) present in the cells (per area) calculated. The length of primary cilia was also calculated Using ImageJ.

2.2.13 SILAC and mass spectrometry

C3H10T1/2 cells were cultured for more than five population doublings (7 days of culture), in DMEM containing either isotopically labelled ^{13}C L-Lysine-2HCl and ^{13}C ^{15}N L-Arginine-HCl (Heavy), or normal Lysine and Arginine (light). Both heavy and light media contained dialysed FBS (all SILAC media, amino acids and FBS from Thermo Scientific). After 7 days of culture, the isotopic incorporation of the heavy amino acids was assessed. More than 99% of the peptides assessed by mass spectrometry contained heavy Arginine and Lysine. Both Heavy and light cells were plated out in 6cm dishes, after 24h the cells were around 50% confluent. Cells were then transfected in 2.5ml of media with 100nM of either miCon2 (non-targeting miRNA mimic) or miR-324-5p mimic in the light and heavy cells respectively using 5μl Dharmafect transfection reagent 1 (DF1) (Dharmacon). After 24h of transfection cells were serum starved for a further 24hr. For stimulated SILAC experiments, both the light (miCon2) and heavy (miR-324-5p) were stimulated with 2μg/μl recombinant Ihh. For the unstimulated SILAC experiment, neither the light (miCon2) nor the heavy (miR-324-5p) were stimulated, instead cells were left in serum free media for a further 48hr, for continuity of experiments. All cells were then lysed using 150μl of lysis buffer, as described for immunoblotting. The Heavy and Light lysates were then mixed at a ratio of

1:1. Protein was boiled and separated upon size by SDS-PAGE, as described for western blot. Nanopure H_2O was used in all gels and buffers to minimize contamination of keratin. Gel preparation, protein digestion, mass spectrometry and initial processing of data was performed by NEPAF (Newcastle). Briefly, gels were then cut into 12 segments. Each segment was then digested with trypsin and peptides separated by liquid chromatography tandem mass spectrometry, (LC-MS/MS and analysed by mass spectrometry, search engine identifications performed using MASCOT (Matrix Science Company). Data was analysed using MaxQuant software as described by Cox *et al.* in 2009 [182]. Subsequent data analysis was performed in Microsoft Excel 2007 and is described in Chapter 5.

2.2.14 Statistical analysis

For real-time RT-PCR the relative amount of target gene and housekeeping gene were calculated using delta ct ($2^{\wedge(gene\ of\ interest\ ct-housekeeper\ ct)}$) or ($2^{\wedge gene\ of\ interest\ ct}/2^{\wedge house\ keeper\ ct}$). Data were normalised against the basal levels of genes, then plotted as the fold induction of gene expression over control levels. Fishers 2-tail exact and Chi squared tests were used to test the significance of categorical data, and were also used to calculate the significance of target enrichment in the sliding window and expanding window analysis of ordered gene lists in Chapter 5. Non-parametric Mann Whitney U test was used test the signigicance of the miRNA screen in chapter 4. Student's two tailed t-test was used to calculate significance of pair wise data. Statistical differences between sample groups across independent experiments were calculated using one way analysis of variance (ANOVA) followed by a Bonferroni post test to account for multiple comparisons. Statistical differences were shown as $p<0.05=*$, $p<0.01=**$ and $p<0.001=***$.

Table 2.1Table of primers

Species	Primer	Sequence
	miR-140 genotyping primers	
Human	miR140_RFLP_F	AGACCTCTCCCGGTGTGGGC
Human	miR140_RFLP_F	TCAGAGTCCTTTTGGGCTTG
Human	miR140_seq_F	TGTAAAACGACGGCCAGTGTGTGGGCATGTCTTCCCGC
Human	miR140_seq_R	CAGGAAACAGCTATGACCGAGTCCTTTTGGGCTTGAGC
	Infusion cloning primers	
Mouse	GPC1_CDS_F	GAGCTCAACCGCGGATATCTAGAATGGAACTCCGGACCCGAGGCT
Mouse	GPC1_CDS_R	CTGGAATTGGGCCCAAATCTAGATTACCGCCACCTGGGCCTGGCT
Mouse	Serpine2_3UTR_F	AAAGCTGCGCACTAGTGACGCAAGTGTTTCTGGTC
Mouse	Serpine2_3UTR_R	ATCCTTTATTAAGCTTGCAAATACTCGAGAGGGTTGTT
Mouse	Gpc1_3UTR_F	AAAGCTGCGCACTAGTGTCCCCAAAAGCCATGTAT
Mouse	Gpc1_3UTR_R	ATCCTTTATTAAGCTTAAAGCGGTAAGCAGCCTTTT
Mouse	Anxa4_3UTR_F	AAAGCTGCGCACTAGTAACTTCATTTTTCTGCACTGCT
Mouse	Anxa4_3UTR_R	ATCCTTTATTAAGCTTCTGAGGAATGTTCAGCACGA
	Sequencing primers	
HaloTag	pHTN F	GGACCTGATCGGCAGCGAG
HaloTag	pHTN R	GGTGTGAAATACCGCACAG
pmirReprot	M13F	TGTAAAACGACGGCCAGT
	Mutagenesis primers	
Mouse	APP_mut324_F	TAAAAATCGATGGGGctacCTTCTTGTGAACGTGG
Mouse	APP_mut324_R	CCACGTTCACAAGAAGgtagCCCCATCGATTTTTA
Mouse	GLI1_mut324_F	CATGAGGTGCCCAGGctacGGAGGTTTGGGCTGGG
Mouse	GLI1_mut324_R	CCCAGCCCAAACCTCCgtagCCTGGGCACCTCATG
Mouse	GPC1_324_mut_site_1 F	AGGAAGCCTGCAAGGctacCCAGTATGTTGCTGTC
Mouse	GPC1_324_mut_site_1 R	GACAGCAACATACTGGgtagCCTTGCAGGCTTCCT
Mouse	GPC1_324_mut_site_2 F	TCACCTGGCCATGGGctacCTGGGTGGCTGGTGAA
Mouse	GPC1_324_mut_site_2 R	TTCACCAGCCACCCAGgtagCCCATGGCCAGGTGA
Mouse	GPC1_324_mut_site_3 F	TTCCAGGGCCTAGGGctacCTGAGTTGCTATATCC
Mouse	GPC1_324_mut_site_3 R	GGATATAGCAACTCAGgtagCCCTAGGCCCTGGAA
	Real time primers	
Mouse	Ptch1_pr56_F	GGAAGGGGCAAAGCTACAGT
Mouse	Ptch1_pr56_R	TCCACCGTAAAGGAGGCTTA
Mouse	Gpc1_pr79_F	ATTGCCGAAATGTGCTCAA
Mouse	Gpc1_pr79_R	GGCCCCAGAACTTGTCAGT
Mouse	Gli1	Applied Biosystems assay
Mouse	Smo	Applied Biosystems assay
Mouse	miR-324-5p	Applied Biosystems assay
Mouse	miR-125b	Applied Biosystems assay
Mouse	18S	Applied Biosystems assay

Chapter 3 Genetic association of miR-140 and its targets

3.1 Introduction

3.1.1   miR-140

miR-140 is one of the most studied miRNAs associated with OA and cartilage. Its expression was first shown to be cartilage specific in zebrafish [138] and then in mouse by Tuddenham *et al.* [139]. Tuddenham *et al.* also found miR-140 targets histone deacetylase 4 (Hdac4), which represses Runx2, a transcription factor that controls chondrocyte hypertrophy and osteoblast differentiation [139]. miR-140 is also increased following differentiation of mesenchymal cells to cartilage [140] and plays a role in both cartilage development and maintenance [142]. miR-140 has been shown to decrease cartilage destruction and increase cartilage growth and development. Mice lacking miR-140 (miR-140-/-) gain an OA like phenotype [142]. Conversely, mice overexpressing miR-140 are resistant to antigen-induced arthritis [142]. This is proposed to be because translation of ADAMTS5, a protease that mediates pathological aggrecan cleavage in cartilage [183], is inhibited by miR-140 [142]. miR-140 has been shown to be decreased in OA [156], which may contribute to some of the altered gene expression (including increased ADAMTS5) observed in OA [156]. Although a study in our laboratory, actually shows miR-140 expression to be increase in OA [155]. miR-140 also decreases IL-1 induced MMP13 production [140, 184], the same MMP family member believed to be involved in cartilage type II collagen cleavage [42].

3.1.2   SNPs and miRNAs

Single nucleotide polymorphisms (SNPs) can create, destroy, or modify miRNA:target interactions. SNPs can occur in the 3'UTR of miRNA targets, in miRNA seeds, in miRNA stem loops and in miRNA flanking regions. The first SNP found in a 3'UTR (target) which was hypothesised to alter a miRNA binding was thought to contribute to Tourette's syndrome [185]. Another interesting example of where a SNP in a 3'UTR creates a miRNA binding site is in Texel sheep where a mutation causes miRNA regulation of myostatin (GDF8) in muscle, leading to muscle hypertrophy and meaty sheep [186]. SNPs in 3'UTRs can also disrupt miRNA binding sites, such as the miR-433 binding site in FGF2, which may

45

contribute to Parkinson's disease [187] and the miR-433 binding site in *HDAC6*, which may cause a chondrodysplasia [188].

SNPs in the miRNA coding region can also alter the function of miRNAs. SNPs within the mature miRNA can change the repertoire of miRNA targets, especially if the SNP occurs in the miRNA seed [189]. Although SNPs in the stem loop do not directly change the mature miRNA sequence, they can change how the miRNA is processed, leading to altered miRNA-5p to miRNA-3p expression ratio [190], meaning SNPs outside of the mature miRNA can alter the expression of targets. miRNAs are processed from larger pieces of RNA meaning it is likely SNPs in miRNA flanking regions can also alter miRNA function by either changing RNA structure, meaning the miRNAs are processed differently, or changing a transcription regulation site. Studies have shown SNPs in miRNAs, their targets and in particular seed regions, are rare [189], as they are under negative selection [191], although some are under positive selection [192].

3.1.3 Previous data on SNPs in miR-140

SNP rs7205289 (A/C), which is in very close proximity to miR-140 is thought to influence the processing of miR-140 and is associated with cleft palate in the Chinese population [190]. rs7205289 is a 'C' to 'A' transition, 'A' being the minor allele, having a prevalence of 11.6% in the control Chinese population, with cleft pallet being associated with an increase in this allele (OR=0.55, p=0.001) [190]. The SNP is located at the miR-140 Drosha cleavage site (Figure 3.1) and the A allele is predicted to alter the processing of the stem loop to the mature miRNA, decreasing miR-140-5p and increasing miR-140-3p [190]. miR-140-5p has been shown to regulate Pdgf signalling, a pathway involved in cleft palate [193]. The function of miR-140-3p is less well understood. Interestingly, miR-140 is transcribed from within WWP2, a gene which is also required for palatogenesis [194]. As miR-140 is an important miRNA in OA and SNP rs7205289 may affect miR-140 processing, it is plausible SNP rs7205289 is associated with OA. In addition, the region of chromosome 16 where miR-140 is located has been suggested to be a possible susceptibility loci for OA [195], and WWP2 is also involved in cartilage biology by interacting with the cartilage transcription factor Sox9 [194]. The aim of this Chapter was to determine if any SNPs which may alter miR-140 function are associated with OA.

Specific aims

Aim 1: Determine if SNPs in 3'UTRs can alter targeting by miR-140.

Aim 2: Determine if rs7205289 is associated with OA.

Aim 3: Identify additional SNPs in and around miR-140.

Aim 4: Determine if rs2102066 (identified in Aim 3) is associated with OA.

Figure 3.1 SNP locations in miR-140

miR140 Stem Loop

■ rs7205289
■ rs2102066
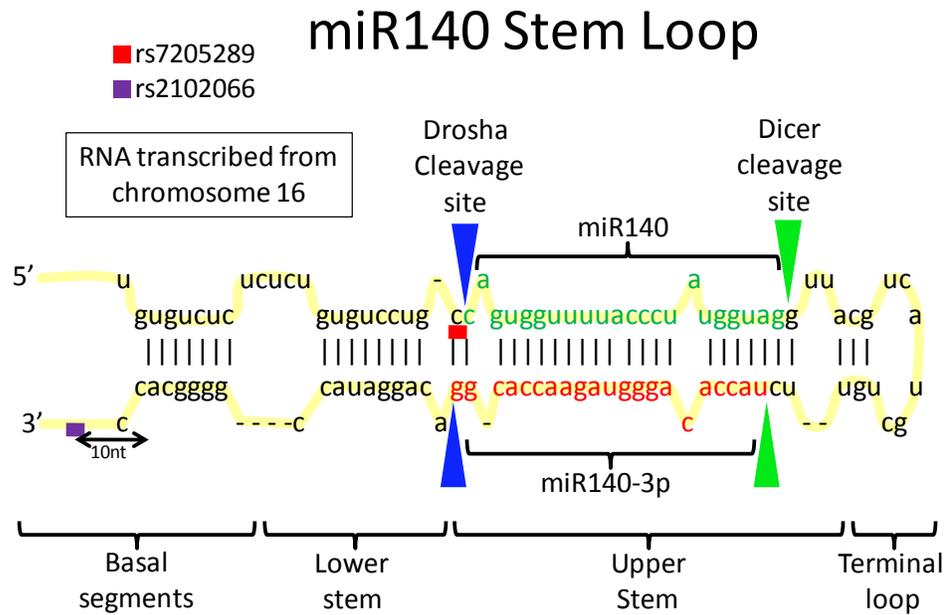
Figure 3.1 The stem loop structure of miR-140. RNA is transcribed from chromosome 16. Green lettering indicates miR-140 (miR-140-5p). Red lettering indicates the miR-140-3p strand. miRNA stem loop can be split into four sections, terminal loop, upper stem (contains mature miRNAs), lower stem and basal segments. Blue arrows indicate the Drosha cleavage site and green arrows indicate the Dicer cleavage site. Red square indicates SNP rs7205289 (A/C) and the purple square indicated the approximate position of rs2102066 (G/A).

3.2 Results

3.2.1 rs35592567 within OA associated gene TP63, may affect targeting by miR-140

Firstly, I investigated SNPs in the targets and predicted targets of miR-140. To determine if there were any SNPs that disrupted validated miR-140 targets I used miRTARBASE (list of validated targets; http://mirtarbase.mbc.nctu.edu.tw/) [134] and miRNASNP (list of SNPs in predicted miRNA targets; www.bioguo.org/miRNASNP/) [189] databases. According to miRTARBASE there were 8 partially or fully validated targets for miR-140-5p and miR-140-3p and according to miRNASNP there were 72 and 70 SNPs which led to the loss and gain of miR-140-5p targets and 120 and 160 SNPs which led to the loss and gain of miR-140-3p targets respectively. There were no SNPS which led to the loss of any of the 8 validated miR-140 targets, I therefore turned our attention to predicted miR-140-5p and miR-140-3p targets that SNPs may alter miRNA binding. Interestingly one (TP63) of the 72 predicted miR-140-5p targets which were predicted to be disrupted by a SNP (rs35592567), has been associated with OA (rs12107036 was signal SNP in the study (arcOGEN) that found the association of TP63 with OA) [39]. The major C allele at rs35592567 has a frequency of around 80% and was predicted to allow miR-140-5p to target TP63, whereas the minor T allele was not predicted to allow miR-140-5p to target TP63 (Figure 3.2).

Figure 3.2 Schematic showing rs35592567 may disrupt miR-140-5p binding to TP63 3'UTR

```
human TP63 wt(A)   5'-AUAGUAAGCAUAGAAACCACUA -3'
                        || :| :      :|||||||||
miR-140-5p         3'-GAUGGUAUCCCAUUUUGGUGAC -5'
                        || :| :      :||||||X|
human TP63 mt(T)   5'-AUAGUAAGCAUAGAAACCAUUA -3'
```

Figure 3.2 Schematic showing rs35592567 may disrupt miR-140-5p binding to TP63 3'UTR. miR-140-5p seed sequence is highlighted in red. rs35592567 minor allele is highlighted in red and underlined.

3.2.2 rs7205289, which is at the miR-140 Drosha cleavage site is not detected in Northern Europeans

Secondly, I wanted to investigate SNPs within the miR-140 locus. rs7205289 is a SNP found in the miR-140 stem loop in the Chinese population [190]. To determine if the minor 'A' allele at rs7205289 was present in Northern Europeans, PCR followed by restriction fragment length polymorphism (RFLP) with restriction enzyme *Bsr*I, was used. PCR was used to amplify a 390bp amplicon which included the rs7205289 SNP. When there is a C allele (major allele) at rs7205289 there is no *Bsr*I restriction site, when there is an A allele (minor allele) at rs7205289 there is a *Bsr*I restriction site. Genomic DNA from 91 male Northern European hip OA suffers and 87 male Northern European controls, who took part in the arcOGEN study cohort [39], were tested. Of the 178 northern Europeans tested none had the minor 'A' allele (Figure 3.3B). To confirm this data and to identify additional SNPs in and around the genomic region encoding miR-140, the region was sequenced in the same 178 individuals. All individuals tested were GG at rs7205289, confirming the RFLP data (Figure 3.3C) and suggesting rs7205289 is not polymorphic in the Western population.

Figure 3.3 rs7205289 is not present in northern Europeans
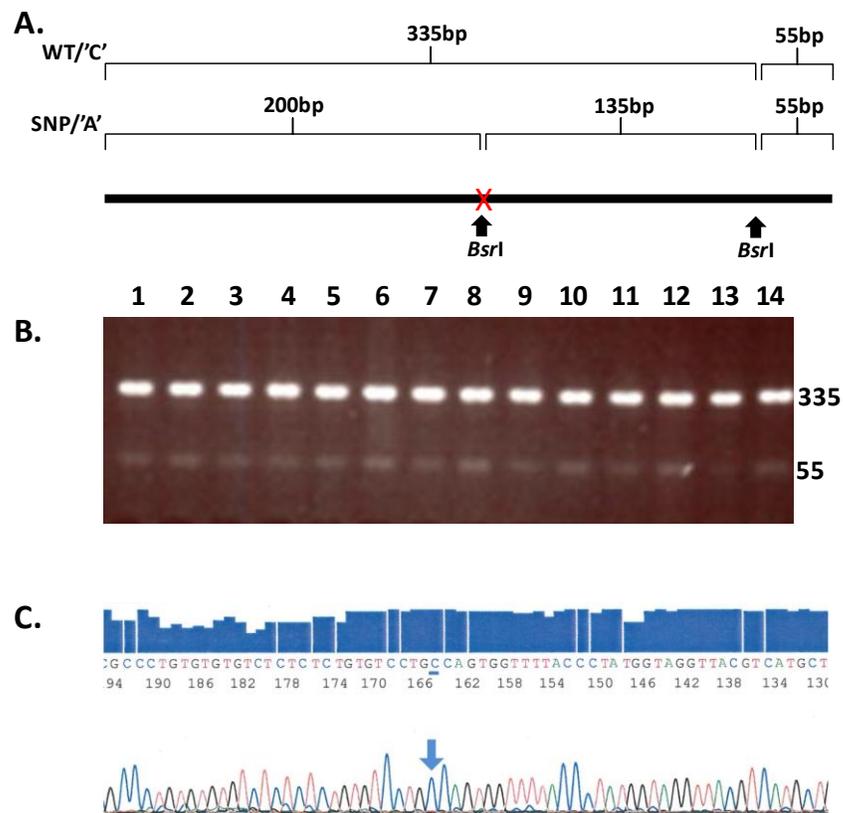
Figure 3.3 rs7205289 is not present in Northern Europeans. (A) Schematic showing RFLP assay on 390bp amplicon, the additional *Bsr*I restriction site (not affected by rs7205289) means wild-type (WT) alleles are cleaved once into 335bp and 55bp fragments, A allele at rs7205289 is cleaved three times into 200bp, 135bp and 55bp fragments. rs7205289 is indicated by red cross. (B) RFLP gel showing all alleles are WT (C allele). (C) Sequencing confirmed all tested alleles were WT (C allele).

3.2.3 Initial data shows rs2102066, which is 10nt downstream of miR-140 stem-loop, is present in Northern Europeans and may be associated with male hip OA

Besides rs7205289, other SNPs may exist in the miR-140 region. Two additional SNPs, rs111366342 and rs112773205, are annotated within the miR-140 stem loop (miRNASNP) however I found neither were polymorphic. There are 10 annotated SNPs in the 1kb upstream of miR-140 and 17 annotated SNPs in the 1kb downstream of miR-140 (miRNASNP). Our data shows one of these SNPs, rs2102066, which is only 10nt downstream of the miR-140 stem loop (purple square in Figure 3.1), is polymorphic. rs2102066 is a 'A' to 'G' transition, 'G' being the minor allele and according to Ensemble has a frequency of 26% in Europeans and 32% in British. Due to its close proximity to miR-140, and the importance of miR-140 in cartilage, I hypothesised rs2102066 may be able to alter miR-140 processing and may be associated with OA.

To genotype rs2102066 an ABI SNP genotyping assay was used, it uses PCR to amplify the genomic region of the SNP and contains two fluorescent probes, one to each of the alleles. Consistent with Ensemble data, the allelic frequency of the minor allele was found to be approximately 30% (Figure 3.4). Sequencing and the ABI genotyping assay gave identical results for all individuals genotyped by both methods. Analysis of the allele frequencies revealed the minor G allele was more common in male hip OA than in controls, 58 (31.9%) and 49 (28.2%) respectively (Figure 3.4), with homozygous GG being over 3 times more common in male hip OA (14 (5.4%)) than controls (4 (4.6%)), ($p<0.05$, fishers 2-tailed exact test) (Figure 3.4). This preliminary data suggested the minor G allele at rs2102066 may contribute to hip OA susceptibility in males.

Figure 3.4 rs2102066 is present in northern Europeans

Male control

Male hip OA

Genotype		Controls (CC12)	OA (CC5)
	AA	50 (57.5%)	47 (51.6%)
	AG	33 (37.9%)	30 (33.0%)
	GG	4 (4.6%)	14 (15.4%)
	Total	87	91

Allele			
	A	133 (76.4%)	124 (68.1%)
	G	49 (28.2%)	58 (31.9%)
	Total	174	182
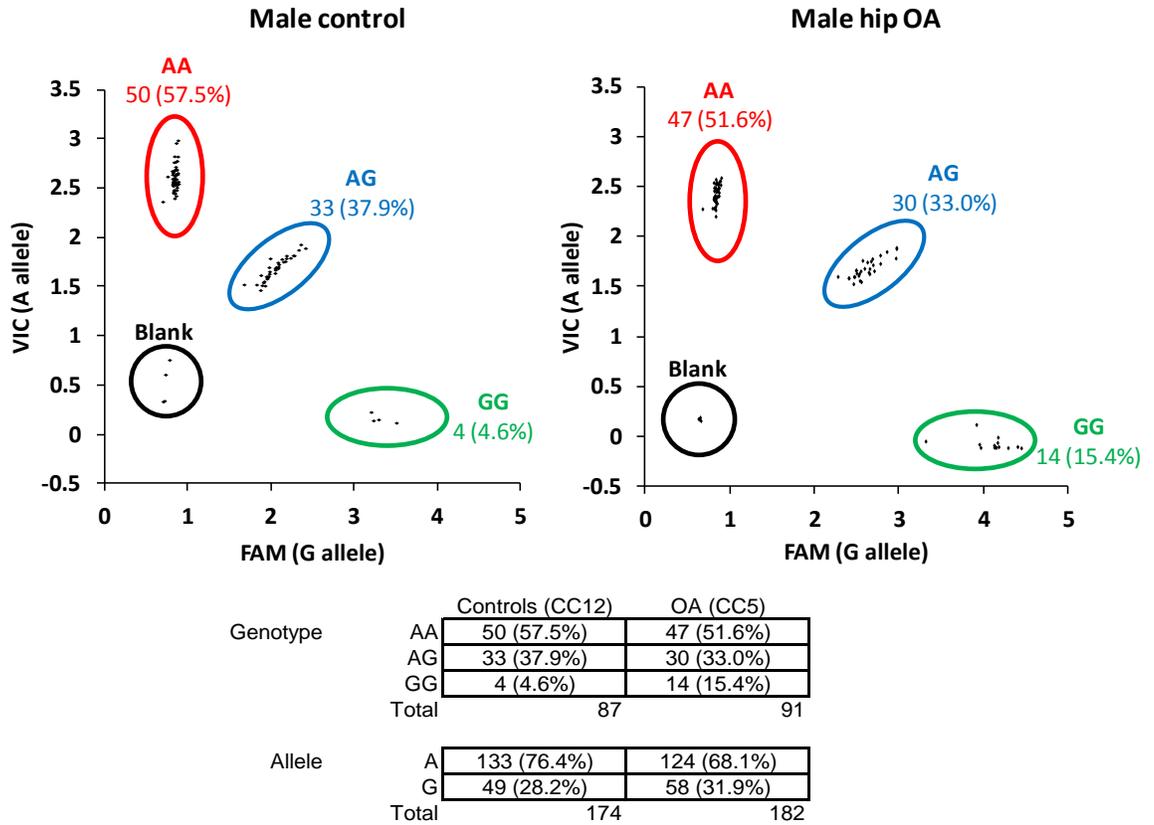
Figure 3.4 rs2102066 is present in Northern Europeans and is associated with male Hip OA. Relative fluorescent intensities of VIC and FAM with correspond to the A and G alleles respectively for the ABI SNP genotyping assay. The minor G allele is associated with OA.

3.2.4   rs2102066 is associated with male hip OA

3.2.4.1 *There is no significant association of rs2102066 with all OA*

Given the relatively small number of individuals genotyped (87 controls and 91 hip OA suffers), I was cautious about concluding rs2102066 was associated with OA. However the data described warranted further investigation. Using samples located in Newcastle as part of the arcOGEN study [39], I genotyped 1096 cases and 821 controls using the ABI SNP assay, which can be stratified as shown in (Figure 3.1)

There was no significant difference in allelic frequencies of rs2102066 for OA compared to control (Table 3.2-upper third). Similarly there were no significant differences of genotypes for rs2102066 for OA compared to controls (Table 3.2-upper third). However, the frequency of the minor G allele was slightly higher in all OA individuals than controls, being 26.92% and 25.64% respectively (p=0.58) (Table 3.2-upper third).

3.2.4.2 *There is no significant association of rs2102066 with OA when stratified on joint affected*

Although the pathologies of hip and knee OA have many similarities there are some differences [196]. In addition, a number of studies have shown many genetic variations do not always have the same level of association with hip and knee OA [39]. Taken together, these studies suggest hip and knee OA are separate diseases, with shared features. For this reason, separate analysis of hip and knee OA was performed. Of the 1096 cases genotyped, 738 had hip only OA, 292 had knee only OA and 66 had hip + knee OA (Table 3.1). This means 2192 alleles of OA sufferers were genotyped, 1476 of which belonged to hip only OA sufferers, 584 belonged to knee only OA sufferers and 132 belonged to hip and knee OA sufferers (Figure 3.1).

There was no significant difference in allelic frequencies or genotypes of rs2102066 for any type (joint effected) of OA compared to control (Table 3.2-upper third). However, the frequency of the minor G allele was higher in hip only OA (28.12%), than in knee only OA (24.14%), although neither were significantly different from controls (25.64%) (Figure 3.2-upper third). Similarly the frequency of the homozygous GG genotype was higher in hip only OA (8.1%), than in knee only OA (5.1%), but again neither were significantly different

from controls (7.4%) (Figure 3.2 upper third). These data show, compared to controls, the minor G allele is increased in hip OA and decreased in knee OA, possibly suggesting the minor G allele may promote hip OA and protect against knee OA.

3.2.4.3rs2102066 is associated with male hip OA

Studies have shown differences in the genetics of OA in males and females [39], with females tending to have a greater heritability for OA than males [197]. For these reasons I stratified our data by gender. There was a slight association (not significant) of the minor G allele and the homozygous GG genotype with male OA (Table 3.2-middle third).

When further stratified by gender and OA joint effected, there was a significant association of the minor G allele with male hip OA (OR 1.28, p=0.04, 95%CI= 1.01-1.62) Table 3.2-green). There was also an association (approaching significance) of the homozygous GG genotype with male hip OA (OR 1.64, p=0.06, 95% CI= 0.98-2.74) (Table 3.2--yellow). There was no significant association of either allele or genotype with male knee OA, female hip OA or female knee OA (Table 3.2--lower third). There was also no significant association of any genotype with male knee, female hip or female knee (Table 3.2).

I have also genotyped individuals which have hip and knee OA, and although their genotypes and allelic ratios were not significantly different from controls, I wanted to add these individuals to the individuals with hip only OA and to individuals with knee only OA, to analyse the total number of hip OA sufferers and total number of knee OA sufferers respectively. This analysis showed a significant association of the minor G allele with male hip OA + hip & knee OA compared with controls (Figure 3.2 -green), but did not show any significant association for male knee OA, female knee OA or female hip OA (Figure 3.2).

Table 3.1 Demographic of individuals genotyped at rs2102066

| | Total | | Gender, N (%) | | | |
| | | | Male | | Female | |
	n	(%)	n	(%)	n	(%)
Cases (N=1096)	1096		418	38	678	62
Controls (N=821)	821		465	57	356	43
Total	1917		883	46	1034	54
Hip only	738	67	278	67	460	68
Knee only	292	27	116	28	176	26
Hip and Knee	66	6	24	6	42	6
Hip only + Hip and Knee	804		302		502	
Knee only + Hip and Knee	358		140		218	

Table 3.2 SNP downstream of miR-140 is associated with male hip OA.

| | AA | | AG | | GG | | | | AG + GG | | AA | | | | AG+AA | | GG | | | | G | | A | | | |
	n	%	n	%	n	%	Total	p-val	n	%	n	%	Total	p-val	n	%	n	%	Total	p-val	n	%	n	%	Total	p-val
Total control	461	56.2	299	36.4	61	7.4	821		360	43.8	461	56.2	821		760	92.6	61	7.4	821		421	25.6	1221	74.4	1642	
Total OA	586	53.5	430	39.2	80	7.3	1096	0.735	510	46.5	586	53.5	1096	0.722	1016	92.7	80	7.3	1096	0.290	590	26.9	1602	73.1	2192	0.580
Total just Hip	383	51.9	295	40.0	60	8.1	738	0.715	355	48.1	383	51.9	738	0.533	678	91.9	60	8.1	738	0.406	415	28.1	1061	71.9	1476	0.535
Total just knee	166	56.8	111	38.0	15	5.1	292	0.834	126	43.2	166	56.8	292	0.550	277	94.9	15	5.1	292	0.574	141	24.1	443	75.9	584	0.568
Total hip and knee	37	56.1	24	36.4	5	7.6	66	0.999	29	43.9	37	56.1	66	0.989	61	92.4	5	7.6	66	0.965	34	25.8	98	74.2	132	0.976
Total hip and hip and knee	420	52.2	319	39.7	65	8.1	804	0.286	384	47.8	420	52.2	804	0.114	739	91.9	65	8.1	804	0.622	449	27.9	1159	72.1	1608	0.142
Total knee and hip and knee	203	56.7	135	37.7	20	5.6	358	0.834	155	43.3	203	56.7	358	0.665	338	94.4	20	5.6	358	0.441	175	24.4	541	75.6	716	0.721
Male control	267	57.4	165	35.5	33	7.1	465		198	42.6	267	57.4	465		432	92.9	33	7.1	465		231	24.8	699	75.2	930	
Male hip	144	51.8	103	37.1	31	11.2	278	0.109	134	48.2	144	51.8	278	0.136	247	88.8	31	11.2	278	0.057	165	29.7	391	70.3	556	0.041
Male knee	72	62.1	37	31.9	7	6.0	116	0.657	44	37.9	72	62.1	116	0.363	109	94.0	7	6.0	116	0.686	51	22.0	181	78.0	232	0.364
Male hip and knee	10	41.7	12	50.0	2	8.3	24	0.305	14	58.3	10	41.7	24	0.129	22	91.7	2	8.3	24	0.819	16	33.3	32	66.7	48	0.187
Total male hip and hip and knee	154	51.0	115	38.1	33	10.9	302	0.089	148	49.0	154	51.0	302	0.081	269	89.1	33	10.9	302	0.065	181	30.0	423	70.0	604	0.027
Total male knee and hip and knee	82	58.6	49	35.0	9	6.4	140	0.951	58	41.4	82	58.6	140	0.809	131	93.6	9	6.4	140	0.785	67	23.9	213	76.1	280	0.757
Total male OA	226	54.1	152	36.4	40	9.6	418	0.347	192	45.9	226	54.1	418	0.317	378	90.4	40	9.6	418	0.183	232	27.8	604	72.2	836	0.165
Female control	194	54.5	134	37.6	28	7.9	356		162	45.5	194	54.5	356		328	92.1	28	7.9	356		190	26.7	522	73.3	712	
Female hip	239	52.0	192	41.7	29	6.3	460	0.409	221	48.0	239	52.0	460	0.471	431	93.7	29	6.3	460	0.386	250	27.2	670	72.8	920	0.825
Female knee	94	53.4	74	42.0	8	4.5	176	0.282	82	46.6	94	53.4	176	0.813	168	95.5	8	4.5	176	0.151	90	25.6	262	74.4	352	0.697
Female hip and knee	27	64.3	12	28.6	3	7.1	42	0.469	15	35.7	27	64.3	42	0.227	39	92.9	3	7.1	42	0.869	18	21.4	66	78.6	84	0.300
Total female hip and hip and knee	266	53.0	204	40.6	32	6.4	502	0.542	236	47.0	266	53.0	502	0.663	470	93.6	32	6.4	502	0.399	268	26.7	736	73.3	1004	0.997
Total female knee and hip and knee	121	55.5	86	39.4	11	5.0	218	0.928	97	44.5	121	55.5	218	0.875	207	95.0	11	5.0	218	0.629	108	24.8	328	75.2	436	0.612
Total female OA	360	53.1	278	41.0	40	5.9	678	0.927	318	46.9	360	53.1	678	0.438	638	94.1	40	5.9	678	0.075	358	26.4	998	73.6	1356	0.427

3.3   Discussion

3.3.1   SNPs in 3'UTRs may change miR-140 targets

SNPs can also occur in individual targets to change how a miRNA regulates that target. A number of SNPs have the potential to alter miR-140 binding to targets, with one of these targets (TP63) being associated with OA. I searched for SNPs which have the potential to alter miR-140 targets, as this miRNA is most highly expressed cartilage. Interestingly in a separate study, TP63 expression was not detected in either OA or NOF cartilage, but was detected in tendon, fat pad and osteophyte [39], perhaps suggesting the high miR-140 expression in cartilage is leading to reduced TP63 expression.

3.3.2   No SNPs found in miR-140 stem loop

Our study did not find any polymorphic SNPs within the miR-140 stem-loop, although if I increased the number of samples there could be. Interestingly, rs7205289 has been found in the Chinese population [190], suggesting rs7205289 is population specific. In the Chinese population rs7205289 is associated with cleft pallet, with CA/AA at rs7205289 contributing to cleft pallet [198], by controlling the processing of miR-140 [199].

3.3.3   rs2102066 is in miR-140 flanking region and is associated with male hip OA

Due to the importance of miR-140 in cartilage and the previously identified OA association of the chromosome 16 region where miR-140 is located [37], it is not surprising, rs2102066, which is only 10nt downstream of the miR-140 stem-loop is associated with male hip. In addition, taken together with previous data, I have shown the genetics of hip OA and knee OA in males and females are different.

3.3.4   rs2102066 may affect miR-140 processing by altering precursor RNA structure.

SNPs within stem loops have previously been shown to alter miRNA processing by changing the RNA structure [190]. To our knowledge no SNP outside of the pre-miRNA has been shown to alter miRNA processing, but as miRNAs are processed from larger sections of RNA and rs2102066 is only 10nt downstream of the miR-140 stem-loop, I hypothesised the minor G allele at rs2102066, may alter miR-140 precursor RNA structure and processing. To test this I modelled the structure of miR-140 with each of the alleles, using

the Centroidfold RNA folding program (www.ncrna.org/centroidfold) (Figure 3.5). The minor G allele alters the structure of the miR-140 flanking region (Figure 3.5).

Figure 3.5 Model of miR-140 with and without SNP rs2102066

Figure 3.5 Model of miR-140 with and without SNP rs2102066. Minor G allele may alter miR-140 flanking region structure. (A) miR-140 stem loop RNA +/- 1kb flanking RNA was inputted into Centroidfold RNA folding program (www.ncrna.org/centroidfold) set to McCaskill. Dicer and Drosha cleavage sites are indicated by green and blue arrows respectively. (B) Zoomed version of A.

3.3.5 Increased power may verify associations

I genotyped 1096 OA and 821 controls and have found a significant association of the minor G allele with male hip OA. I found a number of other associations although none were significant. I therefore wanted to increase the number of individuals genotyped in the study. I have genotyped all of the samples available in Newcastle, however there are many more OA and control individuals in the arcOGEN cohort [39]. I do not easily have access to the whole of the cohort, but there has been a genotyping array performed on all of the samples [39]. ~500,000 SNPs were genotyped on the array, but unfortunately neither of our studied SNPs were included. I therefore wanted to find SNPs in high LD with both rs2102066 (in miR-140 flanking region) and rs35592567 (predicted to prevent miR-140 binding to TP63), which were included on the array. According to SNAP (Broad institute-proxy search) [200], there are 13 SNPs in perfect LD with rs2102066, and 9 SNPs in perfect LD with rs35592567, unfortunately none of these were included on the array (Table 3.3). There were however a number SNPs in high LD with rs2102066 and rs35592567 (Table 3.3). 4 of the SNPs in high LD with rs2102066, and 1 of the SNPs in high LD with rs35592567 were included on the array (Table 3.3). The genotypes of these SNPs will be a good representation of the genotype of the rs2102066 and rs35592567 and can be used to indicate if these SNPs are associated with OA in the whole arcOGEN cohort. Haploview may be a useful tool to determine which alleles are in LD with each other. However, because none of the proxy SNPs are in perfect LD with the studied SNPs the exact genotype in OA and controls can only be estimated. A technique called imputation could be used for this type of analysis, which we are currently beginning.

Table 3.3 SNPs in LD with rs2102066 and rs35592567

SNP	Proxy SNP	Distance	RSquared	DPrime	On array?
rs2102066	rs2102066	0	1	1	
rs2102066	rs10048090	731	1	1	
rs2102066	rs10048088	776	1	1	
rs2102066	rs2270841	804	1	1	
rs2102066	rs8048678	1108	1	1	
rs2102066	rs8048590	1274	1	1	
rs2102066	rs2270840	2069	1	1	
rs2102066	rs1566452	3738	1	1	
rs2102066	rs11864678	4612	1	1	
rs2102066	rs1983015	6101	1	1	
rs2102066	rs904804	6488	1	1	
rs2102066	rs904805	6601	1	1	
rs2102066	rs904806	6829	1	1	
rs2102066	rs904807	7105	1	1	
rs2102066	rs8049004	1303	0.961	1	
rs2102066	rs2291961	1501	0.961	1	
rs2102066	rs8052727	1812	0.961	1	
rs2102066	rs12932078	1634	0.925	1	
rs2102066	rs2270842	2907	0.925	1	yes
rs2102066	rs1983016	3236	0.925	1	yes
rs2102066	rs904803	3436	0.925	1	
rs2102066	rs9302605	5393	0.925	1	
rs2102066	rs4985461	6562	0.925	1	
rs2102066	rs3748388	7355	0.925	1	
rs2102066	rs1052429	8267	0.925	1	yes
rs2102066	rs3762178	5329	0.922	0.96	
rs2102066	rs12932286	1518	0.89	1	
rs2102066	rs3762177	4872	0.89	1	yes
rs2102066	rs7206222	223	0.885	1	
rs2102066	rs8047818	639	0.857	1	
rs2102066	rs7499251	10272	0.857	1	
rs2102066	rs7184994	8996	0.826	1	
rs35592567	rs35592567	0	1	1	
rs35592567	rs11709791	1218	1	1	
rs35592567	rs60960402	5416	1	1	
rs35592567	rs1515487	7338	1	1	
rs35592567	rs11713848	10226	1	1	
rs35592567	rs73202009	16773	1	1	
rs35592567	rs61064225	19144	1	1	
rs35592567	rs73202033	29317	1	1	
rs35592567	rs73202034	31379	1	1	
rs35592567	rs73202036	32266	1	1	
rs35592567	rs9840360	23811	0.941	1	
rs35592567	rs58312266	32589	0.941	1	
rs35592567	rs11708746	444	0.938	1	
rs35592567	rs11708753	495	0.938	1	yes

Table 3.3 SNPs in high LD with rs2102066 and rs35592567 according to SNAP (Broad institute-proxy search). Rsquared of 1 indicates SNPs are in perfect LD with both alleles at each SNP are always inherited with the alleles at the other SNP. Dprime of 1 indicates SNPs are in complete LD and one, but not both, of the alleles at each SNP will always be inherited with one of the alleles at the other SNP.

3.4 Summary

I have found a SNP (rs2102066) which is just downstream of the miR-140 (the most studied miRNA in cartilage) locus, and another SNP (rs35592567) in a predicted, OA-associated, miR-140 target (TP63), are likely to affect miR-140 function and to be associated with OA. Increased sample size is needed to verify and test the association of rs2102066 and rs35592567 respectively. One must note, TP63 is only a predicted target of miR-140 and is yet to be validated, further work is continuing to validate TP63 as a miR-140 target and test if rs35592567 can affect the miR-140-TP63 interaction. In addition, the creation of a constructs would be needed to investigate the role of rs2102066 on miR-140 processing.

If this shows rs2102066 is able to effect miR-140 processing and rs35592567 is able to affect miR-140 regulation of TP63, it will be interesting to test if having the minor allele at both of these SNPs further increases an individual's susceptibility to OA, compared to only having one or none of the SNPs.

There are a number of factors that need to be considered for rs2102066 and rs35592567 in miR-140 biology in OA. Firstly, at the time blood was taken for genotyping, the controls did not have OA, however they may go on to develop the disease. Secondly, rs2102066 is located 10nt downstream of the miR-140 stem loop and may not affect miR-140. To our knowledge the function of a SNP 10nt downstream of any miRNA stem loop has not been characterised. However, if the SNP alters the secondary structure of the miRNA stem loop basal segments, then processing of the miRNA may be altered. This work increases our understanding of OA genetics, may partly explain how the expression of miR-140 becomes altered in OA and may also partly explain the genetic association of TP63 with OA. Taken together, this data may suggest both the miR-140 encoding region and a miR-140 target are associated with OA.

Chapter 4 miR-125b and miR-324-5p regulate Hh signalling

4.1   Introduction

4.1.1 miR-125b-5p and miR-324-5p are increased in end-stage osteoarthritic cartilage

miR-140 has previously been shown to be differentially expressed in OA and is the most extensively studied miRNA in cartilage. In Chapter 3 I investigated possible genetic associations of miR-140 and its target with OA. In chapter 4 I aim to investigate the functions of other miRNAs differentially expressed in OA. To identify miRNAs differentially expressed during Osteoarthritis (OA), a TaqMan® low density array of 365 miRNAs was performed on RNA obtained from total hip replacements for either OA or fracture to the neck of femur (NOF) cartilage by Hannah Elliott (previous laboratory member). Data were normalised, using a published method [201]. A number of miRNAs were differentially expressed in OA cartilage compared to NOF (unpublished data), such as miR-455 [155].

Two differentially expressed miRNAs of particular interest were miR-125b-5p and miR-324-5p as they have both previously been shown to regulate the Hh signalling pathway [127]. The Hh pathway is important for both cartilage developemt and maintance (Section 1.7).

In our labs study miR-125b-5p was significantly increased in the OA cartilage compared to NOF and miR-324-5p was only detectable in the OA cartilage (Figure 4.1). miR-125b-5p has previously been shown to target the 3'UTR of the Hh activator SMO, while miR-324-5p has been shown to target both the 3'UTR of SMO and the 3'UTR of the Hh transcription factor GLI1 [127].

This chapter will start by investigating the structure, processing and evolutionary history of miR-125b-5p and miR-324-5p, I will then use a number of computational prediction methods to predict the targets and functions of miR-125b-5p and miR-324-5p and finally investigate the functions of miR-125b-5p and miR-324-5p, predominately focusing on their role in Hh signalling, which will form the main aim of this chapter.

Specific aims

Aim 1: Determine the structure and processing of miR-125b and miR-324

Aim 2: Investigate the evolutionary history of miR-125b and miR-324

Aim 3: Predict targets and functions of miR-125b-5p and miR-324-5p

Aim 4: Validate miR-125b-5p and miR-324-5p target human SMO and GLI1 3'UTRs

Aim 5: Characterise Hh signalling in mouse C3H10T1/2 cell line

Aim 6: Characterise miR-125b-5p and miR-324-5p effect on Hh signalling in mouse C3H10T1/2 cell line

Aim 7: Identify factors controlling miR-125b-5p and miR-324-5p expression

Aim 8: Determine the mechanism by which miR-125b-5p and miR-324-5p regulate Hh signalling in mouse

Figure 4.1 miR-125b-5p and miR-324-5p are increased in OA.
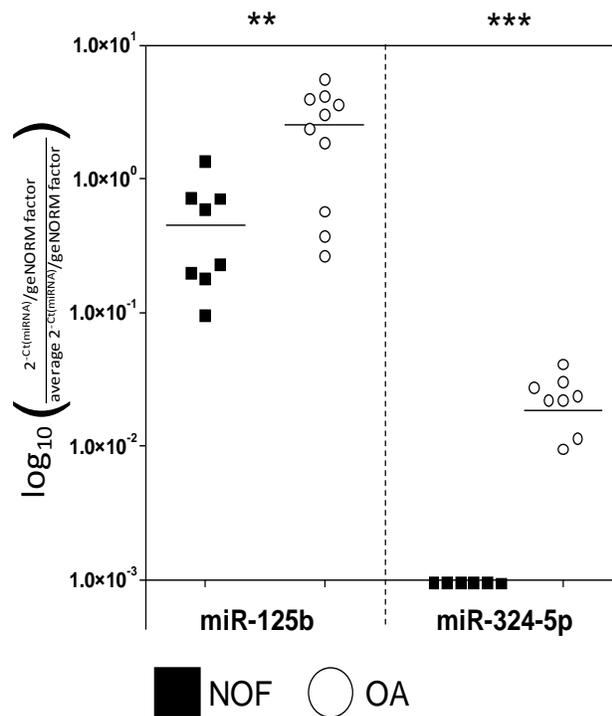
Figure 4.1 miR-125b-5p and miR-324-5p are increased in OA. Real-time RT-PCR was used to identify miRNAs differentially expressed in OA and normal (NOF) cartilage. The expression of both miR-125b-5p and miR-324-5p are increased during end-stage OA, $p<0.01$ (**) and $p<0.001$ (***) respectively (non-parametric Mann Whitney U test).

4.2 Results

4.2.1 Structure and processing of miR-125b and miR-324 stem-loops

Mature miRNAs form from stem-loops, their processing involves cleavage by the enzymes Drosha and Dicer (Chapter 1), which determine mature miRNA sequences and ultimately their targets and functions. Cleavage sites are based on stem-loop sequence and structure, although the exact criteria are unknown. miRBase (the microRNA registry) contains annotation for all miRNA structures and sequences [112].

miR-125b was originally cloned from mouse brain [202], and can be encoded from two genomic loci, one on chromosome 9 and one on chromosome 16 for miR-125b-1 and miR-125b-2 respectively. Each locus produces identical '-5p' strands (miR-125b-5p) but different '-3p' strands (miR-125b-1-3p and miR-125b-2-3p). Both '-3p' antisense miRNAs have also been cloned [203]. Homology searches suggested both homologues of miR-125b are present in human [204], being encoded from chromosome 11 and chromosome 21 for miR-125b-1 and miR-125b-2 respectively, their expression was later validated in human BC-1 cells [205].

miR-324-5p was originally cloned from rat cerebro-cortical dissociated cultures of E18 rat embryos [206]. Expression of miR-324-5p and miR-324-3p was later verified in mouse [207], homology searches suggested miR-324 was also present in human [204], and it has now been cloned from human BC-1 cells [205]. miR-324 is encoded downstream of DVL2 from chromosome 17 and chromosome 11 in human and mouse respectively.

More recently, the use of high-throughput sequencing has allowed further investigation into miRNA sequence, structure and expression [208, 209] with data available on miRBase [112]. Alignment of sequencing reads for miR-125b shows mature miRNA sequences and stem-loop structure are in agreement with the miRBase annotation and miR-125b-5p is more highly expressed than the antisense strands (miR-125b-1-3p and miR-125b-2-3p) in mouse (Figure 4.2A and B) and human (Figure 4.2D and E).

Similar to miR-125b-5p, the mature miR-324-5p sequence is as annotated in miRBase and is more highly expressed than the antisense '-3p' strand for both mouse (Figure 4.2C) and

human (Figure 4.2F). The miR-324-3p mature sequence is different from miRBase annotation in human and mouse. In human, high-throughput sequencing shows miR-324-3p actually starts 2nt downstream of the miRBase annotated miR-324-3p, meaning the 'actual' human miR-324-3p seed is likely to be 'CACUGCC' rather than 'CUGCCCC' to (Figure 4.2F). The 'actual' human miR-324-3p sequence is also shown in Table 4.1. Mouse high-throughput sequencing shows two variants of mature miR-324-3p sequence, around half of the reads start at the miRBase predicted site, the other half start 2nt downstream, meaning there are two possible seeds; 'CACUGCC' (the mouse miRBase annotated/canonical miR-324-3p) and 'CUGCCCC' (non-canonical)(Figure 4.2 C). Canonical mouse miR-324-3p is shown as 'mouse miR-324-3p' and non-canonical mouse miR-324-5p is shown as 'non-canonical miR-324-3p' in Table 4.1. miRBase annotation used to describe the miR-125b and miR-324 stem-loops, the mature miRNAs, their antisense miRNAs and the sequences of each in human and mouse is also shown in Table 4.1.

Table 4.1 miR-125b and miR-324 annotation

Description	Sequence
miRBase mouse miR-125b-1 stem-loop (Chr 9)	UGCGCUCCCCUCAG**UCCCUGAGACCCUAACUUGUGA**UGUUUACCGUUUAAAUCC**ACGGGUUAGGCUCUUGGGAGCU**G
miRBase human miR-125b-1 stem-loop (Chr 11)	UGCGCUCCUCUCAG**UCCCUGAGACCCUAACUUGUGA**UGUUUACCGUUUAAAUCC**ACGGGUUAGGCUCUUGGGAGCU**GCGAGUCGUGCU
miRBase mouse miR-125b-2 stem-loop (Chr 16)	GCCUAG**UCCCUGAGACCCUAACUUGUGA**GGUAUUUUAGUAACAU**CACAAGUCAGGUUCUUGGGACCU**AGGC
miRBase human miR-125b-2 stem-loop (Chr 21)	ACCAGACUUUUCCUAG**UCCCUGAGACCCUAACUUGUGA**GGUAUUUUAGUAACA**UCACAAGUCAGGCUCUUGGGAC**CUAGGCGGAGGGGA
miRBase mouse mature miR-125b-5p	**UCCCUGAGACCCUAACUUGUGA**
miRBase human mature miR-125b-5p	**UCCCUGAGACCCUAACUUGUGA**
miRBase mouse mature miR-125b-1-3p	**ACGGGUUAGGCUCUUGGGAGCU**
miRBase human mature miR-125b-1-3p	**ACGGGUUAGGCUCUUGGGAGCU**
miRBase mouse mature miR-125b-2-3p	**ACAAGUCAGGUUCUUGGGACCU**
miRBase human mature miR-125b-2-3p	**UCACAAGUCAGGCUCUUGGGAC**
miRBase mouse miR-324 stem-loop (Chr 11)	AACUGACUAUGCCUCCU**CGCAUCCCCUAGGGCAUUGGUGU**AAAGCUGGAGAC**CCACUGCCCCAGGUGCUGCU**GGGGGUUGUAGUCUGAC
miRBase human miR-324 stem-loop (Chr 17)	CUGACUAUGCCUCCCC**CGCAUCCCCUAGGGCAUUGGUGU**AAAGCUGGAGACCC**ACUGCCCCAGGUGCUGCUGG**GGGUUGUAGUC
miRBase mouse mature miR-324-5p	**CGCAUCCCCUAGGGCAUUGGUGU**
miRBase human mature miR-324-5p	**CGCAUCCCCUAGGGCAUUGGUGU**
miRBase mouse mature miR-324-3p	**CCACUGCCCCAGGUGCUGCU**
Non-canonical mouse mature miR-324-3p	**ACUGCCCCAGGUGCUGCUGG**
miRBase human mature miR-324-3p	**ACUGCCCCAGGUGCUGCUGG**
Actual human mature miR-324-3p	**CCACUGCCCCAGGUGCUGCUGG**

Table 4.1 Annotation used to describe miR-125b and miR-324 sequences. Bold indicates mature miRNA sequences, underlined indicates miRNA seed sequence.

Figure 4.2 Alignment of deep sequencing reads and stem-loop structures of miR-125b-1, miR-125b-2 and miR-324 in human and mouse
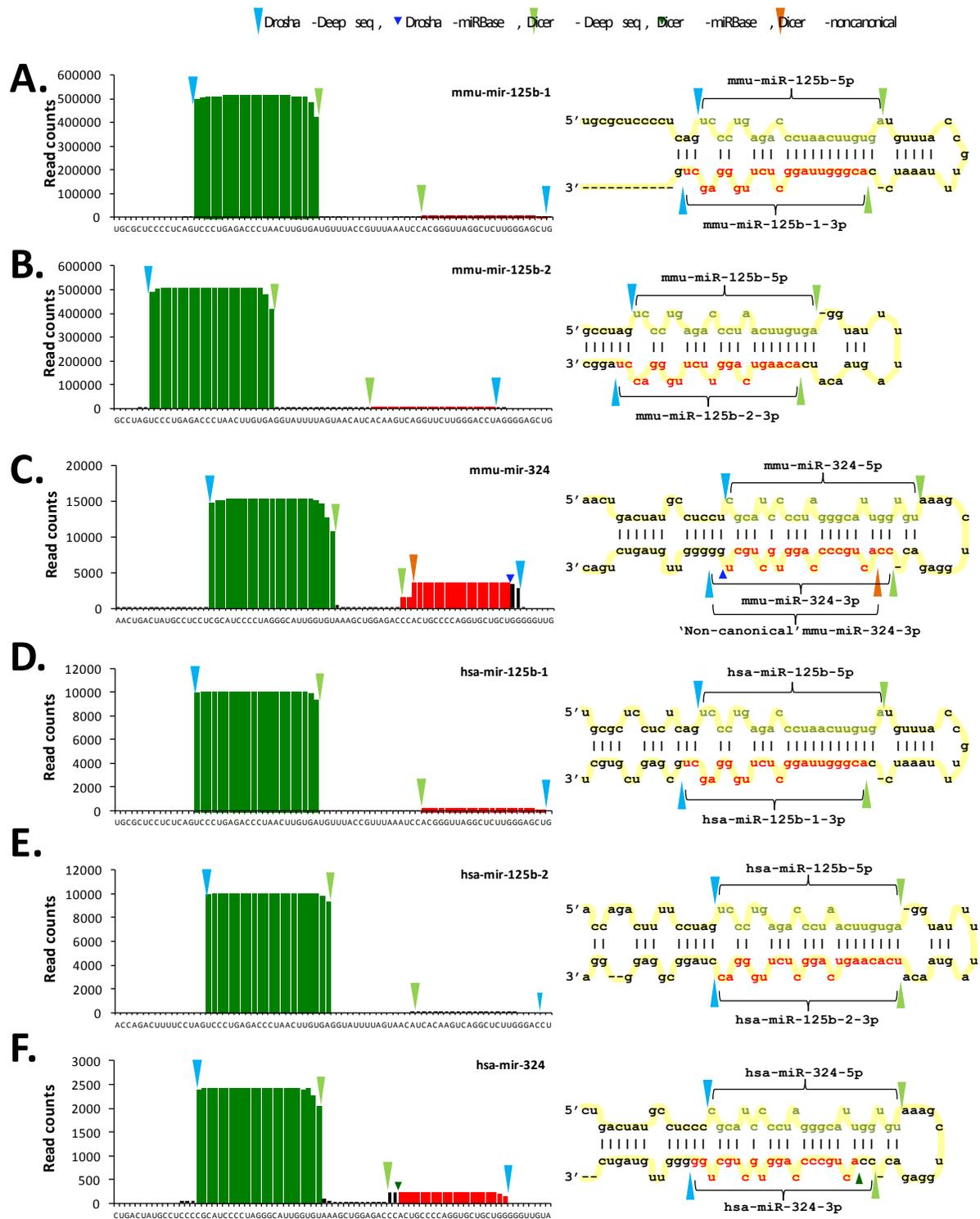
Figure 4.2 Alignment of deep sequencing reads and stem-loop structures of miR-125b-1, miR-125b-2 and miR-324 in human and mouse. Deep sequencing reads were obtained from miRBase and aligned with the total number of reads for each base (read counts) displayed on the y axis and the stem-loop sequence on the x axis. Stem-loop structures according to miRBase and predicted from deep sequencing alignment. Mature '-5p' sequences in green, mature '-3p' sequences in red as suggested by miRBase with reads not predicted by miRBase in black. Drosha and Dicer cleavage site shown by blue and green triangles respectively, as suggested by deep sequencing (large triangles) or as suggested by miRBase (darker small triangles) if different from deep seq. Non-canonical Dicer cleavage site for mouse miR-324-3p shown by orange triangle. (A) Deep sequencing reads and stem-loop for mouse miR-125b-1. (B) Deep sequencing reads and stem-loop for mouse miR-125b-2. (C) Deep sequencing reads and stem-loop for mouse miR-324-5p. (D) Deep sequencing reads and stem-loop for human miR-125b-1. (E) Deep sequencing reads and stem-loop for human miR-125b-2. (F) Deep sequencing reads and stem-loop for human miR-324-5p.

4.2.2 Evolution of miR-125b and miR-324

miRNAs are generally encoded from genomic regions which are more conserved than the surrounding DNA [192]. In humans and mice miR-125b can be encoded from two genomic loci, a feature conserved in many species. Alignment of all miR-125b loci using Clustal OMEGA [210] (www.clustal.org), shows miR-125b is highly conserved between species (Figure 4.3). Phylogenic analysis of miR-125b sequences using Clustal2 Phylogeny (http://www.ebi.ac.uk/Tools/phylogeny/clustalw2_phylogeny/), separates them into three clades, named here as 'a' 'b' and 'c' (Figure 4.4). Times of divergence of species which possess miR-125b were extracted from Time Tree (www.timetree.org) and a phylogenic tree of these species created using Phylodendron (http://iubio.bio.indiana.edu/treeapp/) (Figure 4.5). Many of the species which possess miR-125b diverged around 800million years ago, indicating miR-125b is at least this old (See '#' on Figure 4.5). Species which diverged from humans more than 661.2 million years ago only have one copy of miR-125b where as many species which diverge from humans less than 400.1 million years ago have two copies of miR-125b (Table 4.2), this suggests the two miR-125b loci seen in many species today have arisen from a single duplication between 400.1 and 661.2 million years ago (see '+' on Figure 4.5). This is further supported by the fact that all species which contain two miR-125b loci have one miR-125b loci belonging to sequence phylogenic clade 'b' and one to clade 'c' (Table 4.2). Those species which contain only one miR-125b locus have miR-125b sequences belonging to clade 'a' (Table 4.2). This suggests a sequence not to dissimilar to those sequences in clade 'a' duplicated to give rise to two miR-125b loci not to dissimilar to those in clade 'b' and 'c'. These sequences have since evolved independently giving rise to 'bi','bii','biii','ci', 'cii etc (Figure 4.5 and Table 4.2). Interestingly zebrafish have 3 genomic loci for miR-125b (Table 4.2). Two of the three miR-125b loci are very similar and part of clade 'bii' suggesting they have arisen from a more recent second duplication in fish. It cannot be excluded that there are additional yet to be annotated miR-125b loci in other species. In addition, it cannot be excluded that miR-125b may have arisen from other miRNAs or other regions of genomic DNA. Genome wide searches for miR-125b sequences or similar sequences were unable to identify the origin of miR-125b.

Alignment of all miR-324 loci shows the majority of miR-324 stem-loops are conserved between species (Figure 4.3B). All species which possess a miR-324 stem-loop diverge within the last 100 million years (see '*' in Figure 4.5) indicating miR-324 is relatively young in comparison to miR-125b, at around 100 million years old, and appears to be conserved in almost all placental animals (Figure 4.3C). A number of species which diverged after the occurrence of miR-324 have no annotated miR-324 sequence (Table 4.2). This could be due to the loss of the miR-324 loci, but is more likely because miR-324 is yet to be annotated. Genome wide searches for miR-324 sequences or similar sequences did not identify the origin of miR-324. Meaning it is unknown if miR-324 has arisen from another miRNA or if it has arisen from a 'random' piece of RNA.

Figure 4.3 miR-125b and miR-324 sequences.

A.

```
aca-mir-125b-1   --------UGCGCCCCUCUCUGUCCCGAGACCCUAACUUGUGACGUUUAGUUUUU-AAGUCCACGGGUUAGGCUCUUGGGAGCUGUGAGUUGUGCCAC-------
age-mir-125b-1   --------UGCGCUCCUCUCAGUCCCUGAGACCCUAACUUGUGAUGUUUACCGUUU-AAAUCCACGGGUUAGGCUCUUGGGAGCUGCGAGUCGUGCU----------
bta-mir-125b-1   --------CGCGCGCCUCUCAAUCCCUGAGACCCUAACUUGUGAUGUUUACCGUUU-AAAUCCACGGGUUAGGCUCUUGGGAGCUGCGAGUCGCGCC----------
bfl-mir-125b     GUGGAGGGCGUAUGUCACCCAGCUCCCAAGAUCCUAACCUGUGAGCCAUGGA-------CAUCACAAGUUAGGGUCUCAGGGAAUGGGG-------------------
cfa-mir-125b-1   ----------------------UCCCUGAGACCCUAACUUGUGAUGUUUACCGUUU-AAAUCCACGGGUUAGGCUCUUGGGAGC-----------------------
ccr-mir-125b     ------------------GGUCCCUGAGACCCUAACUUGUGAGCUUUGUGUGCUGAAAAUCACAGGUUAAGCUCUUGGGACC--------------------------
dre-mir-125b-2   ---------GUGCCCCUCUCCUUCCCUGAGACCCUAACUUGUGACGUUCUGCUUCG-AUGUCCACGGGUUGGGUUCUCGGGAGCUGUGAGAGGCAC-----------
eca-mir-125b     --------UGCGCUCCUCUCAAUCCCUGAGACCCUAACUUGUGAUGUUUACCGUUU-AAAUCCACGGGUUAGGCUCUUGGGAGCUGCGAGUCGUGCC----------
fru-mir-125b     ---------------CUCUCAUUCCCUGAGACCCUAACUUGUGACGUUGUGCUGUG-AUGUCCACGGGUUGGGUUCUCGGGAGCUG------------------
ggo-mir-125b-1   --------UGCGCUCCUCUCAGUCCCUGAGACCCUAACUUGUGAUGUUUACCGUUU-AAAUCCACGGGUUAGGCUCUUGGGAGCUGCGAGUCGUGCU----------
hsa-mir-125b-1   --------UGCGCUCCUCUCAGUCCCUGAGACCCUAACUUGUGAUGUUUACCGUUU-AAAUCCACGGGUUAGGCUCUUGGGAGCUGCGAGUCGUGCU----------
lla-mir-125b-1   --------UGCGCUCCUCUCAGUCCCUGAGACCCUAACUUGUGAUGUUUACCGUUU-AAAUCCACGGGUUAGGCUCUUGGGAGCUGCGAGUCGUGCU----------
lca-mir-125b     ------ACCAGACUUUUCCUAGUCCCUGAGACCCUAACUUGUGAGGUA--UUUUAGUAACAUCACAAGUCAGGCUCUUGGGACCUAGGCGGAGGGGA----------
mml-mir-125b-2   ------ACCAGACUUUUCCUAGUCCCUGAGACCCUAACUUGUGAGGUA--UUUUAGUAACAUCACAAGUCAGGCUCUUGGGACCUAGGCGGAGGGGA----------
mne-mir-125b-2   ------ACCAGACUUUUCCUAGUCCCUGAGACCCUAACUUGUGAGGUA--UUUUAGUAACAUCACAAGUCAGGCUCUUGGGACCUAGGCGGAGGGGA----------
mdo-mir-125b-2   -------------------AAUCCCUGAGACCCUAACUUGUGAUGUUUACCGUUU-AAAUCCACGGGUUAGGCUCUUGGGAGC-----------------------
mmu-mir-125b-2   ----------------GCCUAGUCCCUGAGACCCUAACUUGUGAGGUA--UUUUAGUAACAUCACAAGUCAGGUUCUUGGGACCUAGGC-------------------
ola-mir-125b-1   ----AACCGUACCUCUCUCAUUCCCUGAGACCCUAACUUGUGACGUUUUGCUCUG-AUGUGCACGGGUUGGGCUCUUGGGAGCUGCGAGGGGCACUCACACAUCU-
ptr-mir-125b-1   --------UGCGCUCCUCUCAGUCCCUGAGACCCUAACUUGUGAUGUUUACCGUUU-AAAUCCACGGGUUAGGCUCUUGGGAGCUGCGAGUCGUGCU----------
ppy-mir-125b-1   --------UGCGCUCCUCUCAGUCCCUGAGACCCUAACUUGUGAUGUUUACCGUUU-AAAUCCACGGGUUAGGCUCUUGGGAGCUGCGAGUCGUGCU----------
rno-mir-125b-1   --------UGCGCUCCCUCUCAGUCCCUGAGACCCUAACUUGUGAUGUUUACCGUUU-AAAUCCACGGGUUAGGCUCUUGGGAGCUGCGAGUCGUGC-----------
sko-mir-125b     -CUCUAUAUGUGUAUUACCUAAUCCCUGAGACCCUAACUUGUGAGAAUUACCCAGCA---UUCACAGGUUAGUGGCUUAGGAAUUGGGUGGUGCAUUCAUA------
sha-mir-125b     GCAUCAAAUAAGACUUUUCCUAGUCCCUGAGACCCUAACUUGUGAGAUU-UUUUAGCAAUAAUCACAGGUCAGGCUCUUGGGACCUAGGCGGAGGGGAACCAGCAGCU
sma-mir-125b     -------------------AUCCCUGAGACUGAUAAUUGCUCUUUUUUAUCAUGAUAAG-----GGAGUUUGAUGAGAGCAAUAUUAUUA-----------------
ssc-mir-125b-2   ------ACCAGACUUUUCCUAGUCCCUGAGACCCUAACUUGUGAGGUA--UUUUAGUAACAUCACAAGUCAGGCUCUUGGGACCUAGGCGGAGGGGA----------
tni-mir-125b     ---------GUACCUCUCUCAUUCCCUGAGACCCUAACUUGUGACGUUGUGCUGUG-AUGUGCACGGGUUGGGUUCUCGGGAGCUGCGAGGGGCGCUCACACAUC--
xtr-mir-125b-2   ----------GACUUUUCCUAGUCCCUGAGACCCUAACUUGUGAGAUUUUUUUAGCAACAAUCACAAGUUAGGCUCUUGGGACCUAGGCGGAGGG-----------
aca-mir-125b-2   --------CAGACUUUUCCUAGUCCCUGAGACCCUAACUUGUGAGGUU--UUUUAGUAACAAUCACAAGUCAGGCUCUUGGGACCUAGGCGGAGGGGAACC------
age-mir-125b-2   -----ACCAGACUUUUCCUAGUCCCUGAGACCCUAACUUGUGAGGUA--UUUUAGUAACAUCACAAGUCAGGCUCUUGGGACCUAGGCGGAGGGGA----------
bta-mir-125b-2   ----------GACUUUUCCUAGUCCCUGAGACCCUAACUUGUGAGGUA--UUUUAGUAACAUCACAAGUCAGGCUCUUGGGACCUAGGCGGAGGGGA----------
cfa-mir-125b-2   ---------------------UCCCUGAGACCCUAACUUGUGAGGUA--UUUUAGUAACAUCACAAGUCAGGCUCUUGGGACCU---------------------
cgr-mir-125b     -------------CUUUUCCUAGUCCCUGAGACCCUAACUUGUGAGGGA--UUUUAGUAACAUCACAAGUCAGGCUCUUGGGACCUAGGCGGA---------------
dre-mir-125b-1   UCUGCAAAUGUGCCUCUCCACAAUCCCUGAGACCCUAACUUGUGACGUUUUCCUGUU-AUGUGCACGGGUUAGGUUCUCUGGGACCUGAGAGGGGUGCUCUGUCAUCAG
dre-mir-125b-3   GCCACCGCUGCGCACUCCUCUCUGGUCCCUGAGACCCUAACUUGUGAGCUUUGUGUGCUCAAAAAUCACAGGUUAAGCUCUUGGGACCUGGGCAGAGGGCAAAAGCACUGG
eca-mir-125b-2   ------------------CCUAGUCCCUGAGACCCUAACUUGUGAGGUA--UUUUGGUAACAUCACAAGUCAGGCUCUUGGGACCUAGGC------------------
gga-mir-125b     ------AUCAGACUUUUCCUAGUCCCUGAGACCCUAACUUGUGAGGUU-UUGUAGCAACAAUCACAAGUCAGGCUCUUGGGACCUAGGCGGAGGGGA----------
ggo-mir-125b-2   ------ACCAGACUUUUCCUAGUCCCUGAGACCCUAACUUGUGAGGUA--UUUUAGUAACAUCACAAGUCAGGCUCUUGGGACCUAGGCGGAGGGGA----------
hsa-mir-125b-2   ------ACCAGACUUUUCCUAGUCCCUGAGACCCUAACUUGUGAGGUA--UUUUAGUAACAUCACAAGUCAGGCUCUUGGGACCUAGGCGGAGGGGA----------
lla-mir-125b-2   ------ACCAGACUUUUCCUAGUCCCUGAGACCCUAACUUGUGAGGUA--UUUUAGUAACAUCACAAGUCAGGCUCUUGGGACCUAGGCGGAGGGGA----------
mml-mir-125b-1   --------UGCGCUCCUCUCAGUCCCUGAGACCCUAACUUGUGAUGUUUACCGUUU-AAAUCCACGGGUUAGGCUCUUGGGAGCUGCGAGUCGUGCU----------
mne-mir-125b-1   --------UGCGCUCCUCUCAGUCCCUGAGACCCUAACUUGUGAUGUUUACCGUUU-AAAUCCACGGGUUAGGCUCUUGGGAGCUGCGAGUCGUGCU----------
mdo-mir-125b-1   -----------ACUUUUCCUAGUCCCUGAGACCCUAACUUGUGAGGCU-UUUCAGCGACAAUCACAGGUCAGGCUCUUGGGACCUAGGCGGAGG-------------
mmu-mir-125b-1   --------UGCGCUCCCUCUCAGUCCCUGAGACCCUAACUUGUGAUGUUUACCGUUU-AAAUCCACGGGUUAGGCUCUUGGGAGCUG--------------------
ola-mir-125b-2   GACUCUGUCGUACUUCACCUGCUCCCUGAGACCCUAACUUGUGAGCUCUCUUGAUAAAAAUCACGGGUUAGGCUCUUGGGACGCGGGCGGAGGGGCACAAUCGAC--
ppa-mir-125b     --------UGCGCUCCUCUCAGUCCCUGAGACCCUAACUUGUGAUGUUUACCGUUU-AAAUCCACGGGUUAGGCUCUUGGGAGCUGCGAGUCGUAAU----------
ptr-mir-125b-2   ------ACCAGACUUUUCCUAGUCCCUGAGACCCUAACUUGUGAGGUA--UUUUAGUAACAUCACAAGUCAGGCUCUUGGGACCUAGGCGGAGGGGA----------
ppy-mir-125b-2   ------ACCAGACUUUUCCUAGUCCCUGAGACCCUAACUUGUGAGGUA--UUUUAGUAACAUCACAAGUCAGGCUCUUGGGACCUAGGCGGAGGGGA----------
rno-mir-125b-2   ------ACCAGACUUUUCCUAGUCCCUGAGACCCUAACUUGUGAGGUA--UUUUAGUAACAUCACAAGUCAGGCUCUUGGGACCUAGGCGGAGAGG----------
sla-mir-125b     --------UGCGCUCCUCUCAGUCCCUGAGACCCUAACUUGUGAUGUUUACCGUUU-AAAUCCACGGGUUAGGCUCUUGGGAGCUGCGAGUCGUGCU----------
sja-mir-125b     -------------------AUCCCUGAGACUGAUAAUUGCUCUUAGUUAUUAUAUCAUUAA-AUGAGUUUACAAUAAGGGCAAUUAUUA------------------
sme-mir-125b     -------GCAAGUUAUUCGACAUCCCUGAGAUCCAUAAUUGGCCUUGAGAUUUGAACGCAAACGGUAGAUUAUUUGUCUCAGAGUAAUUGAUAAAC-----UGCA----
ssc-mir-125b-1   --------------CCUCUCAGUCCCUGAGACCCUAACUUGUGAUGUUUACCGUUU-AAAUCCACGGGUUAGGCUCUUGGGAGCUGCGAGCCGUG----------
xtr-mir-125b-1   --------UGCACCCCUCUCAAUCCCUGAGACCCUAACUUGUGAUGUUUAGCUUUAAAAAUCCACGGGUUAGGCUCUUGGGAGCUGUGAGUUGUGC-----------
                                       **  .***   *    **              .        .         *.
```

B.

```
hsa-mir-324   --CUGACUAUGCCUCCCGCAUCCCCUAGGGCAUUGGUGUAAGCUGGAGACCCACUGCCCCAGGUGCUGCUGGGGGGUUGUAGUC----
ptr-mir-324   --CUGACUAUGCCUCCCGCAUCCCCUAGGGCAUUGGUGUAAGCUGGAGACCCACUGCCCCAGGUGCUGCUGGGGGGUUGUAGU-----
ppy-mir-324   --CUGACUAUGCCUCCCGCAUCCCCUAGGGCAUUGGUGUAAGCUGGAGACCCACUGCCCCAGGUGCUGCUGGGGGGUUGUAGUC----
mml-mir-324   --CUGACUAUGCCUCCCGCAUCCCCUAGGGCAUUGGUGUAAGCUGGAGACCCACUGCCCCAGGUGCUGCUGGGGGGUUGUAGUC----
mmu-mir-324   AACUGACUAUGCCUCCUGCAUCCCCUAGGGCAUUGGUGUAAGCUGGAGACCCACUGCCCCAGGUGCUGCUGGGGGGUUGUAGUCUGAC
rno-mir-324   --CUGACUAUGCCUCUGCAUCCCCUAGGGCAUUGGUGUAAGCUGGAGACCCACUGCCCCAGGUGCUGCUGGGGGGUUGUAGUC----
bta-mir-324   AACUGGCUAUGCCUCCCGCAUCCCCUAGGGCAUUGGUGUAAGCUGGAGACCCACUGCCCCAGGUGCUGCUGGGGGGUUGUAGUCUGAC
ssc-mir-324   ------CUAUGCCUCCCGCAUCCCCUAGGGCAUUGGUGUAAGCUGGAGACCCACUGCCCCAGGUGCUGCUGGGGGGUUGUAGUCU---
cfa-mir-324   ----------------GCAUCCCCUAGGGCAUUGGUGUGGAGCUGGAGACCCACUGCCCCAGGUGCUGCUGGGGGGUUGUAGUC----
eca-mir-324   --CUGACUAUGCCUCCCGCAUCCCCUAGGGCAUUGGUGUAAGCUGGAGACCCACUGCCCCAGGUGCUGCUGGGGGGUUGUAGUC----
                              ********************** .*********  ************************* *********
```

Figure 4.3 miR-125b and miR-324 sequences. miR-125b and miR-324 stem-loop sequences were obtained from miRBase and aligned using clustal OMEGA [210] (www.clustal.org). Green and red boxes represent the approximate position of the '-5p' and '-3p' mature miRNAs respectively. (A) Alignment of miR-125b sequences (B) Alignment of miR-324 sequences. See Table 4.2 for list of species abbreviations.

Figure 4.4 Phylogeny analysis of miR-125b sequences

Figure 4.4 Phylogeny analysis of miR-125b sequences. Following sequence alignment using Clustal OMEGA, phylogeny analysis was performed using ClustalW2 Phylogeny (http://www.ebi.ac.uk/Tools/phylogeny/clustalw2_phylogeny/), with the following settings; tree format=Default, Distance correction=off (default), Exclude gaps-on (not default), clustering method=neighbour joining (default), P.I.M=off (default). Cluster grouping (clades) were named 'a','b' and 'c' for reference in the text. See Table 4.2 for list of species abbreviations.

Figure 4.5 Evolution of miR-125b and miR-324
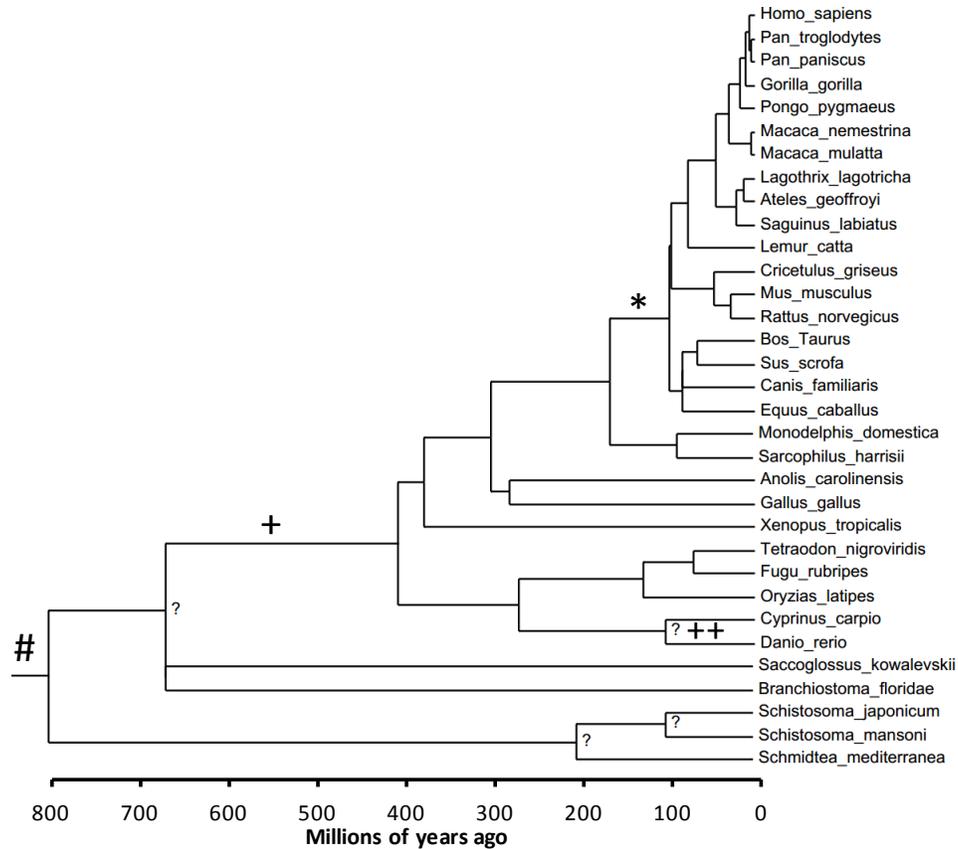
Figure 4.5 Evolution of miR-125b and miR-324. Information on the species which contain miR-125b and miR-324-5p stem-loops was obtained from miRBase. Times of divergence of each of these species were obtained from Time Tree (www.timetree.org). Species containing the stem-loops were arranged in a phylogenic tree using Phylodendron (http://iubio.bio.indiana.edu/treeapp/), and the time of divergence displayed on the x axis as an indicator of miRNA age. '*' indicates miR-324 origin, '#' indicates miR-125b origin, '+' and '++' indicates duplication of the miR-125b locus and '?' where there is uncertainty of the time of divergence of two species.

Table 4.2 Evolution of miR-125b and miR-324-5p

Species	Abs	Years since divergence from human lineage (millions)	Number of miR-125b loci in miRBase	miR-125b loci type				Number of miR-324-5p loci in miRBase
				miR-125b	miR-125b-1	miR-125b-2	miR-125b-3	
Homo sapiens	hsa	0	2		bi	ci		1
Pan troglodytes	ptr	6.3	2		bi	ci		1
Pan paniscus	ppa	6.3	1	bi				
Gorilla gorilla	ggo	8.8	2		bi	ci		
Pongo pygmaeus	ppy	15.7	2		bi	ci		1
Macaca nemestrina	mne	29	2		bi	ci		
Macaca mulatta	mml	29	2		bi	ci		1
Lagothrix lagotricha	lla	42.6	2		bi	ci		
Ateles geoffroyi	age	42.6	2		bi	ci		
Saguinus labiatus	sla	42.6	1	bi				
Lemur catta	lca	74	1	ci				
Cricetulus griseus	cgr	92.3	1	ci				
Mus musculus	mmu	92.3	2		bi	ci		1
Rattus norvegicus	rno	92.3	2		bi	ci		1
Bos taurus	bta	94.2	2		bi	ci		1
Sus scrofa	ssc	94.2	2		bi	ci		1
Canis familiaris	cfa	94.2	2		bi	ci		1
Equus caballus	eca	94.2	2	bi		ci		1
Monodelphis domestica	mdo	162.6	2		cii	bi		
Sarcophilus harrisii	sha	162.6	1	cii				
Anolis carolinensis	aca	296	2		bii	cii		
Gallus gallus	gga	296	1	cii				
Xenopus tropicalis	xtr	371.2	2		biii	cii		
Tetraodon nigroviridis	tni	400.1	1	bii				
Fugu rubripes	fru	400.1	1	bii				
Oryzias latipes	ola	400.1	2		bii	ciii		
Cyprinus carpio	ccr	400.1	1	civ				
Danio rerio	dre	400.1	3		bii	bii	civ	
Saccoglossus kowalevskii	sko	661.2?	1	a				
Branchiostoma floridae	bfl	713.2?	1	a				
Schistosoma japonicum	sja	792.4	1	a				
Schistosoma mansoni	sma	792.4	1	a				
Schmidtea mediterranea	sme	792.4	1	a				

Table 4.2 Evolution of miR-125b and miR-324-5p. Table showing list of species which have an annotated miR-125b loci, the number of years (millions) since divergence from humans, the number of annotated miR-125b loci, the number of annotated miR-324 loci and the clade to which each miR-125b loci belongs in Figure 4.4.

4.2.3 Predicted targets and functions of miR-125b-5p and miR-324-5p

A number of prediction sites can be used to predict miRNA targets; each has a slightly different algorithm and therefore predicts a different repertoire of targets (Chapter 1). Figure 4.6 shows the number and cross-over of targets predicted by each prediction site for mouse miR-125b-5p (Figure 4.6A), mouse miR-324-5p (Figure 4.6B), human miR-125b-5p (Figure 4.6C) and human miR-324-5p (Figure 4.6D).

DAVID pathway analysis is an online tool used to provide functional analysis of large gene lists. It works by using a modified version of the Fishers exact test [211]. In this study I used the DAVID pathway analysis functional annotation chart. Pathway analysis was performed separately on the predicted targets of each prediction site. Where more than 3,000 targets were predicted for a target prediction site, only the most highly scored 3,000 targets were input into DAVID pathway analysis, as this is the limit of the tool. Pathways/terms were considered enriched if they have an EASE score (modified version of the Fishers exact test) >0.1 and a gene count of at least 2. Figure 4.7 shows the number and cross-over of terms/pathways considered to be enriched for each prediction site for mouse miR-125b-5p (Figure 4.7A), mouse miR-324-5p (Figure 4.7B), human miR-125b-5p (Figure 4.7C) and human miR-324-5p (Figure 4.7D). Venn diagrams for all terms/pathways, GO terms only and Kegg pathways analysis are shown.

Figure 4.6 Predicted targets of miR-125b-5p and miR-324-5p.

A. mmu-miR-125b-5p

B. mmu-miR-324-5p

C. hsa-miR-125b-5p

D. hsa-miR-324-5p

E. miR-125b-5p
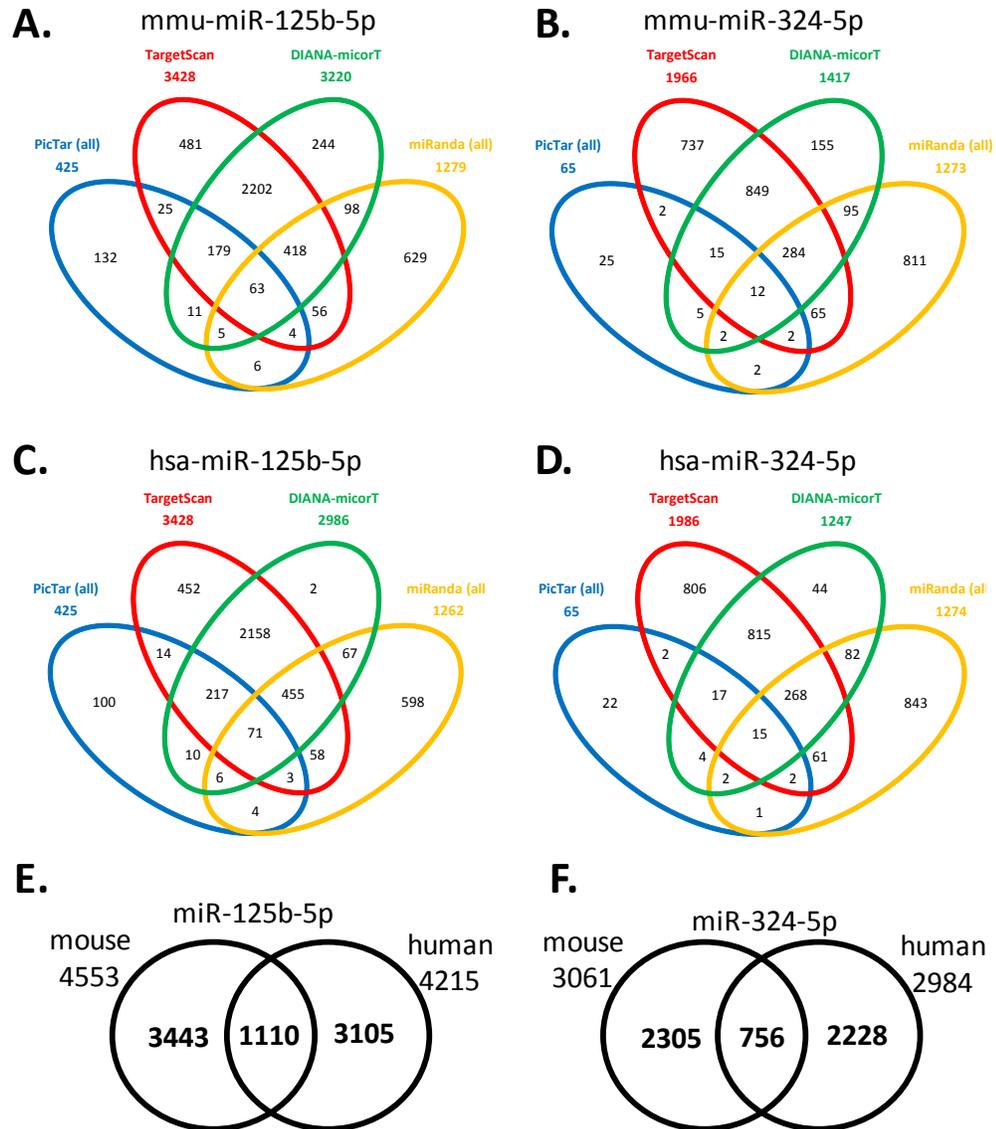
F. miR-324-5p

Figure 4.6 Predicted targets of miR-125b-5p and miR-324-5p. The number and crossover of targets predicted by four online databases. (A) Predicted targets of mouse miR-125b-5p. (B) Predicted targets of mouse miR-324-5p. (C) Predicted targets of human miR-125b-5p. (D) Predicted targets of human miR-324-5p. (E) Common predicted targets of human and mouse miR-125b-5p. (F) Common predicted targets of human and mouse miR-324-5p.

Figure 4.7 Predicted functions of miR-125b-5p and miR-324-5p.
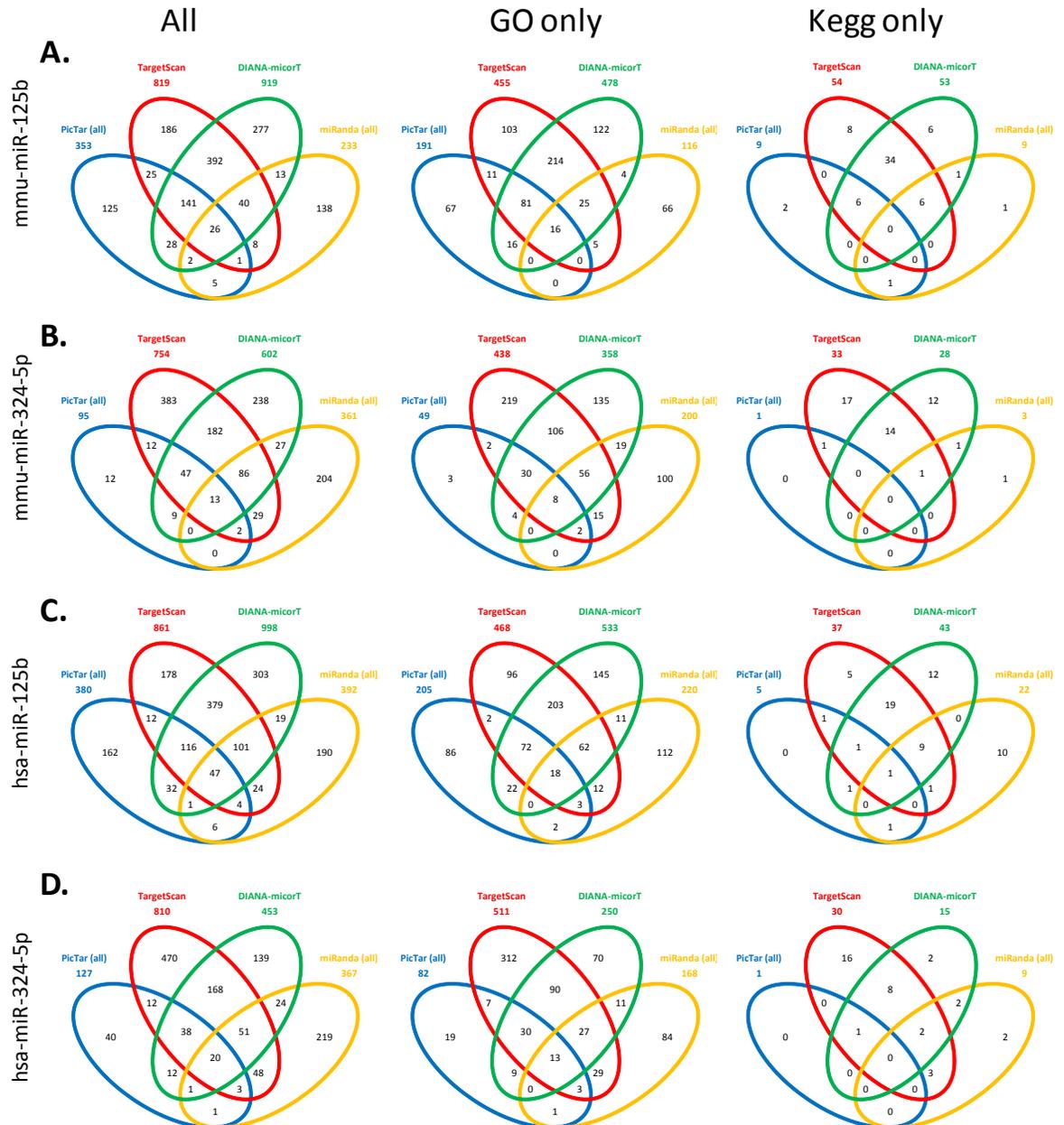
Figure 4.7 Predicted functions of miR-125b-5p and miR-324-5p. Predicted targets for each online database were in put into DAVID pathway functional annotation chart, where more than 3000 targets were predicted only the top 3000 predicted target were input. Cut off for pathways being enriched was left at the DAVID default (0.1). The number and crossover of enriched pathways is shown. 'All' refers to all pathways and terms in the DAVID pathway analysis tool, which includes GO, Kegg and others. (A) Pathway analysis of mouse miR-125b-5p predicted targets. (B) Pathway analysis of mouse miR-324-5p predicted targets. (C) Pathway analysis of human miR-125b-5p predicted targets. (D) Pathway analysis of human miR-324-5p predicted targets.

4.2.4 miR-324-5p but not miR-125b-5p targets human SMO and GLI1 3'UTRs

In 2008, Ferretti *et al.* showed both miR-125b-5p and miR-324-5p regulate the Hh signalling pathway in human medulloblastoma [127]. The authors used luciferase constructs to show miR-125b-5p directly targets the Hh activator SMO, and miR-324-5p directly targets both SMO and the Hh transcription factor GLI1. To confirm this, luciferase reporter constructs were generated by cloning the human 3'UTR of either SMO or GLI1 downstream of luciferase, placing luciferase under their control. In agreement with Ferretti *et al.*, co-transfection of miR-324-5p with the constructs containing the SMO or GLI1 3'UTR reduces luciferase expression, indicating miR-324-5p can directly regulate human SMO and GLI1 3'UTRs (Figure 4.8A and B). Using site-directed mutagenesis we created constructs where the miR-324-5p binding site is no longer complementary to miR-324-5p (Figure 4.8D and E). The extent to which miR-324-5p reduced the mutant construct luciferase was decreased compared to the wild-type construct (Figure 4.8A and B), indicating miR-324-5p exerts its action on the SMO and GLI1 3'UTR, through the binding sites proposed by Ferretti *et al.* Unexpectedly, miR-125b-5p significantly increased luciferase activity for the SMO wild-type and mutant 3'UTRs (Figure 4.8C and F), suggesting the human SMO 3'UTR is not a target of miR-125b-5p.

Figure 4.8 miR-324-5p and miR-125b-5p target components of the human Hh signalling pathway.
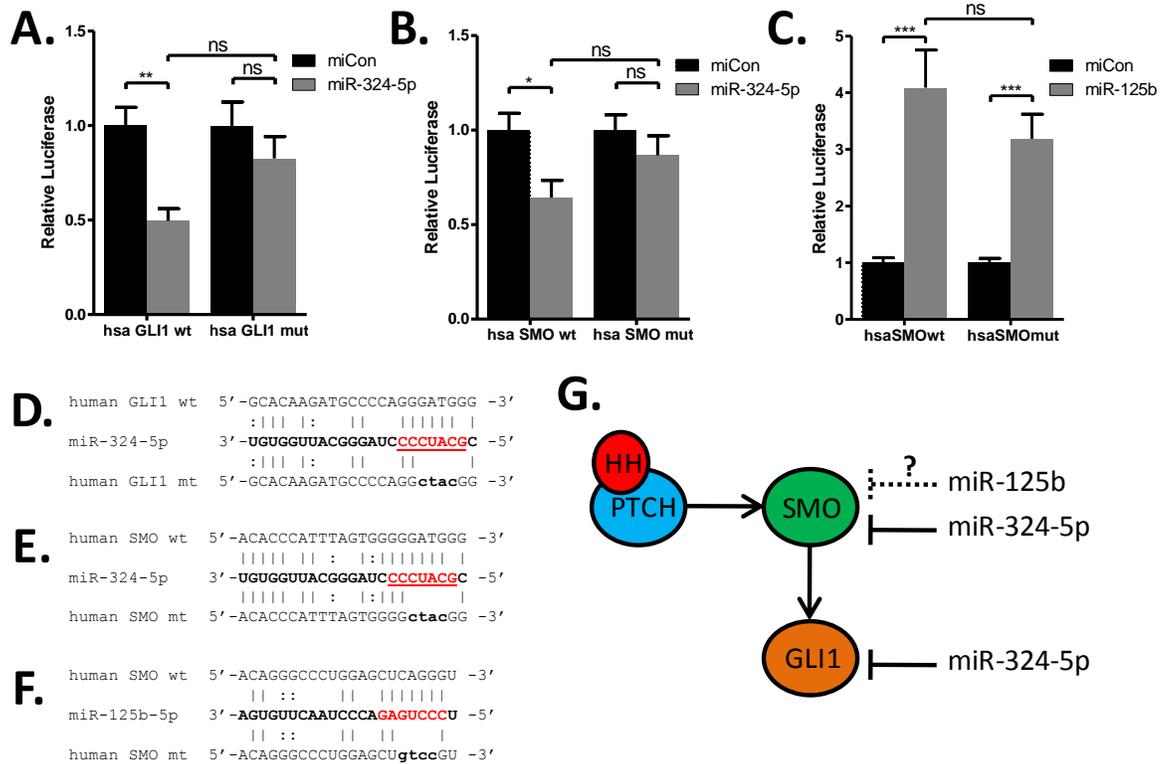
Figure 4.8 miR-324-5p and miR-125b-5p target components of the human Hh signalling pathway. Cells were transfected with either GLI1wt, GLI1mt, SMOwt or SMOmt 3'UTR luciferase constructs (in pMIR-Report-Ambion) with either miCon, miR-324-5p or miR-125b-5p. (A) miR-324-5p targets GLI1. (B) miR-324-5p targets SMO. (C) miR-125b-5p increases SMO 3'UTR luciferase. Schematic of miRNA interaction with wild-type (wt) (upper) and mutant (mt) (lower) 3'UTRs for (D) miR-324-5p:GLI, (E) miR-324-5p:SMO and (F) miR-125b-5p:SMO interactions. (G) Schematic of miR-125b-5p and miR-324-5p targeting Hh signalling in human. Data normalised to miCon for each plasmid and plotted as relative luciferase light units. All data are presented as mean + SEM, statistical difference were calculated using Student's t-test, where, *p<0.05, **<0.01, ***p<0.001. Data combined from 5 independent experiments, each n=6.

4.2.5 Activation of Hh signalling in C3H10T1/2

The Hh signalling pathway is known to be involved in cartilage formation, bone formation [178, 212] and OA [48]. I hypothesised that miR-125b-5p and miR-324-5p are also involved in these processes, especially given their altered expression in OA (Figure 4.1). Isolated human articular chondrocytes (HAC) are commonly used as an *in vitro* system to study the effects of molecules on chondrocytes and C3H10T1/2 cells are commonly used to the study the effects of molecules on osteoblastogeneisis and chondrogenesis. To test the ability of both cells to respond to Hh signalling I stimulated each cell type with Hh ligand (Ihh) for 48h. Gene expression analysis of the well characterised Hh response gene and transcription factor Gli1 confirmed both cell types had active Hh signalling (Figure 4.9). C3H10T1/2 had a more robust Hh response than HAC (Figure 4.9). Itherefore decided to use C3H10T1/2 cells to observe any modulatory effects of miR-125b-5p and miR-324-5p on Hh signalling. C3H10T1/2 is a mouse pluripotent mesenchymal cell-line, isolated from mouse embryonic C3H cells by Reznicoff *et al.* in 1973 [178]. In addition, C3H101/2 have a number of properties making them ideal for these studies. They are easily transfected and, under the correct stimuli, have the ability to undergo differentiation processes such as osteogenesis and chondrogenesis [180, 213].

In addition to activation of the Hh signalling pathway by ligands (Ihh, Shh and Dhh), it can also be activated by small molecule agonists of SMO such as Purmorphamine (Pur) [73] and SAG. [74] Following Ihh or Pur stimulation of C3H10T1/2 cells, Gli1 mRNA was robustly induced (Figure 4.10A) and increased in a time-dependent manner (Figure 4.10B and C). Similar to Gli1, Ptch is also a Hh response gene and part of the Hh signalling pathway, and as expected was also induced following Ihh and Pur stimulation, although to a lesser extent (3-4 fold) (Figure 4.10D), in a time dependent manner (Figure 4.10E and F). Ihh increased Gli1 protein expression as indicated by Immuno-blotting (Figure 4.10G). Part of Hh signalling activation also includes Smo relocalisation to the primary cilia. Following stimulation with Ihh, Smo increases around the primary cilium (Figure 4.10H). Quantification of the level of Smo, shows maximum localisation around the cilia following 1h of Ihh stimulation (Figure 4.10I).

Figure 4.9 Hh response in HAC and C3H10T1/2
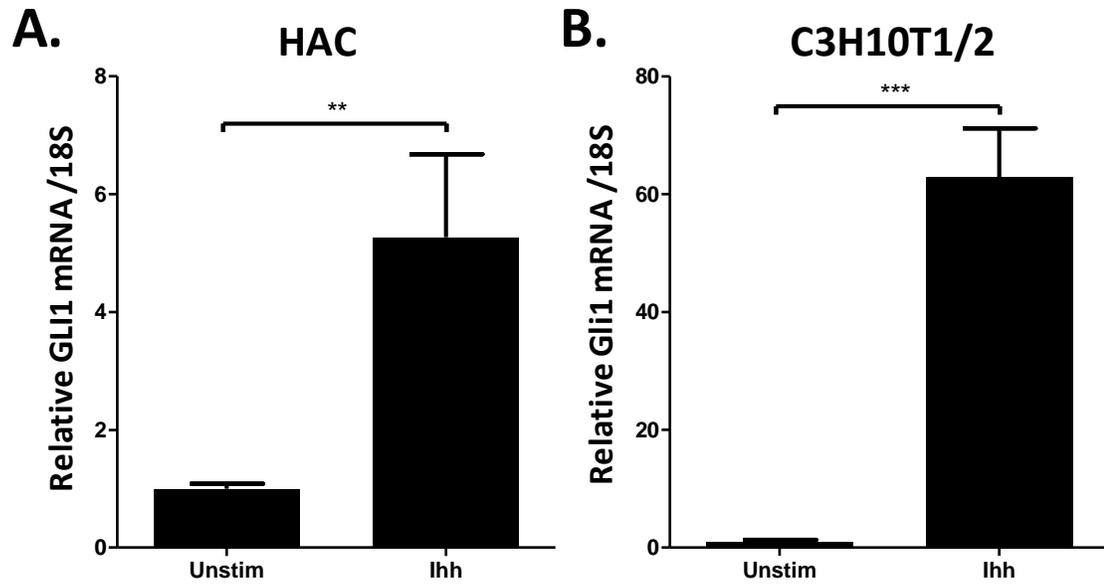
Figure 4.9 Hh response in HAC and C3H10T1/2. (A) Human articular chondrocytes (HACs) were stimulated with 2μg/ml recombinant Indian hedgehog (Ihh) for 48h. Data combined from two patient donors, each n=4. (B) C3H10T1/2 cells were stimulated with 2μg/ml Ihh for 48h. Gli1 mRNA levels were then assessed using real time RT-PCR as described. Data from one n=4 experiment. All data are presented as mean + SD, statistical difference were calculated using Student's *t*-test, where **<0.01, ***<0.001.

Figure 4.10 Activation of Hedgehog signalling in C3H10T1/2.

H.

DAPI SMO Ac-tubulin

I.
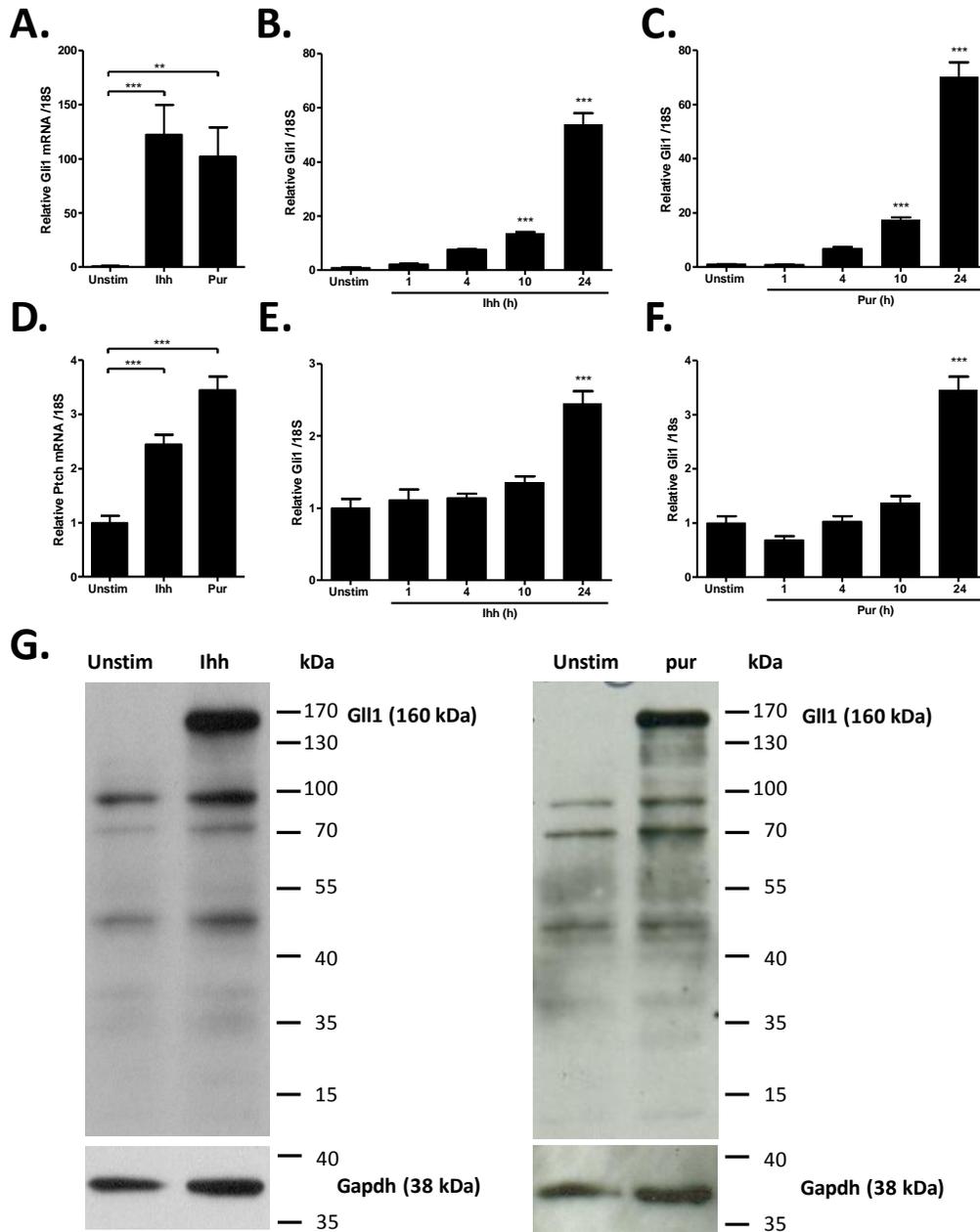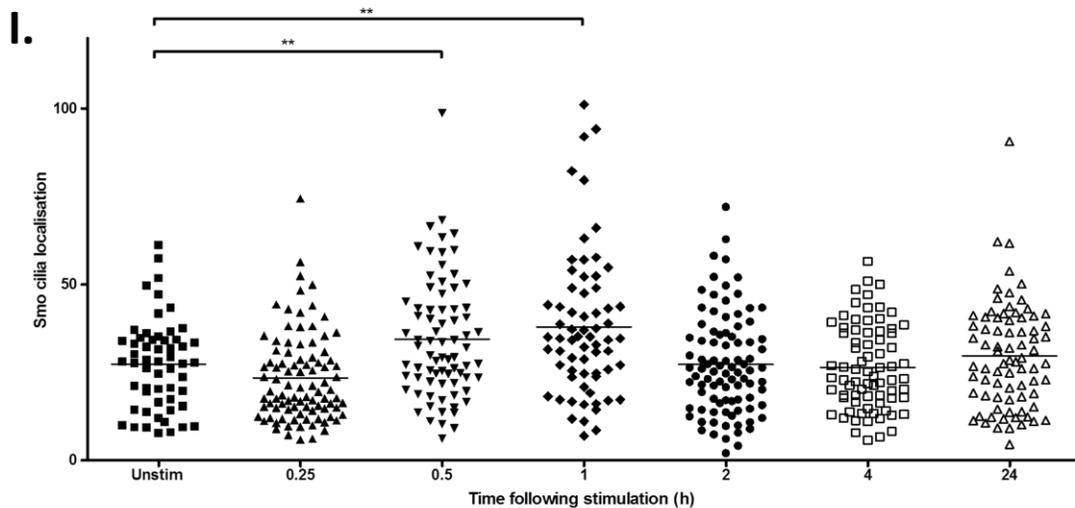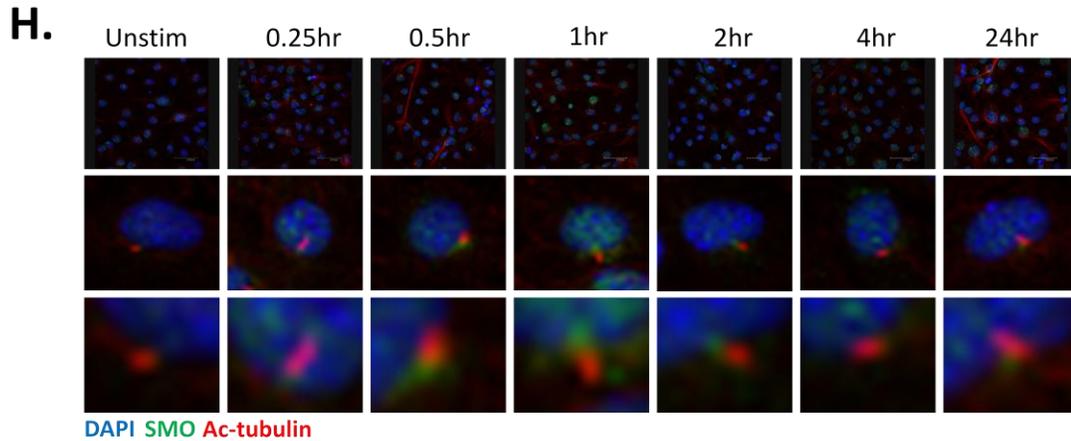
Figure 4.10 Activation of Hh signalling in C3H10T1/2. C3H10T1/2 cells stimulated with recombinant Ihh (2µg/ml) or Purmorphamine (Pur) (2uM) for 24h or the time shown. Real-time RT-PCR was used to assess Gli1 and Ptch expression. Western blotting was used to assess Gli1 Protein. Acetylated α-tubulin and Smo immunostaining was used to assess Smo localisation around the primary cilia. (A) Ihh and Pur both increase Gli1 mRNA after 24h (data combined). (B) Ihh causes time dependent increase in Gli1 mRNA. (B) Pur causes time dependent increase in Gli1 mRNA. (D) Ihh and Pur increase Ptch mRNA after 24h (data combined). (E) Ihh causes time dependent increase in Ptch mRNA. (F) Pur causes time dependent increase in Ptch mRNA. (A-F) All data are presented as mean + SEM, statistical difference were calculated using ANOVA followed by a Bonferroni post test, where **<0.01, ***<0.001. Date from one representative experiment (n=6), of at least two independent experiments (G) Ihh and Pur increases Gli1 protein. Representative blot of at least 5 independent experiments (H) Ihh stimulation causes increased Smo localisation around the primary cilia. Representative cells are shown. Blue=DAPI, Green =SMO, Red= acetylated α-tubulin (I) Quantification of Smo localisation around the cilia with ImageJ (see materials and methods). Statistical difference were calculated using ANOVA followed by a Bonferroni post test, where **<0.01. Date combined from two independent experiments.

4.2.6 Inhibition of Hh signalling components in C3H10T1/2

Ptch, Smo and Gli1 are key Hh signalling proteins. Depletion of these components with RNA interference should demonstrate the role of each in Hh signalling in C3H10T1/2. I therefore depleted each component with a pool of specific siRNAs (see materials and methods). As expected siRNA against Ptch, Smo and Gli1 all have the potential to alter Ihh and/or Pur induced Gli1 and Ptch expression (Figure 4.11), indicating all these components were involved in Hh signalling. There are however a number of additional effects of depleting these components on Hh signalling, some expected and some unexpected, these are described further in Appendix A.

Figure 4.11 RNA interference of Hedgehog signalling components.
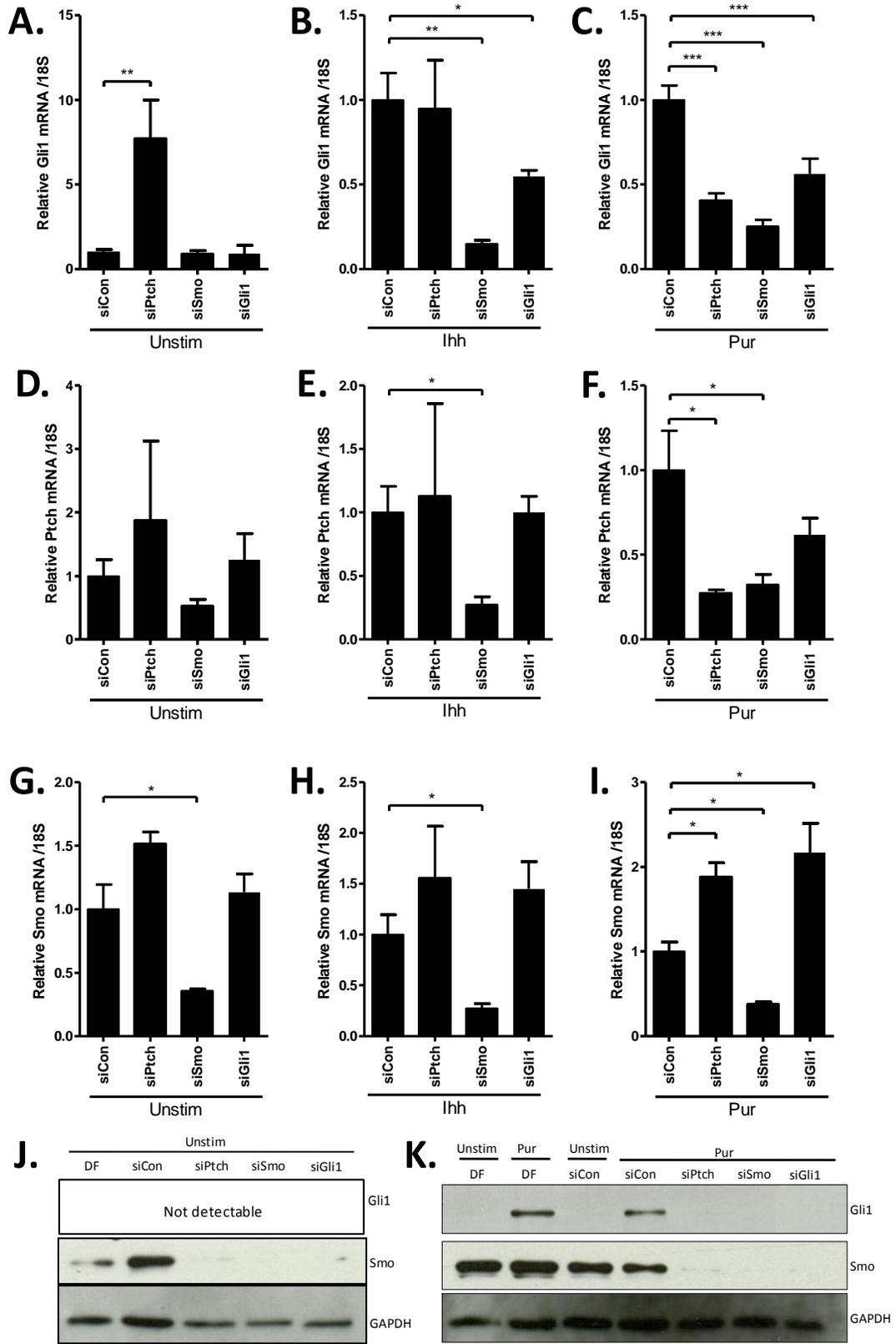
Figure 4.11 RNA interference of Hh signalling components. C3H10T1/2 cells were transfected with non-targeting siCon or siRNA against Ptch, Smo and Gli1 for 24 h. Cells were then serum starved for 24h and either left unstimulated, stimulated for Ihh (2µg/ml) for 24h or stimulated with Purmorphamine (Pur) (2µM) for 24hr. Gli1, Ptch and Smo were assessed by real-time RT-PCR. Gli1 and Smo protein expression were assessed by Immunoblotting. (A) siRNA effect on basal Gli1 expression. (B) siRNA effect on Ihh induced Gli1 expression. (C) siRNA effect on Pur induced Gli1 expression. (D) siRNA effect on basal Ptch expression. (E) siRNA effect on Ihh induced Ptch expression. (F) siRNA effect on Pur induced Ptch expression. (G) siRNA effect on basal Smo expression. (H) siRNA effect on Ihh induced Smo expression. (I) siRNA effect on Pur induced Smo expression. Data combined from three independent experiments, each n=4. Data are presented as mean + SEM, statistical difference were calculated using ANOVA followed by a Bonferroni post test, where $*p < 0.05$, $**<0.01$, $***<0.001$. (J) Effect of siRNA on basal Gli1 and Smo protein. (K) Effect of siRNAs on Pur induced Gli1 and Smo protein levels. GAPDH was used as a loading control.

4.2.7   Hh signalling increases osteogenesis and chondrogenesis in C3H10T1/2

Kinto *et al.* first showed Hh activation can induce osteoblast differentiation of C3H10T1/2 [214]. Nakamura *et al.* then showed Ihh can synergise with Bmp2 in C3H10T1/2 to produce higher levels of alkaline phosphatase [179]. Indeed, I observed that stimulation of C3H10T1/2 with Bmp2 or Ihh increased alkaline phosphatase, and their combination produced a further increase, indicative of osteoblastogenesis (Figure 4.12). Since the original studies, Pur has been identified as an activator of Hh signalling and as expected increased alkaline phosphatase and caused a further increase when in combination with Bmp2 (Figure 4.12A). However siRNA against Gli1 did not reduce Bmp2 + Ihh induced alkaline phosphatase (Figure 4.12B).

C3H10T1/2 can also undergo a type of chondrogenic differentiation when seeded into high density micromass cultures and stimulated with Bmp2 [180]. Chondrogenesis can be indicated by alcian blue, which stains glycosaminoglycans (GAG), and can be quantified by extracting alcian blue using guanidine hydrochloride (GHCl) and measuring absorbance at 600nm (see materials and methods). As expected, Bmp2 increased alcian blue staining and the addition of Ihh to Bmp2 further increased alcian blue staining (Figure 4.13A and C). siRNA against Gli1 decreased basal (unstim), Bmp2 stimulated and Bmp2 + Ihh stimulated alcian blue (Figure 4.13B and D), indicating a role of Hh signalling in these micro masses, as previously suggested [215].

Figure 4.12 Effect of Hh signalling on Osteoblastogenesis.
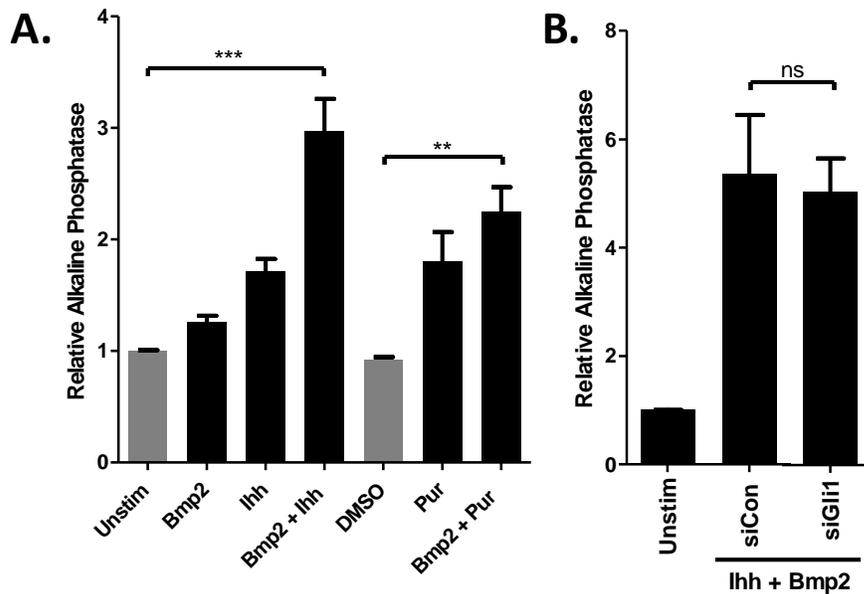
Figure 4.12 Effect of Hh signalling on Osteoblastogenesis. (A) C3H10T1/2 cells were seeded in mono-layer and were stimulated with either Ihh (2µg/ml), Bmp2 (100ng/ml) or Pur (2uM) either alone or in combination for 5 days. C3H10T1/2 cells were transfected with siCon or siGli1, then stimulated with Ihh (2µg/ml) and Bmp2 (100ng/ml) for 5 days. Alkaline phosphatase activity was used as a marker of osteogenesis. Relative alkaline phosphatase indicates relative absorbance at 405nm of alkaline phosphate yellow liquid substrate following incubation with cells. Data combined from three independent experiments, each n=3. All data are presented as mean + SEM, statistical difference were calculated using ANOVA followed by a Bonferroni post test, where **<0.01 and ***<0.001.

Figure 4.13 Effect of Hh signalling on Chondrogenesis.
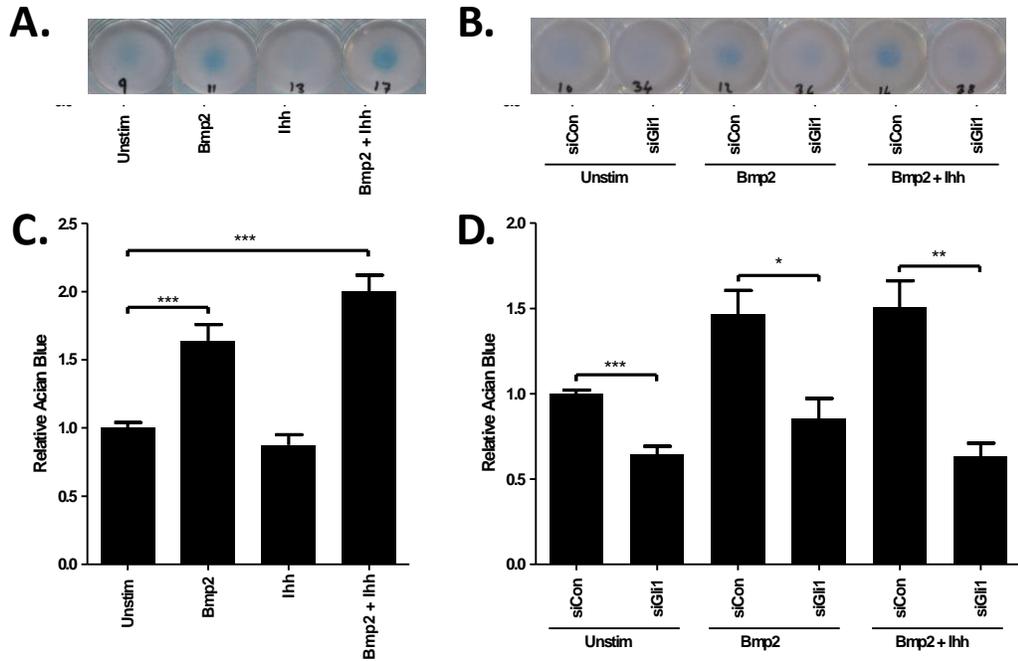
Figure 4.13 Effect of Ihh signalling on chondrogenesis. C3H10T1/2 cells were either left untransfected or pre transfected with either siCon or siGli1, seeded in micromasses and were stimulated with either Ihh (2µg/ml), Bmp2 (100ng/ml) either alone or in combination for 5 days. Alcian Blue was used as a marker of chondrogenesis. Alcian blue was quantified by extracting using GHCl and measuring absorbance at 600nm. (A) Bmp2 and Ihh effect on Alcian blue staining, (representative micromasses), (B) siGli1 effect on Alcian blue staining (representative micromasses). (C) Bmp2 and Ihh effect on extracted Alcian blue, data combined from 6 independent experiments, each n=2. (D) siGli1 effect on extracted Alcian blue, data combined from 2 independent experiments, each n=2. All data are presented as mean + SEM, statistical difference were calculated using ANOVA followed by a Bonferroni post test, where *p<0.05, **<0.01, ***<0.001.

4.2.8 miR-125b-5p and miR-324-5p regulate Hh signalling in C3H10T1/2

Having established readouts of Ihh signalling, osteogenesis and chondrogenesis I next investigated the effects of miR-125b-5p and miR-324-5p on these processes. Our luciferase constructs suggest miR-324-5p directly regulates components of the Hh pathway in humans (Figure 4.8) and previous data by Ferretti *et al.* suggest both miR-125b-5p and miR-324-5p are regulators of Hh signalling in humans [127]. I hypothesised that miR-125b-5p and miR-324-5p may regulate Hh signalling in mouse C3H10T1/2. Real-time RT-PCR and western blot analysis of Gli1, showed pre transfection of either miR-125b-5p or miR-324-5p inhibited Ihh induced Gli1 mRNA and protein expression respectively (Figure 4.14A and B), demonstrating miR-125b-5p and miR-324-5p can both regulate Hh signalling in the mouse C3H10T1/2 cell line.

As shown in Figure 4.10H and I, Ihh stimulation causes Smo to move toward the primary cilia, with the greatest amount of Smo surrounding the primary cilium after 1h (Figure 4.10H and I). I wanted to see if either miR-125b-5p or miR-324-5p have an effect on Smo localisation to the primary cilium, I choose 1h of stimulation to maximize any effect miR-125b-5p and miR-324-5p may have. miR-125b-5p and miR-324-5p both reduced the total amount of Smo surrounding the primary cilia (Figure 4.14C) as quantified by ImageJ (Figure 4.14D), however, this may be due to a reduction in the overall level of Smo staining (Figure 4.14C). Interestingly preliminary data suggest miR-125b-5p but, not miR-324-5p, reduces primary cilia length (Figure 4.14E). Regardless, these data overall provide further evidence for the involvement of miR-125b-5p and miR-324-5p in Hh signalling.

miR-125b-5p and miR-324-5p inhibit Ihh induced Gli1 in C3H101/2 cells, (Figure 4.14). However, in Pur (2µM) treated cells, pre-transfection of miR-125b-5p increased, while pre-transfection of miR-324-5p did not consistently alter, Ihh induced Gli1 expression (Figure 4.15A). I then speculated this was due to differences in the way Ihh and Pur activate the Hh pathway and investigated this further. Both Ihh and Pur induce Hh pathway activity (Gli1 expression) in a similar time-dependent manner (Figure 4.10), indicating timing is unlikely to have a role in the differences in the effects of miR-125b-5p and miR-324-5p on Hh signalling. An interesting observation was that stimulation of

C3H10T1/2 cells with increasing concentrations of Pur and SAG caused increased Hh pathway activity, however at higher concentrations of both, Hh activity decreased (Figure 4.15C-F), similar to previously reported observations [74, 216] and indicating a possible negative feedback loop; something not observed with the examined increasing concentrations of Ihh (Figure 4.15G). The addition of miR-125b-5p to Pur stimulated cells caused an increase in Hh pathway activity (Figure 4.15A), whereas the addition of miR-125b-5p to Ihh stimulated cells caused a decrease in Hh activity (Figure 4.14). I therefore hypothesised miR-125b-5p disrupted this negative feedback loop and allowed for the equivalent level of Hh signalling seen from lower level concentrations of Pur (see discussion). miR-125b-5p increased 1000nM SAG-induced Gli1 but did not increase 50nM SAG induced Gli1 (Figure 4.15H and I). This is likely to be because at 1000nM SAG miR-125b-5p disrupts the negative feedback loop and allows high Hh, whereas at 50nM SAG there is little activation of the negative feedback loop meaning miR-125b-5p is unable to modulate Hh signalling.

Figure 4.14 miR-125b-5p and miR-324-5p inhibit Ihh induced Hh signalling in C3H10T1/2.
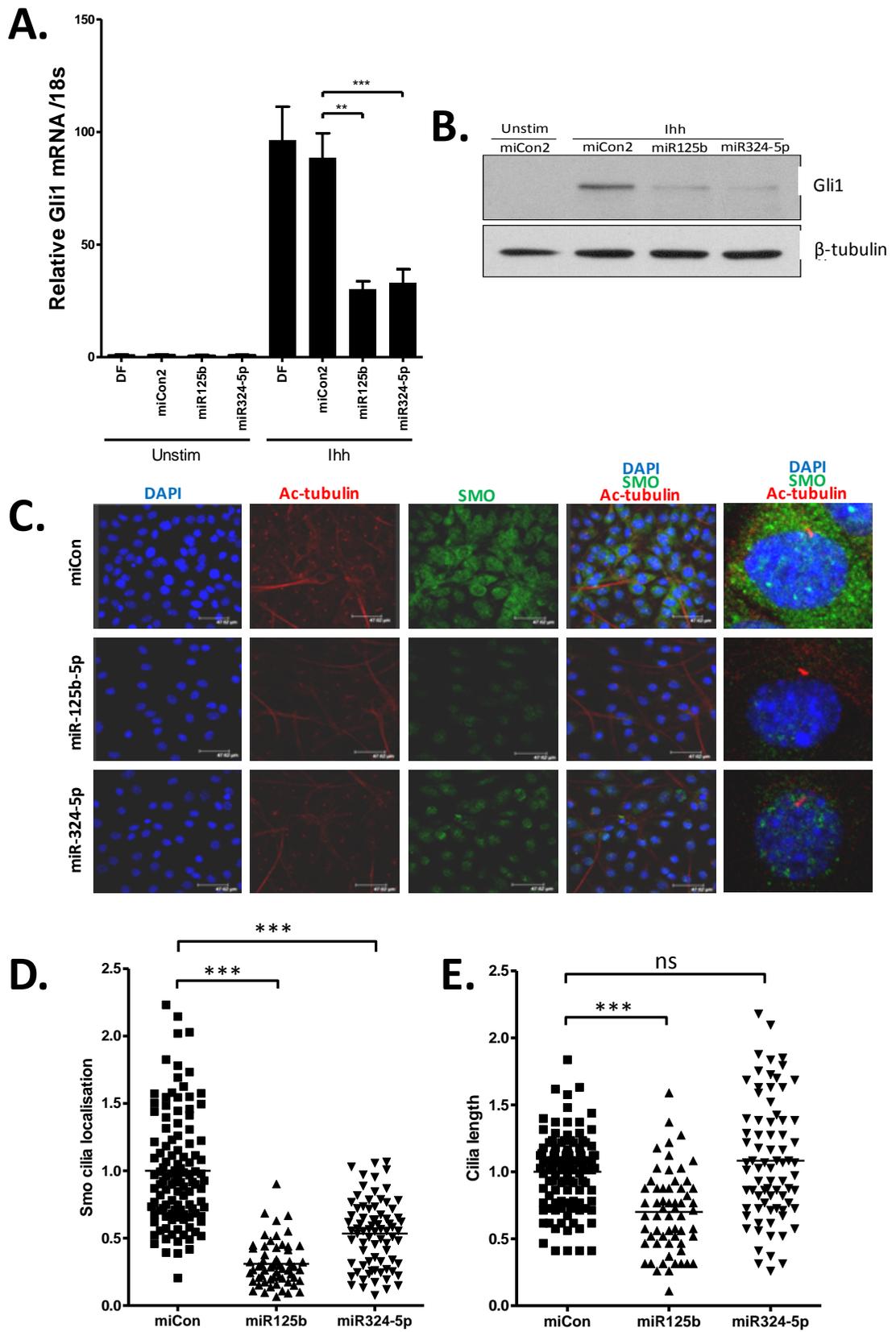
Figure 4.14 miR-125b-5p and miR-324-5p inhibit Ihh induced Hh signalling in C3H10T1/2. C3H10T1/2 were transfected with miCon, miR-125b-5p or miR-324-5p and stimulated with recombinant Indian hedgehog (Ihh) 2µg/ml for either 24h (A and B) or 1h (C and D). (A) Effect of miR-125b-5p and miR-324-5p on Gli1 mRNA, as measured by real-time RT-PCR, data combined from 8 independent experiments, each at least n=4. All data are presented as mean + SEM, statistical difference were calculated using ANOVA followed by a Bonferroni post test, where $**<0.01$, $***<0.001$. (B) Effect of miR-125b-5p and miR-324-5p on Gli1 protein, as measured by and immunoblotting. Representative blot of 5 independent experiments. (C) Smo localisation around the primary cilia was visualised by staining with a Smo Ab and an acetylated α-tubulin Ab to show the primary cilia. (D) Images were quantified using Image J (see materials and methods). (E) Cilia length was also quantified using ImageJ (see materials and methods). (D and E) statistical differences were calculated using Stuvents ttest, where $***<0.001$

Figure 4.15 miR-125b-5p and miR-324-5p have differing effects on Pur and SAG induced Gli1 in C3H10T1/2.
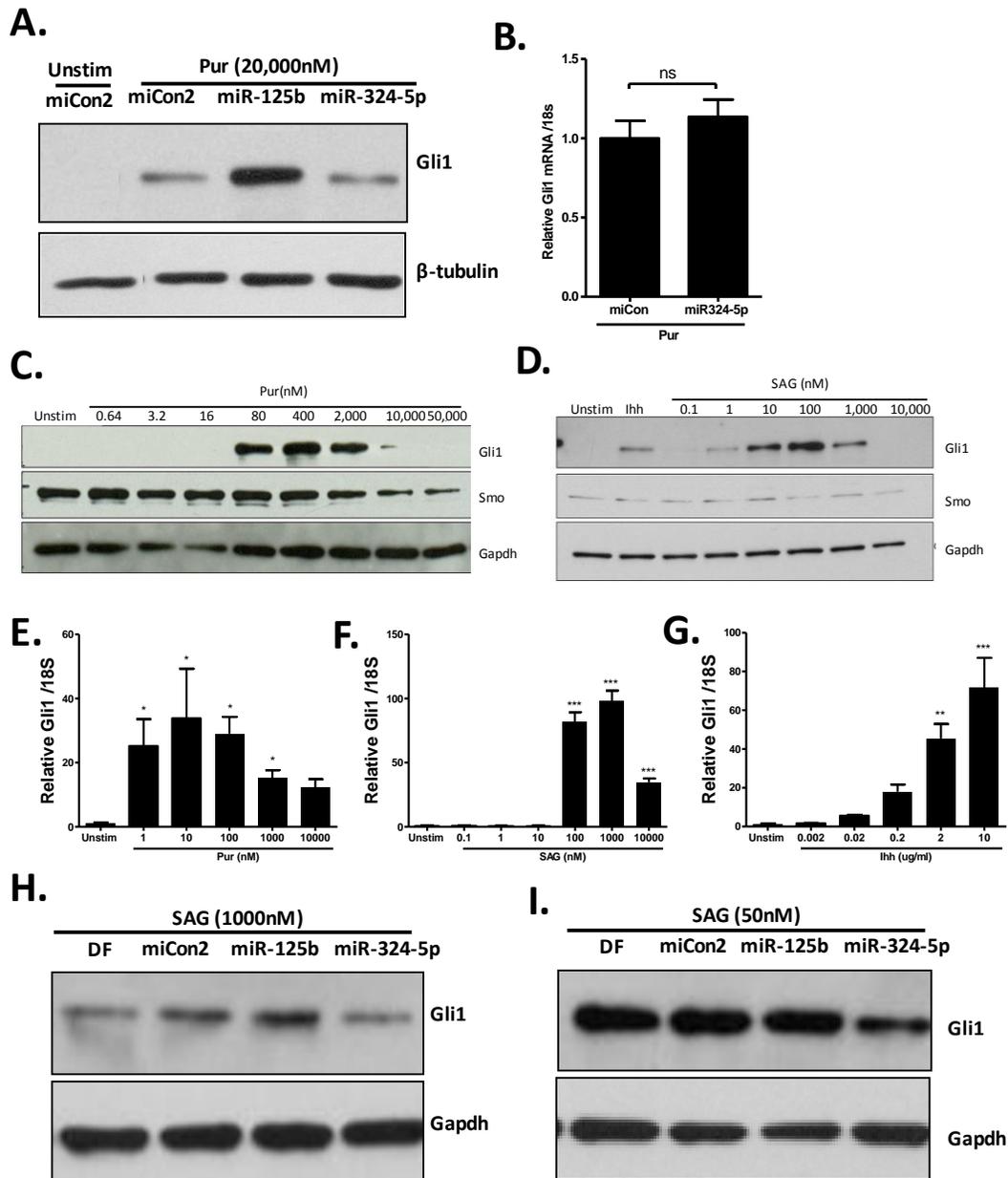
Figure 4.15 miR-125b-5p and miR-324-5p have differing effects on Pur and SAG induced Gli1 in C3H10T1/2. (A) Gli1 protein following transfection with miR-125b-5p or miR-324-5p, serum starving for 24h and stimulated with Pur (2µM) for 48 h. Representative blot from 3 independent experiments (B) Effect of miR-324-5p on Pur induced Gli1 mRNA, Representative data from two independent experiments, n=4. (C-G) C3H10T1/2 were stimulated with Pur, SAG or Ihh at the shown concentration for 48h and assessed for Gli1 mRNA and protein. (H and I) C3H10T1/2 were transfected with miR-125b-5p or miR-324-5p serum starved for 24h and stimulated with SAG (1000nM or 50nM) for 48 h. Gli1 was measured using immunoblotting. Statistical difference were calculated using ANOVA followed by a Bonferroni post test.

4.2.9 miR-125b-5p and miR-324-5p effect on alkaline phosphatase and alcian blue

As activation of Hh signalling increased alkaline phosphatase (Figure 4.12), and miR-125b-5p and miR-324-5p decreased Hh signalling (Figure 4.14), I wanted to test if inactivation of Hh signalling by miR-125b-5p and miR-324-5p alters the level of alkaline phosphatase production. Transfection of miR-125b-5p or miR-324-5p reduced Ihh + Bmp2 induced alkaline phosphatase (Figure 4.16). When I used Pur in place of Ihh I observed an increase in alkaline phosphatase with miR-125b-5p (Figure 4.16), similar to the effect observed for Gli1 mRNA and protein (Figure 4.15A). As expected miR-324-5p decreased Pur and Bmp2 induced alkaline phophatase.

The addition of Ihh to Bmp2 stimulated micromass cultures caused an increase in the level of alcian blue staining (Figure 4.13A and C), likewise pre transfection of siRNA against Gli1 caused a decrease (Figure 4.13B and D). As miR-324-5p decreases Hh signalling (Figure 4.14), I hypothesised miR-324-5p would affect alcian blue staining. To test this I transfected C3H10T1/2 cells with miR-324-5p and seeded these into micromass cultures. Unexpectedly, miR-324-5p did not consistently cause an increase or decrease in the amount of extracellular matrix produced, as indicated by alcian blue staining, for either cells stimulated with Bmp2, the combination of Bmp2 + Ihh, or unstimulated cells (Figure 4.17A and B).

Figure 4.16 miR-125b-5p and miR-324-5p effect on osteoblastogenesis.
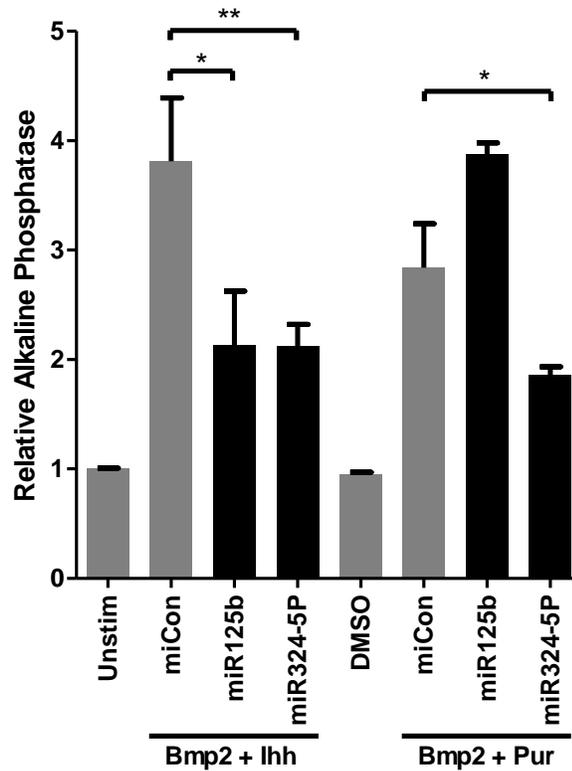
Figure 4.16 miR-125b-5p and miR-324-5p effect on osteoblastogenesis. Alkaline phosphatase was used as a marker of bone formation, C3H10T1/2 cells were transfected with miR-125b-5p and miR-324-5p, then stimulated with Ihh (2μg/ml) + Bmp2 (100ng/ml) or Purmorphamine (Pur) (2uM) and Bmp2 (100ng/ml) for 5 days. P.Nitrophenol was measured to determine the level of alkaline phosphatase. Data combined data from 5 independent experiments each n=3, All data are presented as mean + SEM, statistical difference were calculated using ANOVA followed by a Bonferroni post test, where *$p<0.05$.

Figure 4.17 miR-324-5p effect on chondrogenesis.
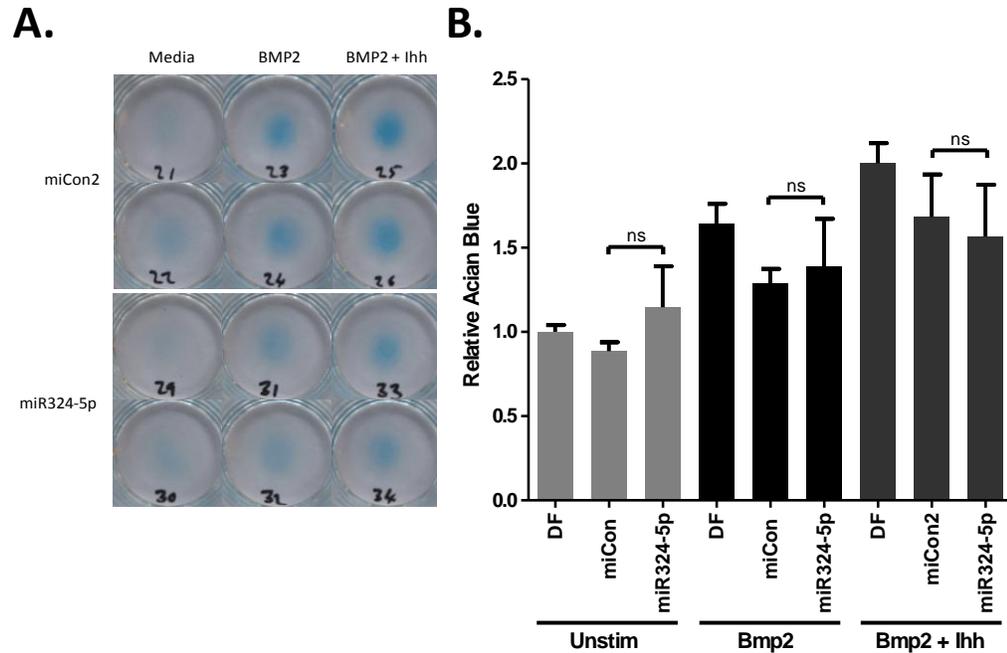
Figure 4.17 miR-324-5p effect on chondrogenesis alcian blue was used as a marker of cartilage formation. (A) C3H10T1/2 cells were transfected with miR-324-5p then plated into micromass cultures, stimulated with Bmp2 (100ng/ml) and/or Ihh (2µg/ml) (B) alcian blue stain was extracted with GHCl and quantified as described previously. Data combined from 3 independent experiments, each n=2. All data are presented as mean + SEM, no statistical difference were found using Student's *t*-test.

4.2.10 Retinoic acid but not Hh stimulation regulates miR-125b-5p and miR-324-5p expression

miRNAs are often part of feedback loops. miR-125b-5p and miR-324-5p regulate Hh signalling (Figure 4.14), therefore I hypothesised miR-125b-5p and miR-324-5p may be regulated by either Ihh or Pur stimulation. To test this I stimulated C3H10T1/2 with Ihh or Pur and measured miR-125b-5p and miR-324-5p expression using real-time RT-PCR. miR-125b-5p expression was slightly decreased following Ihh and Pur stimulation (Figure 4.18A). miR-324-5p was not significantly altered in expression by either Ihh or Pur stimulation (Figure 4.18B).

Retinoic acid (RA) has previously been show to increase the expression of miR-125b-5p and miR-324-5p [127], I hypothesised it may have the same effect in C3H10T1/2. Unexpectedly, RA decreased miR-125b-5p and miR-324-5p expression, although its effect on miR-125b-5p was abolished following the addition of Ihh (Figure 4.18C and D). RA also inhibited Hh signalling (Figure 4.18E and F), opposite of the effect that would be expected from its effect on miR-324-5p.

Figure 4.18 Retinoic acid effects miR-125b-5p and miR-324-5p expression.
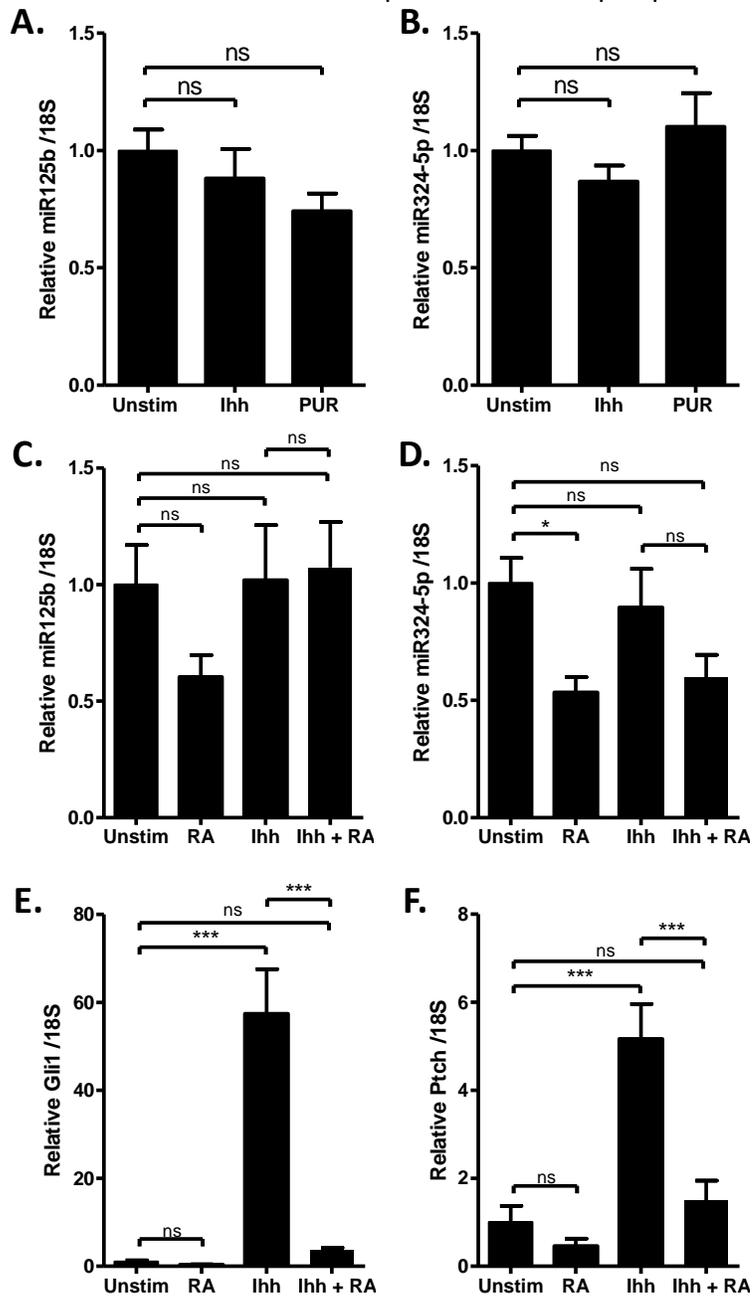
Figure 4.18 Retinoic acid effects miR-125b-5p and miR-324-5p expression. C3H10T1/2 cells were stimulated with either Ihh (2µg/ml), Purmorphamine (Pur) (2µM), or Retinoic acid (RA) (2.5µM) or in combination for 48hr. miR-125b-5p, miR-324-5p, Gli1 and Ptch were then assessed with real-time RT-PCR and normalised to 18S. (A) miR-125b-5p expression following Ihh and Pur stimulation. (B) miR-324-5p expression following Ihh and Pur stimulation. (C) miR-125b-5p expression following retinoic acid and Ihh stimulation. (D) miR-324-5p expression following retinoic acid and Ihh stimulation. (E) Gli1 expression following Retinoic acid and Ihh stimulation. (F) Ptch expression following Retinoic acid and Ihh stimulation. Data combined data from 2 independent experiments each n=5, All data are presented as mean + SEM, statistical difference were calculated using Student's t-test, where *p<0.05, **<0.01 and ***<0.001.

4.2.11 miR-125b-5p and miR-324-5p targets are different in human and mouse

miR-125b-5p and miR-324-5p can regulate Hh signalling in human [127] and mouse (Figure 4.14). *Ferretti et al.*, and ourselves show miR-324-5p can directly regulate GLI1 and SMO in human (Figure 4.8). They suggest miR-125b-5p directly regulates SMO, although our results fail to support this observation (Figure 4.8C). 3'UTR sequences differ between species more than the corresponding protein coding sequence (CDS), therefore I aligned the human and mouse 3'UTR sequences and performed miR-125b-5p and miR-324-5p target searches to attempt to identify if miR-125b-5p can target mouse Smo and if miR-324-5p can target both mouse Smo and Gli1 (Figure 4.19). The mouse target sites, which are equivalent to the target sites in human are shown in Figure 4.19C. The miR-125b-5p and miR-324-5p binding sites in the mouse Smo 3'UTR are completely unconserved which, due to the lack of seed binding (Figure 4.19C), suggests it is unlikely Smo is a target of either miR-125b-5p or miR-324-5p. The miR-324-5p binding site in the mouse Gli1 3'UTR is however similar to the human counterpart, although complementarity is reduced at the 3' region of the miRNA. I hypothesised that mouse Gli1 is therefore a target of miR-324-5p. To test this I created a luciferase construct containing the mouse Gli1 3'UTR, however co-transfection of miR-324-5p with this construct did not cause a reduction in the level of luciferase activity, indicating mouse Gli1 is not a target of miR-324-5p (Figure 4.19D). As subsequently expected, miR-324-5p was also unable to alter the level of luciferase following mutation of putative the miR-324-5p binding site (Figure 4.19D). The combination of 3'UTR alignment, target searches and luciferase constructs shows miR-125b-5p and miR-324-5p are unable to directly target the 3'UTRs of Smo and Gli1 in mouse (Figure 4.19), yet they both are regulators of Hh signalling in mouse (Figure 4.14). To understand how miR-125b-5p and miR-324-5p regulates Hh in mouse, additional miR-125b-5p and miR-324-5p targets need to be identified in mouse cell line C3H10T1/2. I hypothesise some of these targets may be involved in Hh signalling.

Figure 4.19 miR-324-5p targets differ in human and mouse

A. Smo

B. Gli1

C.

miR-125b-5p:SMO **non-conserved**

```
human SMO    5'-ACAGGGCCCUGGAGCUCAGGGU -3'
                ||  ::      ||  |||||||
miR-125b-5p  3'-AGUGUUCAAUCCCAGAGUCCCU -5'
                |||  :      |    |:||||
mouse Smo    5'-UCAGGACCCUGCACUUCAGGAC -3'
```

miR-324-5p:SMO **non-conserved**

```
human SMO    5'-ACACCCATTTAGTGGGGGATGGG -3'
                |||||  ||  :  |:|||||||  |
miR-324-5p   3'-UGUGGUUACGGGAUCCCCUACGC -5'
                |||:|  ||  |  |:||||  |
mouse Smo    5'-ACATCCATCCAGTGGGGAGGACC -3'
```

miR-324-5p:GLI1 **conserved**

```
human GLI1   5'-GCACAAGATGCCCCAGGGATGGG -3'
                :|||  |:    ||    |||||||  |
miR-324-5p   3'-UGUGGUUACGGGAUCCCCUACGC -5'
                ::|  |:  |      ||||||  |
mouse Gli1   5'-TACATGAGGTGCCCAGGGATGGG -3'

                ::|  |:  |      ||          |
mouse Gli1 mt5'-TACATGAGGTGCCCAGGctacGG -3'
```

D.

E.

104

Figure 4.19 miR-324-5p targets differ in human and mouse. (A) Alignment of human and mouse Smo 3'UTRs, miR-125b-5p (pink) and miR-324-5p (green) binding sites are highlighted. (B) Alignment of human and mouse Gli1 3'UTRs, miR-324-5p binding sites are highlighted in green. (C) Detailed view of miR-125b-5p and miR-324-5p potential binding sites in Smo and Gli1 3'UTRs. (D) pMIR-Report plasmid containing mouse either wild-type (wt) or mutant binding site (mt) Gli1 3'UTR downstream of luciferase was transfect into C3H10T1/2 cells with either miCon2 or miR-324-5p, data normalised to miCon and plotted as relative luciferase light units, statistical difference were calculated using Students *t* test. Data combined from 5 independent experiment for mutant and wild-type constructs respectively, each n=6. (E) Schematic showing the mechanism miR-125b-5p and miR-324-5p regulate Hh signalling in mouse is unknown.

4.3   Discussion

4.3.1   miRNA screens in OA

This work is based on miR-125b-5p and miR-324-5p being up regulated in OA. To identify miRNAs differentially expressed in OA, Dr. Hannah Elliott (previous laboratory member) screened end-stage OA femoral head cartilage and used NOF (Neck of femur fracture) cartilage as a control. Both end-stage OA and NOF cartilage have previously been used to identify differences between OA and healthy cartilage [217] and were the best available for the study. miR-125b and miR-324-5p may both have the potential to alter developmental pathways, and even the final 'developed' articular cartilage composition. However, using end-stage OA cartilage does not allow the dynamics of miRNA expression during the disease to be determined, for example the presence or absence of a miRNA both before or during OA, may contribute to OA onset and progression and would be missed by this study. In addition by using NOF cartilage as a control I may have inadvertently identified miRNAs differentially expressed during osteoporosis, a bone degrading disease which can lead to fracture of the neck of femur.

To date there are two published screens of miRNA expression on OA cartilage, each observing different miRNA expression profiles to our laboratory's screen (unpublished data and [155]) and from each other [156, 157]. These variations may be due to differences in the severity or stage of OA, different methods of RNA isolation, or differences in detection techniques. Neither miR-125b-5p nor miR-324-5p were found to be differentially expressed in the other screens. In our study miR-125b-5p was significantly increased and miR-324-5p was only detected in OA cartilage (Figure 4.1). The expression of miR-125b-5p was higher than miR-324-5p in the screened cartilage, consistent with deep sequencing data combined from many tissues (Figure 4.2). Deep sequencing in cartilage suggest miR-125b-5p is in the top 2%, and miR-324-5p is in the top 20% of most abundantly expressed miRNAs in cartilage (Prof Ian Clark, University of East Anglia, personal communication). miRNA expression level does not always correlate to potency, how stable a miRNA target is, how sensitive a target is to changes in gene expression, and the number of targets or pseudo targets a miRNA has, are all determinants of miRNA

potency [124, 125]. A miRNA may also act in concert with other miRNAs on the same target or pathway to potentiate their effects [218, 219].

4.3.2 miR-125b-5p and miR-324-5p sequence and evolution

miR-125b is encoded from two loci (miR-125b-1 and miR-125b-2) in mouse and human, the mature miR-125b-5p sequences are identical in both loci but the -3p sequences are not, meaning the number of miR-125b-5p sequences will be the combined number from both loci, whereas the number of miR-125b-3p sequences will be different for each loci. The total expression of miR-125b-5p is far greater than for each of the miR-125b-3p sequences (Figure 4.2). The relative contribution to miR-125b-5p expression from each loci is unknown. A number of primates only have one loci for miR-125b, perhaps indicating redundancy between the loci in humans and mice as the miR-125b-5p mature sequences are identical. Danio rerio (zebra fish) have three loci for miR-125b, perhaps indicating its importance to that species. In humans and mice, where deep sequencing data is available, one might expect little or no function of miR-125b-1-3p and miR-125b-2-3p due to their low expression; deep sequencing data for *Danio rerio* is not available. miR-125b is an ancient miRNA, being present in most animals and suggesting it has a functional relevance in these animals. These functions however may be different as the repertoire of protein coding genes and their 3'UTRs will be different.

miR-324-5p is expressed in fewer species than miR-125b-5p, but due to the rate at which 3'UTRs evolve, may have divergent functions in these species. Deep sequencing data from human suggests the miR-324-3p sequence is different from that annotated in miRBase, resulting in a different seed sequence and implying target prediction sites for miR-324-3p are incorrect (Figure 4.2). It is not surprising that there are differences between miRBase and deep sequencing as this information is often gained from cloning miRNAs from one species with predictive homology searches performed in another. Deep sequencing data for mice illustrates there are both canonical and non-canonical versions of miR-324-3p, each of which will target its own repertoire of genes, perhaps having distinct functions, the relative expression of non-canonical and canonical '-3p' may be tissue specific and dependent on expression of RNA binding proteins which alter miRNA processing. Deep sequencing also suggests miR-324-5p is up to 10 times more abundant than miR-324-3p in

mouse and humans (Figure 4.2). The differences in '-5p' and '-3p' expression may indicate the role of miR-324-5p is more important than that of miR-324-3p.

4.3.3 Predicted targets and functions of miR-125b-5p and miR-324-5p

There is little crossover between databases with only 63, 12, 71 and 15 targets being predicted by all databases for mouse miR-125b-5p, mouse miR-324-5p, human miR-125b-5p and human miR-324-5p respectively (Figure 4.6). Due to the variation in miRNA targets predicted by each database it was not surprising there was also little crossover in the predicted functions of miR-125b-5p and miR-324-5p (Figure 4.7). Luciferase reporter constructs, functional assays and DAVID pathway analysis of a list of genes following over and under expression of miRNAs are required to further investigate and validate these functions.

4.3.4 Validated targets and functions of miR-125b-5p and miR-324-5p

Previous studies have shown miR-125b-5p and miR-324-5p are involved in a number of processes. miR-125b-5p is the more well studied of the two miRNAs, according to miRecords (mirecords.biolead.org/), evidence suggests miR-125b-5p targets 65 different genes including Snail, an important gene for cartilage and Smo [127]. Recently miR-125b has been shown to target important enzymes for cartilage such as ADAMTS-4 [220] and MMP13 [221]. miR-324-5p on the other hand only has two published validated targets; Smo and Gli1 [127].

Ferretti *et al.* used both luciferase constructs and functional studies to show miR-125b-5p and miR-324-5p regulate Hh signalling. The majority of our data confirms the findings by Ferretti *et al.* apart from our observation that miR-125b-5p increases human SMO 3'UTR luciferase (Figure 4.8C). This could suggest that miR-125b-5p may not target Smo or that miR-125b-5p may target other proteins in the cell, in addition to Smo, leading to the observed increase in luciferase, masking any inhibitory effect of mi125b on Smo. Transfection of miR-125b-5p with other luciferase constructs may clarify this. Ferretti *et al,* did not report miR-125b-5p to have this effect on the 3'UTR construct [127], however I used a construct with a CMV promoter and different cells. In agreement with Ferretti *et al.* [127]*,* miR-324-5p targets GLI1 and SMO in human. In contrast to Ferretti *et al*, our data suggests human SMO is not a target of miR-125b-5p. Western blot analysis of Hh response

genes by Ferretti *et al,* show both miR-125b-5p and miR-324-5p are regulators of Hh signalling [127], although our data now casts doubt over the mechanism by which miR-125b-5p acts in human.

4.3.5 Hh signalling in C3H10T1/2

The data in Figure 4.11 shows the effects of inhibiting components the Hh pathway on basal, Ihh induced and Pur induced Hh signalling. The data is described fully in Appendix A. In summary the data suggest the Hh activator Smo is always required for Hh signalling. Ptch inhibits basal levels of Hh signalling but is also required for Pur induced Hh signalling, and Gli1 is required for Ihh and Pur induction of itself, but not of Ptch.

The Hh signalling pathway involves the primary cilium, and during stimulation Smo moves toward it. In contrast to other studies, Ihh stimulation in our experiments did not cause Smo and acetylated α-tubulin co-localisation, but instead does result in Smo localising near the primary cilium, perhaps at its base. Staining for basal body protein such as γ-tubulin may verify this. To our knowledge unlike Shh and Pur, Ihh has not been shown to cause Smo localisation previously, which may suggest that Ihh signalling differs to Shh and Pur.

4.3.6 miR-125b-5p and miR-324-5p in Hh signalling

miR-125b-5p and miR-324-5p both reduce Ihh induced Hh signalling activity (Figure 4.14), although the level of Hh activity remains higher than basal levels, indicating miR-125b-5p and miR-324-5p are not 100% effective in diminishing Hh pathway activity in mouse C3H10T1/2. The addition of miR-125b-5p to Pur (2μM) stimulated cells actually caused an increase in Hh pathway activity (Figure 4.15). The reason for this is unclear, one hypothesis could be it acts in a similar way to the effect it has on the luciferase reporter of Smo; miR-125b-5p may target a transcription inhibitor, pathway inhibitor or protease which allows increased Hh activity. This hypothesis however does not explain why miR-125b-5p is able to reduce Ihh induced Gli1 (Figure 4.14). I therefore hypothesised that abnormally high stimulation from Pur or SAG, results in the Hh pathway regulating itself by turning on a natural, unknown, protective, negative feedback loop and which could itself be sensitive to miR-125b-5p levels, resulting in decreased negative feedback and

therefore an increase in pathway activity (Figure 4.20B). To investigate this C3H10T1/2 cells were stimulated with a wide range of Pur and SAG concentrations. High concentrations of both Pur and SAG are inhibitory to pathway activity indicating the negative feedback loop to protect the cell from high Hh signalling. At the concentration of Pur used generally is this study (2000nM/2μM) the pathway is beginning to be inhibited (Figure 4.15C and E), indicating the negative feedback loop is active. Our hypothesis suggests miR-125b-5p increases Hh activity due to inhibition of the negative feedback loop. In support of this, transfection of miR-125b-5p slightly decreased Hh activity in lower (50nM SAG) stimulated cells, similar to the effect on Ihh stimulation (Figure 4.15I), increased Hh activity in highly (1000nM SAG) stimulated cells (Figure 4.15H), similar to the effect on high Pur stimulation.

The exact mechanism of how high levels of Pur can lead to activation of the regulator mechanism and how the regulatory mechanism can lead to reduced Gli1 is unknown. In Wnt signalling high levels of ligand are predominantly associated with activation of β-catenin and low levels of ligand are predominantly associated with activation of Ca^{2+} signalling, with canonical β-catenin and non-canonical Ca^{2+} signalling being able to reciprocally inhibit each other [102]. A similar mechanism of high and low reciprocal inhibition may exist in Hh signalling.

There was limited evidence for high concentrations of Ihh causing decreased Hh signalling. However preliminary data showed high concentrations (5μg/ml) of Ihh did not cause reduced Gli1 expression but did cause reduced Ptch1 expression (Figure 4.20C and D). This high Ihh induced reduction of Ptch may involve Smo, as Smo mRNA is also reduced with high Ihh concentration. This shows some sort of regulatory mechanism may be active in high Ihh stimulated cells but this mechanism may have differing effects on different Hh response genes.

An alternative possibility is that high concentrations of Pur and SAG are toxic to the cells, with miR-125b-5p somehow preventing this toxicity. Supporting this notation, pathway analysis of predicted miR-125b-5p targets suggest miR-125b-5p may have a possible role as a negative regulator of apoptosis and cell death. The effect of miR-125b-5p on different

Hh stimuli is interesting and warrants further investigation, but will not form the focus of this study.

miR-125b-5p, but not miR-324-5p, reduces the length of the primary cilia, possibly indicating an additional mechanism by which miR-125b-5p may regulate Hh signalling. The mechanism by which miR-125b-5p decreased primary cilia length is unknown, however miR-125b-5p is predicted to target components of the cytoskeleton, which may also have a role in cilia structure organisation.

Gli1 is a widely used readout of active Hh signalling, and is also responsible for the transcriptional activation of a number of genes. Gli2 and Gli3 are however also important regulators of Hh-induced gene expression whose function differs from Gli1 [222]. Activation of Gli1, Gli2 and Gli3 may differ depending on the level of Hh activation [222]. Further work is needed to determine the role miR-324-5p has in the regulation of Gli2 and Gli3. Neither Gli2 nor Gli3 have a miR-324-5p seed binding site in mouse, suggesting any effect miR-324-5p has on them will be directed through miR-324-5p's ability to inhibit Hh. Interestingly, Gli3 is predicted to have a miR-324-5p seed binding site in human, and preliminary data from our lab suggests its expression is also decreased following miR-324-5p overexpression.

Figure 4.20 Role of miR-125b-5p and miR-324-5p in high and low Ihh and purmorphamine Hh signalling in mouse.
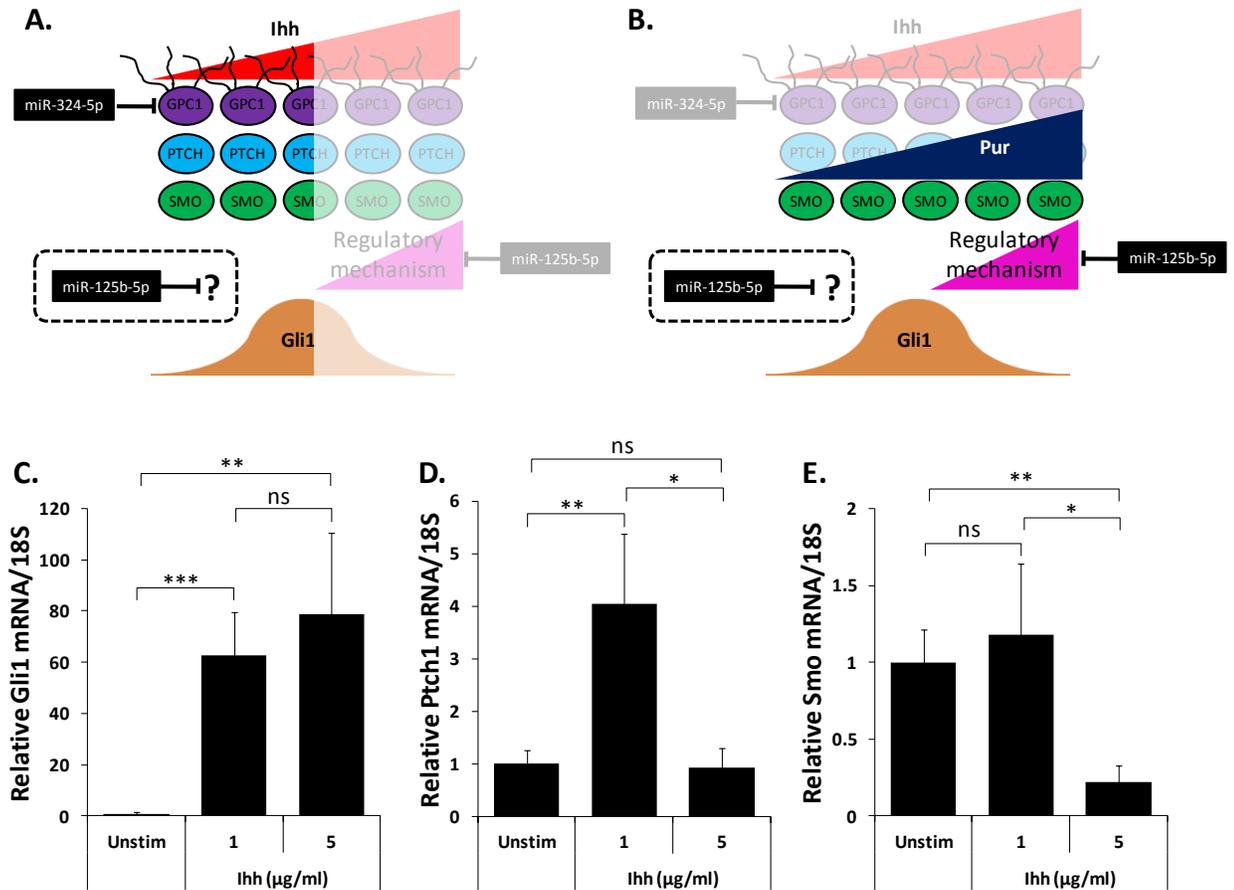
Figure 4.20 Role of miR-125b-5p and miR-324-5p in high and low Ihh and Pur stimulated Hh signalling in mouse. (A) Role of miR-324-5p and miR-125b-5p in Ihh signalling. Non faint; increasing concentration of Ihh leads to increased Hh signalling. miR-324-5p regulates Gpc1 to reduce Ihh induced Hh signalling. miR-125b-5p regulates an unknown target to regulate Ihh induced Hh signalling. Faint; there is limited evidence for high concentrations of Ihh causing decreased Hh signalling. (B) Role of miR-324-5p and miR-125b-5p in Pur induced Hh signalling. High concentrations of Pur cause the activation of a regulatory mechanism which functions to inhibit high Hh signalling. miR-125b-5p regulates and unknown target to decrease low Pur induced Hh signalling. miR-125b-5p also inhibits this regulatory mechanism to allow to higher levels of Hh signalling following high levels of Pur. Faint; Pur acts downstream of Gpc1 meaning miR-324-5p does not alter Pur induced Hh signalling. (C-E) Preliminarily data shows high concentrations (5µg/ml) of Ihh does not cause reduced Gli1 expression but does cause reduced Ptch1 expression, possibly via a mechanism involving Smo. This shows a regulatory mechanism may be active in high Ihh stimulated cells but this mechanism may have differing effects on different Hh response genes. Data from one experiment of n=4. All data are presented as mean + SEM, statistical difference were calculated using ANOVA followed by a Bonferroni post test, where *<0.05, **<0.01, ***<0.001.

4.3.7 miR-125b-5p and miR-324-5p in OA and cartilage

Down regulation of miR-125b-5p and miR-324-5p allows high Hh signalling [127]. Ferretti *et al.* found these miRNAs are increased upon cell differentiation, inhibiting the Hh pathway and therefore holding the Glioma cells in their differentiated state [127]. miR-125b-5p and miR-324-5p are increased in end-stage OA and may be holding chondrocytes in a hypertrophic state. miR-125b-5p and miR-324-5p may have been decreased at the inception of OA allowing high Hh and promoting a change in chondrocyte phenotype from a prehypertrophic type II collagen secreting cell into a hypertrophic type X collagen secreting chondrocyte.

Interestingly, in end stage OA cartilage there is an increased expression of Ihh (high Hh signalling), and increased expression of miR-125b-5p and miR-324-5p. I predict miR-125b-5p and miR-324-5p are up regulated in OA as part of an ineffective protective mechanism to decrease Hh pathway activity (Figure 4.21). In this study I show miR-125b-5p and miR-324-5p regulate alkaline phosphatase, a marker of bone formation and although miR-324-5p inhibited alcian blue in initial experiments, there were both increases and decreases in subsequent experiments indicating more work is needed to understand the role, if any, miR-324-5p has on chondrogenesis. Interestingly, miR-324-5p regulates Pur and Bmp induced alkaline phosphatase (Figure 4.16), yet miR-324-5p does not consistently regulate Pur induced Hh signalling (Figure 4.15), suggesting miR-324-5p might also play a role in Bmp signalling. Interestingly, miR-125b has been shown to regulate Smad4 [223]. These results indicate miR-125b-5p and miR-324-5p may play a role during development and OA.

Retinoic acid is the oxidized form of vitamin A and can alter mi125b and miR-324-5p expression (Figure 4.18) and has previously been shown to have a role in cartilage. This suggests, retinoic acid may work in part by altering miR-125b-5p and miR-324-5p expression to affect Hh pathway activity and alter chondrocyte hypertrophy and eventual OA. I have shown RA to inhibit Hh signalling and the expression of miR-125b-5p and miR-324-5p (Figure 4.18). This seems contradictory as one would expect deceased miR-125b-5p and miR-324-5p to allow higher Hh signalling, suggesting as well as inhibiting active Hh signalling, miR-125b-5p and miR-324-5p may also be required for active Hh signalling.

Direct inhibition of miR-125b-5p and miR-324-5p will shed light on this. Further work is also need to show what induces miR-125b-5p and miR-324-5p expression during OA.

Figure 4.21 miR-125b-5p and miR-324-5p may play a role in OA pathogenesis.

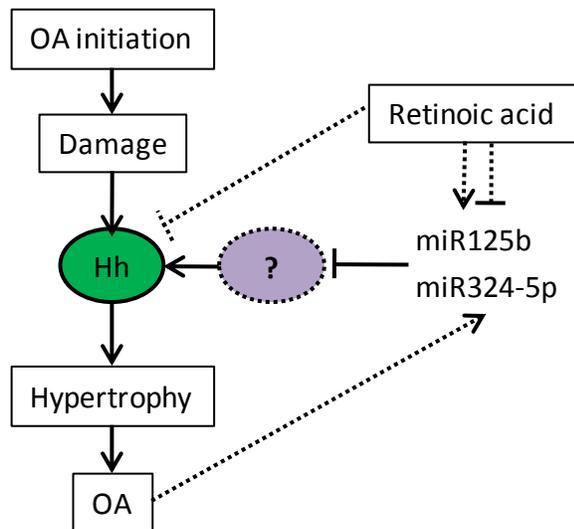Figure 4.21 miR-125b-5p and miR-324-5p may play a role in OA pathogenesis. A schematic showing miR-125b-5p and miR-324-5p regulate the Hh pathway, to control OA pathogenesis. The mouse cells the mechanism by which miR-125b-5p and miR-324-5p regulate Hh signalling is not known. Retinoic acid may attenuate miR-125b-5p and miR-324-5p expression levels to regulate Hh signalling and cartilage metabolism.

4.3.8   Different targets in human and mice

3'UTR alignment, target searches and luciferase studies (Figure 4.19) suggest the mechanism by which miR-125b-5p and miR-324-5p regulate Hh signalling in human and mouse may differ. There is evidence that suggests miR-125b-5p and miR-324-5p act up stream of Smo in the mouse Hh pathway, and therefore not on Gli1 as in the human Hh pathway. Firstly miR-125b-5p and miR-324-5p are able to inhibit Smo translocation around the cilia. Secondly, miR-324-5p does not consistently alter Pur induced Hh signalling. Identification of mouse miR-125b-5p and miR-324-5p targets may show the mechanism by which miR-125b-5p and miR-324-5p regulate Hh in mouse differs from human.

4.4   Summary

miR-125b-5p and miR-324-5p are increased in OA (Figure 4.1). They are both the predominantly expressed miRNA from their stem-loops (Figure 4.2) and are both conserved in many species (Figure 4.3). Target prediction databases predict many targets (Figure 4.6) and functions (Figure 4.7) of miR-125b-5p and miR-324-5p although there is little consistency between the target prediction databases. In agreement with Ferretti *et al* [127], I show miR-324-5p can bind and regulate both the Hh transcription factor GLI1 and the Hh activator SMO in human (Figure 4.8A and B). In contrast with Ferretti *et al.* I show miR-125b-5p increases human SMO luciferase (Figure 4.8C). I do however show miR-125b-5p and miR-324-5p are both important regulators of Hh signalling (Figure 4.14) and osteogenesis (Figure 4.16). RA alters miR-125b-5p and miR-324-5p expression and may form a possible mechanism to alter miR-125b-5p and miR-324-5p expression *in vivo* (Figure 4.18). In contrast to suggestions by Ferretti *et al.* [127], our results indicate miR-125b-5p and miR-324-5p do not regulate mouse Smo and Gli1 (Figure 4.19). To understand how miR-125b-5p and miR-324-5p regulates Hh in mouse, additional miR-125b-5p and miR-324-5p targets need to be identified in mouse cell line C3H10T1/2. I hypothesise some of these targets may be involved in Hh signalling.

Chapter 5 Identification of novel miR-324-5p targets

5.1   Introduction

To elucidate the mechanism by which miR-324-5p regulates Hh in mouse I wanted to identify direct downstream targets of miR-324-5p which are required for Hh signalling. In addition to online target prediction websites (Chapter 4) there are a number of different approaches that can be used to identify miRNA targets, including proteome profiling [224], transcriptome profiling [225] and pull-down of various members of the miRNA-RISC-mRNA complex [226, 227]. The advantage of experimental target identification is it avoids the inherent bias of computational target prediction. miRNAs function by inhibiting translation and/or causing mRNA degradation, leading to reduced protein and/or mRNA expression, because of this miRNA targets can be found by high-throughput analysis of mRNA (transcriptome microarrays or high-throughput sequencing (HT-seq)) and protein (proteomics) following miRNA overexpression [224, 225].

Transcriptome profiling can be used to quantify the relative expression of mRNA transcripts and has previously been used to identify miRNA targets [225]. In initial experiments, miRNAs were overexpressed and a significant enrichment of miRNA seed binding sites (of the miRNAs that was overexpressed) was found in mRNAs whose expression decreased, suggesting these mRNAs were targets [119]. In those studies the miRNAs were increased above physiological levels and would arguably lead to the miss-identification of some targets which are not relevant at miRNA physiological levels. To avoid this some studies have used antisense oligonucleotides (antagomirs) to inhibit the action of the miRNA and observed increased expression of target mRNAs [228]. The increased expression of target mRNA following miRNA inhibition is often small, perhaps limiting the usefulness of miRNA inhibitors in target identification experiments. Advances in next generation sequencing techniques means RNA sequencing is likely to replace transcriptome microarrays in these types of experiments [229], their higher sensitivity may also be useful for miRNA inhibitor experiments. A number of miRNA targets have been shown to be decreased at a protein level but not an mRNA level [121, 230], transcriptome profiling will fail to identify these targets. Proteomic profiling however is able to identify this type of target.

There are a number of different types of quantitative proteomics. Stable Isotope Labelling with Amino acids in Cell culture (SILAC) proteomics has the ability to compare two populations of cells in the same mass spectrometry (MS) experiment and has previously been used to screen for miRNA targets [230]. Briefly, SILAC involves labelling one population of cells with heavy amino acids (Lysine and Arginine) and leaving the other population of cells unlabelled, the populations are then treated differently (transfection of miRNA of interest or control miRNA in this case), lysed, mixed at a ratio of 1:1, then analysed using MS. In order to achieve SILAC label incorporation cells need to undergo cell division. C3H10T1/2 divide relatively quickly, undergoing population doubling around every 18h [231], meaning cells can be readily labelled. Types of SILAC include 'traditional', 'pulse', 'super' and 'in vivo'. Traditional SILAC is as described above and has been used for miR-1 target identification [224]. In pulse-SILAC, cells are only transferred to labelled medium once treated or transfected, meaning it identifies differences in newly synthesized proteins, this technique has been used to identify miR-34a targets [232] and a slight variation used to detect miR-223 targets from a miR-233 knock-out mouse [233]. Super-SILAC involves pooling a number of heavy labelled cell-lines to investigate non dividing cells [234]. In-vivo-SILAC (the SILAC mouse) is now being developed and may prove useful for studying the effects of miRNAs *in vivo* experiments [235].

Ribosomal mRNA is an indicator of the level of translation [236]. Guo *et al.,* analysed the amount of mRNA bound to ribosomes following miRNA overexpression to determine the effect a miRNA has on translation [121]. They found an enrichment of miRNA seed binding sites in genes whose translation was decreased, however they also used RNA sequencing of total mRNA and found almost all genes whose translation was decreased also had decreased total mRNA, suggesting for the majority of targets the miRNA functions to destabilise the mRNA rather than inhibit translation [121].

miRNAs target mRNAs through the action of a miRNA-RISC-mRNA complex, which is made up of many proteins including AGO proteins [237]. Pull-down of components of this complex followed by profiling of mRNA can identify direct miRNA targets. miRNA-mRNA complexes have been identified following the pull-down of labelled miRNAs, however initially the technique produced controversial results as there was no enrichment of

miRNA binding sites in the 3'UTRs [238]. A similar technique called LAMP (Labelled miRNA pull-down assay system) can be used to pull-down labelled miRNAs from an *in vitro* setting after being mixed with cell lysates [239]. Another miRNA pull-down called miR-TRAP assay involves cross linking with UV radiation to generate a covalent bond, allowing for a more stringent pull-down of miRNA targets, but involves pull-down from an *in vitro* setting [226]. The reverse can also be performed to identify miRNAs which target a specific mRNA, by pull-down of a labelled transcript and identifying the targeting miRNAs [227].

It is also possible to pull-down protein members of the RISC complex such as Ago1 [237] [240] and Ago2 [241] and profile the mRNA to detect targets. In 2009 Chi *et al.* published a miRNA target identification method where they cross-linked the mRNA and protein, meaning the interactions are stronger and more readily survive the IP process, this method is called HITS-CLIP (High-throughput sequencing of RNAs isolated by crosslinking immunoprecipitation) [242]. A modification of the HITS-CLIP method has allowed more precise mapping of miRNA-target interactions, this modified method is called PAR-CLIP (Photoactivatable-Ribonucleoside-Enhanced crosslinking and immunoprecipitation)[243]. HITS-CLIP and PAR-CLIP are both technically challenging and require a large amount of bioinformatic analysis to map the miRNA-RISC-mRNA interactions.

Target identification strategies can feedback to further define properties of miRNA-target interactions. Previous target identification publications have shown that genes which have a conserved miRNA seed sequence binding site, are more likely to be miRNA targets [130]. In addition, the more base pairing that there is between the miRNA seed and its target, the more likely a gene is to be a miRNA target [119]. There are four different types of binding at the seed; 8mers, 7mer m8s, 7mer a1s and compensatory binding (Figure 5.1) [129]. 8mer target sites are an exact reverse complement match to positions 2-8 of the miRNA followed by an 'A' at position 1. 7mer m8 target sites also have an exact reverse complement match to positions 2-8, but do not have an 'A' at position 1. 7mer A1 sites have an exact reverse complement match to positions 2-7 of the miRNA, followed by an 'A' at position 1. Compensatory binding is where seed miss matches are compensated for by additional binding at the 3' end of the miRNA.

Here I will develop a strategy using proteome and transcriptome profiling to identify miRNA targets and use it to identify miR-324-5p targets required for Hh signalling. The main aim of this chapter is to identify miR-324-5p targets in mouse which have a possible role in Hh signalling.

Specific aims

Aim1: Determine if online target prediction programs identify mouse miR-324-5p targets involved in Hh signalling.

Aim2: Develop a strategy to identify novel miRNA targets.

Aim 3: Identify potential miR-324-5p targets which have a role in Hh signalling.

Aim 4: Validate potential miR-324-5p targets identified in Aim3.

Figure 5.1 Different types of seed

A. 8mer seed match

```
                                      87654321
miR-324-5p    3'-UGUGGUUACGGGAUCCCCUACGC -5'
                              ||||||||
mRNA          5'-TACAACTGTAGACCAGGGATGCA -3'
```

B. 7mer m8 seed match

```
                                      87654321
miR-324-5p    3'-UGUGGUUACGGGAUCCCCUACGC -5'
                              |||||||
mRNA          5'-TACAACTGTAGACCAGGGATGCC -3'
```

C. 7mer a1 seed match

```
                                      87654321
miR-324-5p    3'-UGUGGUUACGGGAUCCCCUACGC -5'
                               ||||||
mRNA          5'-TACAACTGTAGACCACGGATGCA -3'
```

D. 3' Compensatory

```
                                      87654321
miR-324-5p    3'-UGUGGUUACGGGAUCCCCUACGC -5'
                     |||||||      ||  ||||
mRNA          5'-TAACCAATGAGACCAGGGATGCC -3'
```

Figure 5.1 Different types of seed. miR-324-5p and hypothetical mRNA targets to illustrate different types of miRNA targets. Complementary bases are shown in red. Seed region is underlined and the first 8 bases of the miRNA are indicated by the numbers. (A) 8mer seed match. (B) 7mer m8 seed match. (C) 7mer a1 seed match. (D) 3'compensatory seed match.

5.2 Results

5.2.1 Online target prediction programs predict targets involved in Hh signalling

3,061 different genes were predicted to be miR-324-5p targets by the four main online prediction programs analysed in Chapter 4, 11 of which are part of the Kegg Hh pathway (Figure 5.2A), representing 20.4% of the 54 genes listed in the Kegg Hh pathway, which is slightly higher than the percentage of all genes predicted to be miR-324-5p targets, ~16.4% (3,061/~18,615) (p=0.44 fishers). Four genes were predicted to be a miR-324-5p targets by two or more databases and are part of the Kegg Hh pathway (Figure 5.2A) (Wnt7a, Csnk1g1, Rab23 and Bmp5). No targets predicted by three or more databases are part of the Kegg Hh pathway (Figure 5.2A). Due to the large number of predicted miR-324-5p targets it is not surprising some predicted miR-324-5p targets are part of the Kegg Hh pathway (Figure 5.2). The large variation (outlined in Chapter 4), and unreliability of prediction [244], means it is unlikely all of these predicted miR-324-5p targets are real and due to the lack of comprehensiveness of the Hh kegg pathway it is likely a number of other predicted miR-324-5p targets may have a role in Hh. I therefore decided to develop a new strategy to identify miR-324-5p targets and determine if any are required for Hh signalling.

Figure 5.2 miR-324-5p is predicted to target parts of the Hh pathway

A.

	PicTar (all)	TargetScan	DIANA-micorT	miRanda (all)
#	0	5	6	4
		Csnk1g1	Bmp5	Bmp5
		Dhh	Bmp8b	Csnk1g2
		Gsk3b	Csnk1g1	Gli1
		Rab23	Rab23	Gli3
		Wnt7a	Wnt7a	
			Wnt9b	
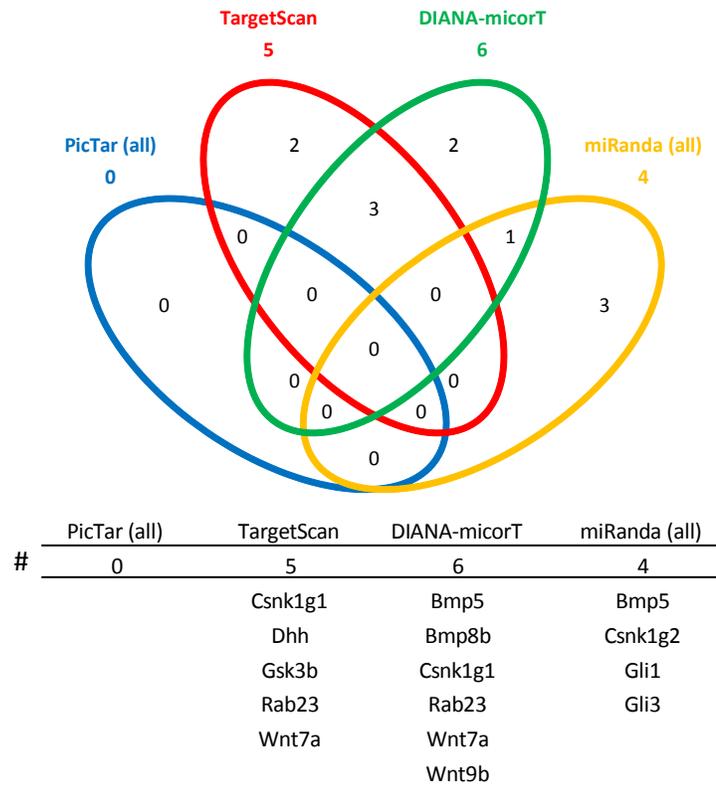
B.
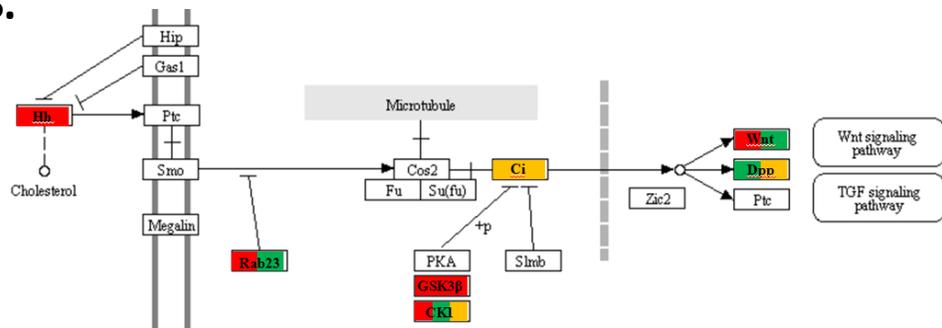
Figure 5.2 miR-324-5p is predicted to target components of the Hh pathway. (A) 11 genes in the Hh Kegg pathway are predicted to be miR-324-5p targets by at least one target prediction program. 4 genes in the Hh Kegg pathway are predicted by two prediction programs. (B) Schematic showing the position of predicted targets in the Kegg Hh pathway. Colour indicates which prediction program predicted the target. Gene names shown are from the Kegg Hh pathway and not necessarily the exact predicted genes. All predicted genes are in the Kegg pathway list of genes (http://www.kegg.jp/) and are orthologs of the genes shown.

5.2.2 Development of a strategy to identify novel miR-324-5p targets

Overexpression of miRNAs followed by transcriptome and proteome profiling has previously been used to screen for miRNA targets [225]. Our eventual goal is to identify direct miR-324-5p targets in Ihh stimulated cells. However as an initial proof of concept experiment I decided to identify miR-324-5p targets in unstimulated C3H10T1/2 cells to avoid the potent effect miR-324-5p has on Ihh signalling from masking direct miR-324-5p targets.

5.2.2.1 Preparation of SILAC cells

I decided to use SILAC to identify proteins whose expression decreased following miR-324-5p overexpression. SILAC relies on the heavy cells being 100% labelled with heavy amino acids. To achieve this I passaged C3H10T1/2 cells 4 times at a split ratio of 1:5, which is the equivalent to >8 cell divisions and should give >99% incorporation of heavy amino acids (Figure 5.3). To check label incorporation I analysed heavy cells using MS, >99% of the peptides were heavy labelled. Unlabelled cells are referred to as 'light'.

Figure 5.3 Preparation of SILAC cells

Normal C3H10T1/2 cells

1/5 split ↘ ↙ 1/5 split

| Light media | | Heavy media |
| ($^{12}C_6$-Lys and $^{12}C_6$ $^{14}N_4$-Arg) | | ($^{13}C_6$-Lys and $^{13}C_6$ $^{15}N_4$-Arg) |

2-3 days (~2-3, 80%) 2-3days

1/5 split ↓   ↓ 1/5 split

2-3 days (~4-5, 96%) 2-3 days

1/5 split ↓   ↓ 1/5 split

2-3 days (~7, 99.2%) 2-3 days

1/5 split ↓   ↓ 1/5 split

2-3 days (~9-10, 99.8%) 2-3 days*

1/5 split ↓   ↓ 1/5 split

2-3 days (~11-12, >99.9%) 2-3 days

Predicted

*=Label incorporation check
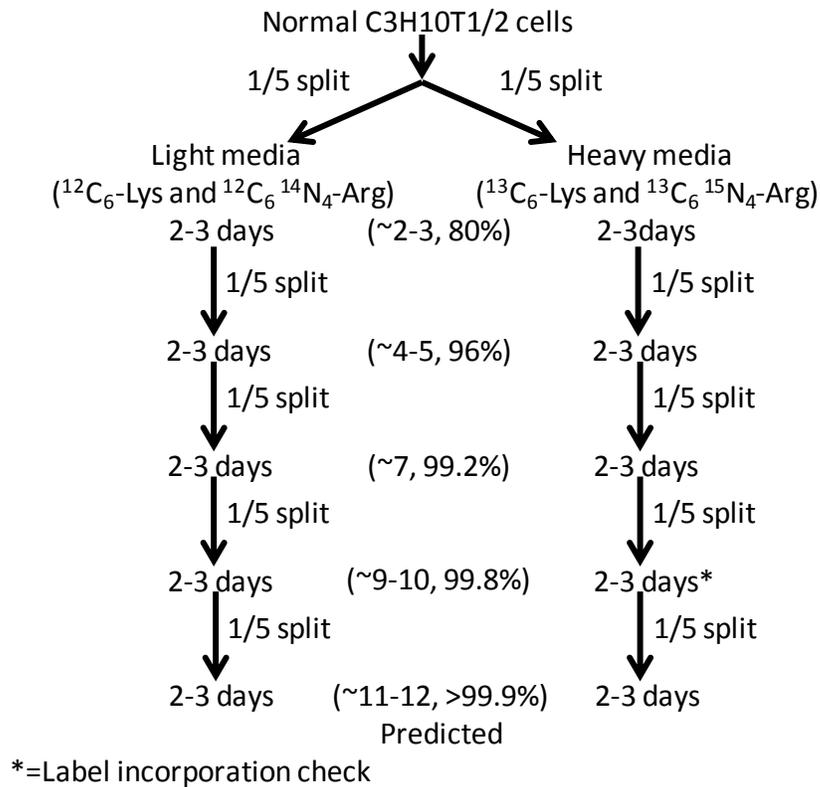
Figure 5.3 Schematic showing labelling of C3H10T1/2 cells with heavy Arginine ($^{13}C_6$ $^{15}N_4$-Arg) and heavy Lysine ($^{13}C_6$-Lys) or light Arginine ($^{12}C_6$ $^{14}N_4$-Arg) and light Lysine ($^{12}C_6$-Lys). C3H10T1/2 cells were maintained in normal DMEM medium and then split into either heavy or light SILAC DMEM at a ratio of 1:5. Cells were then split at a ratio of 1:5 every 2-3 days, 3 more times before label incorporation was confirmed to be >99%. Cells were maintained in their respective SILAC DMEM medium until required, passaging every 2-3 days at a ratio of 1:5. Approximate number of cell doublings and predicted heavy amino acid incorporation are shown in parentheses.

5.2.2.2 SILAC 1

In the first SILAC experiment (SILAC1), 'heavy' labelled mouse C3H10T1/2 cells were transfected with mature miR-324-5p and 'light' C3H10T1/2 were transfected with miCon (see materials and methods) (Figure 5.4A). Following data processing (see materials and methods), there were 1,735 proteins identified in SILAC1. Protein expression data was calculated from the ratio of miR-324-5p (heavy) to miCon (light) and converted into a log2 value. Of the 1,735 proteins identified, 249 (14.4%) decreased, 159 (9.2%) increased, by a fold change of log2 0.2 and 1327 (76.5%) remained unchanged (Figure 5.4B) (From this point onwards a gene, protein or mRNA transcript will be referred to as increased when the fold change in expression is increased by >log2 0.2 and decreased when its expression is decreased by >log2 0.2, all other genes, proteins or mRNA transcripts will be referred to as unchanged). Log2 0.2 is equivalent to a 1.15 fold change, (see Appendix C for why a log2 0.2 cut off was chosen). 183 (10.5%) of the 1,735 genes whose protein was identified contain a miR-324-5p binding site in their 3'UTR, of which a greater number decreased than increased, 35 (19.1%) and 16 (8.7%) respectively (Figure 5.4C). 132 (72.1%) proteins whose genes contain a miR-324-5p seed binding site in their 3'UTR remained unchanged in expression (Figure 5.4C). A larger (1.48 fold, p=0.051) fraction of proteins whose expression decreased (35/249=14.0%) have a miR-324-5p seed binding site in their 3'UTR, than proteins whose expression either remained unchanged (132/1327=9.9%) or increased (16/159=10.0%) (Figure 5.4D).

Of the 183 proteins identified whose genes 3'UTR contained a miR-324-5p seed binding site, 9 contained a conserved miR-324-5p seed binding site, a greater number of which decreased (three) than increased (one), five remained unchanged. The fraction of genes containing a conserved miR-324-5p seed binding site, was 3.01 fold higher in genes whose protein expression decreased (3/249), compared to genes whose protein either remained unchanged (5/1,327) or increased (1/159) (p=0.103). Additionally both 'cumulative fraction' (Figure 5.4E) and 'normalised fraction' (Figure 5.4F) plots showed a greater enrichment of conserved miR-324-5p seeds than non-conserved miR-324-5p seeds, in proteins whose expression decreased. See Appendix B for how 'cumulative fraction' and 'normalised fraction' plots were made.

Of the 183 identified proteins whose gene contained a miR-324-5p seed binding site in their 3'UTR, 34 contained an 8mer, 70 contained a 7mer m8 and 92 contained a 7mer A1 miR-324-5p seed binding site. The fraction of genes which contained an 8mer, 7mer m8 and 7mer A1 type of miR-324-5p seed binding site is 1.86, 1.52 and 1.27 fold higher in genes whose protein expression decreased, compared to genes whose protein either remained unchanged or increased (p=0.123, p=0.169 and p=0.393) respectively. The cumulative fraction plot (Figure 5.4G) illustrates this. The normalised fraction plot (Figure 5.4H) illustrates an enrichment of 8mer and 7mer m8 seeds, but not 7merA1 seeds.

Figure 5.4 SILAC 1
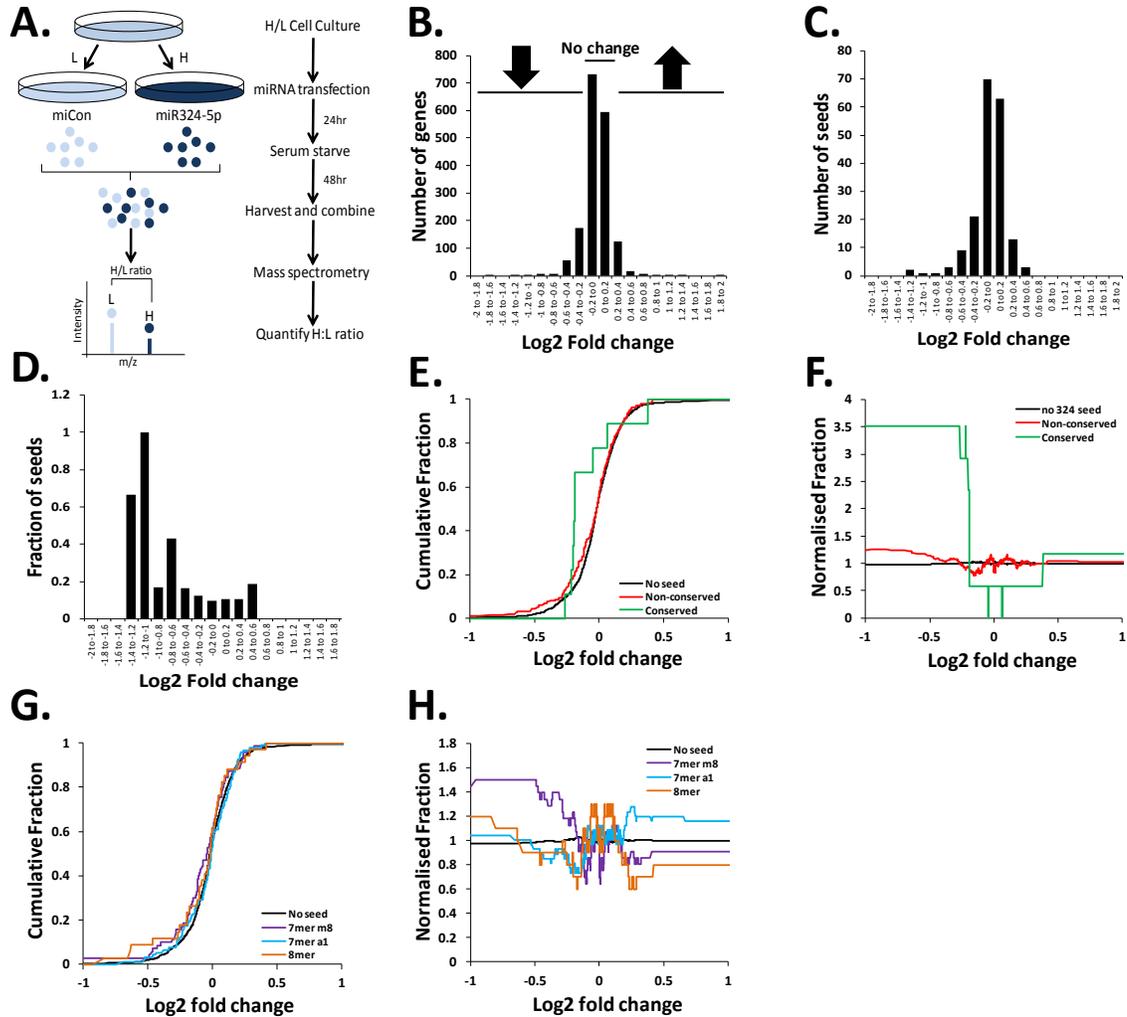
Figure 5.4 SILAC 1. miR-324-5p targets are decreased at a protein level in miR-324-5p transfected heavy unstimulated cells. (A) Schematic showing experimental approach for SILAC 1. (B) Histogram showing expression changes of 1,735 proteins following miR-324-5p transfection; 249 decreased, 159 increased by a fold change of log2 0.2 and 1327 remained unchanged. (C) 183 of the 1,735 genes whose protein was identified contain a miR-324-5p binding site in their 3'UTR; 35 decreased, 16 increased and 132 did not alter by a fold change of log2 0.2. (D) Fraction of genes whose protein was identified which contain a miR-324-5p seed binding site; 14.1%, 10.0% and 9.9% of the genes whose protein expression decreased, increased and did not alter by a fold change of log2 0.2 respectively, contain a miR-324-5p binding site. The fraction of genes containing a miR-324-5p binding site is 1.48 fold higher in genes whose protein expression decreased compared to genes whose protein either remained unchanged in expression or increased (p=0.051). (E) Cumulative fraction of conserved and non conserved miR-324-5p seed binding sites. Of the 183 seed binding sites 9 are conserved, 3 of which are decreased. (F) Normalised fraction plot of conserved and non conserved miR-324-5p seed binding sites. (G) Cumulative fraction of 8mer, 7mer m8 and 7mer a1 miR-324-5p seed binding sites. (H) Normalised fraction plot of 8mer, 7mer m8 and 7mer a1 miR-324-5p seed binding sites.

5.2.2.3 SILAC2

To eliminate any effect of heavy amino acids on protein expression, a second SILAC experiment (SILAC2) was performed, where the labels were reversed, 'light' cells were transfected with miR-324-5p and the 'heavy' cells with miCon (Figure 5.5A). There were 1457 proteins identified in SILAC2. In SILAC2, protein expression data was calculated from the ratio of miR-324-5p (light) to miCon (heavy), then converted into a log2 fold change in expression. Similar to SILAC1, the majority of proteins remained unchanged in expression (827, 56.7%) and a larger number decreased (423, 29.0%), than increased (207, 14.2%) (Figure 5.5B). Interestingly, there are more genes whose protein expression either decreased or increased in SILAC2 (423+207) than in SILAC1 (249+159). 114 of the 1,457 genes (7.8%) whose protein was identified in SILAC2, contain a miR-324-5p binding site in their 3'UTR, of which a larger number of proteins whose genes contain a miR-324-5p seed binding site were decreased (56 or 13.2%), than were increased (11 or 5.3%) (Figure 5.5C). Additionally the fraction of genes containing a miR-324-5p seed binding site was 2.57 fold higher in genes whose protein expression decreased 13.2% (56/423), compared to genes whose protein either increased 5.3% (11/207) or remained unchanged in expression 5.7% (47/827) (p=8.55x10^{-7}) (Figure 5.5D).

Of the 114 proteins whose genes contained a miR-324-5p seed binding site, 6 contained a conserved miR-324-5p binding site. Of these 4 decreased, 2 remained unchanged and none increased, meaning as in SILAC1 there is also a stronger enrichment of conserved miR-324-5p targets than non-conserved miR-324-5p targets, in proteins whose expression decreased. The fraction of genes which contained a conserved miR-324-5p seed binding site was 4.93 fold higher in genes whose protein expression decreased, compared to genes whose protein either remained unchanged or increased in expression (p=0.042) (Figure 5.5E and F).

Of the 114 genes whose protein was identified in SILAC2 and contained a miR-324-5p binding site in their 3'UTR, 20, 50 and 55 of the genes contained an 8mer, 7mer m8 and 7mer A1 miR-324-5p binding site respectively. The fraction of genes which contained an 8mer, 7mer m8 and 7mer A1 miR-324-5p seed binding site was 2.02, 3.87 and 2.10 fold

higher in genes whose protein expression decreased compared to genes whose protein either remained unchanged or increased (p=0.113, p=9.15X10^{-7} and p=0.006) respectively. Suggesting 7mer m8 sites were more enriched than 8mer and 7mer A1 sites in genes whose protein expression decreased by >log2 0.2 (Figure 5.5G and H).

Figure 5.5 SILAC2
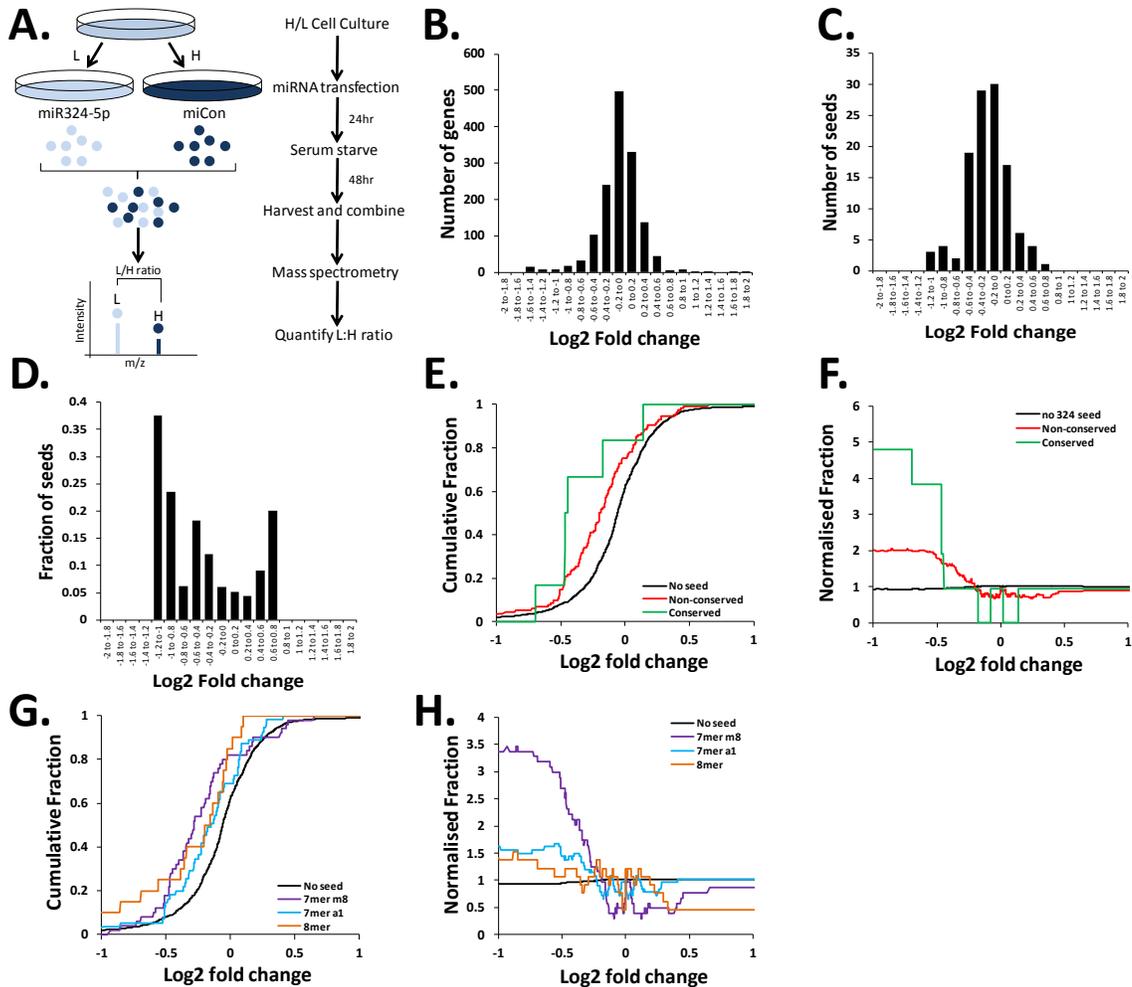
Figure 5.5 SILAC 2. miR-324-5p targets are decreased at a protein level in miR-324-5p transfected light unstimulated cells. (A) Schematic showing experimental approach for SILAC 2. (B) Histogram showing expression changes of 1,457 proteins following miR-324-5p transfection; 423 decreased, 207 increased by a fold change of log2 0.2 and 827 remained unchanged (C) 114 of the 1,457 genes whose protein was identified contain a miR-324-5p binding site in their 3'UTR; 56 decreased, 11 increased and 47 did not alter by a fold change of log2 0.2. (D) Fraction of genes whose protein was identified which contain a miR-324-5p seed binding site; 13.2%, 5.3% and 5.7% of the genes whose protein expression decreased, increased and did not alter by a fold change of log2 0.2 respectively, contain a miR-324-5p binding site. The fraction of genes containing a miR-324-5p binding site is 2.57 fold higher in genes whose protein expression decreased compared to genes whose protein either remained unchanged in expression or increased (p=8.55x10[-7]). (E) Cumulative fraction of conserved and non conserved miR-324-5p seed binding sites. Of the 114 seed binding sites 6 are conserved, 4 of which are decreased. (F) Normalised fraction plot of conserved and non conserved miR-324-5p seed binding sites. (G) Cumulative fraction of 8mer, 7mer m8 and 7mer a1 miR-324-5p seed binding sites. (H) Normalised fraction plot of 8mer, 7mer m8 and 7mer a1 miR-324-5p seed binding sites.

5.2.2.4 Combine SILAC1 and SILAC2

There were 1,227 proteins identified in both SILAC1 and SILAC2 experiments. Figure 5.6A shows data from SILAC1 correlates with data from SILAC2 (R2=0.4715). For example, 79.4% of proteins decreased in SILAC1 and identified in SILAC2, were also decreased in SILAC2 with the addition of miR-324-5p (104/131) (Figure 5.6A). The average of both SILAC1 and SILAC2 was used to perform the seed analysis shown in Figure 5.6B-H. Following miR-324-5p transfection the majority (920/1227=75.0%) of proteins remained unchanged and a greater number decreased (198/1227=16.1%) than increased (109/1227=8.9%) (Figure 5.6B). 93 of the 1,227 genes (7.6%) whose protein was identified in SILAC1 and SILAC2 contained a miR-324-5p binding site in their 3'UTR. A larger number of proteins whose gene contained a miR-324-5p binding site were decreased (30/93=32.3%) than were increased (8/93=8.6%) (Figure 5.6C). The fraction of genes which contained a miR-324-5p seed binding site was a 2.74 fold higher in genes whose protein expression decreased (30/198=15.2%) compared to genes whose protein either increased (8/109=7.3%) or remained unchanged in expression (55/920=6.0%), (total not decreased; 63/1029=6.1%) (p=1.10x10^{-5}) (Figure 5.6D).

5 of the 93 proteins whose gene contained a miR-324-5p seed binding site, contained a conserved miR-324-5p seed binding site. 3 of these genes decreased, only 1 remained unchanged and only 1 increased, meaning there was a greater enrichment of conserved miR-324-5p targets than non conserved miR-324-5p targets in proteins whose expression decreased. The fraction of genes which contained a conserved miR-324-5p seed binding site was 7.9 fold higher in genes whose protein expression decreased (3/198) compared to genes whose protein either remained unchanged or increased in expression (2/1029) (p=0.008) (Figure 5.6E and F).

Of the 93 genes whose proteins were identified in both SILAC1 and SILAC2 and contained a miR-324-5p binding site in their 3'UTR, 17, 38 and 46 contained an 8mer, 7mer m8 and 7mer A1 miR-324-5p binding site respectively. The fraction of genes which contained an 8mer, 7mer m8 and 7mer A1 miR-324-5p seed binding site is 3.73, 4.51 and 1.67 fold higher in genes whose protein expression decreased compared to genes whose protein

either remained unchanged in expression or increased, p=0.005, p=1.12×10^{-6} and p=0.144 respectively, suggesting 7mer m8 sites are more enriched than both 8mer and 7mer A1 sites, and 8mer sites are more enriched than 7mer A1 sites in genes whose protein expression decreased by >log2 0.2 (Figure 5.6G and H). Although there were some differences between SILAC1 and SILAC2, the similarities suggest there was a miR-324-5p transfection dependent enrichment of decreased proteins whose genes contain miR-324-5p seed binding site.

Figure 5.6 SILAC1 and SILAC2 correlate
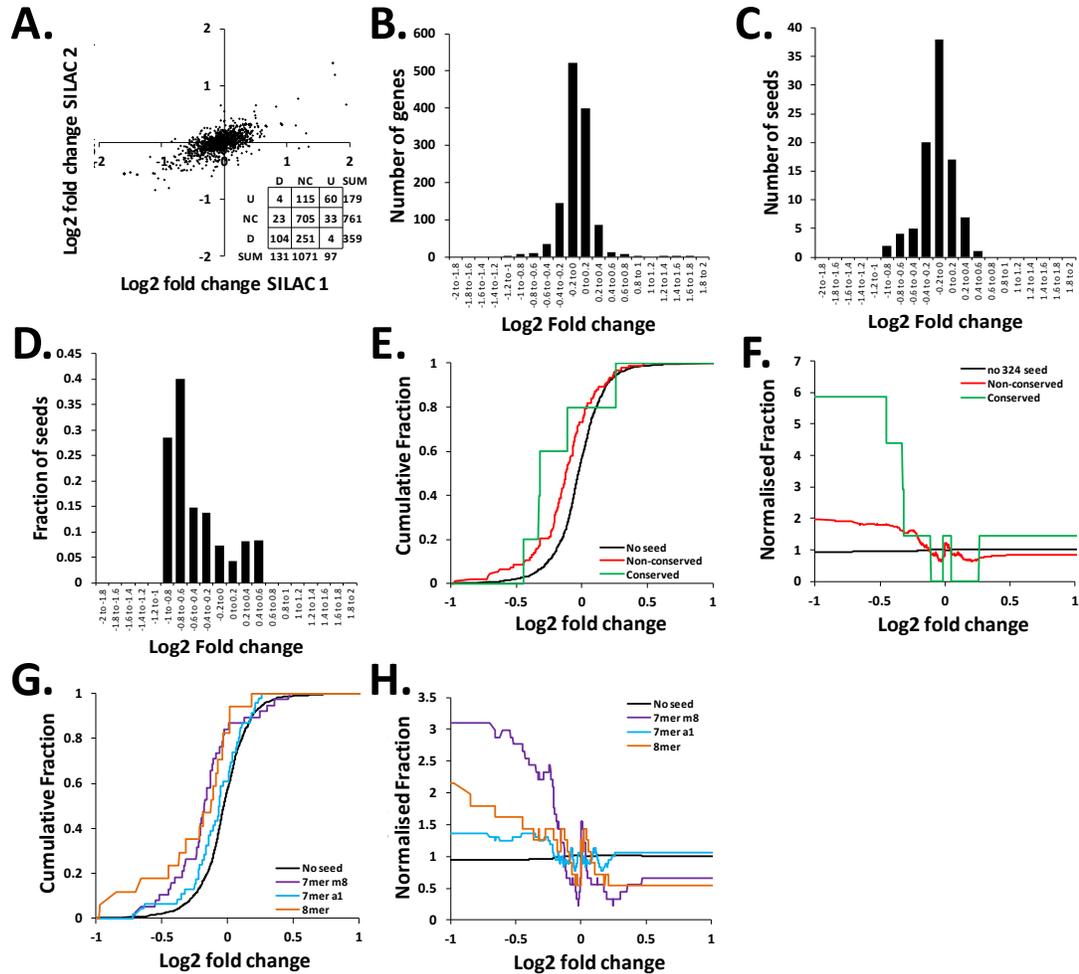
Figure 5.6 SILAC 1 and SILAC 2 combined. (A) SILAC 1 data correlates with SILAC 2 data (R2 =0.4715). Inset; numbers of proteins increased 'U', decreased 'D' and not changed 'NC' by log2 0.2 in SILAC1 (columns) and SILAC 2 (rows). (B) Histogram showing average expression changes of 1,227 proteins identified in SILAC 1 and SILAC 2 following miR-324-5p transfection; 198 decreased, 109 increased by a fold change of log2 0.2 and 920 remained unchanged (C) 93 of the 1,227 genes whose protein was identified contain a miR-324-5p binding site in their 3'UTR; 30 decreased and 8 increased by a fold change of log2 0.2, 55 remained unchanged. (D) Fraction of genes which contain a miR-324-5p seed binding site; 15.2%, 7.3% and 6.0% of the genes whose protein expression decreased, increased and remained unchanged respectively, contain a miR-324-5p binding site. The fraction of genes containing a miR-324-5p binding site is 2.74 fold higher in genes whose protein expression decreased compared to genes whose protein either remained unchanged in expression or increased (p=1.10x10^{-5}). (E) Cumulative fraction of conserved and non conserved miR-324-5p seed binding sites. Of the 93 seed binding sites 5 are conserved, 3 of which are decreased. (F) Normalised fraction plot of conserved and non-conserved miR-324-5p seed binding sites. (G) Cumulative fraction of 8mer, 7mer m8 and 7mer a1 miR-324-5p seed binding sites. (H) Normalised fraction plot of 8mer, 7mer m8 and 7mer a1 miR-324-5p seed binding sites.

5.2.2.5 Transcriptome microarray 1 (unstim)

Recently Guo *et al*. showed the majority (>84%) of miRNA targets are reduced at mRNA, level as well as at the protein, level [121]. This meant transcriptome microarrays could be used to search for the majority of miRNA targets. Here a similar experiment to SILAC1 and SILAC2 was performed, but this time extracting the RNA for transcriptome analysis (Figure 5.7A). I obtained mRNA expression data for 25,697 probes corresponding to 18,097 different genes. Similar to protein expression the majority (18,367/25,697=71.5%) of mRNA probes changed less than log2 0.2 fold. In contrast to protein, a similar number of mRNAs were decreased and increased (3,708/25,697=14.4% and 3,622/25,697=14.1% respectively) (Figure 5.7B). 2,297 (8.9%) of the 25,697 probes I obtained mRNA expression data for were to genes which contained at least one miR-324-5p seed binding site, the majority (1,602/2,297=69.7%) of which remained unchanged in expression. A greater number of genes which contained miR-324-5p seed binding sites decreased (477/2,297) than increased (218/2,297) (Figure 5.7C), meaning the fraction of genes which contain a miR-324-5p seed binding site in their 3'UTR was 1.64 fold higher in genes whose expression decreased (477/3708=12.9%), compared to genes whose expression remained the same (1,602/18,367=8.7%) or increased (9218/3,622=6.0%) ($p=1.34 \times 10^{-19}$) (Figure 5.7D).

72 of the 2,297 mRNAs which contained a miR-324-5p seed binding site, contained a conserved miR-324-5p seed binding site, 24 of which decreased, 44 remained unchanged and only 4 increased, representing a 2.98 fold ($p=4.85 \times 10^{-6}$) fold enrichment in decreased mRNAs. Consistent with protein expression (SILAC) this was a greater enrichment than for mRNAs containing non-conserved miR-324-5p seed binding sites (Figure 5.7E and F).

Of the 2,297 mRNAs which contained a miR-324-5p seed binding site 420 contained an 8mer, 1,039 contained a 7mer m8 and 1,057 contained a 7mer A1. In contrast to SILAC1 and 2, but as should be expected, 8mer seeds (1.88 fold, $p=3.47 \times 10^{-8}$) were more enriched than 7mer A1 (1.58 fold, $p=4.49 \times 10^{-9}$) and 7mer m8 seeds (1.52 fold, $p=1.06 \times 10^{-7}$) (Figure 5.7G and H) in those genes whose expression decreased compared to genes whose expression either increased or remained unchanged following miR-324-5p transfection.

Figure 5.7 Transcriptome microarray 1 (unstim)
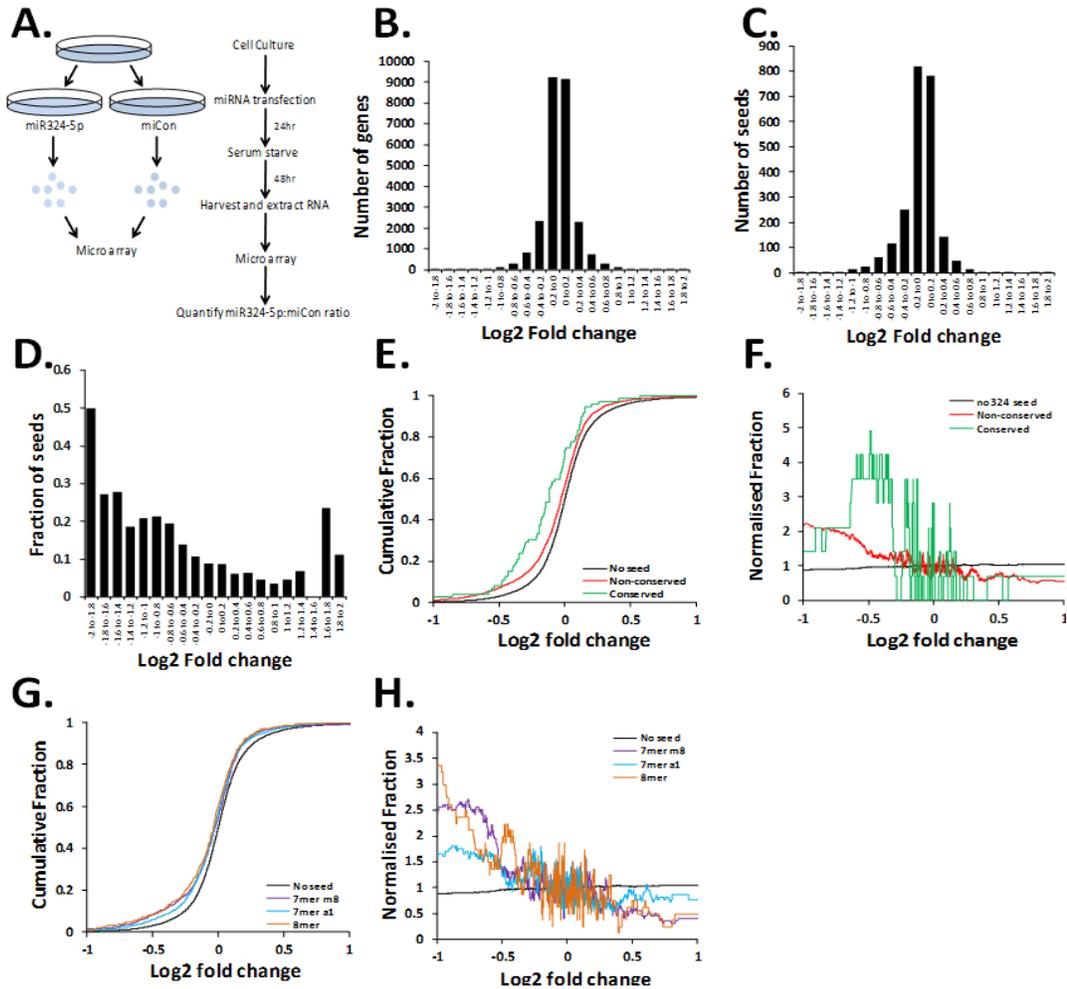
Figure 5.7 Transcriptome microarray 1 (unstim). miR-324-5p targets are enriched with the decreased transcripts of miR-324-5p transfected unstimulated cells. (A) Schematic showing experimental approach for the unstimulated array. (B) Histogram showing expression changes of 25,697 transcripts corresponding to 18,097 genes, following miR-324-5p transfection; 3,708 decreased, 3,622 increased by a fold change of log2 0.2 and 18,367 remained unchanged. (C) 2,297 of the 25,697 identified transcripts contain a miR-324-5p binding site in their 3'UTR; 477 decreased, 218 increased by a fold change of log2 0.2 and 1,602 remained unchanged. (D) Fraction of transcripts which contain a miR-324-5p seed binding site; 12.9%, 6.0% and 8.7% of transcripts decreased, increased and remained unchanged by a fold change of log2 0.2 respectively, contain a miR-324-5p binding site. The fraction of genes containing a miR-324-5p binding site is 1.64 fold higher in genes whose transcript decreased compared to genes whose transcript either remained unchanged in expression or increased (p=1.34x10[-19]). (E) Cumulative fraction of conserved and non-conserved miR-324-5p seed binding sites. Of the 2,297 seed binding sites 72 are conserved, 24 of which are decreased, 4 increased and 44 remained unchanged. (F) Normalised fraction plot of conserved and non conserved miR-324-5p seed binding sites. (G) Cumulative fraction of 8mer, 7mer m8 and 7mer a1 miR-324-5p seed binding sites. (H) Normalised fraction plot of 8mer, 7mer m8 and 7mer a1 miR-324-5p seed binding sites.

5.2.2.5 SILAC protein data correlates with array mRNA data

Of the 18,097 different genes identified at the mRNA level, 1,299 also had proteomic expression data. Log2 fold changes in mRNA correlated with log2 fold change in protein following miR-324-5p transfection ($R2=0.1948$) (Figure 5.8A). 93 of the 1299 genes contained a miR-324-5p seed binding site in their 3'UTR, mRNA and protein data for these 93 genes is shown in Figure 5.8B ($R2=0.3089$). 18 of these 93 decreased at a protein and mRNA level, representing a 3.67 fold ($p=1.68 \times 10^{-6}$) enrichment of miR-324-5p seed containing genes, in genes whose expression decreased at a protein and mRNA level, compared to other genes 975/1,207. This was a greater enrichment than in genes which decreased at only either an mRNA or protein level (1.6 fold and 2.7 fold respectively). Suggesting using both mRNA and protein data to predict miRNA targets is better than using either alone. The 18 potential targets whose protein decreased, mRNA decreased and contained a miR-324-5p seed binding site are shown in Figure 5.8C.

Figure 5.8 Combine SILAC 1 and 2 with unstim array
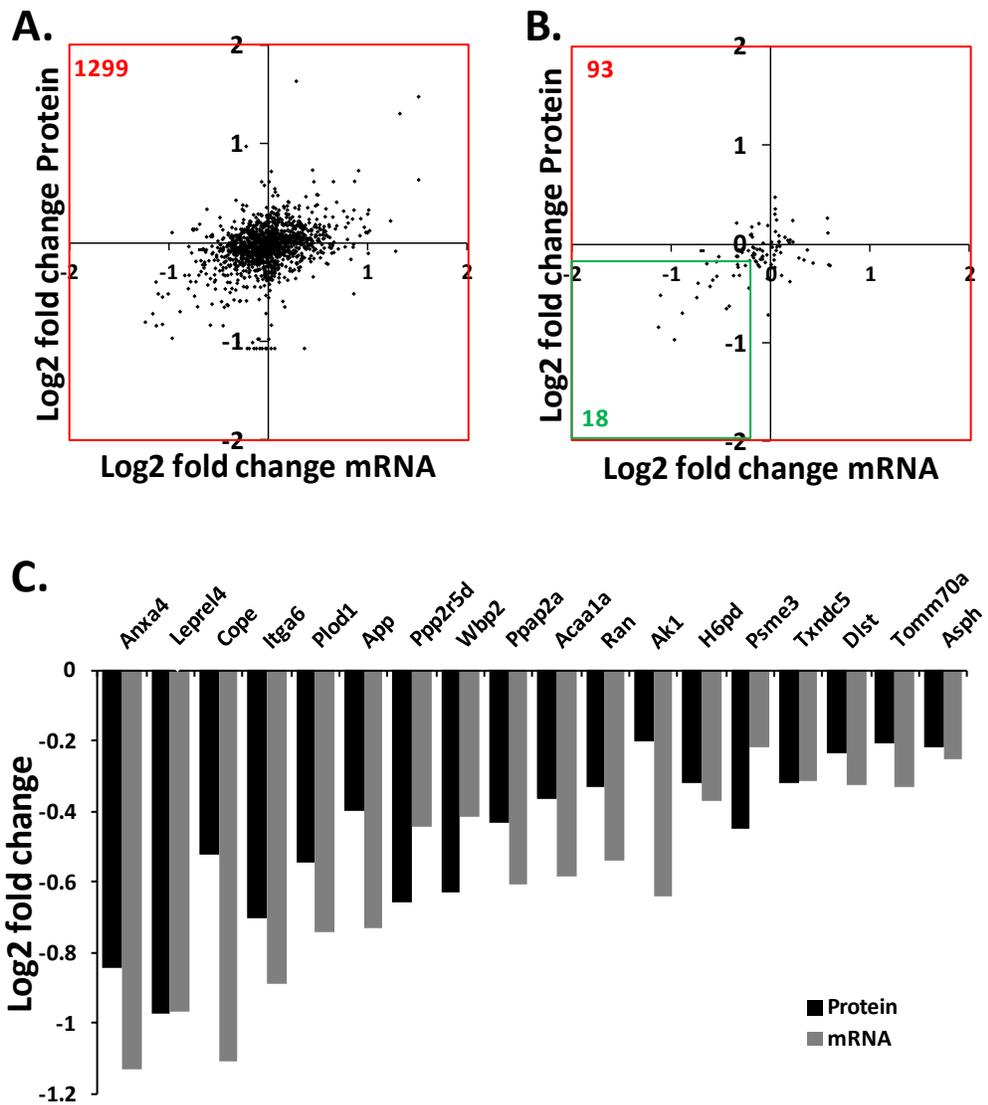
Figure 5.8 SILAC 1 and SILAC 2 average combined with unstimulated array. (A) Change in protein expression correlates with change in mRNA expression following miR-324-5p transfection. 1,299 genes were identified at a protein and mRNA level, 92 of which decrease at a protein and mRNA level. (B) 93 genes were identified at a protein and mRNA level and also contain a miR-324-5p seed binding site, 18 of which decreased at an mRNA and protein level, meaning 74 out of 1206 genes lacking miR-324-5p seed binding sites also decrease. (C) Bar chart showing log2 fold change in protein and mRNA expression for 18 potential targets following miR-324-5p transfection.

5.2.2.6 *Target identification strategy and validation*

I created constructs where the 3'UTR of 2 potential targets were cloned down-steam of a luciferase reporter gene placing luciferase under the control of the predicted targets 3'UTR. When co-transfected with miR-324-5p both of the potential targets had reduced luciferase expression (Figure 5.9A), indicating both of them were valid miR-324-5p targets.

Our results indicate SILAC and transcriptome microarrays can be integrated to form an effective miRNA target identification strategy (outlined in Figure 5.9B).

Figure 5.9 Target identification strategy and validation

A.

B.

Appropriate cells

↓

(1) Transfect miRNA of interest/miCon for24hr

↓

(2) Manipulate the cells in an appropriate way to
ensure potential target transcripts are present

↙ ↘

(3a) SILAC proteomics   (3b) Micro array

↓   ↓

(4a) Check for enrichment   (4b) Check for enrichment
of targets in proteins   of targets in transcripts
decreased >log2 0.2   decreased >log2 0.2

↓   ↓

(5) , Identify potential targets whose protein and mRNA expression
decrease and contain a miR324-5p seed binding site in their 3'UTR

↓

(6) Validate potential targets using
luciferase constructs

↓

(7) Investigate function of
validated miRNA targets

Figure 5.9 Target identification strategy and validation. (A) C3H10T1/2 cells were transfected with either Anxa4 or App 3'UTR luciferase constructs (in pMIR-Report) with either miCon or miR-324-5p. Data combined from 3 independent experiments each n=6, data shown is mean and SEM, statistical difference was calculated using student *t* test where ***<0.001 and *<0.05. (B) Schematic showing miRNA target identification strategy.
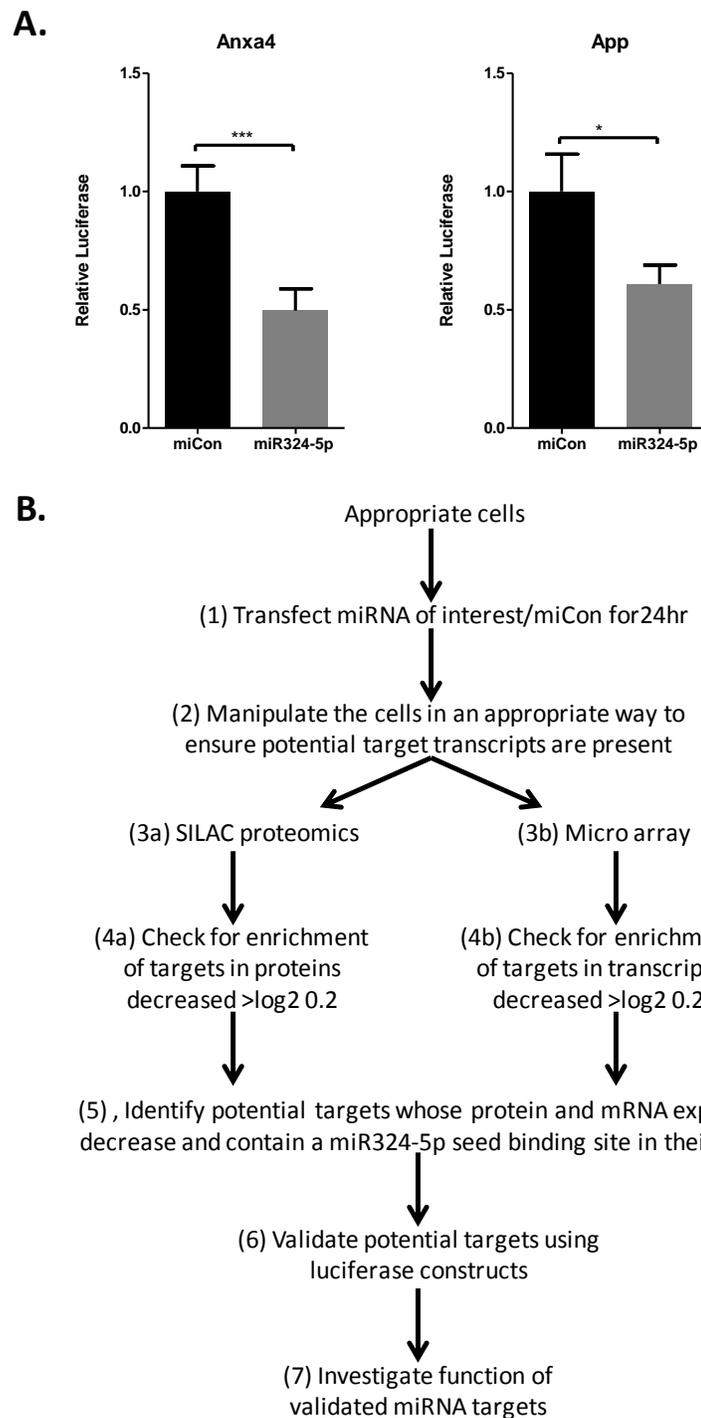
5.2.3 Identification of potential miR-324-5p targets with a role in Hh signalling.

Here I have outlined an effective, comprehensive strategy to identify novel miRNA targets, this was then employed to identify miR-324-5p targets responsible for regulating Hh signalling (the main aim of this Chapter and discussed in the following section). miRNA-target interactions are dependent on both miRNA and target mRNA abundance [245]. Because I was searching for miR-324-5p targets responsible for loss of Hh signalling in Ihh stimulated cells, it is probably necessary to search for miR-324-5p targets in Ihh stimulated cells. The same experimental approach used in Chapter 4, was used here, as it is know miR-324-5p has the ability to inhibit Ihh induced Hh signalling in that experimental approach.

5.2.3.1 SILAC3 (Ihh stimulated)

To determine proteins whose expression decreased following miR-324-5p transfection in Ihh stimulated cells, a third SILAC experiment (SILAC3) was performed. In SILAC3, as in SILAC1, heavy labelled C3H10T1/2 cells were transfected with miR-324-5p and light C3H10T1/2 cells were transfected with miCon (Figure 5.10A). 2,086 proteins were identified, 1,604 of which remained unchanged, 255 decreased and 227 increased (Figure 5.10B). Of the 2,086, 176 (8.4%) contained a miR-324-5p seed in their 3'UTR. Twice as many of these seed containing genes deceased at a protein level than increased, 36 (20.5%) and 18 (10.2%) respectively (Figure 5.10C). Similar to SILAC1 and SILAC2, there was a significant enrichment (1.99 fold, p=0.0005) of miR-324-5p binding sites in proteins whose expression decreased (36/255=14.1%) compared to proteins whose expression increased (18/227=7.9%) or remained unchanged (122/1604=7.6%) (Figure 5.10D).

8 of the 176 proteins whose gene contained a miR-324-5p seed binding site, contained a conserved miR-324-5p seed binding site, 3 of which decreased, 4 remained unchanged and only 1 increased. Consistent with SILAC1 and SILAC2, there is a greater enrichment (4.35 fold, p=0.03) of proteins whose gene contained conserved miR-324-5p binging sites, than those which contained non-conserved miR-324-5p seed binding sites (Figure 5.10 E and F).

Of the 176 genes whose protein was identified in SILAC3 and contained a miR-324-5p binding site in their 3'UTR, 26 contained an 8mer, 70 contained a 7mer m8 and 92 contained a 7mer A1 miR-324-5p binding site. The fraction of genes which contained an 8mer, 7mer m8 and 7mer A1 miR-324-5p seed binding sites is 3.90, 3.03 and 1.19 fold higher in genes whose protein expression decreased compared to genes whose protein either remained unchanged or increased (p=0.0005, p=2.16X10^{-5} and p=0.568) respectively. Suggesting 8mer and 7mer m8 sites are more enriched than 7mer A1 sites in genes whose protein expression decreased (Figure 5.10G and H).

Figure 5.10 SILAC 3 (Ihh stimulated)
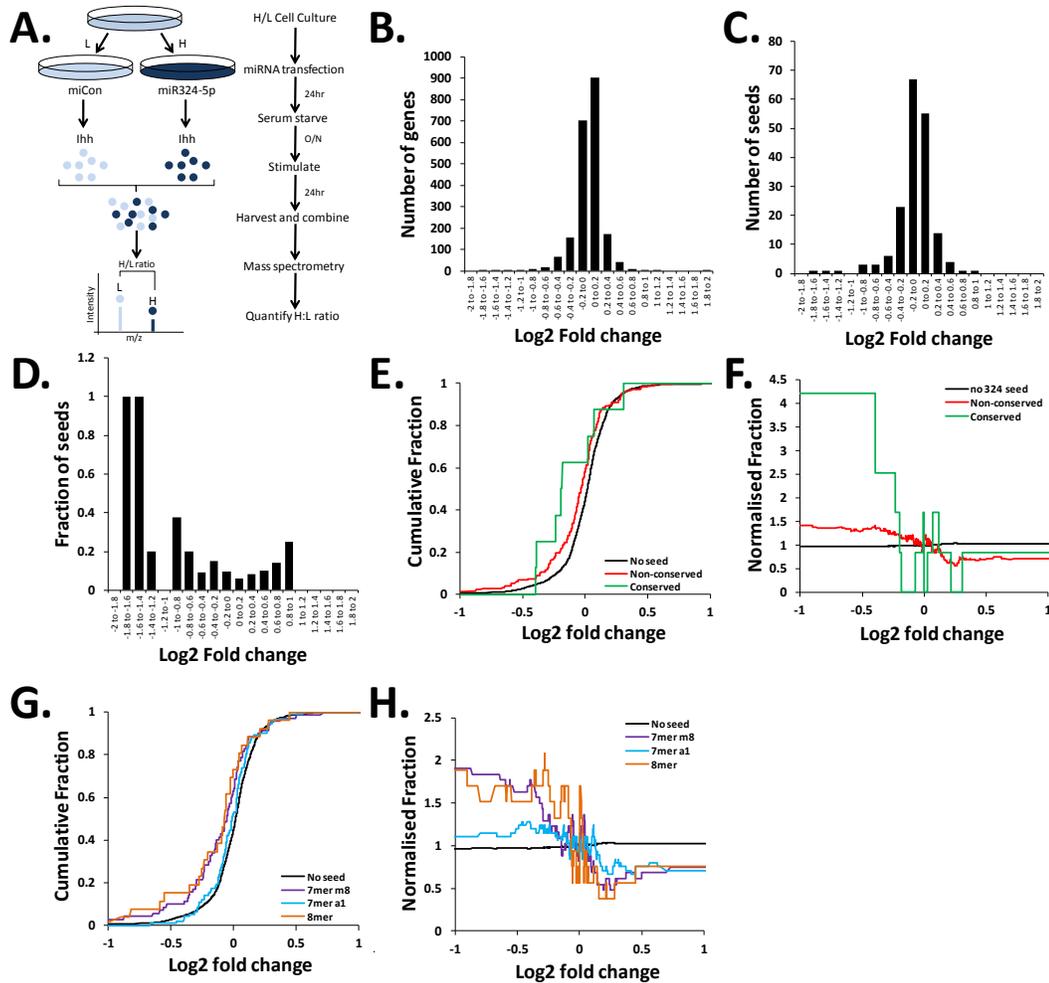
Figure 5.10 SILAC 3 (Ihh stimulated). miR-324-5p targets are decreased at a protein level in miR-324-5p transfected heavy Ihh stimulated cells. (A) Schematic showing experimental set up for SILAC 3. (B) Histogram showing expression changes of 2,086 proteins following miR-324-5p transfection; 255 decreased, 227 increased by a fold change of log2 0.2 and 1,604 remained unchanged. (C) 176 of the 1,086 genes whose protein was identified contain a miR-324-5p binding site in their 3'UTR; 36 decreased, 18 increased by a fold change of log2 0.2 and 122 remained unchanged. (D) Fraction of genes whose protein was identified which contain a miR-324-5p seed binding site; 14.1%, 7.9% and 7.6% of the genes whose protein expression decreased, increased and remained unchanged by a fold change of log2 0.2 respectively, contain a miR-324-5p binding site. The fraction of genes containing a miR-324-5p binding site is 1.99 fold higher in genes whose protein expression decreased compared to genes whose protein either remained unchanged in expression or increased (p=0.0005). (E) Cumulative fraction of conserved and non conserved miR-324-5p seed binding sites. Of the 176 seed binding sites 8 are conserved, 3 of which are decreased, 1 increased and 4 remained unchanged. (F) Normalised fraction plot of conserved and non conserved miR-324-5p seed binding sites. (G) Cumulative fraction of 8mer, 7mer m8 and 7mer a1 miR-324-5p seed binding sites. (H) Normalised fraction plot of 8mer, 7mer m8 and 7mer a1 miR-324-5p seed binding sites.

5.2.3.2 Transcriptome microarray 2 (Ihh stimulated)

To determine transcripts whose expression decreased following miR-324-5p transfection in stimulated cells a second transcriptome microarray was performed (transcriptome microarray 2 stim). In the stimulated transcriptome microarray, similar to SILAC3, C3H10T1/2 cells (normal) were transfected with miR-324-5p or miCon and stimulated with Ihh (Figure 5.11A). As with the unstimulated transcriptome microarray, mRNA expression data was obtained for 25,697 probes corresponding to 18,097 different genes (note: all transcriptome microarrays were performed and analysed simultaneously). Similar to the unstimulated microarray the majority (19,746/25,697=76.9%) of mRNAs remained unchanged in expression and a similar number of mRNAs were decreased and increased (2,975/25,697=11.6% and 2,958/25,697=11.5% respectively) (Figure 5.11B). Again as with the unstimulated array, 2,297 (8.9%) of the 25,697 probes I obtained mRNA expression data for, contained at least one miR-324-5p seed binding site. The majority (1,668/2,297) of which remained unchanged in expression. A greater number of the miR-324-5p seed containing genes decreased (409/2,297) than increased (220/2,297)(Figure 5.11C). The fraction of genes which contained a miR-324-5p seed binding site in their 3'UTR is 1.75 fold higher in genes whose expression decreased (409/2,975=13.7%), than genes whose expression remained the same (1,668/19,764=8.4%) or increased (220/2,958=7.4%) (p=1.41×10^{-22}) (Figure 5.11D).

Of the 2,297 mRNAs which contained a miR-324-5p seed binding site, 72 contained a conserved miR-324-5p seed binding site, 19 of which decreased, 46 remained unchanged and only 7 increased. Consistent with all previous experiments, there was a greater enrichment (2.75 fold, p=8.37×10^{-5}) of mRNAs that contained conserved miR-324-5p binding sites than mRNAs that contained non-conserved miR-324-5p seed binding sites in those mRNAs whose expression decreased (Figure 5.11E and F).

Of the 2,297 mRNAs which contained a miR-324-5p seed binding site, 420, 1,039 and 1,057 contained an 8mer, 7mer m8 and a 7mer A1 respectively. 8mer seeds (1.85 fold, p=6.41×10^{-7}) and 7mer m8 seeds (2.03 fold, p=1.10×10^{-19}) were more enriched than 7mer A1 seeds (1.42 fold, p=6.64×10^{-5}) (Figure 5.11G and H), in those genes whose expression

decreased, compared to genes whose expression either increased or remained unchanged following miR-324-5p transfection.

**Figure 5.11 Transcriptome microarray 2 (Ihh stimulated)**
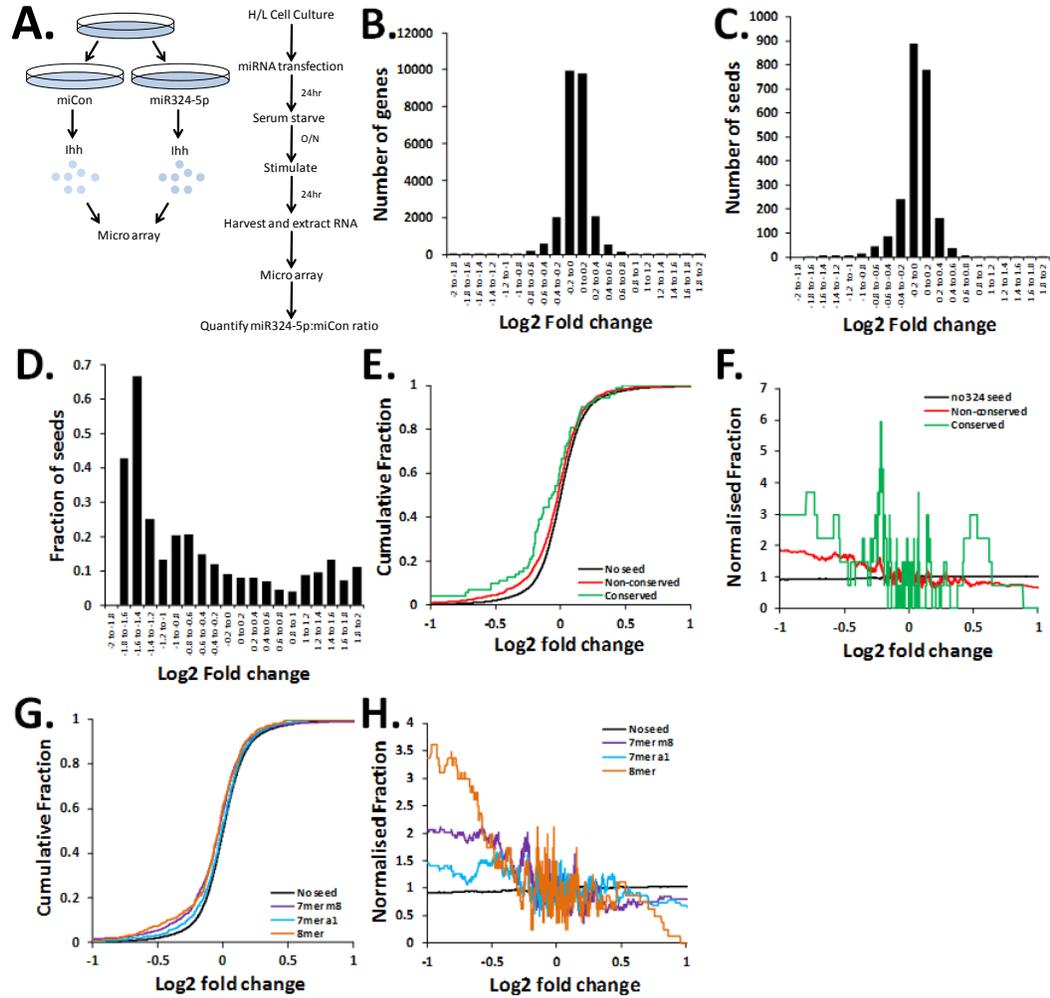
Figure 5.11 Transcriptome microarray 2 (Ihh stimulated). miR-324-5p targets are decreased at an mRNA level in miR-324-5p transfected and Ihh stimulated cells. (A) Schematic showing experimental approach for the Ihh stimulated array. (B) Histogram showing expression changes of 25,697 transcripts corresponding to 18,097 genes, following miR-324-5p transfection; 2,975 decreased, 2,958 increased and 19,746 remained unchanged. (C) 2,297 of the 25,697 identified transcripts contain a miR-324-5p binding site in their 3'UTR; 409 decreased, 220 increased by a fold change of log2 0.2 and 1,668 remained unchanged. (D) Fraction of transcripts which contain a miR-324-5p seed binding site; 13.7%, 7.4% and 8.4% of transcripts decreased, increased and remained unchanged by a fold change of log2 0.2 respectively, contain a miR-324-5p binding site. The fraction of genes containing a miR-324-5p binding site is 1.76 fold higher in genes whose transcript decreased compared to genes whose transcript either remained unchanged in expression or increased (p=1.41x10^{-22}). (E) Cumulative fraction of conserved and non conserved miR-324-5p seed binding sites. Of the 2,297 seed binding sites 72 are conserved, 19 of which are decreased, 7 increased and 46 remained unchanged. (F) Normalised fraction plot of conserved and non conserved miR-324-5p seed binding sites. (G) Cumulative fraction of 8mer, 7mer m8 and 7mer a1 miR-324-5p seed binding sites. (H) Normalised fraction plot of 8mer, 7mer m8 and 7mer a1 miR-324-5p seed binding sites.

5.2.3.3Combine SILAC3 (stim) and transcriptome microarray 3(stim)

Of the 18,097 different genes identified on the array, 2,086 were also identified in SILAC3 and the fold change in mRNA correlated with the fold change in protein (Figure 5.12A). 176 of the genes contained a miR-324-5p seed binding site in their 3'UTR. mRNA and protein expression data for these 176 genes is shown in Figure 5.12B. 18 of the 176 genes are decreased at a an mRNA and protein level and can be considered as potential miR-324-5p targets in Ihh stimulated cells according to our miRNA target identification strategy (Figure 5.12).

Figure 5.12 Combine SILAC 3 with Ihh stimulated array
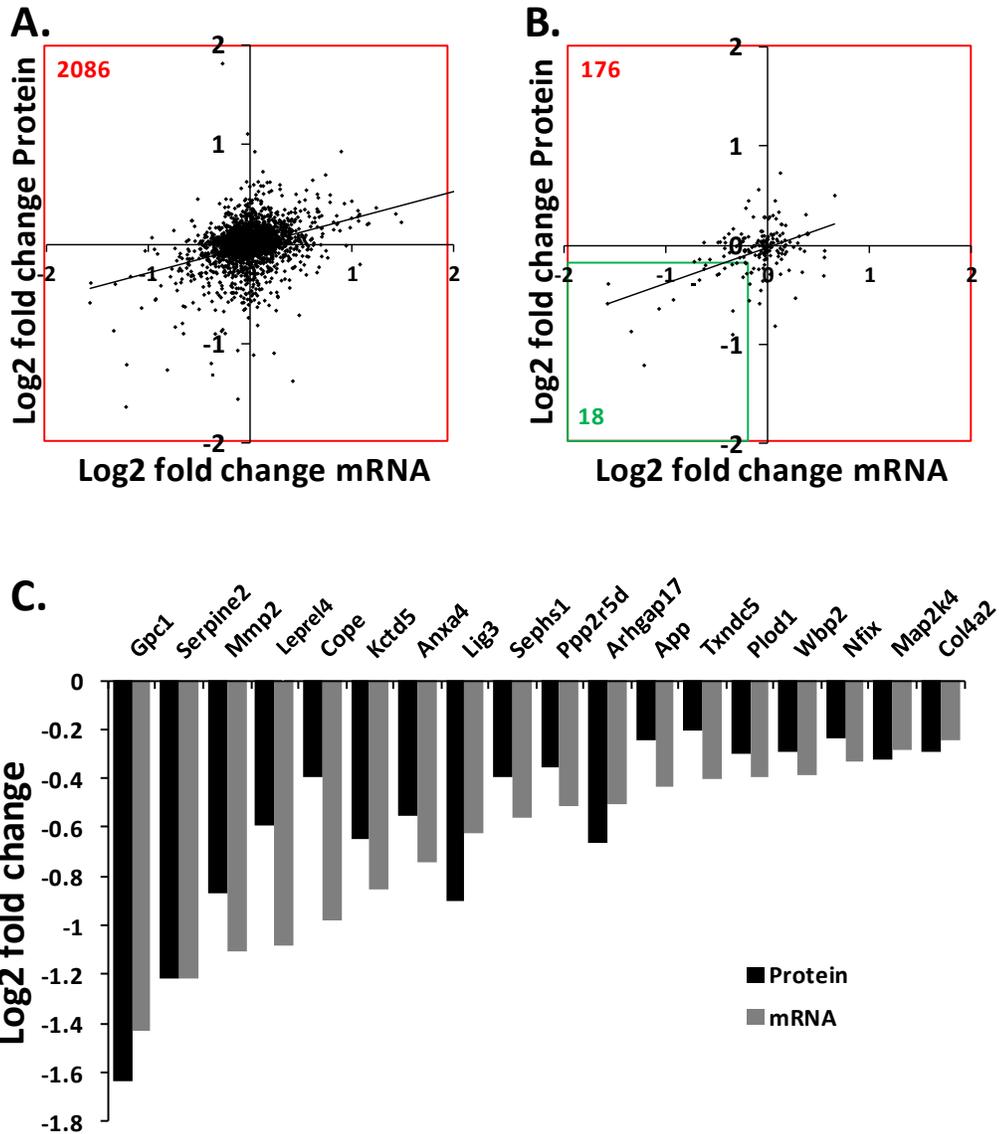
Figure 5.12 SILAC 3 combined with Ihh stimulated transcriptome array. (A) Change in protein expression correlates with change in mRNA expression following mR324-5p transfection. 2,086 genes were identified at a protein and mRNA level (B) 176 genes were identified at a protein and mRNA level and also contain a miR-324-5p seed binding site, 18 of which decreased at an mRNA and protein level. (C) Bar chart showing log2 fold change in protein and mRNA expression for 18 potential targets following miR-324-5p transfection.

5.2.3.4 Potential targets which may have a role in Hh signalling

Two strategies were used to identify which of potential targets were involved in Hh signalling. Firstly, the Hh Kegg pathway contains a number of genes which have previously been shown to be involved in Hh signalling. However, of the 18 potential targets none were part of the Hh KEGG pathway (Figure 5.13), although Hh Kegg pathway is not entirely comprehensive. Secondly, literature searches showed relatives of the potential target Glypican 1 (Gpc1) have previously been shown to play a role in Hh signalling [246, 247]. I therefore hypothesised Gpc1 was involved in Hh signalling and is a miR-324-5p target in mouse. The overall strategy used to identify potential miR-324-5p targets involved in Hh signalling is summarised in Figure 5.13.

Figure 5.13 Finding Gpc1

Appropriate C3H10T1/2 cells

↓

(1) Transfect miR324-5p/miCon for24hr

↓

(2) Serum starve for 24hr then stimulate with Ihh for 24hr

(3a) SILAC proteomics identifies 2,086 proteins

(3b) Micro array identifies 25,697 transcripts (18,097 genes)

↓

(4a) 1.99 (p=0.0005) fold enrichment of targets in 255 proteins decreased >log2 0.2

(4b) 1.76 (p=1.41×10^{-22}) fold enrichment of targets in 2,958 transcripts decreased >log2 0.2

↓

(5) 176 genes contain miR324-5p seed binding sites, 18 (the potential targets) of which decreased

Hh kegg pathway →   ← Literature

No genes part of the Hh kegg pathway

Gpc1 involved in Hh pathway?

↓

(6) Gpc1 validated as a real miR324-5p target using luciferase

↓

Chapter 6: Investigate role of Gpc1 in Hh pathway

Figure 5.13 Experimental flow used to identify Gpc1 as a miR-324-5p target with a possible role in Hh signalling. Numbers correspond to equivalent parts of the miRNA target identification strategy (Figure 5.8B)
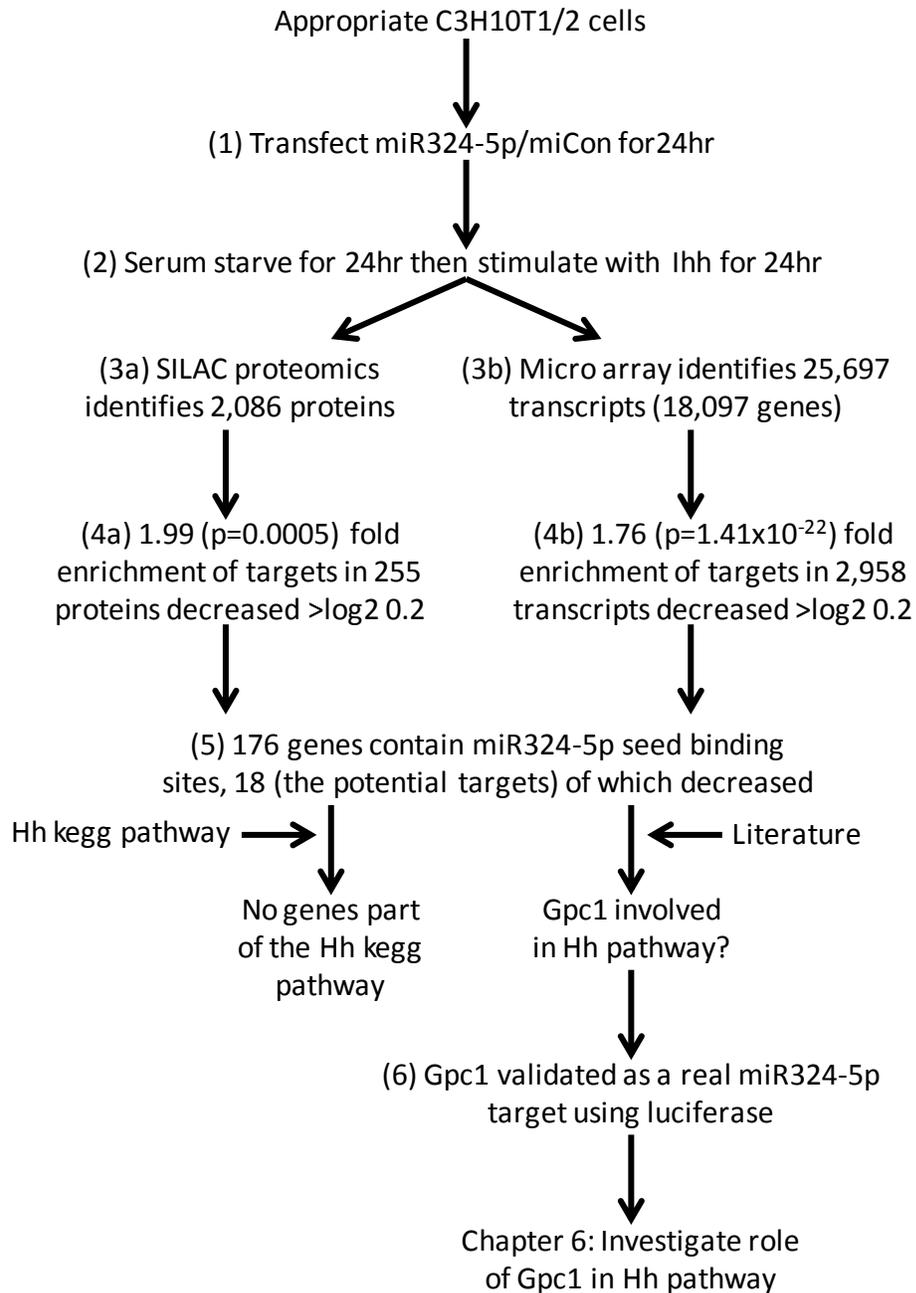
5.2.4   Validation of Gpc1 as a miR-324-5p target

Gpc1 was identified in Ihh stimulated proteomics (SILAC 3) by 6 peptides, 5 of which decrease with the addition of miR-324-5p (Figure 5.14A). The array data showed Gpc1 mRNA was decreased in both stimulated and unstimulated cells with the addition of miR-324-5p (Figure 5.14B), indicating Gpc1 was a miR-324-5p target irrespective of Ihh stimulation. Immunoblotting of Gpc1 protein following miR-324-5p transfection was used to confirm miR-324-5p reduces Gpc1 at a protein level (Figure 5.14C). Real-time RT-PCR of Gpc1 mRNA following miR-324-5p transfection was used to confirm miR-324-5p reduced Gpc1 mRNA (Figure 5.14D). To validate Gpc1 as being a direct target of miR-324-5p, I created a construct where the mouse Gpc1 3'UTR was cloned downstream of a luciferase gene placing luciferase under the control of the Gpc1 3'UTR. Co-transfection of this construct with miR-324-5p showed miR-324-5p can reduce the amount of luciferase produced by this construct (Figure 5.14E), indicating Gpc1 was a direct miR-324-5p target in mouse. Further analysis of the mouse Gpc1 3'UTR showed it has 3 miR-324-5p seed matches (Figure 5.14F). Site directed mutagenesis of each of the 3 potential miR-324-5p binding sites indicated the most downstream site (site 3), which is a '12mer site', was the only functional miR-324-5p binding site in the mouse Gpc1 3'UTR, as only when this site was mutated, was there a rescue of miR-324-5p mediated reduction of luciferase levels (Figure 5.14E).

Figure 5.14 Validation of Gpc1 as a direct miR-324-5p target
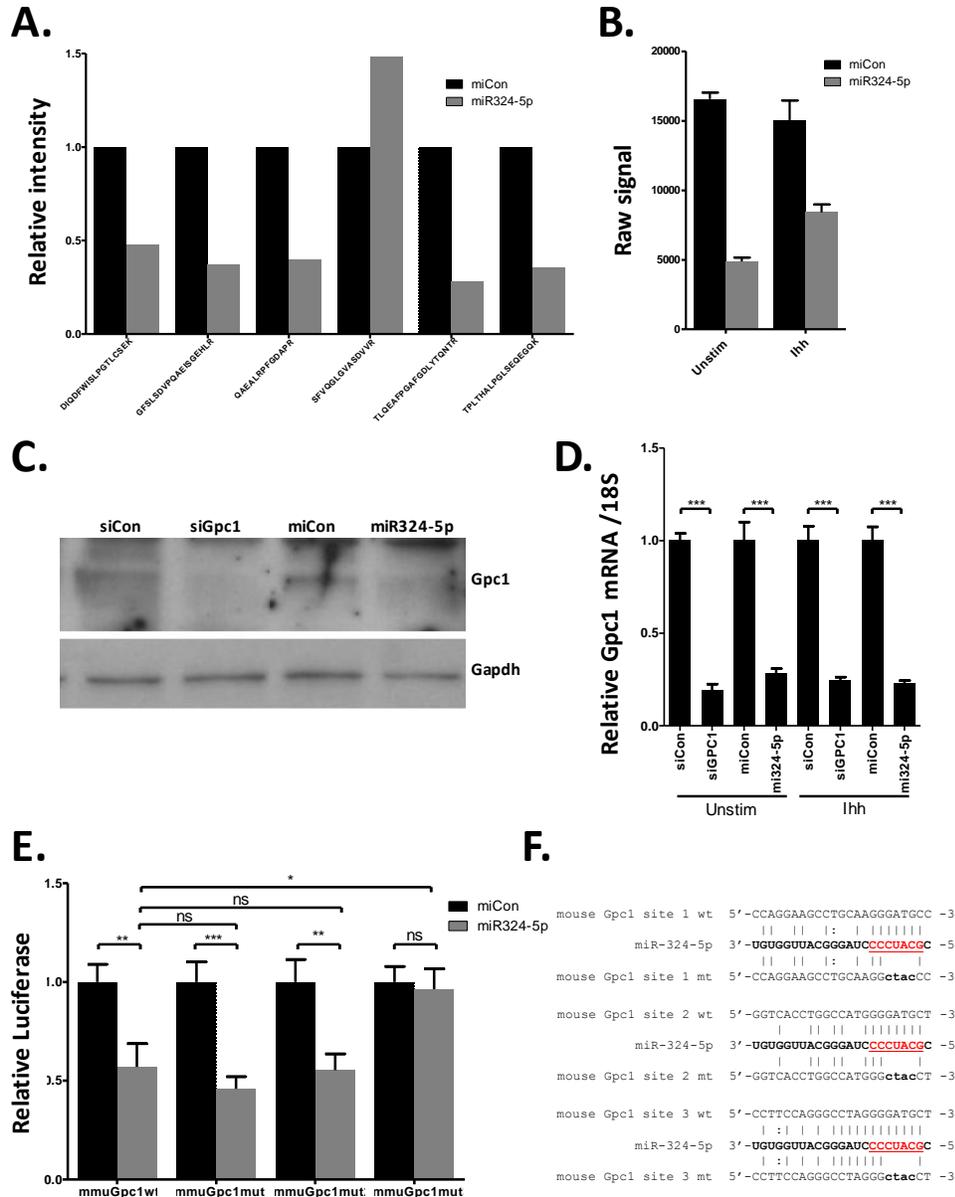
Figure 5.14 Gpc1 is a miR-324-5p target in mouse. (A) Effect of miR-324-5p on Gpc1 peptides identified in SILAC 3. (B) Effect of miR-324-5p on Gpc1 mRNA expression as identified in microarray 1 (unstim) and microarray 2 (Ihh stimulated). (C) Western blot analysis of Gpc1 protein following siGPC1 and miR-324-5p transfection. (D) real-time RT-PCR analysis of Gpc1 mRNA following siGpc1 and miR-324-5p transfection. (E) pMIR-Report plasmid containing either wild-type or mutant mouse Gpc1 3'UTR downstream of luciferase was transfect into C3H10T1/2 cells with either miCon2 or miR-324-5p, data normalised to miCon and plotted as relative luciferase light units, statistical difference were calculated using ANOVA. Data are combined from 5 independent experiments, each n=6-8. (F) Potential miR-324-5p binding site in wild-type and mutant Gpc1 3'UTRs. miRNA seed sequence is shown in red and underlined, mutated miR-324-5p seed binding site are shown in lower case.

5.3 Discussion

The main aim of this chapter was to identify miR-324-5p targets in mouse which have a possible role in Hh signalling. 11 predicted mouse miR-324-5p targets were predicted to have a role in Hh signalling (Figure 5.2), but due to the unreliability of prediction algorithms, a strategy to identify miRNA targets was developed. C3H10T1/2 cells were transfected with mature miR-324-5p mimic, increasing miR-324-5p incorporation into RISC and its interaction with targets, ultimately caused miR-324-5p targets to have reduced expression. SILAC proteomics and transcriptome microarrays were then used to detect which genes had reduced protein and mRNA expression. To validate our experiments I analysed the 3'UTR of genes whose protein and mRNA expression decreased following miR-324-5p transfection and found these 3'UTRs had a significantly higher number of miR-324-5p binding sites than the 3'UTRs of genes whose expression remained unchanged or increased; these genes were therefore considered as potential miR-324-5p targets. Literature searches found one potential target, Gpc1, had a possible role in Hh signalling (Chapter 6). Gpc1 was then validated as a direct miR-324-5p target.

5.3.1 Analysis of target identification methodology

5.3.1.1 Cell culture

During strategy development, C3H10T1/2 cells were left unstimulated, keeping the experimental design a simple as possible and ensuring the majority of changes in gene expression were due to miR-324-5p overexpression rather than due to the blocking of Hh signalling.

In the SILAC experiments heavy labelled lysine and arginine were used because the preparation of samples for MS involved the cleavage of proteins into peptides with trypsin, an enzyme that cleaves after lysines and arginines. This means every peptide will contain at least one heavy labelled amino acid. More specifically $^{13}C_6$-Lys and $^{13}C_6$ $^{15}N_4$-Arg were used to label cells as heavy, where the carbon in lysine and both the carbon and nitrogen in arginine contained one extra neutron in the nucleus. Lysine contains 6 carbons meaning any peptide that contained a lysine was 6 mass units heavier, while arginine

contains 6 carbons and 4 nitrogens meaning any peptide that contained an arginine was 10 mass units heavier. Standard $^{12}C_6$-Lys and $^{12}C_6$ $^{14}N_4$-Arg were used in light media.

Because two populations of cells are needed for SILAC experiments I wanted to avoid any differences between the populations being misinterpreted as real changes in expression or as miRNA targets. Labels in SILAC2 were therefore the reverse of the labels in SILAC1, miR-324-5p targets were enriched irrespective of label orientation. Proteins whose expression correlated with the label rather than the miRNA were likely to be due to stochastic variations between the experiments. Interestingly, a larger proportion of proteins changed in expression in SILAC2 than in SILAC1, this may be due to differences in transfection efficiency between the experiments, although this was not tested.

5.3.1.2 Timing

Some studies have shown miRNAs reduce protein and mRNA expression in as little as 4h [248, 249]. In our experiments I analysed much later (72h post transfection), meaning there was likely to be many secondary effects of miRNA targets, making it difficult to determine direct miR-324-5p targets from the indirect effects of targets.

5.3.1.3 The potential targets

18 potential targets were identified in unstimulated cells, luciferase constructs were made for 2 of these and both were validated as real miR-324-5p targets. This limited data suggests the strategy is effective, however it is likely some genes will have decreased in expression due to secondary effects of the miRNA and are not direct targets. During luciferase validation both the target (luciferase construct) and miR-324-5p itself were vastly increased in the cell, increasing the chance of miRNA-target interaction and probably leading to an over estimate of the number of physiologically relevant targets.

5.3.1.4 Targets in Hh

miRNA-target interactions are dependent on the miRNA and mRNA expression profile [245]. I wanted to identify miR-324-5p targets involved in Ihh signalling, therefore I stimulated the cells with Ihh, to ensure potential target mRNAs were present. The stimulation with Ihh was akin to that used in Chapter 4, where the mechanism by which miR-324-5p inhibits Ihh induction of Gli1 was unknown. Similar to the unstimulated cells

18 potential targets were identified in stimulated cells, although many of them were different (Figure 5.15).

To assess if any potential miR-324-5p targets are required for Hh signalling, the list of potential targets identified in stimulated cells was compared to the list of genes in the Kegg Hh pathway. No potential miR-324-5p targets were part of the Hh kegg pathway meaning more detailed analysis and literature searches were required. Literature searches, showed some family members of one potential target, Gpc1, can regulate Hh signalling [246, 250-253] (discussed further in Chapter 6). I therefore hypothesised Gpc1 could also regulate Hh signalling (Chapter 6).

5.3.1.5 Gpc1 validation

Gpc1 was only identified as a potential target in stimulated cells (SILAC3 and transcriptome microarray 2), as SILAC1 and 2 (unstim) did not identify Gpc1. Further experiments indicated Gpc1 was a target independent of stimulation (Figure 5.14E), perhaps suggesting Gpc1 was decreased with the addition of miR-324-5p in SILAC1 and 2, but just not detected, the reason this was unclear, but is possibly due to the inherent bias of proteomics to identify some peptides and not others. This was unlikely however as Gpc1 was identified by 6 peptides in SILAC 3 none of which were identified in either SILAC1 or SILAC2. An alternative possibility is Ihh stimulation 'allowed' Gpc1 protein to be identified, there is no evidence Gpc1 mRNA expression changed following Ihh stimulation (Figure 5.14E), but it is possible its cellular location did, meaning it became more abundant in extracts. Gpc1 location and its effects on Hh signalling warrants further investigation. Unfortunately, direct comparison of Gpc1 protein expression using immunoblotting on the actual lysates used in SILAC 1, 2 and 3 was not performed. In addition Gpc1 is a highly modified, membrane-bound, protein and its detection may be susceptible to slight variations in sample preparation between experiments.

Of the 6 Gpc1 peptides in SILAC 3, 5 were decreased following miR-324-5p transfection (Figure 5.14A). The reason why one peptide increased following miR-324-5p transfection is unknown, although it may be due to miss identification of the peptide or stochastic

variation. Either way it was largely ignored as the reduction in Gpc1 protein was validated using immunoblotting.

A luciferase construct was used to show miR-324-5p directly bound the Gpc1 3'UTR (Figure 5.14D). Further analysis of the mouse Gpc1 3'UTR showed it has 3 miR-324-5p seed matches (Figure 5.14F) referred here to as site one, site two and site three. Site directed mutagenesis of each of the 3 potential miR-324-5p binding sites indicated only the most downstrem site (site 3) is functional (Figure 5.14D). This is not surprising as there is more base pairing of this site with miR-324-5p (12mer) than site 1 (7mer) and site 2 (8mer)(Figure 5.14). In addition site 3 is closer to the end of the 3'UTR than sites 1 and 2 (Figure 5.14F), with sites closer to the ends of the 3'UTR having previously been shown to be more effective targets [218, 219].

5.3.1.6 Target ID in unstimulated cells vs. stim cells

A total of 28 potential targets were identified in both the stimulated and unstimulated experiments, 18 of which were identified in the unstimulated experiments, 18 of which were identified in the stimulated experiments and 8 of which were identified in both (Figure 5.15), this indicated these 8 are miR-324-5p targets independent of stimulation. Similar to Gpc1, many of the other potential targets may also be stimulation independent. Equally, there are likely be a number of miR-324-5p targets which were not identified in either of the experiments, and a number of the identified potential targets may not be genuine targets.

5.3.1.7 Summary of the methodology used

Here I have created an effective miRNA target identification strategy (Figure 5.9B). Primarily it involves transfection of a mature miRNA of interest (or control) for 24h to increase the amount of that miRNA within the cell, this increases the amount of that miRNA within the RISC complex and subsequently increases the targeting effects of that miRNA. The cells are then treated in an appropriate way to ensure potential target transcripts are present. Proteomics (SILAC) and transcriptome microarrays are then performed on the protein and mRNA obtained from the transfected cells to detect those proteins and mRNAs whose expression decrease following miRNA transfection. The

enrichment of miRNA seed binding sites in the 3'UTR of genes whose protein and mRNA decreased is then validated. Luciferase constructs are used to validate individual targets before being investigated further.
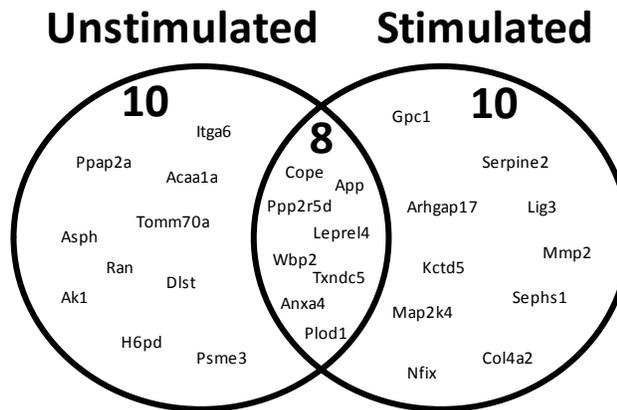
Figure 5.15 Venn diagram of potential targets

Figure 5.15 Venn diagram of potential targets. 18 potential targets were identified in the unstimulated experiments (SILAC1, SILAC2, unstimulated array) and 18 potential targets were identified in the stimulated experiments (SILAC3 and stimulated array). 8 potential targets were identified in both the unstimulated and stimulated experiments.

5.3.2 Analysis of what makes a good target

5.3.2.1 Enrichment of miR-324-5p seed containing genes

The transcriptome microarray and SILAC proteomics results showed miR-324-5p transfection reduced the expression of miR-324-5p targets. The conclusion is two fold. Firstly, a greater number of miR-324-5p seed containing genes decreased than increased following miR-324-5p transfection (Figure 5.4C, Figure 5.5C, Figure 5.6C, Figure 5.7C, Figure 5.8B, Figure 5.10C, Figure 5.11C and Figure 5.12B) and secondly of the genes whose expression decreased, a greater number than expected by chance contained a miR-324-5p seed binding site following miR-324-5p transfection (Figure 5.4D-F, Figure 5.5D-F, Figure 5.6D-F, Figure 5.7D-F, Figure 5.10D-F and Figure 5.11D-F).

5.3.2.2 Conserved sites are more enriched than non-conserved sites

In all experiments (SILAC1, SILAC2, transcriptome microarray 1 (unstim), SILAC3 and transcriptome microarray 2 (stim)) there was a greater enrichment of conserved miR-324-5p seed binding sites than non conserved miR-324-5p seed binding sites (Figure 5.4E & F, Figure 5.5E & F, Figure 5.6E & F, Figure 5.7E & F, Figure 5.10E & F and Figure 5.11E & F). In addition, there was a 2.73 fold enrichment of conserved miR-324-5p binding sites in the 28 potential targets compared with miR-324-5p seed binding site containing genes not classified as a potential target. This suggested that generally the more species a given miR-324-5p binding site is present in, the greater the chance of it being a real target. The reason why conserved seed binding sites are more likely to be real targets is probably linked to their evolution, but must also involve currently unknown mechanisms which allow increased binding to conserved targets. Robust miRNA-Target interactions would have been subject to a greater selection pressure than weak interactions and therefore have a greater chance of spreading through the gene pool. The effect the miRNA-Target interaction has on the organism will of course be determined either positive or negative selection pressure. Nevertheless a miRNA and target do need to interact in order for them to become conserved.

5.3.2.3 Match at position 8 of the miRNA

The most important region for miRNA-target interaction is known as the seed region as it is predicted to base pair with the target mRNA 3'UTR in the RISC complex [254]. There are

a number of different types of seed, the three main ones being 8mers, 7mer m8s and 7mer A1s [129] and each has a different level of binding (Figure 5.1). 8mers, 7mer m8s and 7mer A1s all were all enriched in genes whose expression decreased, but 7mer A1s were consistently the least enriched (Figure 5.4G & H, Figure 5.5G & H, Figure 5.6G & H, Figure 5.7G & H, Figure 5.10G & H and Figure 5.11G & H). This was expected as they are the shortest type, and lack a match at position 8 (Figure 5.1C). Consistent with our results nucleotide number 8 has previously been shown to play a large role miRNA-mRNA interactions [255]. In contrast, Lewis *et al*. previously suggested nucleotide number 8 did not play a large role in determining if a gene is targeted, although they only analysed conserved miRNA seed binding sites [129]. In our experiments 8mers and 7mer m8s were found to be the most enriched but neither was consistently enriched more than the other (Figure 5.4G & H, Figure 5.5G & H, Figure 5.6G & H, Figure 5.7G & H, Figure 5.10G & H and Figure 5.11G & H), in contrast to previous studies [129] this suggest the 'A' at position 1 (Figure 5.1A) does not always provide any additional benefit when positions 2-8 are complementary (Figure 5.1B).

In addition, the consensus binding site for the whole of the miRNA was calculated to determine if regions outside of the seed are important. Consistent with analysis of enrichment of different seed types, there was a preceding 'G' (match at position 8) in 68% of the miR-324-5p 6mer seed binding sites in the 28 potential targets (Figure 5.16A), which was 1.86 fold higher than expected, further suggesting this base is beneficial for targets. Consensus sequence analysis of 7mers (Figure 5.16B), shows a slight preference for a 'T' ('U') to match the 'A' at position 9, and interestingly also at position 17 of the miRNA (Figure 5.16B).

5.3.2.4 Positions in 3'UTR

The position of the miRNA binding site within the 3'UTR has also been shown to be important for miRNA mediated reduction in gene expression [219]. The position of the miR-324-5p seed binding sites in the 28 potential targets seemed randomly distributed, and in contrast to previous studies [218, 219], showed no preference for the ends of the 3'UTR (Figure 5.16C).

5.3.2.5 Number of miR-324-5p binding sites

Some potential targets had more than one potential miR-324-5p seed binding site. Predicted targets whose protein and mRNA decreased (the potential targets) had an average of 1.52 seed binding sites per gene, where as all other genes whose 3'UTR contained a miR-324-5p seed binding site had an average of 1.25 seed binding sites per gene. This suggested having additional binding sites increases the likelihood of a gene being a miR-324-5p target. However, although a target may have many miR-324-5p seed binding sites not all of them need to be functional in order for a gene to be a target (Figure 5.14D). One potential target identified in Ihh stimulated cells is Serpine2 which had six consecutive miR-324-5p seed binding sites, strongly suggesting Serpine2 was a 324 target. I validated Serpine2 as a target using luciferase (Figure 5.16D). Interestingly, the addition of miR-324-5p to the Serpine2 construct did not reduce luciferase to greater amount than for other validated targets (Figure 5.9A, Figure 5.14D, Figure 5.16D), suggesting having additional miR-324-5p seed binding sites does not influence how much a given target is reduced in expression (but may increase the probability of a gene being a miR-324-5p target). It has been suggested some miRNA targets bind to miRNAs to sequester the miRNA itself rather than to be regulated, these targets are called sponges [124]. Because Serpine2 has 6 7mer and 5 6mer miR-324-5p seed binding sites (Figure 5.16D) it may function effectively in this way.

5.3.2.6 Summary

In summary a gene is required to contain a miR-324-5p seed in its 3'UTR and be reduced at a protein and mRNA expression in order for it to be classified as a potential miR-324-5p target. In addition our data has shown any gene that is conserved, has a match at position 8 or has more than one miR-324-5p seed binding site has an increased the probability of a gene being a miR-324-5p target. Other studies, but not necessarily this one, have shown the position within the 3'UTR [218, 219], and additional binding away from the seed also increases the probability of a gene being a miR-324-5p target.

5.3.2. False positives and negatives

There were however many genes that contained miR-324-5p seed binding sites that did not decrease and some genes which lack miR-324-5p seed binding sites did decrease. The reason why some miR-324-5p seed binding site containing genes do not decrease would have been partly due to the lack of the properties outlined above. Some miRNA seed containing genes might not be targets because they were detrimental to the organism and have evolved to escape the RISC complex, possibly by increased binding to RNA binding proteins [256, 257]. It is likely decreased genes which lack miRNA seed binding sites were due to secondary effects of real miRNA targets, although some miRNAs can interact with targets which lack a full seed binding site [127].

Figure 5.16 What makes a good target?

Figure 5.16 What makes a good target? (A) Consensus seed binding site analysis of surrounding 6mer binding sites in the 28 potential targets. (B) Consensus seed binding site analysis of surrounding 7mer binding sites in the 28 potential targets. (C) Relative position of all miR-324-5p seed binding sites in all 28 potential targets combined. (D) Serpine2 has 6 7mer miR-324-5p binding sites (blue) and 5 6mer miR-324-5p binding sites (green) and is a miR-324-5p target. Schematic shows miR-324-5p binding only to 7mer sites. Data combined from 3 independent experiments each n=6, data shown is mean and SEM, statistical difference was calculated using student *t* test where ***<0.001.

5.3.3 Analysis of online target prediction algorithms

Online databases are a useful tool to aid with target identification, but when used alone offer little guidance as to which genes are real targets. For example, the combination of unstimulated SILAC and array experiments showed 8/93 proteins (Figure 5.6C) and 218/2,297 mRNAs (Figure 5.7C), whose genes 3'UTR contained a miR-324-5p seed binding site actually increased in expression, strongly indicating they were not miR-324-5p targets, and would be falsely identified by any prediction program. Each prediction program has a slightly different algorithm and therefore predicts a different repertoire of targets. In an attempt to analyse which of the algorithms is the best, or more precisely which algorithm agreed most with our data, I analysed the number of predicted targets from each algorithm that decreased in transcriptome microarray 1 following miR-324-5p overexpression. The total number of probes on the array that were predicted to be a target by each algorithm is shown in Figure 5.17 along with the number of which decreased following miR-324-5p overexpression, the positive predictor value, the 'fold better than by chance' (how many fold higher, the fraction of predicted targets that decreased, is than the fraction of non-predicted targets that decreased) and the p-value for each program. Cumulative fraction plot, normalised fraction plot, sensitivity and specificity are shown in Figure 5.17. The cumulative fraction plot and normalised fraction plot indicated PicTar is the most specific algorithm (Figure 5.17C) and has the highest positive predictor value, however it predicted the fewest number of targets (Figure 5.17). TargetScan was the most sensitive program (Figure 5.17C) and predicted the largest fraction of targets (Figure 5.17A and B), TargetScan did however predict the largest number of targets (Figure 5.17). Consistent with previous studies [244, 258], there was an inverse relationship between sensitivity and specificity for all prediction programs, and no one program was best at both. Choosing which program to use should be dependent on the application it is being used for, for example the number of targets it is possible to further investigate.

Figure 5.17 Analysis of online target prediction algorithms

A.

B.

C.

D.

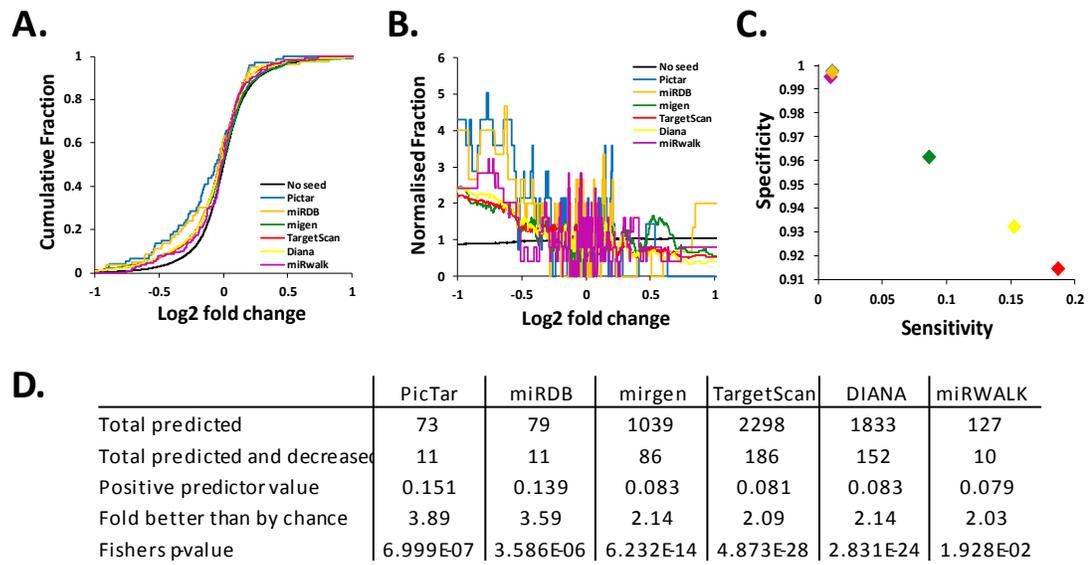|	PicTar	miRDB	mirgen	TargetScan	DIANA	miRWALK
Total predicted	73	79	1039	2298	1833	127
Total predicted and decreased	11	11	86	186	152	10
Positive predictor value	0.151	0.139	0.083	0.081	0.083	0.079
Fold better than by chance	3.89	3.59	2.14	2.09	2.14	2.03
Fishers p-value	6.999E-07	3.586E-06	6.232E-14	4.873E-28	2.831E-24	1.928E-02

Figure 5.17 Analysis of online prediction sites. (A) Cumulative fraction of genes predicted to be miRNA targets. (B) Normalised fraction of genes predicted to be targets by online prediction sites. (C) Sensitivity and specificity of online prediction sites. (D) Numbers show the total number of probes whose gene is a predicted target (out of 25,697 probes on the array) and number of predicted targets whose gene decreased (out of 996 decreased probes). The 'fold better than by chance' indicates how many fold higher, the fraction of predicted targets that decreased, is than the fraction of non-predicted targets that decreased. Positive predictor value (where '1' is perfect prediction and '0' is no prediction) and Fishers p-value for each prediction program is also shown.

5.3.4 miR-324-5p may regulate translation

Traditionally miRNAs and RISC are thought to bind to their target mRNA causing either mRNA degradation, reducing the mRNA available for translation, or directly inhibit translation leaving the mRNA expression unaffected [110] (Figure 5.18A). For the vast majority of targets their reduction in protein correlates with the reduction in mRNA [121]. Recent data in Drosophila cells [248] and zebrafish [249] suggested the reduction in protein translation may actually precede the reduction in mRNA. More specifically they suggest the studied miRNAs function by reducing translation initiation and are not dependent on mRNA degradation [248, 249](Figure 5.18B). mRNA degradation does however still occurs for the majority of targets (due to decreased translation reducing mRNA stability) and correlates with protein expression at later time points as seen by Guo *et al.* [121]. With further validation in mammalian cells [259] these experiments will shed more light on how miRNAs function, in that miRNAs directly target protein translation with mRNA degradation being a secondary effect for the vast majority of targets.

For a gene to be classified as a potential miR-324-5p target in our experiments it had to be decreased by log2 0.2 at an mRNA and protein level, however studies have shown up to 16% of miRNA targets are only decreased at a protein level [121] and therefore such a target would not be classified as a potential target in this study. Here I used both microarray and proteomics meaning it was possible to distinguish between targets regulated at a protein and mRNA level, from targets only regulated a protein level. 30 miR-324-5p seed containing genes whose mRNA and protein were identified decreased at a protein level. Of these, 18 (60%) also decreased at an mRNA level (Figure 5.18C), slightly lower than the 84% percentage suggested by the Bartel lab [121]. 40% of the miR-324-5p seed containing genes which decreased at a protein level did not decrease at an mRNA level (Figure 5.18C) suggesting miR-324-5p can regulate translation without altering mRNA.

The reason why some targets were only decreased at a protein level is unclear. In the traditional model it would have been said the miRNA is targeting translation rather than mRNA degradation (Figure 5.18A). In the new model, where most miRNAs target translation, causing the mRNA to be degraded (due to less protection of RNA by

ribosomes) and a further reduction in translation (due to there being fewer transcripts) [248, 249], it is likely to be because the mRNA mediated reduction in translation does not destabilise the mRNA or reduce the protection by ribosomes.

In an attempt to determine if this type of target was real (or existed in our data) I further analysed our data by investigating the enrichment of miR-324-5p targets in genes whose protein expression could not be explained by changes in mRNA expression (Figure 5.18D-E). There was no enrichment of miR-324-5p seed binding sites in proteins whose expression decreased and their decrease could not be accounted for by a concomitant decrease in mRNA expression (fold change in protein expression/fold change in mRNA expression) (Figure 5.18D). Interestingly, and possibly opposing this, there was a significant enrichment of miR-324-5p seeds in genes whose protein expression decreased but whose mRNA expression changed less than log2 0.05 (Figure 5.18E). There was no enrichment of miR-324-5p seeds in genes whose mRNA decreased but did not also decrease at the protein level (their mRNA decrease cannot be accounted for by a reduction in protein; fold change mRNA expression/fold change protein expression) (Figure 5.18F). Taken together, these data suggested; A) there may have been some targets whose translation was regulated but there was no effect on mRNA, B) there was unlikely to have been targets where the miRNA affects mRNA but does not affect protein. This supports the hypothesis that the majority of miRNAs initially target translation and then depending on the type of target may or may not then lead to a reduction in mRNA [230]. Guo *et al.* concluded from their studies that 84% of miRNA targets were decreased due to mRNA destabilisation [121]. They performed their studies by comparing the total mRNA with ribosomal mRNA following miRNA overexpression. They found little difference in the level of mRNA being translated compared to the total mRNA and eluded (possibly incorrectly) to mRNA destabilisation preceding the reduced translation [121]. Taken together, the more recent data [248, 249], and the work by Guo *et al.* suggests miRNAs work by initially targeting translation and in 84% (60% in our study) of targets there is a subsequent reduction in mRNA.
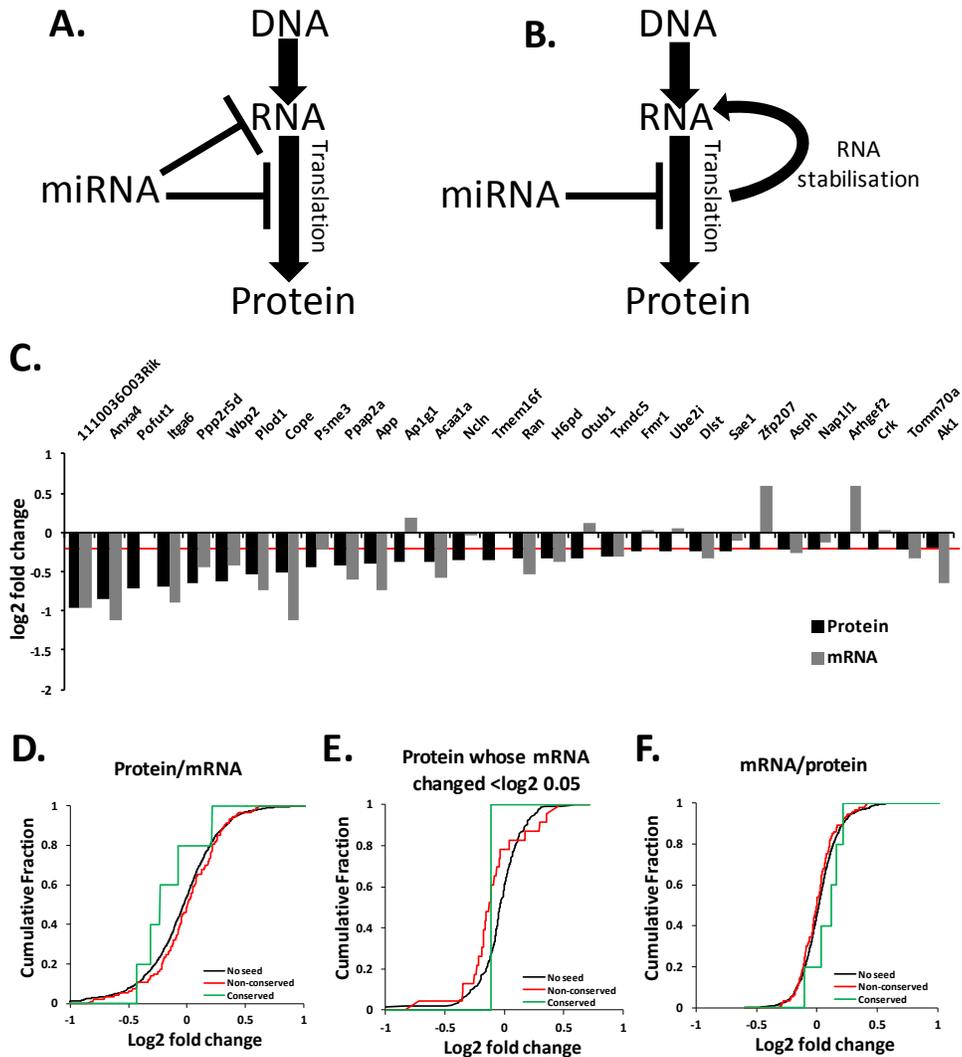
Figure 5.18 miR-324-5p may regulate translation

Figure 5.18 miR-324-5p may regulate translation (A) Traditional dogma of how miRs work, (B) New theory on how miRs work (C) Bar chart showing log2 fold change in protein and mRNA expression for 30 genes whose protein expression decreased following miR-324-5p transfection. Red line indicated log2 fold change 0.2. (D) No enrichment of miR-324-5p seed binding sites in proteins whose expression decreased and their decreased expression cannot be accounted for by the respective decrease in mRNA expression (fc protein expression/fc mRNA expression). Best p value is 0.088 at a fold change of log2 fc 0.18 (E) significant enrichment of miR-324-5p seeds in genes whose protein expression decreased but whose mRNA expression changed less than log2 0.05. Best p value is 0.0007 at a fold change of log2 fc -0.11 (F) No enrichment of miR-324-5p seeds in genes whose mRNA decreased and their mRNA decrease cannot be accounted for by a reduction in protein (fc mRNA expression/fc protein expression). The most significant enrichment is at a fold change of log2 fc -0.09 (p value is 0.08).

5.3.5 Estimate total number of miR-324-5p targets

miRNAs exert their actions by targeting the 3'UTR of many targets. Our target identification strategy did not allow us to give the definite identity of all miR-324-5p targets. It does however allow us to estimate the total number of targets and give the identity of a fraction of these (only with a certain degree of confidence).

During the strategy development around 5.7% (92-18=74, 74/1299=5.7%) of genes which lack miR-324-5p seed binding sites decreased following miR-324-5p overexpression (Figure 5.8), this strongly suggested they are a secondary effects of direct miR-324-5p targets. 19.35% (18/93) of miR-324-5p seed binding site containing genes decreased, this suggested 29.46% (5.7/19.35) of genes whose 3'UTR contains a miR-324-5p binding site and decreased at both an mRNA and protein level are decreased due to secondary effects of a real miRNA targets and just over 70% of the potential targets are real. This equated to just over 5 of the 18 potential targets being false positives, leaving 13 genuine miRNA targets. In this experiment (SILAC1, SILAC2 and unstim array combined), 1,299 genes are identified, which equates to only around 7% of the genome (1,299/18,615). I predicted 13 of the 1,299 are real targets, which would suggest ~185 targets in the whole genome. These calculations assumed a target needs to contain a miR-324-5p seed binding site in its 3'UTR and needs to be decreased log2 0.2 fold at an mRNA and protein level. This of course is not always the case.

5.3.6 Assessment of strategy effectiveness

Many groups use either SILAC or microarray to identify miRNA targets, by using both I was able to assess the relative effectiveness of each. Both the proteomics and microarray approaches enabled us to enrich for miR-324-5p targets (Figure 5.10 and Figure 5.11). However, a combination of both proteomics and microarray gave a greater enrichment than either alone. More precisely I observed a 1.6 fold enrichment using microarray, a 2.7 fold enrichment using the proteomics and a 3.7 fold enrichment of miR-324-5p seed binding site containing genes when both were combined. This suggested SILAC proteomics was more specific than transcriptome microarrays and their combination was more specific than both individually. More targets were enriched in the transcriptome microarrays than the proteomics, meaning

a greater number of potential targets could be generated from the array alone than the proteomics alone, although the certainty of any given target being real will be less.

It is difficult to draw conclusions about the effectiveness of microarrays and SILAC in predicting targets. In order to properly assess the effectiveness of microarrays, SILAC and target prediction algorithms, the selection of genes used to make luciferase constructs should have been independent of any results obtained, independent on whether or not the gene contains a miR-324-5p seed sequence binding site and enough should have been made to gain statistical significance. If this were done a proper independent comparison of genes which decreased with real miR-324-5p targets could have been performed. However, this would have proven too costly and time consuming. Instead, only estimates of effectiveness could be made.

I have shown transcriptome microarrays and SILAC proteomics can form an effective methodology to identify miRNA targets. However neither alone nor in combination are 100% effective, in addition there are a number of other strategies that could have been chosen. In this section I will explore some of the benefits provided by these other strategies and what might benefit this strategy.

5.3.7.1 miRNA inhibitors and target site protectors

Transfection of miRNAs increased their concentration above their normal physiological level. Arguably this may have led to an over-estimate of the number of miRNA targets. The use of antagomirs (miRNA inhibitors) and a search for genes whose expression increased following their transfection may have given a more physiological indication of miRNA targets in these cells [228]. However attempting to decrease miR-324-5p and observing any effects is likely to have proven difficult. Target site protectors for each miRNA-target interaction could also be used to determine if a miRNA-target interaction was in fact functional at physiological levels [260].

5.3.7.2 Not only 3'UTR targets

In addition to being decreased by log2 0.2 at an mRNA and protein level, a gene also needed to contain a miR-324-5p seed sequence binding site in its 3'UTR in order for it be considered as a potential target in this study. Genes which lack a full seed binding site in their 3'UTR have also been shown to be real miRNA targets [127]. In addition miRNAs can also target the 5'UTR or coding sequence of genes [261]. However, I

decided not to take types of targets into account as the vast majority of miRNAs act on the 3'UTR of genes [242].

5.3.7.3 Secondary effects

Here I was searching for direct miRNA targets which can positively affect Hh signalling, in order to explain miR-324-5p inhibition of the pathway. It is possible miR-324-5p may regulate Hh signalling through the secondary effects of miR-324-5p. Due to the long 72h post transfection time point chosen many of the secondary effects of the miRNA would be obtainable from this data. A downside of such a long time point is some real direct and indirect targets may have returned back to normal expression by 72h, meaning even though their expression changed they would not be identified. However in other studies we are able to measure sustained overexpression following transfection for up to 14days. To elucidate a mechanism by which a miRNA acts, all of the miRNA effects should be considered not only its perceived direct targets.

5.3.7.4 Non traditional actions of miRNAs

In a rare number of cases miRNAs can actually increase the expression of their target, possibly by recruiting transcription factors [261]. If this type of target was responsible for the miR-324-5p effect on Hh signalling, a negative regulator of Hh signalling should have been sought after.

As well as binding mRNA, it is also possible for miRNAs to interact with receptors, such as Toll-like receptors (TLRs)[262, 263]. TLR7 and 8 preferentially bind miRNAs rich in GU at nucleotides 18-21[263], something which the miR-324-5p sequence contains (cgcauccccuagggcau**uggu**gu). However further work is needed to determine if miRNAs with UGGU at nucleotides 18-21 or if miR-324-5p itself can bind TLR receptors. Interestingly, activation of TLR3 inhibits Shh induction of Hh signalling [264], indicating miRNAs may regulate Hh signalling through direct binding of TLRs. Both TLRs and miRNAs have been found in exosomes [263], suggesting they have the potential to interact.

5.3.7.5 Strengths and weaknesses of strategy used

Although my strategy is effective at identifying miRNA targets in C3H10T1/2 cells, it is not necessarily going to be effective or practical for all miRNA target identification studies. This strategy is expensive, mainly due to the cost of culturing the cells in $^{13}C_6$ $^{15}N_4$-Arg and $^{13}C_6$-Lys labelled medium and the cost of enough mass spectrometry runs to allow sufficient coverage of the proteome. The strategy is also unlikely to work in cells which either do not divide or divide slowly. I have also been unable to avoid the bias of a miRNA needing a seed binding site, ideally decreased genes should be considered as potential targets irrespective of seed binding sites but then it would be too hard to distinguish between direct and indirect effects especially as I analysed my cells quite late (72h post-transfection).

miRNA pull-down, or HITS/PAR-CLIP would have offered an alternative strategy, and may have increased the ratio of direct to indirect targets identified, but will still face many of the same problems as proteomics and transcriptomics. Interestingly, a strategy based on literature and computational analysis rather than experimental profiling would have been cheaper (although not necessarily less time consuming) and may have still been effective in determining targets which have a potential role in Hh signalling, although it is unlikely to have identified Gpc1.

In my opinion the use of transcriptome microarray following both over and under expression of a miRNA combined with online prediction sites would be the most cost and time effective for a general miRNA target identification strategy. However, depending on the project context other strategies should also be considered, for example if miRNA and either mRNA or protein expression is already known, a computational approach looking for inverse correlations could be used. Due to the speed at which this field is moving my opinion on this is likely to quickly change.

5.4   Summary

The main aim of this chapter was to identify miR-324-5p targets in mouse which have a possible role in Hh signalling. Due the large number of predicted miR-324-5p targets, there were many predicted miR-324-5p targets involved in Hh signalling. However, due to the large variation and unreliability of prediction programs it was unlikely any given predicted miR-324-5p target is a real miR-324-5p target. For that reason I developed a

strategy to identify miRNA targets. I went on to assess the strengths and weaknesses of the strategy and employed it to identify miR-324-5p targets. More specifically I have identified Gpc1 as a miR-324-5p target which I believe to have a role in Hh signalling and may be involved in the mechanism by which miR-324-5p inhibits Hh signalling in mouse C3H10T1/2.

Chapter 6 Novel miR-324-5p target, Gpc1, regulates Hh

6.1   Introduction

miR-324-5p was increased in OA hip cartilage (Figure 4.1) and has previously been shown to regulate the Hh signalling pathway in humans by targeting the Hh pathway activator SMO and transcription factor GLI1 [127]. I have shown miR-324-5p regulation of Hh is conserved in mouse, yet the mechanism by which it does so is not (Chapter 4). In Chapter 5 I identified Gpc1 as a mouse miR-324-5p target and hypothesised it may play a role in Hh signalling and form part of the mechanism by which miR-324-5p regulates Hh signalling in mouse.

6.1.1   Glypicans

Gpc1 is a member of the Glypican family of heparin sulphate proteoglycans (HSPGs). In mouse there are 6 different Glypicans (1-6), and each has a homologue in human. The 6 Glypicans can be divided into two groups; Gpc1/2/4/6 and Gpc3/5 [265]. There are only two Glypicans in Drosophila, Dally and Dally-like protein (dlp), which are thought to be homologues of Gpcs 3/5 and Gpcs 1/2/4/6 respectively [265]. The mammalian Glypicans 1/2/4/6 can be further subdivided into two groups Gpc1/2 and Gpc4/6. Interestingly a number of miRNAs are encoded from within close proximity to the Glypicans, which may be involved in similar processes (General discussion). Glypicans are anchored to the cell membrane by a Glycophosphatidylinositol (GPI) anchor which can be cleaved by an enzyme called Notum in mammals [266].

6.1.2   Glypicans can positively regulate Hh

HSPGs have previously been shown to regulate a number of morphogens to produce many different types of morphogen gradient, for a review see [267]. The role of HSPGs in Hh signalling has been known for over 10 years. The addition of heparin, changes in the composition of PGs, and some of the enzymes involved in HSPG synthesis, can all alter Hh signalling [268-270].

Much of the work involving Glypicans in Hh signalling has been performed in Drosophila. In 2003 a group used RNAi to screen for genes involved in Drosophila Hh signalling and identified the Drosophila Glypican homologue Dally-like protein (dlp) [271]. Further investigation showed dlp plays a critical role in Hh signalling, either at the level of patched (ptc) or up stream of ptc [250], and may also contribute to the

movement of Hh ligand from the Hh producing cell to the Hh receiving cell. Dlp is required in Hh receiving cells for Hh signal transduction [251, 252]. More specifically, Glypicans are thought to contribute to the internalisation of Hh and Ptch [268], as dlp is often found with hh and ptc in endocytic vesicles [252]. Further evidence is suggested by Yan *et al*. for the role of dlp as a Hh co-receptor [272].

6.1.3 Glypicans can negatively regulate Hh

As well as positively regulating Hh, some Glypicans have been shown to negatively regulate Hh signalling. Gpc3 can interact with Shh, causing the endocytosis and lysosomal degradation of Shh [246]. As mentioned previously, dlp also aids in the endocytosis of Hh ligand, yet is a positive regulator of Hh signalling. One crucial difference appears to be that unlike dlp, Gpc3 does not require Ptch for Hh ligand internalisation. In drosophila, dlp activates Hh signalling by removal of ptc from the membrane leading to active smo, where as in mouse, Gpc3 will act to endocytose Hh ligand without Ptch, leaving Ptch on the membrane and still able to inhibit Smo. It is suggested Gpc3 contributes to the degradation of Hh ligand in a similar way to what has previously been described for Ptch [273].

Recently a study has suggested Gpc1 to have a role in zebrafish Hh signalling. Depletion of Gpc1 in zebrafish causes developmental defects, and increased expression of Hh signalling target genes such as *gli2a* and *ptch1* [274, 275]. The addition of cycopamine (a Hh antagonist) to these fish partially rescues the phenotype, and injection of Shh gave a similar phenotype to Gpc1 depletion [274], suggesting Gpc1 plays a negative role in Hh signalling, possibly acting as a sink for Hh ligand. Zebrafish Gpc1 is likely to be the ortholog of human and mouse Gpc1 as there are 6 Gpcs and synteny in both species [274].

6.1.4 HS chains and GPI anchor

Glypicans and other proteoglycans can regulate Hh signalling in both a positive and negative manner and exert much of their action though their negatively charged HS chains interacting with positively charged Na^+ ions, H_2O and morphogens [276]. HS-GAG chains form on the serine residues in Ser-Gly sequences, of which Gpc1 has 3. HS chain formation is catalysed by a number of enzymes which sequentially add sugar residues to the non-reducing end of the growing chain [276]. The insertion sites for the HS chains are located close to the carboxyl terminus suggesting these chains are close

to the cell surface and may mediate interactions with molecules on the cell surface [276]. EXT family proteins then catalyse the HS polymerisation process by alternatively adding GlcA and GlcNAc residues. During the chain polymerisation process, a number of modifications take place, including sulphation [276]. It is possible to inhibit the formation of these chains with the addition of sodium chlorate, which reduces the amount of sulphate in HS GAG chains, impairing their synthesis and ability to function [277]. Sodium chlorate works by inhibiting ATP-sulphurylase, the first enzyme involved in the biosynthesis of 3'-phosphoadenosine 5'-phosphosulphate (PAPs), which are the active form of sulphate and sulphur donor in protein sulphation [278]. The addition of sodium chlorate to cells will show the function of HSPGs, although it has been reported Gpc3 core protein can also directly interact with Hh ligands [246].

The GPI anchor on glypicans is thought to play a role in the endocytosis of morphogens by acting as an endocytic signal [246, 252]. Notum, whose function is to cleave the GPI anchor [266], has itself been shown to be involved in Hh signalling [279].

I hypothesised miR-324-5p regulates Hh signalling through a mechanism involving Gpc1. The main aim of this chapter is to investigate the role of Gpc1 in Hh signalling.

Specific aims

Aim 1: Determine if HSPGs are important for Hh signalling

Aim 2: Determine the effect of Gpc1 depletion on Hh signalling

Aim 3: Determine the effect of Gpc1 overexpression on Hh signalling

Aim 4: Elucidate the mechanism by which Gpc1 regulates Hh signalling

Aim 5: Compare the mechanisms by which miR-324-5p regulates Hh signalling in human and mouse

Figure 6.1 Figure showing schematic structure of Glypican1
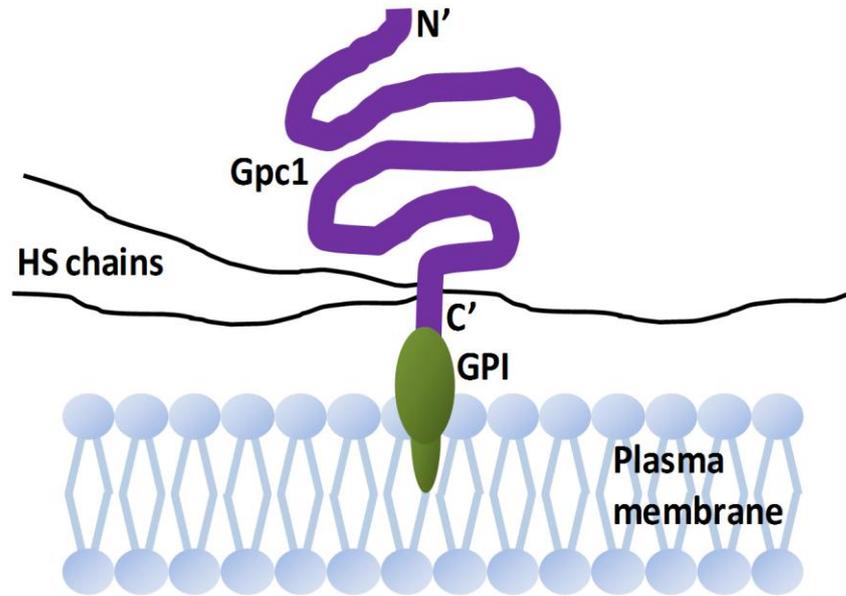

Figure 6.1 Structure of Glypicans (Gpc1). Gpc1 is attached to plasma membrane via Glycophosphatidylinositol (GPI) anchor. Heparin sulphate (HS) chains form near the C' terminal of Gpc1.

6.2 Results

6.2.1 Heparin Sulphate chains play a role in Hh signalling

Since Gpc1 is a HSPG I therefore wanted to determine if HSPGs are important for Hh signalling in C3H10T1/2 cells. HSPGs exert much of their function through their HS chains. Sodium chlorate can inhibit HS chain formation [277]. Pre-treatment of cells with 2.5mM sodium chlorate for 48h decreased the level of Ihh (2µg/ml for 24h) induced Hh signalling (Figure 6.2A). HS chains chelate morphogen, therefore their lack, or lack of HSPG, would reduce morphogen levels. Consistent with this in my MS analysis of Hh signalling where I identified miR-324-5p targets (Chapter 5, SILAC3), I observed less Ihh in miR-324-5p transfected cells than in control transfected cells (Figure 6.2B). This was consistent for all seven of the peptides identified for Ihh.

Figure 6.2 Heparin sulphate plays a role in Hh signalling
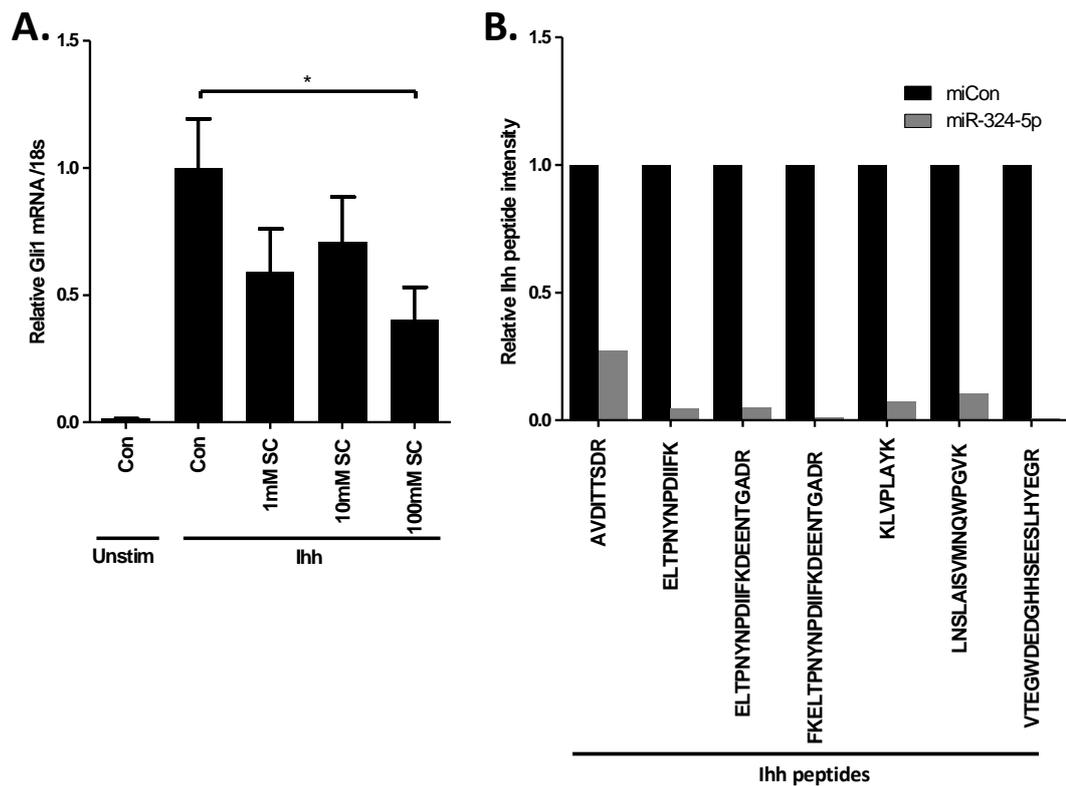
Figure 6.2 Effect of HS chain depletion on Hh signalling. (A) C3H10T1/2 cells were treated with Sodium chlorate (SC) for 48h at the concentration shown then serum starved and stimulated with Ihh (2µg /ml) for 24h. Data combined from 2 independent experiments, each n=3. Statistical differences were calculated using ANOVA followed by a Bonferroni post test, $p < 0.05 =$*. (B) Relative Ihh protein levels detected within cell lysates by mass spectrometry (SILAC3) following transfection of either miCon or miR-324-5p for 24h as indicated, serum starvation for 24h and stimulation with Ihh (2µg/ml) for 24h.

6.2.2 Depletion of Gpc1 inhibits Hh signalling

To test if Gpc1 was important for Hh signalling, C3H10T1/2 cells were transfected with siRNA against Gpc1 (siGpc1). As expected, and similar to miR-324-5p transfected cells, Gpc1 mRNA and protein was reduced in siGpc1 transfected cells (Figure 5.14). Similar to stimulation of siGli1 (Figure 6.3) and miR-324-5p (Figure 4.14) transfected cells, stimulation of siGpc1 transfected cells with Ihh resulted in less or no induction of Gli1 mRNA or protein in comparison to siCon transfected cells (Figure 6.3), indicating Gpc1 is required for Ihh signalling. In contrast to siGli1 (Figure 6.3), but similar to miR-324-5p (Figure 4.14) siGpc1 did not cause a reduction in basal Gli1 mRNA (Figure 6.3A).

Again, consistent with miR-324-5p (Figure 4.16), but in contrast to siGli1 (Figure 6.4A), transfection of siGpc1 decreased Ihh and Bmp2 induced alkaline phosphatase (Figure 6.4A). Also consistent with miR-324-5p (Figure 4.17) and in contrast with siGli1 (Figure 4.13), siGpc1 did not cause a significant difference in Bmp2 or Bmp2 + Ihh induced alcian blue (Figure 6.4B). siGpc1 did however increase basal levels of alcian blue (Figure 6.4B).

Figure 6.3 Effect of Gpc1 depletion on Hh signalling
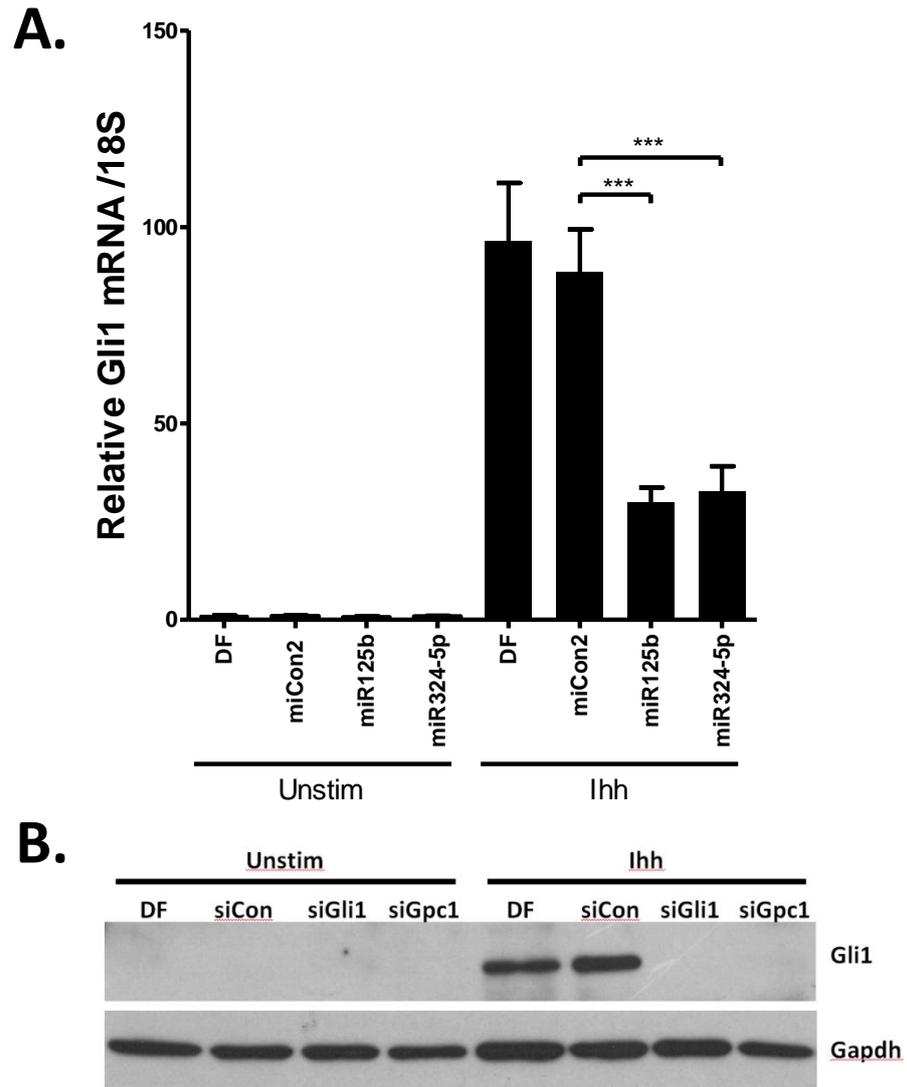
Figure 6.3 Effect of Gpc1 depletion on Indian Hedgehog signalling. C3H10T1/2 cells were transfected with non-targeting siCon or siRNA against Gpc1 or Gli1 for 24h. Cells were then serum starved for 24h and either left unstimulated, stimulated with Ihh (2μg/ml) for 24h. Gli1 and 18S were assessed by real-time RT-PCR. Gli1 and Gapdh protein expression were assessed by Immunoblotting (validation of Gpc1 siRNA is shown in (Figure 5.14). (A) siGli1 and siGpc1 effect on basal and Ihh stimulated Gli1 mRNA expression. Data combined form 4 independent experiments, each n=4 Statistical differences were calculated using ANOVA followed by a Bonferroni post test where $p<0.01=$** and $p<0.001=$***. (B) siGli1 and siGpc1 effect on basal and Ihh stimulated Gli1 protein expression. Representative blot of 3 independent experiments.

Figure 6.4 siGpc1 on alkaline phosphatase and aclian blue
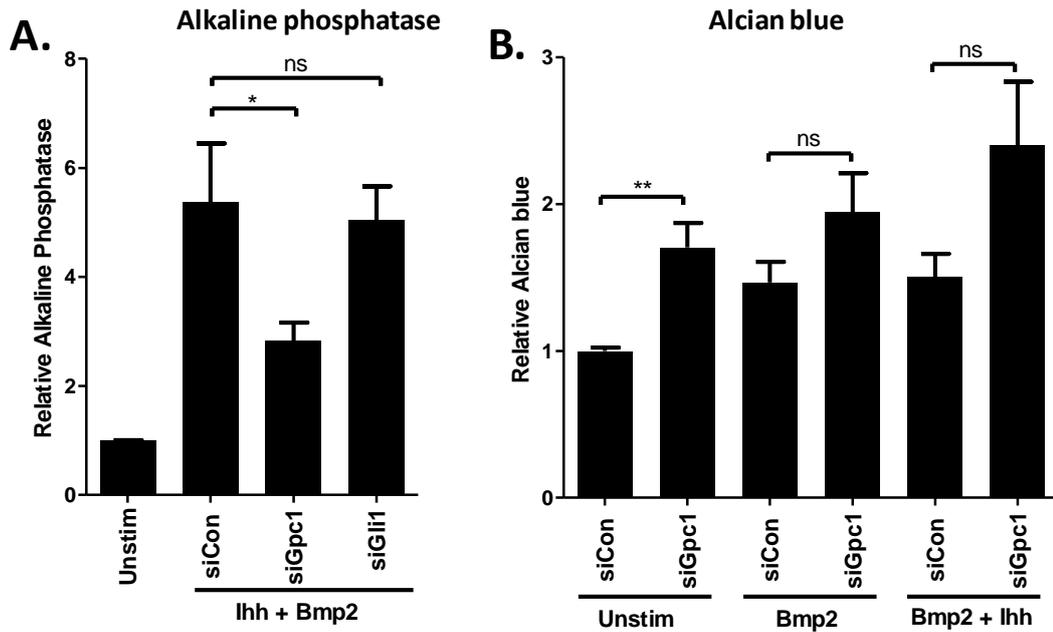
Figure 6.4 Effect of Gpc1 depletion on osteoblastogenesis and chondrogenesis. (A) Alkaline phosphatase was used as a markers of bone formation, C3H10T1/2 cells were transfected with siCon, siGpc1 or siGli1, then stimulated with Ihh (2µg/ml) and BMP2 (100ng/ml) for 5 days. P.Nitrophenol was measured to determine the level of alkaline phosphatase. Data are combined data from 3 independent experiments each n=4, SEM. (B) Alcian blue was used as a marker of cartilage formation. C3H10T1/2 cells were transfected with siCon, siGpc1 or siGli1 then plated into micromass cultures, stimulated with Bmp2 (100ng/ml) and/or Ihh (2µg/ml) alcian blue stain was extracted with GuHCl and quantified, data are combined data from 3 independent experiments each n=2, SEM. Statistical differences were calculated using ANOVA followed by a Bonferroni post test where p<0.05=*and p<0.01=**.

6.2.3   Gpc1 overexpression inhibits Hh signalling

Because a reduction in Gpc1 (siGpc1) reduced Hh signalling, I hypothesised an increase in Gpc1 would increase Hh signalling. I therefore created a Gpc1 overexpression plasmid (see methods). Transfection of my Gpc1 overexpression plasmid increased Gpc1 protein, as detected by western blot (Figure 6.5B and C). In contrast to my hypothesis, overexpression of Gpc1 reduced basal Gli1 mRNA and slightly (not significantly) reduced Ihh and Pur induced Gli1 mRNA (Figure 6.5A). Overexpression of Gpc1 also reduced basal Gli1 protein (Figure 6.5C). This data indicated optimal levels of Gpc1 are required for Hh signalling.

Figure 6.5 Increased Gpc1 expression inhibits Hh signalling
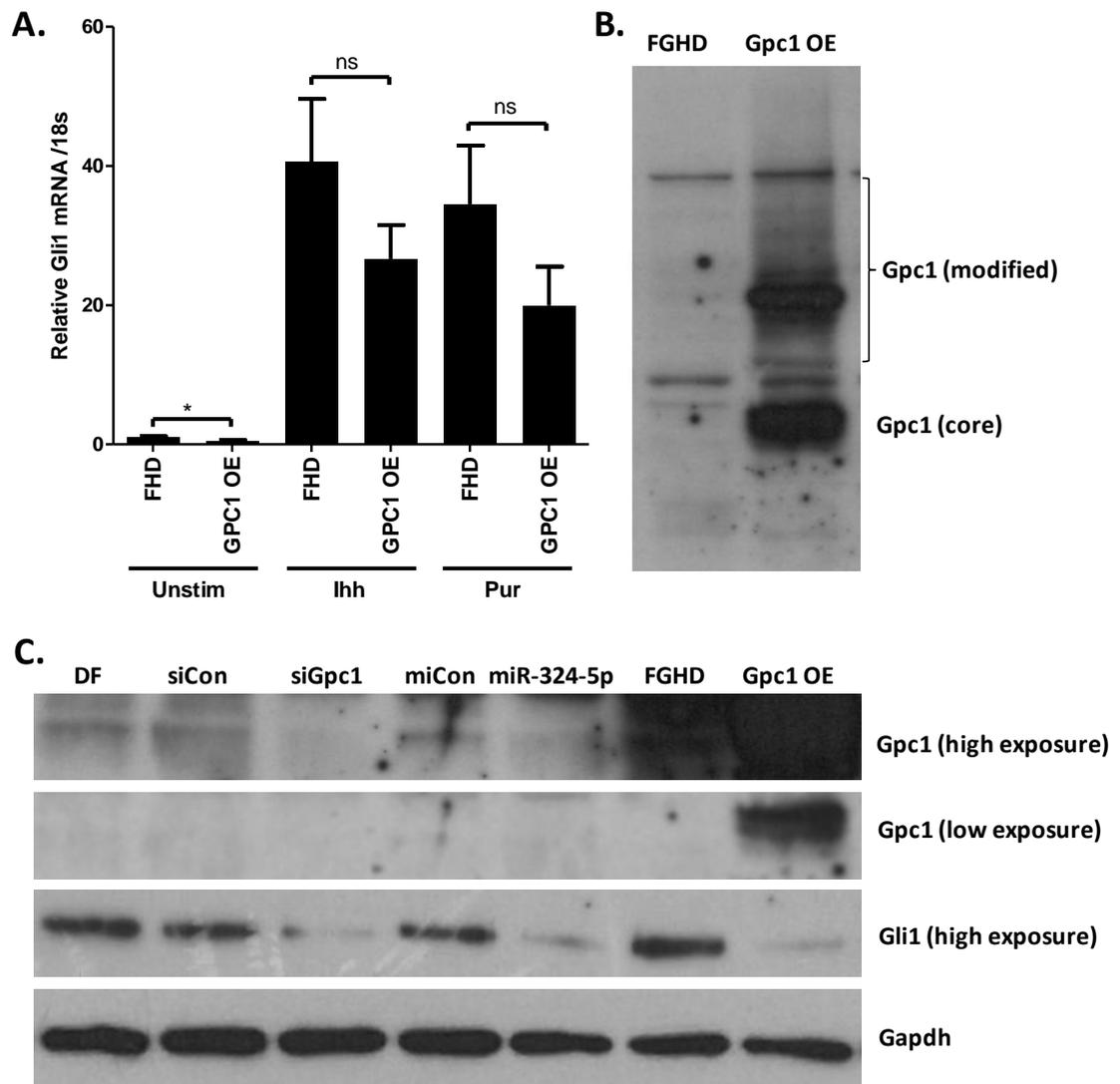
Figure 6.5 Increased Gpc1 expression inhibits Hh signalling. (A) Effect of Gpc1 overexpression (OE) on Gli1 mRNA compared with Fugene HD transfection reagent alone (FHD), following serum starvation and stimulation with Ihh (2µg/ml) or Pur (2µM) for 24h. Data combined from two independent experiments each n=4. Statistical difference were calculated using Student's t-test, where *p<0.05. (B) Example Western blot showing Gpc1 overexpression (OE) compared with Fugene HD transfection reagent alone (FGHD). (C) Representitive Western blot of 3 independent experiments showing Gpc1 overexpression increases Gpc1 protein expression and decreases Gli1 protein compared with Fugene HD transfection reagent alone (FGHD). Blot also shows effect of Dharmafect transfection reagent (DF), siCon, siGpc1, miCon and miR-324-5p. At low exposure no Gli1 protein was detected.

6.3 Discussion

As hypothesised, siRNA depletion of Gpc1 reduced Ihh induced Hh signalling (Figure 6.3 and Figure 6.5C) and phenocopied the effect of miR-324-5p (Figure 4.14 and Figure 6.5C). This suggested the mechanism by which miR-324-5p regulated Ihh induced Hh signalling in mouse was at least in part, by directly targeting Gpc1 (Figure 6.6). In contrast to siGli1, but similar to miR-324-5p (Figure 4.14), siGpc1 did not cause a reduction in basal Gli1 mRNA (Figure 6.3), suggesting Gpc1 was not involved in maintaining basal levels of Gli1 mRNA, but is involved in Ihh induction of Gli1 expression. This further suggests the mechanism by which miR-324-5p regulated Hh in mouse is through targeting Gpc1 and not Gli1. In addition, siRNA against Gpc1 phenocopied many of the downstream effects of miR-324-5p, where as siRNA against Gli1 did not, this reinforced the hypothesis that miR-324-5p functioned through Gpc1 rather than Gli1 (Figure 6.6).

Figure 6.6 miR-324-5p regulation of Hh in mouse
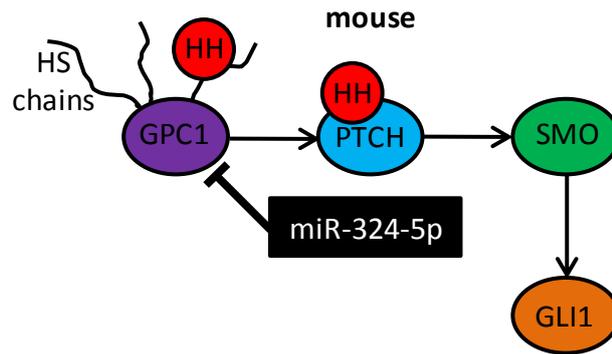
Figure 6.6 miR-324-5p regulation of Hh in mouse. Gpc1 plays a role in Hh signalling and is likely to act via its HS chains to effect the reception of Hh ligand. Gpc1 is a target of miR-324-5p in mouse, meaning miR-324-5p is likely to inhibit Hh signalling through its action on Gpc1. This model may also explain why siGli1 (Figure 6.3) but not miR-324-5p (Figure 4.14) or siGpc1 (Figure 6.3) can regulate basal Gli1 levels and why miR-324-5p, siGpc1 and Gli1 can all regulate Ihh induced Gli1.

6.3.1 Possible mechanisms by which Gpc1 depletion inhibits Hh signalling

The mechanism by which reduced Gpc1 inhibited Hh signalling was not completely clear, however it is possible Gpc1 in this system acts in a similar way to dlp in Drosophila [252], or Gpc5 in mouse [247], which are involved in the internalisation of hh/Shh with ptc/Ptch respectively. This is supported by the data presented in Figure 6.2B, where it is likely the miR-324-5p induced loss of Gpc1 is responsible for the loss of Ihh internalisation and the subsequent decrease in Ihh present in lysate, as measured in SILAC3. I hypothesised that the siGpc1/miR-324-5p reduction of Gpc1 contributed to the reduction of Hh signalling by both decreasing the concentration of Ihh at the cell surface (there are fewer HS chains to bind Ihh (Figure 6.7A) and also preventing the internalisation of Ihh with Ptch (as there is no Gpc1 to facilitate the formation of a Hh:Gpc1:Ptch complex and mediate vesicle endocytosis (Figure 6.7B).

Figure 6.7 Decreased Gpc1/miR-324-5p reduces Hh signalling
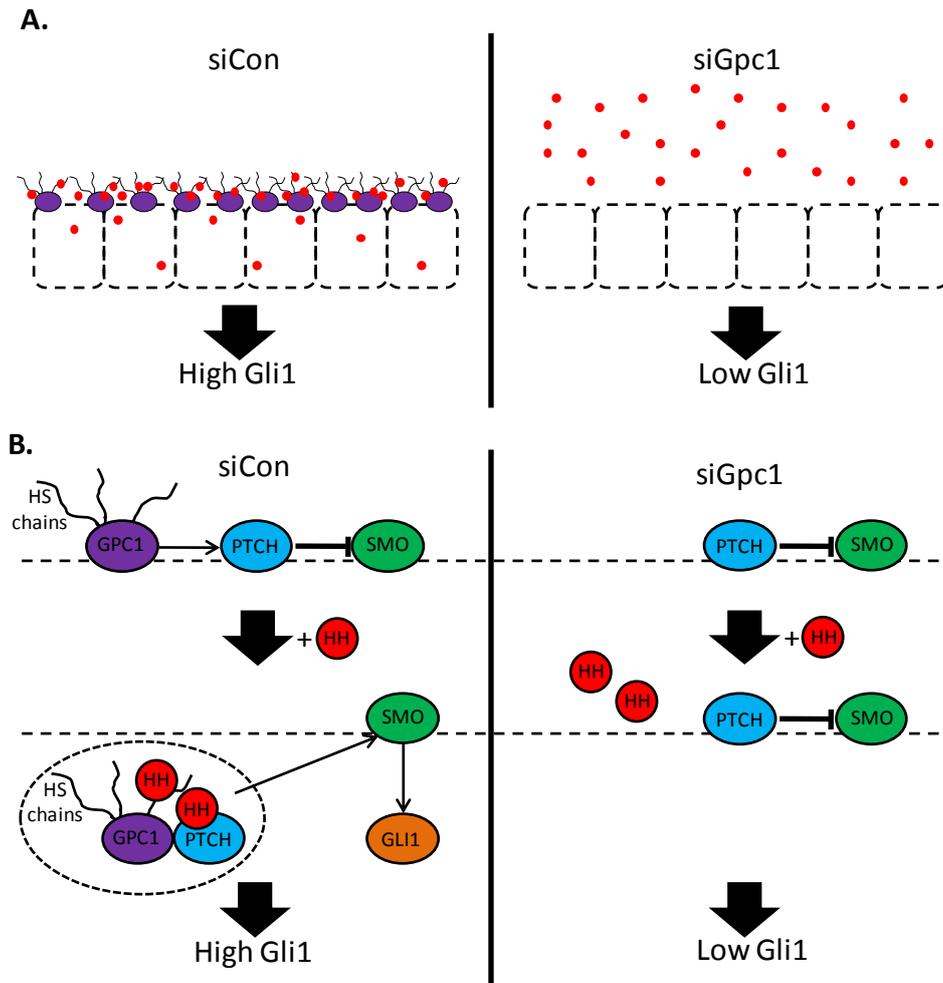
Figure 6.7 Schematic showing possible mechanisms by which decreased Gpc1/miR-324-5p reduces Hh signalling. (A) Left; normal levels of Gpc1, Gpc1 increases the concentration of Hh ligand at the cell surface. Right; reduced levels of Gpc1 (siGpc1/miR-324-5p), Hh ligand does not accumulate at the cell membrane and is lost in the media. (B) Left; normal levels of Gpc1, Gpc1 facilitates and stabilises the interaction between Hh ligand and Ptch. Gpc1 binds Hh ligand. Gpc1 is endocytosed with Hh ligand and Ptch (Hh:Gpc1:Ptch complex), meaning Ptch is removed from the membrane, leaving active Smo, to trigger downstream Hh signalling. Right; reduced levels of Gpc1 (siGpc1/miR-324-5p), Hh ligand has reduced binding to Ptch meaning Hh signalling remains inactive.

6.3.2 Possible mechanisms by which Gpc1 overexpression inhibits Hh signalling

Unexpectedly, Gpc1 overexpression also reduced the level of Hh signalling (Figure 6.5), indicating excessive Gpc1 may act as a negative regulator of Hh signalling and illustrating the importance for tight Gpc1 regulation. I hypothesised this could be due to a number of reasons.

Hypothesis A, the primary cilium is an essential cell organelle for Hh signalling [84], therefore I wanted to determine if the mechanism by which Gpc1 overexpression decreased Hh signalling is by regulation of the primary cilia. Gpc1 was overexpressed and the number of Gpc1 positive cells with primary cilia was compared to the number of Gpc1 negative cells with primary cilia. I found Gpc1 overexpression decreased the number of primary cilia, more precisely around 83% of cells not transfected had a detectable primary cilium, compared to around 27% of those transfected (Figure 6.8), perhaps suggesting Gpc1 overexpression reduced Hh signalling by decreasing the number of primary cilia, meaning Ihh was unable to activate signalling in these cells (Figure 6.9A). The reason why Gpc1 overexpression disrupted the primary cilia was not investigated, however Gpc1 maybe be involved in the correct localisation of the proteins required for ciliogenesis. Alternatively the lack of primary cilia may be due to increased cell stress, from transient transfection of the Gpc1 overexpression construct, or due to the cell being at different points in the cell cycle, both are yet to be examined. However, due to the low transfection efficiency and low number of cells which will lack primary cilia, it is unlikely primary cilia disruption alone can account for the reduction in Hh signalling following Gpc1 overexpression.

Figure 6.8 Gpc1 disrupts primary cilia
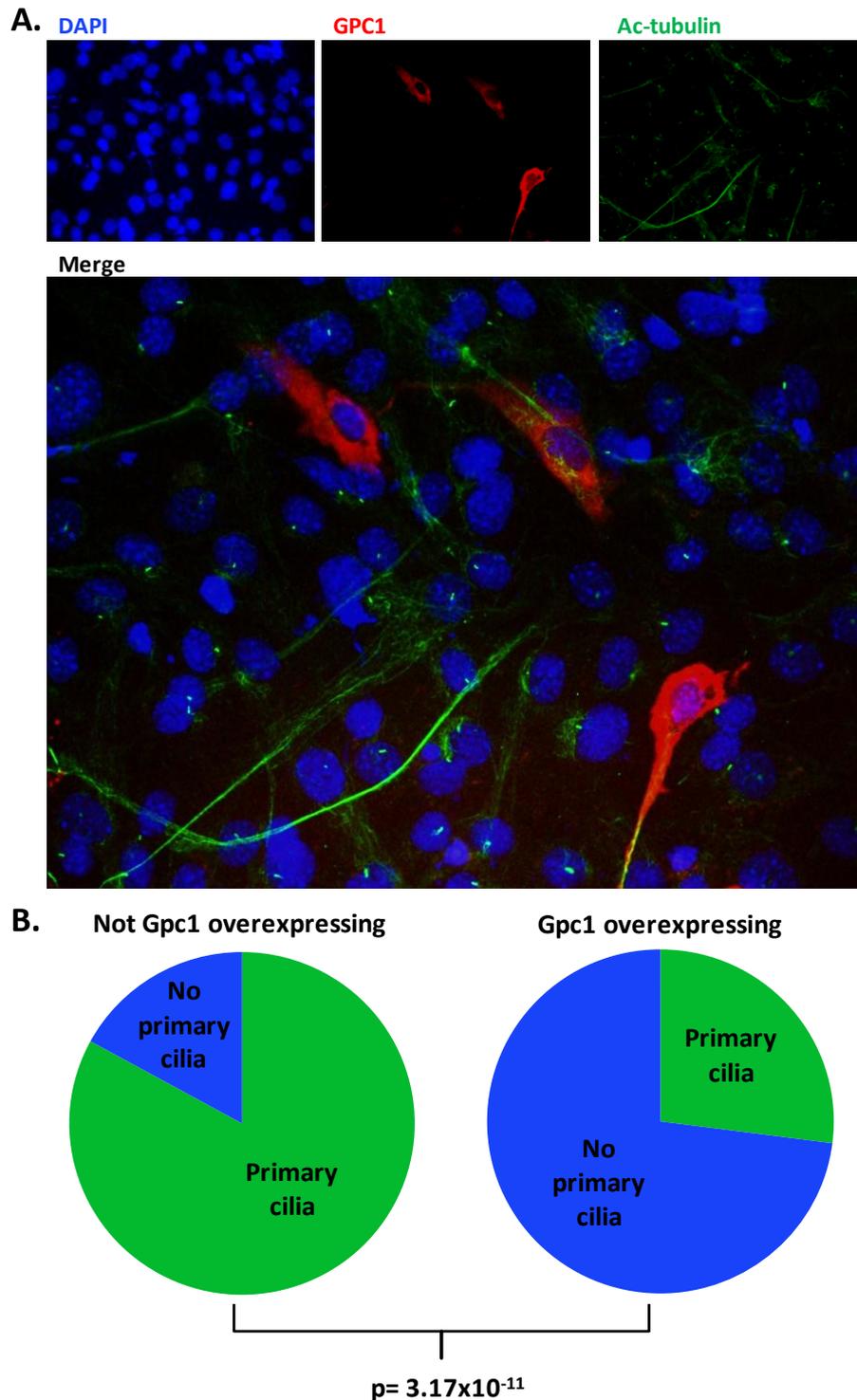
Figure 6.8 Gpc1 disrupts the primary cilium. (A) Example immunoflourescence image showing Gpc1 overexpressing cells lack primary cilia. (B) Combined data from two independent experiments showing the number of Gpc1 overexpressing cells with a primary cilia is less than the number of cells not overexpressing Gpc1 with a primary cilia, 112/135 Gpc1 negative have primary cilia and 10/37 Gpc1 positive cells have primary cilia (fishers exact test, p=3.17x10^{-11}).

Hypothesis B, the amount of morphogen, both *in vivo* and *in vitro*, is finite, therefore for every increase in morphogen there must also be a decrease in morphogen at other locations. The transfection efficiency to overexpress Gpc1 was low (Figure 6.8), meaning the majority of cells were not transfected and did not overexpress Gpc1. It was therefore possible almost all of the Ihh ligand accumulated at and activated Hh signalling in only the small number of transfected cells and was unavailable to the majority of cells. Then when cells were lysed and analysed together there was an apparent overall reduction in Hh signalling in the cells (Figure 6.9B-right) compared to mock transfected cells (Figure 6.9B-left).

Hypothesis C, in the previous section I hypothesised Gpc1 facilitated the formation of a Hh:Gpc1:Ptch complex in vesicles to activate Hh signalling, however I had no evidence for this, it is therefore unknown if Gpc1 can endocytose Hh ligand with or without Ptch, meaning it is possible Gpc1 may actually endocytose Hh without Ptch (Gpc1:Hh complex) (Figure 6.9C). If this were the case it is likely overexpression of Gpc1 inhibited Hh signalling as Hh was unable to bind and activate Ptch, similar to the way Gpc3 negatively regulates Hh in mice [246]. Both the second and third hypothesis suggest overexpressed Gpc1 acts as a sponge to either bind Hh on the 'wrong cell' or internalise it in the 'wrong vesicle'. The type of complex formed with Gpc1 may be dependent on the level of Gpc1. It is possible basal levels Gpc1 promotes the formation and endocytosis of a 'Gpc1:ptch:Hh' complex (akin to Dlp model of Hh activation signal [252]) (Figure 6.7B-left), at high Gpc1 levels it promotes the formation and endocytosis of a 'Gpc1:Hh' complex (akin to Gpc3 inhibition of Hh signalling [246]) (Figure 6.9C-right) and when Gpc1 levels are reduced below basal no complex forms (Figure 6.7B-right). The reason for high levels of Gpc1 favouring a 'Gpc1:Hh' complex over a 'Gpc1:ptch:Hh' complex is unknown, casting doubt over this model.

Hypothesis D, I have previously shown high levels of both Pur and SAG are inhibitory to the Hh pathway (Chapter 4). It is therefore possible Gpc1 overexpression actually increased Hh:Ptch interactions but, in doing so, activated negative feedback loops (Figure 6.9D), similar to those I hypothesised to explain high Pur/SAG being inhibitory in Chapter 4.

Figure 6.9 Increased Gpc1 reduces Hh signalling
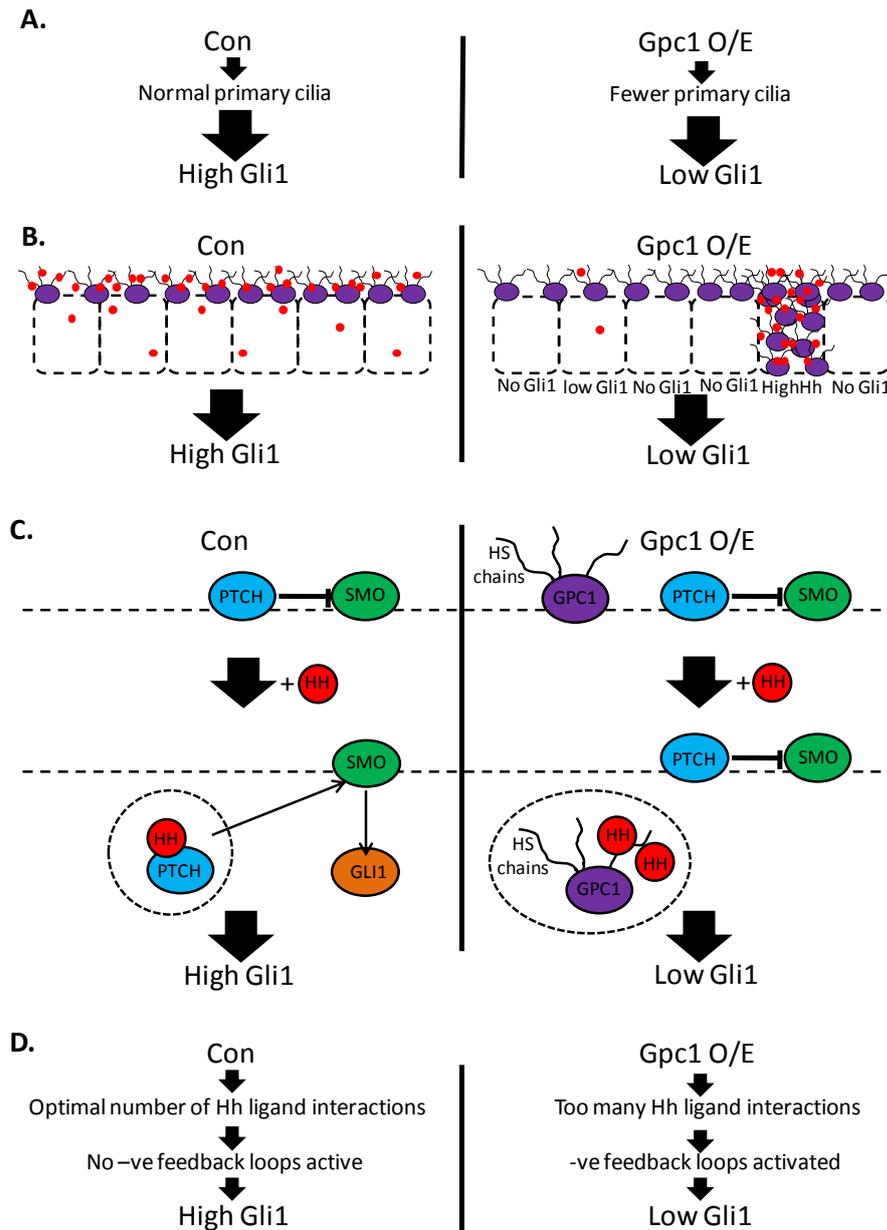
Figure 6.9 Schematic showing possible mechanisms by which increased Gpc1 reduces Hh signalling. (A) Gpc1 overexpression (O/E) reduces the number of primary cilia. (B) Left; normal levels of Gpc1, Gpc1 increases the concentration of Hh ligand at the cell surface. Right; Gpc1 overexpression, Hh ligand will accumulate at and activate Hh signalling in only the small number of transfected cells and is unavailable to the majority of cells, meaning there will be an apparent overall reduction in Hh signalling in the cells. (C) Left; normal levels of Gpc1, the lack of high Gpc1 allows Hh ligand and ptch to interact and are endocytosed, leaving active Smo, to trigger downstream Hh signalling. Right; increased levels of Gpc1, Hh ligand interacts and is endocytosed with Gpc1 meaning Hh ligand has reduced binding to Ptch and leaving Hh signalling inactive. (D) High Gpc1 increases Hh:ptch interactions but activates negative feedback loops.

6.3.3 miR-324-5p and Gpc1 have activator and repressor properties and may have differing roles in high and low Hh signalling

Gpc1 depletion and overexpression reduces the level of Hh signalling, illustrating the importance of tight Gpc1 regulation. HSPGs have previously been shown to have positive and negative effects on cell signalling. Gpc3 can interact with Wnt to promote Wnt signalling [280], and also directly interact with Shh, to inhibit Hh signalling by influencing the endocytosis and lysosomal degradation of Shh [246]. In addition increased miR-324-5p also reduced Hh signalling (via decreased Gpc1). I therefore hypothesised decreased miR-324-5p would also inhibit Hh (due to increased Gpc1). Indeed preliminary data suggested miR-324-5p inhibition decreased Hh signalling (Figure 6.10A), although the effect of miR-324-5p inhibition on Gpc1 has not been examined. These data suggested miR-324-5p is likely to be inversely correlated with Gpc1 and both an increase and decrease in either miR-324-5p or Gpc1 leads to decreased Hh signalling (Figure 6.10B).

As well as being both positive and negative regulators of signalling pathways, Glypicans can also have differing effects on high and low levels of Hh signalling [279], for example Dally and the GPI cleavage enzyme Notum are involved in high levels Hh signalling, but are dispensable for low levels of Hh signalling [279]. Similarly I have shown siGpc1 does not affect basal levels of Gli1 mRNA, but does effect Ihh induced Gli1 mRNA. I have however shown siGpc1 decreases both basal and Ihh induced Gli1 protein (Figure 6.3B and Figure 6.5C). In contrast, Cui *et al*. have very recently shown Gpc1 depletion to promote basal Hh signalling in zebrafish [274], although the system I examined is obviously different. In addition, the Hh response genes of low, medium and high level signalling are different [279, 281], suggesting Gpc1 may control some Hh response genes and not others. For example dpp, which is the Drosophila homologue for Bmp, is responsive to low levels of Hh signalling [281], and as Gpc1 depletion does not appear to effect basal levels of Hh (Figure 6.3) this may partly explain the differing effects of siGpc1 on alkaline phosphatase and alcian blue.

Figure 6.10 miR-324-5p and Gpc1 have activator and repressor properties
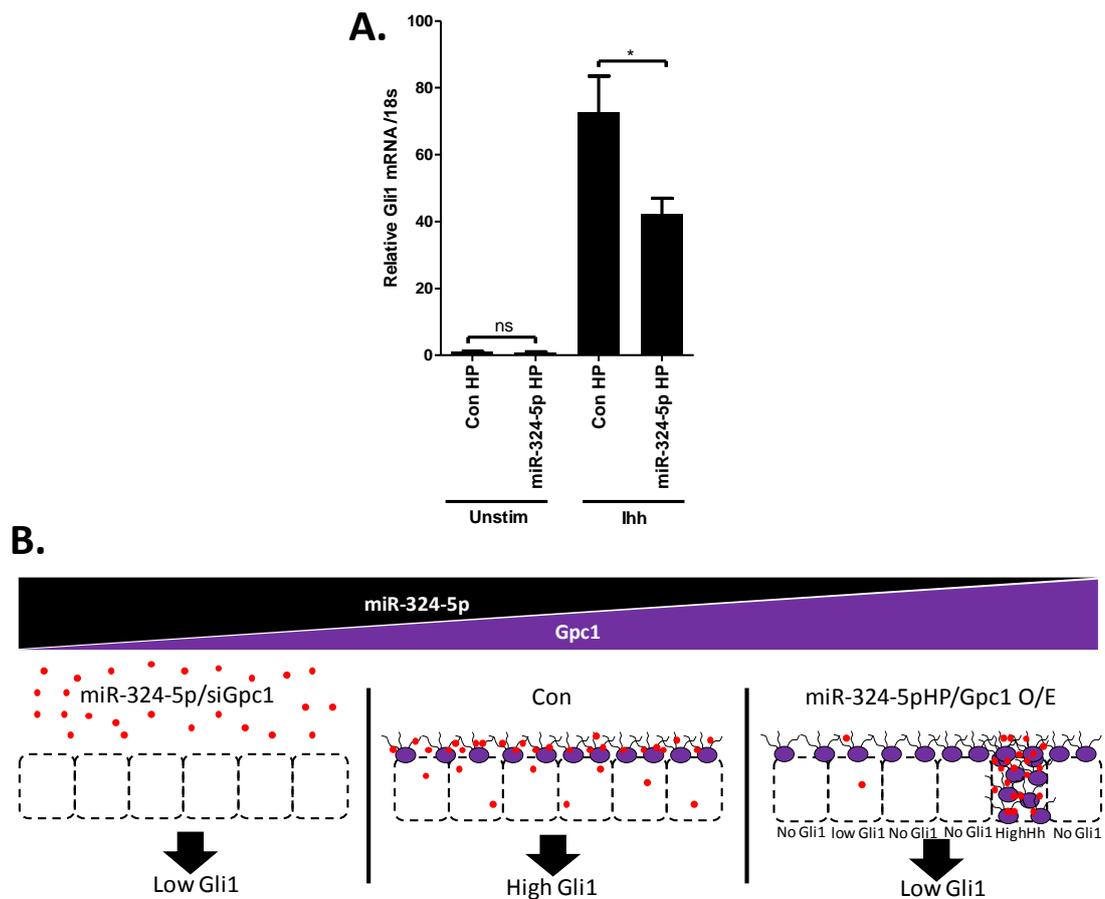
Figure 6.10 miR-324-5p and Gpc1 have activator and repressor properties. (A) miR-324-5p inhibition decreases Hh signalling. C3H10T1/2 cells were transfected with Control hairpin inhibitor (Con HP) or miR-324-5p hairpin inhibitor (miR-324-5p HP) for 24h. Cells were then serum starved for 24h and either left unstimulated, stimulated for with Ihh (2µg/ml) for 24h. Gli1 and 18s were assessed by real-time RT PCR, data from one n=6 experiment. Statistical difference were calculated using Student's t-test, where *=p<0.05. (B) Possible mechanism by which decreased Gpc1 (due to miR-324-5p) and increased Gpc1 (due to transfection of Gpc1 overexpression plasmid) can both lead to decreased Hh signalling. Left; loss of Gpc1 means Hh ligand does not accumulate at the cell membrane and is lost in the media (or may also involve reduced endocytosis). Middle; normal levels of Gpc1, Gpc1 acts as a key component of Hh signalling, by increasing the concentration of Hh ligand at the cell membrane and or promoting Hh binding to ptch and or facilitating endocytosis of ptch. Right; Overexpession of Gpc1 in some cells means the majority of cells do not receive a Hh signal and therefore results in a lower overall level of Hh activity (may also involve loss if primary cilia, incorrect endocytosis and activation of negative feedback loops).

6.3.4 Different mechanisms in human and mouse

In humans miR-324-5p regulates SMO and GLI1 to regulate Hh signalling whereas in mouse miR-324-5p does not regulate Gli1, is unlikely to regulate Smo, and instead regulates Gpc1, which itself is a regulator of Hh signalling (Figure 6.11). To determine if GPC1 is a target in human I analysed the human GPC1 3'UTR. The Human GPC1 3'UTR also contains a miR-324-5p binding site (Figure 6.11), suggesting it may also be a miR-324-5p target. If miR-324-5p can regulate human GPC1, it will mean miR-324-5p can regulate Hh in mouse through SMO, GLI1 and GPC1. Quite why miR-324-5p would need to regulate Hh by multiple mechanisms is unknown. However, the Hh pathway is still relatively poorly understood and regulation of all three molecules (GPC1, SMO and GLI1) may be required to 'fine tune' different types of Hh signalling.

From an evolutionary perspective it is interesting that miR-324-5p regulates Hh in humans and mouse by differing mechanisms. This is either due to Hh requiring different regulation in human and mouse or because the regulation of miR-324-5p is different in human and mouse, the latter is unlikely as miR-324-5p is encoded from just downstream of Dvl2 in both species (Figure 7.1). It is more possible that the Hh pathway works in slightly different ways in humans and mouse and therefore requires slightly different regulation. Perhaps miR-324-5p functions in human is to completely shut down Hh, where as in mouse it only regulates Gpc1 dependent Hh signalling (Figure 6.11). On the other hand inhibition of Gpc1 in mice may contribute to the regulation of morphogens such as Bmps. In summary I have shown although miR-324-5p regulates Hh in both mouse and human, the mechanism by which it does so is different, suggesting using mice to study the role of miR-324-5p in human OA is not the most appropriate model.

Figure 6.11 Different mechanisms by which miR-324-5p regulates Hh signalling in human and mouse

Figure 6.11 Different mechanisms by which miR-324-5p regulates Hh signalling in human and mouse. (A) miR-324-5p target sites differ in human and mouse Gpc1, alignment of human and mouse Gpc1 3'UTRs, full miR-324-5p binding sites are highlighted in bright green, miR-324-5p seed binding sites with single mismatch are highlighted with dull yellow (B) miR-324-5p complementarity to full miR-324-5p binding sites, Figure 5.14 of course showed only site 3 is function in mouse, mouse sites 1 and 2 are shown for completeness. (C) In humans miR-324-5p regulates SMO and GLI1. It remains to be determined if miR-324-5p regulates GPC1 in human and if GPC1 is a regulator of Hh in human. (D) In mouse miR-324-5p does not regulate Smo or Gli1 and instead regulates Gpc1, which is a regulator of Hh signalling.

6.4   Summary

In Chapter 5 I showed miR-324-5p directly targets the mouse Gpc1 3'UTR, and here (Chapter 6) I show that by doing so it is able to regulate Hh signalling. I have shown siRNA depletion of Gpc1 phenocopied miR-324-5p in its inhibition of Hh signalling and effect on alkaline phosphatase and alcian blue. I suggested the mechanism by which siGpc1/miR-324-5p reduced Hh signalling was due to loss of Gpc1 HS chains, as global HS chain inhibition (sodium chlorate) decreased Hh signalling and miR-324-5p transfection lead to decreased Ihh retention in cell lysate as detected by SILAC MS. Interestingly, I also showed Gpc1 overexpression inhibited Hh signalling and suggested a number of theories for this, including effects on the primary cilia, Ihh localisation and activation of negative feedback loops. Preliminary data shows reduced miR-324-5p may also inhibit Hh signalling, suggesting optimal levels of both Gpc1 and miR-324-5p are required for active Hh signalling. Taken together with Chapter 4, this work illustrates a miRNA function can be conserved between species yet the mechanism by which it works is potentially not.

Chapter 7 Additional functions of miR-324-5p

7.1   Introduction

miR-324-5p is increased in OA cartilage (Figure 4.1) and regulates Hh signalling by targeting GLI1 and SMO in human (Chapter 4) and Gpc1 in mouse (Chapter 6). miRNAs have many targets and functions, in C3H10T1/2 cells I have estimated miR-324-5p has ~185 targets (Chapter 5). In addition to regulation of Hh, these targets are likely to have a number of downstream functions. The main aim of this chapter is to investigate these additional functions.

Specific aims

Aim 1: Use pathway analysis of genes whose expression decreased following miR-324-5p transfection to predict additional functions

Aim 2: Further investigate the additional functions identified in Aim 1, namely miR-324-5p regulation of Wnt signalling.

7.2   Results

7.2.1   Pathway analysis of genes decreased by miR-324-5p

As well as identifying specific miR-324-5p targets, analysis of genes differentially regulated by miR-324-5p can be used to predict further pathways regulated by the miRNA. Genes whose expression decreased more than log2 0.2 in Array1 (unstim) were input into the online pathway analysis tool, DAVID [211], to search for enrichment of pathways or common gene functions. My results indicated a significant enrichment for a number of GO (Gene Ontology http://www.geneontology.org/) terms (Table 7.1), biological process GO terms (Table 7.2) and Kegg signalling pathways (Table 7.3). These included Wnt signalling, TLR signalling, TGFβ signalling, actin cytoskeleton, development, cancer and brain diseases.

Table 7.1 GO terms enriched for following miR-324-5p overexpression.

Term	Count	PValue
Golgi apparatus	143	2.21209E-11
proteinaceous extracellular matrix	73	3.73872E-09
extracellular matrix	74	9.50518E-09
organelle lumen	200	1.55491E-08
intracellular organelle lumen	199	2.02717E-08
membrane-enclosed lumen	203	5.05597E-08
extracellular matrix structural constituent	17	1.09742E-07
cell cycle	121	2.67275E-07
extracellular matrix part	30	6.74597E-07
skeletal system development	66	9.76055E-07
regulation of transcription from RNA polymerase II promoter	119	1.26925E-06
Wnt receptor signaling pathway	37	2.74489E-06
metal ion binding	571	2.92749E-06
cation binding	574	4.42565E-06
actin cytoskeleton organization	43	4.60382E-06
DNA binding	287	5.99061E-06
embryonic organ development	56	6.40928E-06
positive regulation of nitrogen compound metabolic process	102	6.74822E-06
enzyme linked receptor protein signaling pathway	61	8.73781E-06
blood vessel development	56	9.43455E-06
vasculature development	57	9.56492E-06
nucleotide binding	341	1.04613E-05
actin filament-based process	44	1.09262E-05
regulation of transcription	343	1.24755E-05
embryonic development ending in birth or egg hatching	85	1.31013E-05
ion binding	576	1.3699E-05
platelet-derived growth factor binding	8	1.51564E-05
chordate embryonic development	84	1.65828E-05
intracellular non-membrane-bounded organelle	290	1.68812E-05
non-membrane-bounded organelle	290	1.68812E-05
cell division	61	2.20698E-05
cytoskeletal protein binding	83	2.3053E-05
small GTPase mediated signal transduction	57	2.51441E-05
positive regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	97	2.62248E-05
intracellular signaling cascade	157	2.99017E-05
extracellular structure organization	38	2.99987E-05
organelle membrane	137	3.14166E-05
positive regulation of gene expression	93	3.53932E-05
positive regulation of transcription	91	3.55922E-05
mitotic cell cycle	54	4.01647E-05
cell cycle process	78	4.08242E-05
structure-specific DNA binding	24	4.37356E-05
chromosome	73	5.43135E-05
negative regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	78	5.74883E-05
small conjugating protein ligase activity	33	6.65439E-05
Golgi apparatus part	49	7.14567E-05
transition metal ion binding	392	7.65778E-05
negative regulation of nitrogen compound metabolic process	78	8.24603E-05
chromosomal part	63	8.83022E-05
negative regulation of gene expression	79	0.00010145

Table 7.1 GO terms enriched for following miR-324-5p overexpression. 3000 (maximum for DAVID pathway analysis) gene whose expression were decreased following miR-324-5p transfection were input into DAVID pathway functional annotation chart. Table shows 50 most enriched GO terms.

Table 7.2 Biological process GO terms enriched for following miR-324-5p overexpression.

Term	Count	PValue
cell cycle	121	2.67275E-07
skeletal system development	66	9.76055E-07
regulation of transcription from RNA polymerase II promoter	119	1.26925E-06
Wnt receptor signaling pathway	37	2.74489E-06
actin cytoskeleton organization	43	4.60382E-06
embryonic organ development	56	6.40928E-06
positive regulation of nitrogen compound metabolic process	102	6.74822E-06
enzyme linked receptor protein signaling pathway	61	8.73781E-06
blood vessel development	56	9.43455E-06
vasculature development	57	9.56492E-06
actin filament-based process	44	1.09262E-05
regulation of transcription	343	1.24755E-05
embryonic development ending in birth or egg hatching	85	1.31013E-05
chordate embryonic development	84	1.65828E-05
cell division	61	2.20698E-05
small GTPase mediated signal transduction	57	2.51441E-05
positive regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	97	2.62248E-05
intracellular signaling cascade	157	2.99017E-05
extracellular structure organization	38	2.99987E-05
positive regulation of gene expression	93	3.53932E-05
positive regulation of transcription	91	3.55922E-05
mitotic cell cycle	54	4.01647E-05
cell cycle process	78	4.08242E-05
negative regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	78	5.74883E-05
negative regulation of nitrogen compound metabolic process	78	8.24603E-05
negative regulation of gene expression	79	0.00010145
negative regulation of transcription, DNA-dependent	63	0.000101558
negative regulation of transcription	73	0.000107003
respiratory tube development	30	0.000112425
cell cycle phase	66	0.000112959
heart development	49	0.00011496
negative regulation of RNA metabolic process	63	0.000123016
negative regulation of cell differentiation	42	0.000125401
phosphate metabolic process	146	0.000138184
phosphorus metabolic process	146	0.000138184
negative regulation of macromolecule metabolic process	93	0.000142492
tissue morphogenesis	51	0.000157824
mitosis	43	0.000162741
nuclear division	43	0.000162741
modification-dependent protein catabolic process	93	0.000162909
modification-dependent macromolecule catabolic process	93	0.000162909
cytoskeleton organization	65	0.000168014
positive regulation of transcription from RNA polymerase II promoter	70	0.000168141
cartilage development	23	0.000170983
transmembrane receptor protein serine/threonine kinase signaling pathway	23	0.000170983
embryonic morphogenesis	70	0.000183206
negative regulation of macromolecule biosynthetic process	79	0.000193342
lung development	29	0.000200887
blood vessel morphogenesis	44	0.000207409

Table 7.2 Biological process GO terms enriched for following miR-324-5p overexpression. 3000 (maximum for DAVID pathway analysis) gene whose expression was decreased following miR-324-5p transfection were inputted into DAVID pathway functional annotation chart. Table shows 50 most enriched biological process GO terms.

Table 7.3 Kegg pathways enriched for following miR-324-5p overexpression.

Term	Count	PValue
Focal adhesion	51	1.16E-06
Ubiquitin mediated proteolysis	37	1.23E-05
Cell cycle	35	1.98E-05
Pathways in cancer	67	5.87E-05
Pancreatic cancer	23	6.11E-05
Valine, leucine and isoleucine degradation	17	1.13E-04
Colorectal cancer	25	1.43E-04
Oocyte meiosis	30	2.16E-04
GnRH signaling pathway	26	4.06E-04
Alzheimer's disease	40	7.50E-04
Glioma	19	8.57E-04
Neurotrophin signaling pathway	31	8.60E-04
ECM-receptor interaction	22	0.001488
Small cell lung cancer	22	0.00205
Wnt signaling pathway	33	0.002097
p53 signaling pathway	19	0.002197
ErbB signaling pathway	22	0.002783
Chronic myeloid leukemia	20	0.002848
Axon guidance	29	0.004166
Prostate cancer	22	0.004285
Renal cell carcinoma	18	0.006359
Melanoma	18	0.007389
Propanoate metabolism	10	0.010736
Bladder cancer	12	0.014762
Dorso-ventral axis formation	8	0.016979
Apoptosis	19	0.026425
TGF-beta signaling pathway	19	0.026425
MAPK signaling pathway	46	0.028951
Ether lipid metabolism	10	0.029459
Keratan sulfate biosynthesis	6	0.034673
Citrate cycle (TCA cycle)	9	0.038313
Phosphatidylinositol signaling system	16	0.05257
Chondroitin sulfate biosynthesis	7	0.054331
Regulation of actin cytoskeleton	37	0.062861
Progesterone-mediated oocyte maturation	17	0.0738
T cell receptor signaling pathway	22	0.07477
Circadian rhythm	5	0.075266
Inositol phosphate metabolism	12	0.079725
Toll-like receptor signaling pathway	19	0.079749
RIG-I-like receptor signaling pathway	14	0.091037
Huntington's disease	31	0.094498

Table 7.3 Kegg pathways enriched for following miR-324-5p overexpression. 3000 (maximum for DAVID pathway analysis) gene whose expression were decreased following miR-324-5p transfection were inputted into DAVID pathway functional annotation chart. Table shows all (41) of the enriched (default DAVID pathway cut off of p=0.1) Kegg pathways.

7.2.2   miR-324-5p is a regulator Wnt signalling

Interestingly, one of the most enriched pathways was the Wnt signalling pathway (Table 7.1, Table 7.2 and Table 7.3). The Wnt pathway is an important player in cartilage homeostasis and osteoarthritis (see Chapter 1). In addition, miR-324-5p is encoded from just downstream (or within one identified transcript- ENSMUST00000102575) of the Dishevelled (Dvl2) gene (Figure 7.1), an important molecule for Wnt signalling [282]. miRNAs have previously been shown to be involved in similar signalling pathways to their host gene [283]. Therefore due to their close proximity I hypothesised miR-324-5p and Dvl2 were controlled by similar mechanisms and involved in similar processes.

There were many genes altered by miR-324-5p which were involved in Wnt signalling. The cumulative fraction plot (Figure 7.2A) and normalised fraction plot (Figure 7.2C) show there is a significant (Figure 7.2B and Figure 7.2D) enrichment of Wnt signalling related genes (Kegg Wnt pathway), in genes whose expression decreased following miR-324-5p overexpression. Figure 7.3 shows the relative positions of these genes (green) within the kegg Wnt signalling pathway. Some of these decreased genes are also predicted to be direct targets of miR-324-5p (yellow).

To test if miR-324-5p is a regulator of canonical Wnt signalling I used the TOP/FOP FLASH luciferase reporter assay system. This is a luciferase based assay containing TCF/LEF binding sites (TOP) or mutant TCF/LEF binding sites (FOP). Canonical Wnt signalling signals through β-Catenin, which binds TCF/LEF binding sites meaning TOP/FOP flash is a good read out canonical Wnt signalling. Transfection of miR-324-5p inhibited Wnt3a induced TOP/FOP FLASH luciferase activity (Figure 7.4A), suggesting miR-324-5p is a regulator of canonical Wnt signalling.

In Chapter 6 I showed miR-324-5p regulated Hh signalling by targeting Gpc1. Glypicans have been shown to be involved in Wnt signalling as well as Hh signalling [246, 280]. To test if Gpc1 is a regulator of Wnt signalling I co-transfected siGpc1 with TOP/FOP flash and stimulated with Wnt3a. siGPC1 caused reduced TOP FLASH luciferase activity (Figure 7.4B).

Bmp2 + Wnt3a has previously been shown to regulate chondrogenesis (as measured by alcian blue) [284], I therefore hypothesised Gpc1 may also regulate Bmp2 + Wnt3a induced alcian blue by inhibition of Wnt signalling. Preliminary data suggested Gpc1, is involved in Bmp2 + Wnt3a induced alcian blue (Figure 7.4). Similar to the results in Chapter 4, where miR-324-5p did not affect Bmp2 + Ihh induced alcian blue (Figure 4.17), it also did not regulate Bmp2 + Wnt3a induced alcian blue (Figure 7.4).

Figure 7.1 miR-324-5p and DVL2
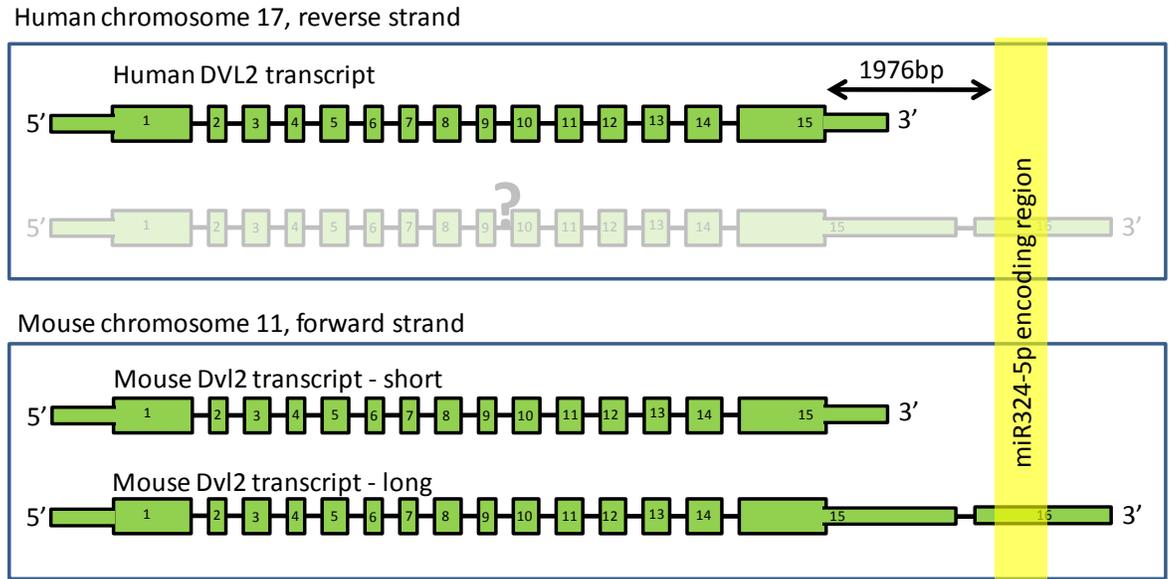
Figure 7.1 miR-324-5p is encoded from within one mouse Dvl2 transcript and ~2kb downstream of most common Dvl2 transcripts. Human DVL2 transcript ID=ENST00000005340, mouse Dvl2 transcript-short (Dvl2-001) ID=ENSMUST00000019362, mouse Dvl2 transcript-long (Dvl2-006) I =ENSMUST00000102575. Question mark indicates it is unknown if a long human DVL2 transcript exists or if miR-324-5p is under the control of the same promoter as DVL2.

Figure 7.2 miR-324-5p decreases Wnt signalling genes
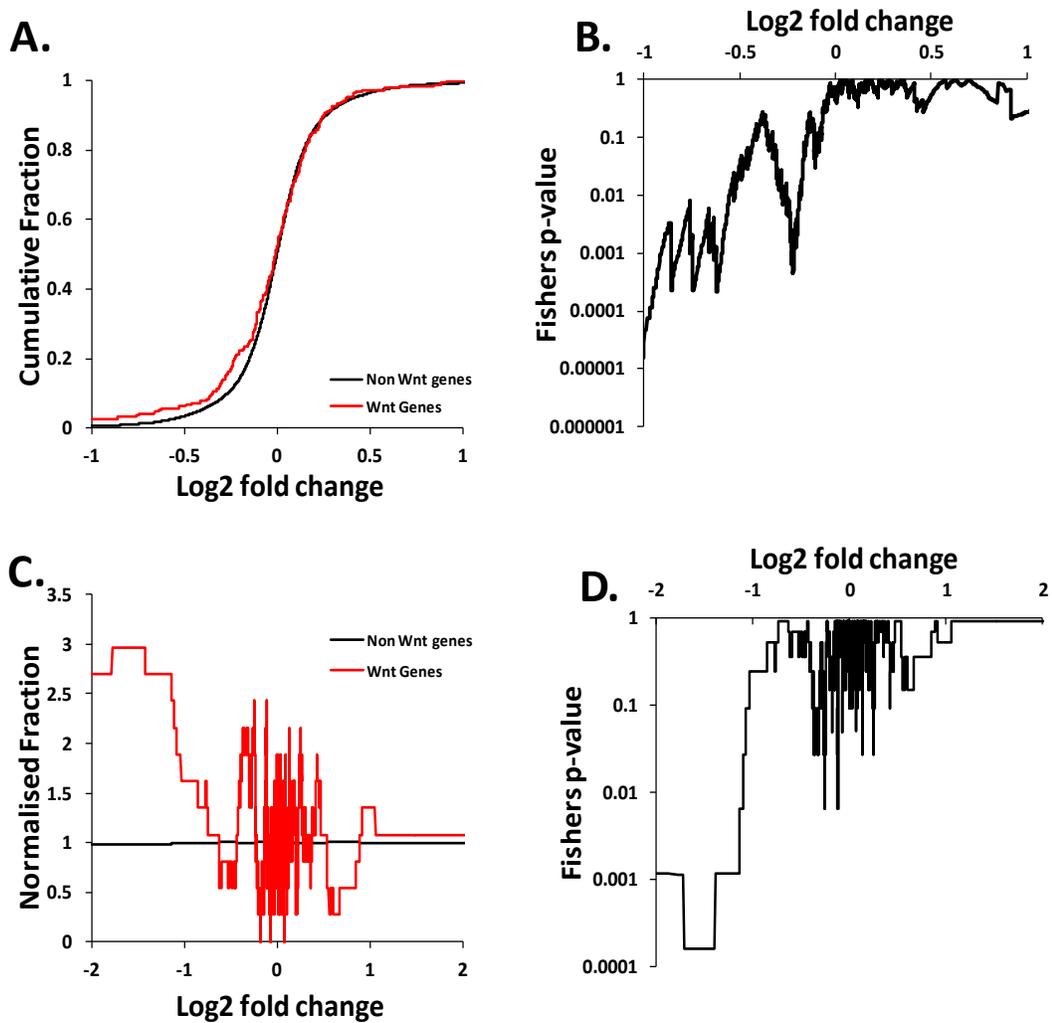
Figure 7.2 miR-324-5p decreases Wnt signalling genes. List of Wnt signalling genes were determined from Wnt signalling Kegg pathway. (A) Cumulative fraction plot showing a higher fraction of Wnt signalling genes are decreased following miR-324-5p overexpression than fraction of non-Wnt signalling genes. (B) Fishers p value for cumulative fraction plot. (C) Normalised fraction plot showing a higher fraction of Wnt signalling genes are decreased following miR-324-5p overexpression than fraction of non-Wnt signalling genes. (D) Fishers p value for normalised fraction plot.

Figure 7.3 Kegg Wnt pathway components are decreased following miR-324-5p transfection
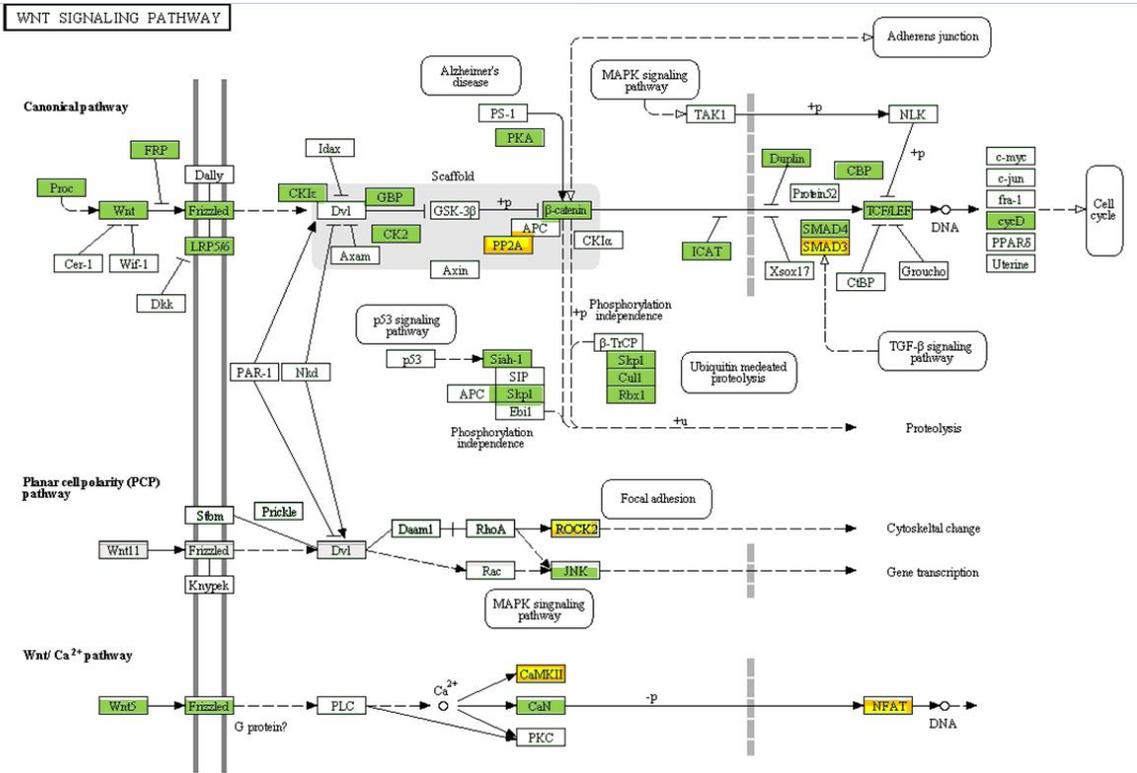
Figure 7.3 Schematic showing the position of miR-324-5p decreased genes (Chapter 4 - microarray 1) in the Kegg Wnt pathway. Gene names shown are from the Kegg Wnt pathway and not necessarily the exact genes decreased. All predicted genes are in the Kegg pathway list of genes (not shown) and are orthologs of the genes shown. Green indicates genes decreased following miR-324-5p transfection. Yellow indicates genes decreased following miR-324-5p transfection and which are predicted miR-324-5p targets.

Figure 7.4 miR-324-5p regulates Wnt signalling
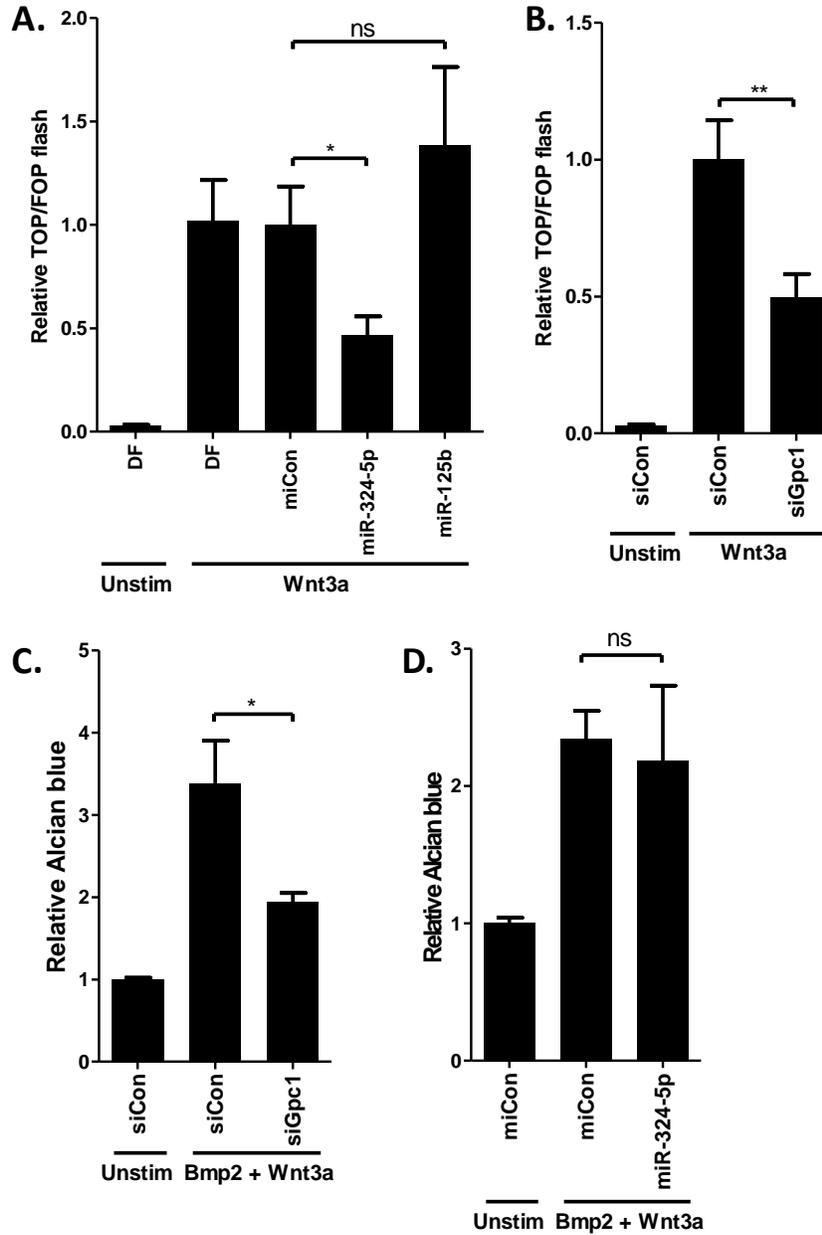
Figure 7.4 Effect of miR-324-5p and siGpc1 on Wnt signalling and chondrogenesis. (A and B) miR-324-5p and siGpc1 decrease Wnt3a induced TOP FLASH. C3H10T1/2 cells were transfected with either FOP or TOP flash luciferase constructs, then after incubation period were transfected with either miCon, miR-324-5p, miR-125b, siCon or siGpc1 for 24hr, cells were then serum starved and stimulated with Wnt3a (100ng/ml) or left unstimulated, luciferase values for TOP FLASH were then normalised to the respective values for FOP FLASH. TOP/FOP values were then normalised to either miCon or siCon. (Data combined from 3 independent experiments, each n=6) (C and D) Alcian blue was used as a marker of cartilage formation and extracted with GuHCl and quantified. C3H10T1/2 cells were transfected with siCon siGpc1, miCon or miR-324-5p then plated into micromass cultures, stimulated with Bmp2 (100ng/ml) and Wnt3a (100ng/ml) (n=6).

7.2.3 miR-324-5p is regulated by Wnt3a

In Chapter 4 I showed miR-324-5p regulated Hh signalling, therefore I tested if Hh signalling could regulate miR-324-5p. I found neither Ihh nor Pur regulated miR-324-5p expression (Figure 4.15B). Here I demonstrated that miR-324-5p regulated Wnt3a induced canonical Wnt signalling, therefore I tested if Wnt3a can regulate miR-324-5p expression. Preliminary data shows miR-324-5p expression was increased following Wnt3a stimulation of C3H10T1/2 (Figure 7.5).

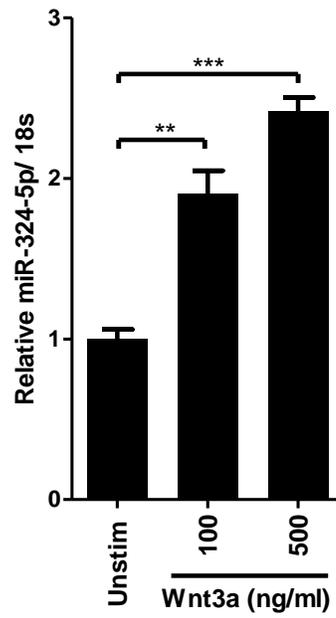**Figure 7.5 miR-324-5p is regulated by Wnt signalling**

Figure 7.5 Wnt3a alter miR-324-5p expression. C3H10T1/2 cells were stimulated with Wnt3a (100ng/ml) for 24h. miR-324-5p, was then assessed with real-time RT-PCR and normalised to 18s. Data combined from 2 independent experiments, each n=4.

7.3 Discussion

7.3.1 Pathway analysis of genes decreased by miR-324-5p

Pathway analysis of genes whose expression decreased following miR-324-5p overexpression suggested miR-324-5p was involved in a number of pathways in addition to Hh, including Wnt signalling, TLR signalling, TGFβ signalling, actin cytoskeleton, development, cancer and brain diseases. A number of these pathways were also enriched following DAVID pathway analysis of predicted mouse miR-324-5p targets (Figure 4.7D-right), suggesting pathway analysis of predicted targets is a good indicator of miRNA function. More specifically of the 41 Kegg pathways enriched (Table 7.3), 23 (57%) were also enriched following DAVID pathway analysis of predicted mouse miR-324-5p targets (Figure 4.7D-right and Figure 7.6), which is over 5 times than that expected by chance (p=2.4 x10^{-23}). In addition, of the 576 GO terms enriched following DAVID pathway analysis of genes decreased with miR-324-5p transfection (Table 7.3), 256 (44.4%) were also enriched following DAVID pathway analysis of predicted mouse miR-324-5p targets (Figure 4.7D-right). Taken together with Figure 5.17 (no single online target prediction algorithm predicted more than 5 times the expected number of decreased genes), this possibly suggests pathway analysis of predicted targets is better at predicting the function of a miRNA than online algorithms are at predicting which genes will decrease following overexpression of that miRNA.

The pathway analysis did not distinguish between positive and negative regulators of a pathway, nor did it distinguish between the genes negatively or positively regulated by the pathway, possibly suggesting genes increased by miR-324-5p should also be considered when investigating miRNA function.

A.   **Kegg only**

B.   **GO only**

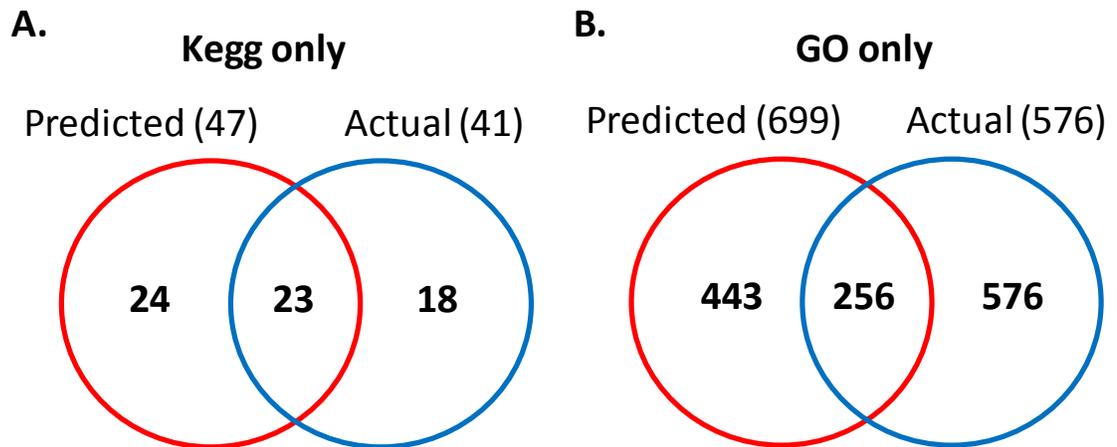Figure 7.6 Venn diagrams showing crossover of number of enriched pathways following DAVID pathway analysis of predicted targets (Predicted) and DAVID pathway analysis of 3000 genes (DAVID max) whose expression decrease >log2 0.2. (A) Number of enriched Kegg pathways, (Fishers p-value=2.4 x10^{-23} based on the total number of possible Kegg pathways being 445). (B) Number of enriched GO pathways.

7.3.2 Role of miR-324-5p in Wnt signalling

Here I have shown miR-324-5p is a negative regulator of canonical Wnt signalling (regulator of TOP/FOP flash), this was in agreement with a screen from a previous study where miR-324-5p was shown to regulate the STF19 reporter (canonical Wnt signalling reporter) [285]. siRNA depletion of miR-324-5p target Gpc1, also reduced Wnt pathway activity (Figure 7.4), suggesting the mechanism by which miR-324-5p regulates Wnt signalling also involves Gpc1. siGpc1 regulated Bmp2 + Wnt3a induced alcian blue (Figure 7.4), probably due to Gpc1 regulation of Wnt signalling.

Glypicans have previously been implemented in the regulation of Wnt signalling. For example, the HSPG gene dally-like (dly), a drosophila glypican, can regulate Wnt gradients [286] and Gpc3 can interact with Wnt to promote Wnt signalling [280].

A number of studies have shown miRNAs form parts of feedback loops. As I have shown miR-324-5p regulates Wnt signalling, I postulated that miR-324-5p may actually be regulated by Wnt signalling. Additionally miR-324-5p is encoded from just downstream (or within one identified transcript) of DVL2, a gene important in Wnt signalling. miR-324-5p expression was found to be Wnt3a responsive, indicating miR-324-5p forms part of a negative feedback loop in the Wnt signalling pathway, and can also regulate Hh.

As siGpc1 decreased Bmp2 + Wnt3a induced alcian blue (Figure 7.4C), it is unclear why miR-324-5p did not (Figure 7.4D). The role of miR-324-5p on Bmp signalling is unknown, meaning it is possible miR-324-5p could act via other targets to positively regulate Bmp signalling, potentially compensating for the inhibitory effects on Hh and Wnt.

This work shows a novel mechanism by which a miRNA links two important signalling pathways, as well as illustrating how a single miRNA can regulate the same pathway, in two species by two distinct mechanisms.

Interestingly, the antisense miRNA, miR-324-3p, has recently been shown to regulate Wnt signalling by directly targeting WNT2B [287].

7.3.3 Role of miR-324-5p in TLR signalling

In Chapter 5 I hypothesised miR-324-5p may be able to regulate Hh signalling by activating TLRs. Interestingly, TLR signalling pathway genes were enriched for in genes decreased following miR-324-5p overexpression (Table 7.3) and is only one of 12 Kegg pathways also enriched for in genes whose expression increased following miR-324-5p overexpression. This possibly suggests miR-324-5p is able to activate TLR signalling as suggested in Chapter 5.

7.3.4 Role of miR-324-5p in TGFβ signalling

TGFβ signalling was also enriched for in decreased genes following overexpression of miR-324-5p (Table 7.3). TGFβ signalling is important for cartilage with its deregulation being an attributing factor to OA [288]. More specifically a switch from Smad2/3 signalling to Smad1/5/8 signalling is thought to play a role in OA, partly via the induction of MMP13 [289]. The Wnt signalling pathway and MAPK signalling pathway have been shown to regulate this switch [288]. Interestingly, the MAPK signallng pathway is another of the enriched pathways in decreased genes following miR-324-5p transfection (Table 7.3).

7.3.5 Role of miR-324-5p in actin cytoskeleton

Pathway analysis also indicated miR-324-5p has a possible role in the actin cytoskeleton (Table 7.2). The actin cytoskeleton plays an important role in primary cilium organization [290], perhaps suggesting an additional mechanism by which miR-324-5p regulates Hh signalling, although no differences were observed in primary cilia following miR-324-5p overexpression (Figure4.14). In addition the actin cytoskeleton is important for maintaining chondrocyte phenotype [291].

7.3.6 Role of miR-324-5p in development

There are a number of enriched development pathways (Table 7.2), reinforcing miR-324-5p plays a role in development. This is also consistent with the role of miR-324-5p in the Hh and Wnt development pathways. Of particular interest to this study is the enrichment of genes involved in the cartilage development pathway (Table 7.2), BMP1, FGFR3, CYTL1, HSPG2, COL2A1, ZEB1, SOX9, TGFB2, ATP7A, DLX2, COL9A1, HIF1A, LECT1, MAPK14, FBXW4, CHST11, PKD1, ROR2, GNAS, COL1A1 and BMPR1B are all genes whose expression decreased following miR-324-5p overexpression and are part of the 'GO' cartilage

development pathway. In addition, my SILAC data indicated miR-324-5p can decrease type II collagen expression, a major constituent of articular cartilage. None of the online databases predict the mouse type II collagen gene, Col2a1, to have a miR-324-5p binding site in its 3'UTR, 5'UTR or coding sequence, suggesting it is not a direct target of miR-324-5p.

7.3.7   Role of miR-324-5p in vascularisation

Healthy cartilage resists vascularisation by secretion of anti-angiogenic proteins, these anti-angiogenic proteins are lost during OA [292], leading to the vascularisation of cartilage [293]. Here I have shown miR-324-5p is a possible regulator of vascularisation (Table 7.1 and Table 7.2), suggesting an addition mechanism by which miR-324-5p may be involved in OA pathogenesis.

7.3.8   Role of miR-324-5p in cancer

There was also an enrichment of a number of cancer pathways (Table 7.1, Table 7.2 and Table 7.3), suggesting miR-324-5p may play a role in some cancers. In addition miR-324-5p has previously been found to be aberrantly expressed in a number of cancers [294]. Some cancers also involve aberrant activation of Hh and Wnt signalling, suggesting modulation of miR-324-5p expression may form part of an effective cancer treatment.

7.3.9   Role of miR-324-5p in brain diseases

A number of studies have profiled miRNA expression in different tissues, one such study showing miR-324-5p was most highly expressed in Brain tissue. More specifically miR-324-5p was most highly expressed in the cortex and cerebellum [206].

Interestingly, according to DAVID pathway analysis some of the most enriched pathways following miR-324-5p transfection were associated with the brain. These include Parkinson's, Alzheimer's, neuron development, Glioma and ErbB signalling (Table 7.1, Table 7.2 and Table 7.3). One of my validated targets of miR-324-5p was APP (Amyloid precursor protein) (Chapter 5), a gene involved in plaque formation and degenerative brain diseases, illustrating an important role for miR-324-5p in brain. In unpublished work, APP has been validated as miR-324-5p target in human [295]. Embryos deficient in Hh signalling often have neurological defects, possibly suggesting some of the effects of miR-

324-5p on brain pathways are due to inhibition of Hh signalling. Further to this, Gpc1 has also been show to play a role in prion metabolism, suggesting there may be a number of mechanisms by which miR-324-5p can regulate brain function and disease [296]. Gpc1 is also found to be increased in glioma [297].

Recently, Stappert *et al*. has shown miR-324-5p (and miR-324-3p) regulates neural differentiation, but they do not show a mechanism by which miR-324-5p works, they also comment on the need to identify bonefied miR-324 targets involved in neural differentiation [298]. My research may lead to the identification of some of these targets.

7.4 Summary

In this chapter I have shown miR-324-5p regulates a number of pathways in addition to Hh signalling. I have further investigated the role miR-324-5p plays in Wnt signalling. I have also discussed how many of these additional pathways may be involved in development, cartilage maintenance and OA pathogenesis. miR-324-5p may also play key roles in other diseases such as cancer and diseases of the brain.

Chapter 8 General discussion

The Hh and Wnt signalling pathways are important in development and diseases of many tissues, particularly cartilage, where their deregulation can lead to OA [48, 299-301]. miRNAs are important regulators of gene expression, and have also been implemented in the regulation of signalling pathways, development and diseases [142, 155, 301]. I hypothesised miRNAs are involved in OA progression and can regulate developmental pathways such as Hh and Wnt signalling. This study investigated the function of two miRNAs upregulated in OA; miR-125b-5p and miR-324-5p (Chapter 4), but focused on miR-324-5p in Hh and Wnt signalling (Chapter 4-7). I also investigated SNPs which may affect the function of the most abundant and extensively studied miRNA in cartilage (miR-140) (Chapter 3).

Focusing on miR-324-5p I have shown it is increased in end-stage OA cartilage and has previously been shown to regulate the Hh signalling pathway in human neural progenitor and tumour cells by targeting the Hh pathway activator SMO and transcription factor GLI1 [127]. I demonstrated miR-324-5p regulation of Hh is conserved in mouse, yet the mechanism by which it does so is not. miR-324-5p has no full potential binding site in the mouse Smo 3'UTR and does not regulate the mouse Gli1 3'UTR, but instead regulates the mouse Gpc1 3'UTR in order to regulate Hh signalling (Chapter 4). I performed SILAC proteomics and microarray analysis following miR-324-5p overexpression to identify Gpc1 as a direct miR-324-5p target in mouse (Chapter 5). Gpc1 is a HSPG and my data shows it is required for Hh signalling, probably by presenting Hh ligand via its HS chains to the Hh receptor Ptch (Chapter 6). I went on to show miR-324-5p also regulates Wnt signalling by a mechanism also likely to involve Gpc1 (Chapter 7). I finally demonstrated that miR-324-5p is a Wnt responsive miRNA (Chapter 7). In summary miR-324-5p is Wnt responsive, increased in OA, and is a regulator of the Hh and Wnt signalling pathways, both important signalling pathways in cartilage biology. In addition I identified other novel miR-324-5p targets such as App, Anxa4 and Serpine2, illustrating the complexity of miRNA biology.

In the general discussion I will; 1) Explain why miR-324-5p is a potent inhibitor of Hh signalling; 2) Discuss how miR-324p-5p links the Hh and Wnt signalling pathways; 3)

Discuss the importance of miRNA regulation of HSPGs, focusing on miR-324-5p regulation of Gpc1; 4) Discuss the role miR-324-5p is likely to have *in vivo*; and 5) finally I will summarise and comment on the wider implications of this work.

8.1   miR-324-5p is a potent inhibitor of Hh signalling

My data show miR-324-5p is an effective inhibitor of Hh signalling. In addition a recent paper actually uses miR-324-5p as an inhibitor of Hh rather than using pathway specific siRNAs to study the effect of oxidative stress [302]. I also found miR-324-5p is more potent than siGpc1 at inhibiting Hh signalling, which suggests miR-324-5p may target multiple components of the Hh signalling pathway.

Drosophila smo mutant clones, which have abolished Hh signalling have reduced dlp (drosophila glypican homologue) and ptc mutant clones which have increased Hh have increased dlp, suggesting glypicans in drosophila may actually be regulated by Hh signalling [252]. If conserved in mammals this mechanism may explain why both miR-324-5p and siGpc1 are potent inhibitors of Hh signalling. For example the miR-324-5p induced reduction of Gpc1 and subsequent reduction of the Hh pathway may further reduce glypican expression, reinforcing the miR-324-5p reduction of Hh. A similar feedback, involving the Hh receptor Ptch, is already known and as expected the expression of Ptch1 in my study was Hh dependent (Figure 4.10). Evidence from my studies does not suggest Gpc1 is under the control of Hh signalling in mammals (Figure 8.1), and is actually under the control of p53, the human tumour suppressor [303]. My microarray data did however indicate Gpc4 expression may be regulated by Hh signalling (Figure 8.1). Gpc4 does not have a miR-324-5p seed binding site suggesting its reduction is indirect, probably due to the miR-324-5p mediated reduction of Hh. If Gpc4 is similarly like Gpc1 and required for Hh signalling, this data may suggest part of the reason why miR-324-5p was a potent Hh inhibitor was because of the miR-324-5p mediated reduction of Hh (via Gpc1 inhibition) caused a reduction in the expression of Gpc4, which further potentiated the reduction of Hh signalling.

In addition to glypicans there are a number of other HSPGs, including Syndecans and Perlecan. In mammals there are 4 different syndecans (Sdc1-4) and one perlecan (Hspg2).

218

Similar to glypicans, perlecan predominantly has HS chains, whereas syndecans have a mix of HS and CS chains. Similar to glypicans, syndecans are normally membrane bound, however unlike glypicans, they contain their own membrane spanning domain, rather than attaching via a GPI anchor. Perlecan is normally secreted from the cell. Sdc3 (Syndecan 3) and Hspg2 are also reduced with the addition of miR-324-5p (Figure 8.1A), suggesting miR-324-5p may also act through these HSPGs to regulate Hh signalling (Figure 8.1B). Neither Sdc3 nor Hspg2 have miR-324-5p seed binding sites or are Hh responsive indicating miR-324-5p regulates their expression by an indirect, Hh independent, mechanism (Figure 8.1).

Figure 8.1 miR-324-5p regulates many membrane proteins to regulates Hh

A.
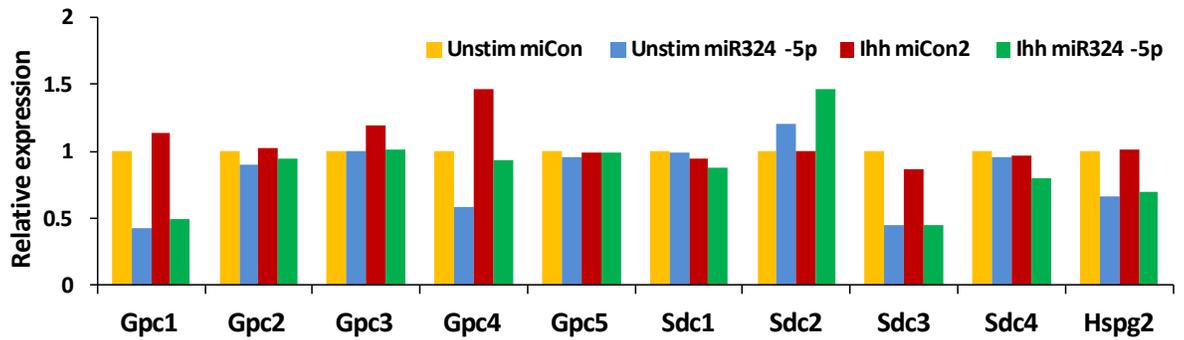
B.
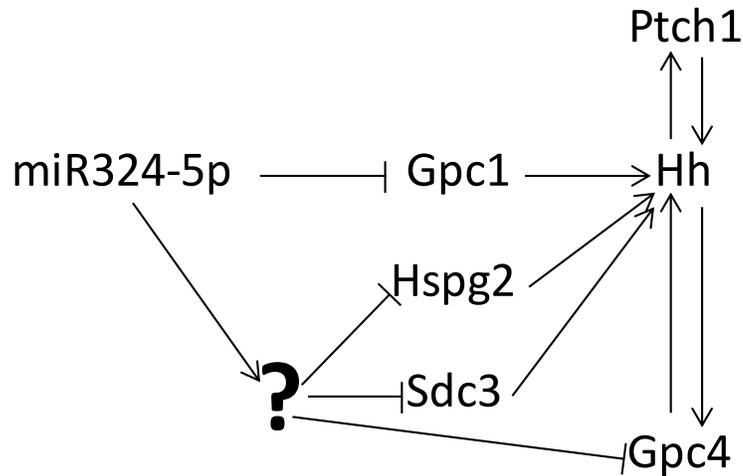
Figure 8.1 miR-324-5p regulates many HSPGs to regulates Hh. In addition to Gpc1 miR-324-5p also regulates a number of other HSPGs which may or may not play a role in Hh signalling. (A) Microarray data for all identified glypicans, Sydecans and Perlecan (Hspg2). (B) Schematic showing miR-324-5p can directly regulate Gpc1, can regulate Hspg2, Sdc3 and Gpc4 via an unknown Hh independent mechanism and can regulate Gpc4, probably via inhibition of Hh signalling.

8.2   miR-324-5p links Hh and Wnt

I have shown miR-324-5p is a regulator of Hh (Figure 4.14) and Wnt (Figure 7.4) signalling, and is Wnt3a (Figure 7.5), but not Ihh responsive (Figure 4.18), indicating miR-324-5p forms a negative feedback loop in the Wnt signalling pathway, and can also inhibit Hh signalling. This is perhaps unsurprising as miR-324-5p is encoded from just downstream of DVL2 (Figure 7.1), a gene involved in Wnt signalling [282]. Wnt ligands can lead to the activation of the canonical, the non-canonical Ca^{2+} and the non-canonical planar cell polarity (PCP) Wnt signalling pathways (Chapter 1). The non-canonical Ca^{2+} pathway has been shown to inhibit the canonical pathway by an unknown mechanism [102]. I know miR-324-5p inhibits canonical Wnt signalling (miR-324-5p decreased TOP/FOP flash expression). Without further investigation it remains unclear if miR-324-5p is canonical or non-canonical responsive, but if it is non-canonical responsive, it may form part of the mechanism by which non-canonical Wnt signalling inhibits canonical Wnt signalling. Wnt3a activates the canonical and non-canonical pathway [102]. Separate inhibition of the canonical and non-canonical Wnt pathways with Dkk1 and CaMKII inhibitor KN93 respectively will determine if miR-324-5p is either canonical or non-canonical responsive

The Hh and Wnt signalling pathways are closely linked and have many interactions. They are both involved in osteogenesis with Ihh being required for Wnt signalling activation [97]. In addition Sufu, which is a Hh regulator, is also involved in Wnt signalling [304]. Likewise GSK3 which is known to regulate Wnt can also regulate Hh signalling [70, 305]. miR-324-5p regulation of both the Hh and Wnt signalling pathways adds to this complexity. From an evolutionary perspective miR-324-5p appears relatively recently (around 100million years ago) in comparison to the Hh and Wnt pathways, which are largely conserved in all vertebrates meaning they are more than 100million years old. This suggests miR-324-5p may function in order to satisfy the differential needs of the Hh and Wnt signalling pathways without major alteration of the core Hh and Wnt signalling molecules. This can also been seen at a more general level where the 3'UTRs are less conserved than the coding sequence, meaning miRNAs may contribute to the evolution of these pathways and allow them to function in many different organisms, without changing the core proteins of the pathways. This makes sense as mutations in 3'UTRs are far more

likely benefit the organism than mutation in the coding sequence. This work shows a novel mechanism by which a miRNA links two important signalling pathways and regulates at least one of these pathways by potentially differing mechanisms.

8.3 The importance of miRNA:HSPG interactions

8.3.1 The importance of GAG chains in 'signalling hub' and other signalling pathways

In agreement with previous publications [306], I showed sodium chlorate can decrease Ihh induced Hh signalling, suggesting part of the mechanism by which Ihh induces Hh signalling involved HS chains. In addition to binding morphogens, HSPGs bind other molecules involved in Hh signalling, including lipoproteins which allows the release of morphogen with HSPGs [307]. Shifted [308] and Ihog (Hh co-receptor) [309], are thought to be involved in the interaction of Hh with HS chains and play roles in long and short range Hh signalling respectively. Another protein in Drosophila, Crossveinless2, allows the interaction of HS chains with Bmp receptors [310]. These interactions mean morphogens, their receptors, and their co receptors, are brought together in a 'signalling hub', each having the potential to regulate one another and ultimately regulate morphogen signalling. Depending on the molecules present in the signalling hub it will either act to promote signalling or as a decoy. Here I have shown miR-324-5p regulation of Gpc1 may play a role in these interactions and the generation of a 'signalling hub'.

8.3.2 HSPGs may explain aspect of Hh biology

An interesting aspect of Hh signalling is Shh, Ihh and Dhh all a have similar capacity to bind the Hh receptor Ptch, yet have different potencies on cells (generally Shh>Ihh>Dhh) [67], indicating there must be accessory molecules in these cells binding differentially to each ligand. The loss of the Shh HSPG interacting domain decreases its potency [311, 312], suggesting HSPGs are responsible for the high Shh potency. Perhaps, each ligand binds differently to each HSPG, explaining the different potencies of Ihh, Shh and Dhh and suggesting the potency in dependent on the HSPG expression profile. Here I show Gpc1 is required for Ihh signalling, it would be of interest to examine if Gpc1 is also required for Shh and Dhh signalling.

8.3.3 HSPGs are important for morphogen gradient regulation

HSPGs, receptors and co-receptors all play major roles in regulating morphogen movement and a cells response to morphogen. To date much of the work looking at morphogen gradient regulation has been performed in Drosophila, with some of the mechanisms being able to be translated to vertebrates. Secreted morphogens interact

with the ECM, particularly HSPGs in order for them to properly from gradients prevent loss to the extracellular space and prevent aberrant signalling to the wrong cells. Baeg *et al*. showed GAG synthesis is required for organisation and extracellular distribution of wingless gradients, more specifically they showed a role for the HSPG gene, dally-like (dly), a drosophila glypican in regulation of Wg (Wnt) morphogen gradient regulation [286]. HSPGs can interact with morphogens to produce many different types of morphogen gradient, for review see [267]. In mice which lack the ability to sulphate CS chain on CSPGs there are defects in Ihh morphogen gradient formation [313]. Both HSPGs and CSPGs are involved in Ihh morphogen gradient formation, illustrating how complex these processes can be.

8.3.4   miR-324-5p regulates morphogen through regulation of Gpc1

My data illustrates miR-324-5p inhibition of Gpc1 is likely to play a role a morphogen gradient generation, regulating both the movement and reception of morphogens. Interestingly, miR-324-5p is encoded from within a transcript variant of DVL2 (part of the Wnt signalling pathway) and miR-324-5p also regulates Gpc1 (also possibly part of the Wnt signalling pathway), this is an example of a miRNA which is encoded from, or co regulated with, a signalling pathway gene which regulates a glypican. The opposite of this may also be true in that miRNAs that are encoded from, or are co-regulated with, the glypicans, regulate other components of the morphogen signalling pathways.

8.3.5   Other miRNA glypican interactions

Interestingly, the GPC5/6 and the GPC3/4 cluster of glypicans are in close proximity (just downstream) of the miR-17~92 and miR-106a~363 miRNA clusters respectively. miR-17~92 is a group of polycistronic miRNAs encoding 6 mature miRNAs (miR-17, miR-18a, miR-19a, miR-20a, miR-19b-1 and miR-92a-1). The miR-17~92 cluster has been suggested to play a role with Hh signalling [314] and their deletion leads to developmental phenotypes including skeletal defects in mice and humans [167, 315, 316]. The miR-106a~363 is also group of polycistronic miRNAs also encoding 6 mature miRNAs (miR-106a, miR-18b, miR-20b, miR-19b-2, miR-92a-1 and miR-363). The miR-17~92 and miR-106a~363 clusters of miRNAs have themselves arisen from a genomic duplication of a precursor miRNA cluster [317]. It is possible the functions of these miRNA clusters and the glypican clusters are intimately linked. In an experiment where miR-17~92 cluster expression was induced by provirus integration, an induction of

GPC5 expression was also observed [318], indicating they are under the control of a common promoter. Indeed UCSC ENCODE data shows there is an active promoter at the start site of miR-17~92 and also an 'inhibitor' site between the miR-17~92 cluster and Gpc5, perhaps indicating why Gpc5 expression is not always increased with a miR-17~92 increase [318]. The expression level of the miR-106a~363 cluster is lower than the miR-17~92 cluster [315], and is encoded from further upstream of GPC3 than miR-17~92 is in relation to GPC5. This indicates GPC3 and miR-106a~363 are potentially more independently controlled than Gpc5 and miR-17~92. There is also a third miR cluster prologue (to the miR-17~92, miR-106a~363 clusters), the miR-106b~25 cluster, which is encodes three miRs (miR-106b, miR-93 and miR-25) it is located in an intron of MCM7, and appears unrelated to glypicans [316].

8.4 miR-324-5p in development, and OA

I have shown miR-324-5p regulates Hh and Wnt signalling, and as glypicans (which I have shown miR-324-5p can target) can regulate many morphogens it is plausible miR-324-5p may also act on other morphogens such as BMPs, FGFs and TGFβ to exert its effects (including BMP2 in the alkaline phosphatase experiments), reinforcing the role of miR-324-5p in skeletogenesis and OA, although this was not tested directly in this study.

Canonical Wnt signalling is increased in OA [319], where it is thought contribute to the loss of cartilage by controlling cartilage degrading enzymes called MMPs [320] and to the formation of osteophytes and thickening of subchondral bone by reactivation of genes involved in development [299]. Both over-activation and inhibition of Wnt signalling can lead to skeletal deformities and an early onset OA [299-301], illustrating how tightly regulated Wnt needs to be in cartilage homeostasis.

Camk2, which is involved in non-canonical Wnt signalling, is involved in skeletal development [321]. miR-324-5p induction in OA may be due to non-canonical Wnt signalling. I show miR-324-5p inhibits canonical Wnt signalling. Interestingly, Nalesso *et al.* have shown there is reciprocal inhibition of canonical and non-canonical Wnt signalling pathways [102]. If my finding are reproducible in human it is possible miR-324-5p may act as the mechanism for non-canonical Wnt signalling to inhibit canonical Wnt signalling in human chondrocytes.

Hh signalling plays a role in anterior-posterior patterning, left-right patterning, patterning of the mammalian limb, limb bud development and endochondral ossification. If miR-324-5p is expressed in these tissues it is likely it will also have a role in these developmental processes. As miR-324-5p inhibits Hh it is likely overexpression of miR-324-5p will have similar effects to reduced Hh. However due to the other actions of miR-324-5p there are likely to be a number of differences between reducing Hh and increasing miR-324-5p expression.

The Hh signalling pathway is thought to contribute to the loss of cartilage by controlling cartilage degrading MMPs [320] and to the formation of osteophytes and thickening of subchondral bone by reactivation of genes involved in development [299]. Similar to Wnt, mouse models have shown both over activation and inhibition of

Hh signalling can lead to OA [48]. Illustrating how tightly regulated both Wnt and Hh are in cartilage homeostasis. Similarly both increases and decreases in miR-324-5p and its target Gpc1 cause decrease Hh signalling, perhaps suggesting intricate links between miR-324-5p, Gpc1, Hh, Wnt and OA. A better understanding of miR-324-5p regulation of these pathways may provide a way to pharmacologically alter Wnt or Hh in such a way which is beneficial to cartilage.

In vivo studies have shown the importance of HSPGs in Hh signalling. The reduction of an enzyme involved in the extension of the HS chains (Ext1) in mice, increased the range of Ihh signalling [322]. In mice which lack the ability to sulphate CS chain on CSPGs there are defects in Ihh morphogen gradient formation [313]. Mice which are mutant for Gpc3 (negative regulator of Hh) have increased size [323], similar to mice which are heterozygous for ptch (and have increased Hh signalling) [324], suggesting Gpc3 mutant mice have increased Hh signalling. In humans, EXT1 and EXT2 mutations can cause hereditary multiple exostoses (multiple osteochondromatosis), a condition involving the growth of multiple spurs consisting of bone and cartilage in children [325-328]. All of these other studies involve a lack of regulation of morphogen gradients. Gpc1 has been genetically associated with Biliary atresia [274], a disease involving the liver. In terms of OA, a study has shown SDC4 is up regulated in OA [329]. Deletion of Sdc4 in mice shows Sdc4 has a role in bone repair [330].

Interestingly, GPC1 gene has previously been identified as a possible candidate for causing Brachydactyly Type E, an inherited condition causing skeletal deformities [331]. Genotype-phenotype correlations show GPC1 can be a cause of skeletal deformities in rodents, and is most highly expressed in developing and mature osteoblasts [332]. Also in two patients which have a loss of the GPC1 encoding region (part of chromosome 2q), there are skeletal deformities and also mental retardation. Additionally Gpc1 is most highly expressed in the developing nervous system in rodents [331], perhaps highlighting a duel role of miR-324-5p/ Gpc1 in the skeleton and the nervous system, also as indicated by pathway analysis following miR-324-5p overexpression (Chapter7). miR-324-5p inhibition of Gpc1 may result in both skeletal and central nervous system abnormalities. However, I do not suggest inhibition of Gpc1 is the sole mechanism by which miR-324-5p exerts its effects (Chapter 7), for example I have shown miR-324-5p can also inhibit App an important gene for brain

disease pathogenesis. Recently, a paper has shown miR-324-3p and miR-324-5p have a role in differentiation of neural cells [298].

8.5   Summary and future direction

Due to the important roles Hh and Wnt play during development it is likely miR-324-5p will also play a role in development. In situ hybridisation of miR-324-5p in developing tissues may shed light on it role during development. The creation of miR-324-5p depleted and miR-324-5p overexpressing mice, both global and cartilage restricted, to investigate the overall role of miR-324-5p and its role in skeletal development respectively will also take this work forward. This work shows a novel mechanism by which a miRNA links two important signalling pathways, as well as illustrating how a single miRNA can regulate the same pathway, in two species by two distinct mechanisms. It also increases understanding of Hh and Wnt signalling, which may lead to new therapeutics for the many diseases in which these pathways are deregulated.

Taken together this body of work demonstrates how miRNAs, their targets and their functions can be linked in their expression and association with OA.

Chapter 9 Appendix

Appendix A   Inhibiting components of the Hh pathway

As expected, siRNA against Smo reduced Smo mRNA (Figure 4.11G). However siRNA against Gli1 did not cause a reduction in Gli1 mRNA (Figure 4.11A) and siRNA against Ptch did not cause a reduction in Ptch mRNA (Figure 4.11D), possibly as these two genes are Hh responsive as well as being part of the pathway. The reduction in mRNA due to the action of the siRNA will be compensated by increased transcription of Ptch.

Depletion of Hh signalling proteins Ptch, Smo and Gli1, prevents Ihh induced Gli1 expression showing, all three molecules are required for Ihh signalling. One might expect depletion of Ptch to further activate the pathway, as its classical function is to prevent Smo localisation to the primary cilia, but this was not observed. Depletion of Smo and Gli1 did not affect basal levels of Gli1, probably due to the very low detectable levels of Gli1, however depletion of Ptch in unstimulated cells significantly increases Gli1 mRNA, presumably by releasing Smo (to the cilia) and allowing downstream Hh activity (Figure 4.11). As anticipated, depletion of both Smo and Gli1 prevented Pur induction of Gli1 mRNA and protein expression (Figure 4.11C and E). Unexpectedly, depletion of Ptch also caused the loss of Pur induced Gli1 induction (Figure 4.11C and E), indicating Ptch, which is thought to act as a negative regulator upstream of Smo (which is the suggested site of Pur action), is actually required for Pur induction of Hh signalling. Similarly Ptch, Smo and Gli1 are also all required for Pur induction of Ptch (Figure 4.11F). siSmo, but not siPtch or siGli1, reduced Smo mRNA in unstimulated (Figure 4.11G), Ihh stimulated (Figure 4.11H) and Pur stimulated (Figure 4.11I) cells. Interestingly, in contrast to Smo mRNA, siSmo, siPtch and siGli all prevent expression of Smo protein (Figure 4.11D and E), perhaps indicating why ptch depletion can prevent Pur induction of Gli1 (Figure 4.11) and indicates Smo localisation is dependent on Ptch and Gli1.

In summary, all three molecules are involved in Hh signalling and their role can differ depending on the type of Hh activation.

Appendix B   How the cumulative fraction and normalised fraction plots were created

To illustrustrate the level enrichment of different miRNA seed binding sites in genes whose expression decreased, I created 'cumulative fraction' (Figure 9.1) and 'normalised fraction' (Figure 9.1) plots using Microsoft Excel.

Figure 9.1 How the cumulative fraction and normalised fraction plots were created
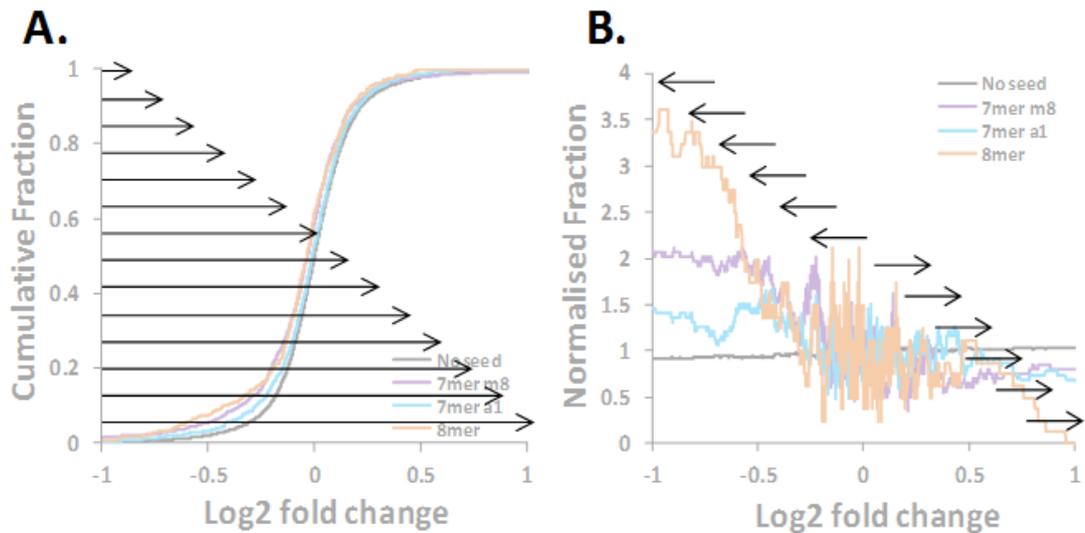
Figure 9.1 How the cumulative fraction and normalised fraction plots were created. (A) Cumulative fraction is calculated, by an increasing window of 1, starting with the most decreased genes. The sum of 'seed type of interest' in each calculated list (represented by arrow) was divided by the sum of 'seed of interest' in whole data set, to give the cumulative fraction and plotted against the respective log2 fold change. Where significance plots are shown (Appendix C), Fishers p-value was calculated at increments of every 5 of the original calculated cumulative fraction values (comparing the site type of interest to the appropriate control e.g. no seed). (B) Normalised fraction is calculated by a moving window of 500, at increasing increments of 1. The window moved outwards from genes whose expression did not change bi-directionally to increased and decreased genes. The sum of 'seed type of interest' in the calculated in each list (represented by arrow) is then divided by the sum of 'seed type of interest' in whole data set, giving the 'fraction' of the 'seed of interest'. This fraction is then normalised to the average of this fraction in the whole data set and plotted against the respective log2 fold change.

Appendix C　　　　　Reason why log2 0.2 cut off was chosen

Log2 0.2 cut off was chosen was because it allowed the most significant enrichment of miR-324-5p seeds in all experiments Figure 9.2.
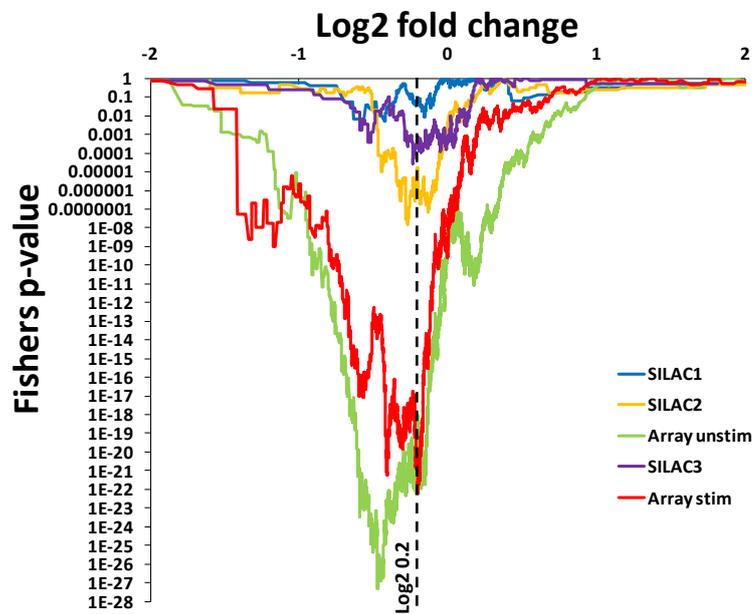
Figure 9.2 Reason why log2 0.2 cut off was chosen

Figure 9.2 Reason why log2 0.2 cut off was chosen. Fishers p-value of the enrichment of non-conserved cumulative frequency plots for each of the target identification experiments (Figure 5.4E, Figure 5.5E, Figure 5.7E, Figure 5.10E and Figure 5.11E). See Appendix B for how this plot was made.

Appendix D Purmorphamine studies

There were three unexpected results concerning HS regulation of Hh and Hh regulation of alkaline phosphatase.

Firstly, sodium chlorate also decreased Pur induced Hh signalling (Figure 9.3A), suggesting the mechanism by which Pur induces Hh signalling also involves HS chains. It is known Pur acts directly on Smo, suggesting HS chains are also involved in Smo localisation or involved in Hh signalling downstream of Smo, strengthening the argument of HS controlling a 'signalling hub'.

Secondly, and consistent with above, siGpc1 decreased Pur induced Gli1 (Figure 9.3B) meaning it is likely HS chains on Gpc1 are involved in Pur induced Hh. Site directed mutagenesis of HS chain attachment sites within the Gpc1 overexpression plasmid could be used to specifically determine the role of Gpc1 HS chains in Hh signalling. Alternatively, HS may also interact with morphogens of other signalling pathways which may have effects on downstream Hh signalling.

Thirdly, and again unexpectedly, siGli1 did not decrease Ihh + Bmp2 induced alkaline phosphatase (Figure 6.4A). The reason for this was unclear, however Bmp2 alone was not included in these experiments, making it possible the addition of Ihh did not actually cause any further increase in Bmp2 induced alkaline phosphatase, meaning siGli1 would have no effect. However, in the same experiments siGpc1 was able to affect the level of alkaline phosphatase (Figure 6.4), perhaps via an as yet untested effect on Bmp signalling. Alternatively, the effect of adding Ihh to Bmp2 on alkaline phosphatase could have been Gli1 independent with the effect of Hh activation on Bmp2 being downstream of Gpc1 but upstream of Gli1.

In contrast to the above, and probably as expected siGli1 does and siGpc1 does not affect Pur and Bmp2 induced alkaline phosphatase (Figure 6.4). If Hh signalling is able to affect Bmp2 signalling independent of Gli1, it is possible regulation of Gpc1 may form an effective mechanism to regulate all effects of Hh. Whereas regulation of Gli1 would only function to regulate Gli1 dependent Hh signalling. Further work may show switching between Gli1 dependent and Gli1 independent may form an effective mechanism to regulate cartilage and bone during development and OA (similar to that shown for the canonical and non-canonical Wnt signalling pathways [102]).

Figure 9.3 Purmorphamine studies
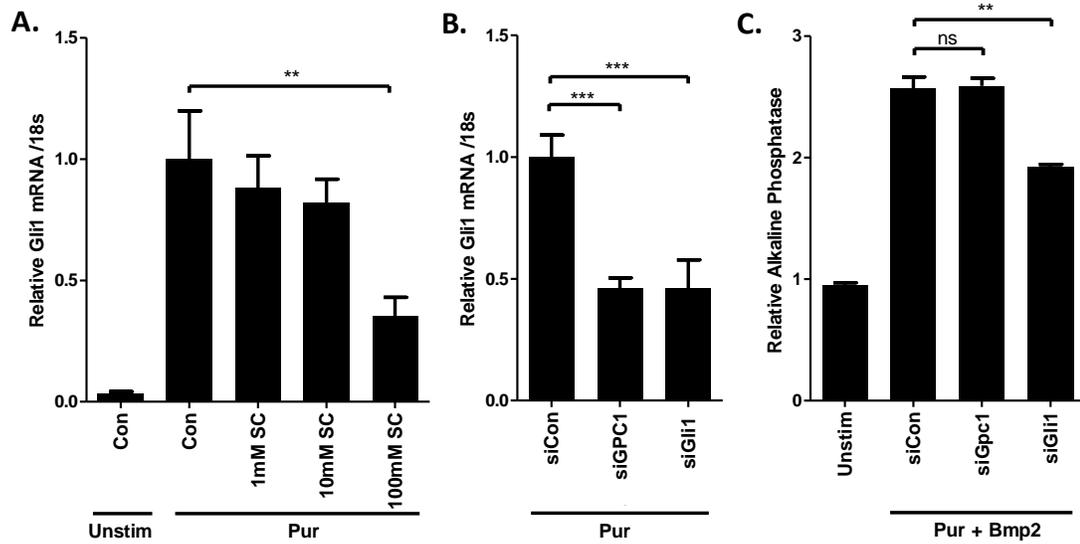
Figure 9.3 Purmorphamine studies. (A) C3H10T1/2 cells were pre-treated with sodium chlorate (SC) for 48h at the concentration shown then serum starved and stimulated with Purmorphamine (Pur) (2μM) for 24h. Data combined from 2 independent experiments, each n=3. (B) C3H10T1/2 cells were transfected with non-targeting siCon or siRNA against Gpc1 or Gli1 for 24h. Cells were then serum starved for 24h and either left unstimulated, stimulated for with Pur (2μM) for 24h. Gli1 and 18s were assessed by real-time RT-PCR. Data combined from 4 independent experiments, each n=4. (C) C3H10T1/2 cells were transfected with siCon, siGpc1 or siGli1, then stimulated with Pur (2μM) and BMP2 (100ng/ml) for 5 days. P.Nitrophenol was measured to determine the level of alkaline phosphatase. Data combined from 3 independent experiments, each n=3. All data are presented as mean + SEM, statistical difference were calculated using Student's *t*-test, where, *p<0.05, **<0.01, ***p<0.001. Data combined from 5 independent experiments, each n=6.

References

1.   Arden, N. and M.C. Nevitt, *Osteoarthritis: epidemiology.* Best Pract Res Clin Rheumatol, 2006. **20**(1): p. 3-25.

2.   Altman, R., et al., *Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association.* Arthritis Rheum, 1986. **29**(8): p. 1039-49.

3.   Goldring, S.G., MB., *Clinical aspects, pathology and pathophysiology of osteoarthritis.* J Musculoskelet Neuronal Interact, 2006. **6**: p. 376-378.

4.   Goldring, M.B. and S.R. Goldring, *Osteoarthritis.* J Cell Physiol, 2007. **213**(3): p. 626-34.

5.   Lin, E.H., *Depression and osteoarthritis.* Am J Med, 2008. **121**(11 Suppl 2): p. S16-9.

6.   Nuesch, E., et al., *All cause and disease specific mortality in patients with knee or hip osteoarthritis: population based cohort study.* BMJ, 2011. **342**: p. d1165.

7.   Reginster, J.Y., *The prevalence and burden of arthritis.* Rheumatology (Oxford), 2002. **41 Supp 1**: p. 3-6.

8.   Fung, H.B. and H.L. Kirschenbaum, *Selective cyclooxygenase-2 inhibitors for the treatment of arthritis.* Clin Ther, 1999. **21**(7): p. 1131-57.

9.   Dohme, M.S. *Merck Announces Voluntary Worldwide Withdrawal of VIOXX(R)*. 2004 [cited.

10.   Baragi, V.M., et al., *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, 2009. **60**(7): p. 2008-18.

11.   Alvarez-Soria, M.A., et al., *Diacerein has a weak effect on the catabolic pathway of human osteoarthritis synovial fibroblast--comparison to its effects on osteoarthritic chondrocytes.* Rheumatology (Oxford), 2008. **47**(5): p. 627-33.

12.   Sawitzke, A.D., et al., *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, 2010. **69**(8): p. 1459-64.

13.   Wieland, H.A., et al., *Osteoarthritis - an untreatable disease?* Nat Rev Drug Discov, 2005. **4**(4): p. 331-44.

14.   Qi, Y., G. Feng, and W. Yan, *Mesenchymal stem cell-based treatment for cartilage defects in osteoarthritis.* Mol Biol Rep, 2012. **39**(5): p. 5683-9.

15.   Ronn, K., et al., *Current surgical treatment of knee osteoarthritis.* Arthritis, 2011. **2011**: p. 454873.

16.   Martel-Pelletier, J., et al., *Cartilage in normal and osteoarthritis conditions.* Best Pract Res Clin Rheumatol, 2008. **22**(2): p. 351-84.

17.   Goldring, M.B., *Update on the biology of the chondrocyte and new approaches to treating cartilage diseases.* Best Practice & Research Clinical Rheumatology, 2006. **20**(5): p. 1003-1025.

18.   Knudson, C.B. and W. Knudson, *Cartilage proteoglycans.* Semin Cell Dev Biol, 2001. **12**(2): p. 69-78.

19.   Wong, M., et al., *Zone-specific cell biosynthetic activity in mature bovine articular cartilage: a new method using confocal microscopic stereology and quantitative autoradiography.* J Orthop Res, 1996. **14**(3): p. 424-32.

20.   Thomas, R.S., et al., *Effects of Wnt3A and mechanical load on cartilage chondrocyte homeostasis.* Arthritis Res Ther, 2011. **13**(6): p. R203.

21. Cawston, T.E. and A.J. Wilson, *Understanding the role of tissue degrading enzymes and their inhibitors in development and disease.* Best Pract Res Clin Rheumatol, 2006. **20**(5): p. 983-1002.

22. Cawston, T.E., *Metalloproteinase inhibitors and the prevention of connective tissue breakdown.* Pharmacol Ther, 1996. **70**(3): p. 163-82.

23. Verbruggen, G., et al., *Influence of aging on the synthesis and morphology of the aggrecans synthesized by differentiated human articular chondrocytes.* Osteoarthritis Cartilage, 2000. **8**(3): p. 170-9.

24. Guilak, F., et al., *The role of biomechanics and inflammation in cartilage injury and repair.* Clin Orthop Relat Res, 2004(423): p. 17-26.

25. Blagojevic, M., et al., *Risk factors for onset of osteoarthritis of the knee in older adults: a systematic review and meta-analysis.* Osteoarthritis Cartilage, 2010. **18**(1): p. 24-33.

26. Grotle, M., et al., *Obesity and osteoarthritis in knee, hip and/or hand: an epidemiological study in the general population with 10 years follow-up.* BMC Musculoskelet Disord, 2008. **9**: p. 132.

27. Hui, W., et al., *Leptin produced by joint white adipose tissue induces cartilage degradation via upregulation and activation of matrix metalloproteinases.* Ann Rheum Dis, 2012. **71**(3): p. 455-62.

28. Srikanth, V.K., et al., *A meta-analysis of sex differences prevalence, incidence and severity of osteoarthritis.* Osteoarthritis Cartilage, 2005. **13**(9): p. 769-81.

29. Cooley, H.M., J. Stankovich, and G. Jones, *The association between hormonal and reproductive factors and hand osteoarthritis.* Maturitas, 2003. **45**(4): p. 257-65.

30. Maleki-Fischbach, M. and J.M. Jordan, *New developments in osteoarthritis. Sex differences in magnetic resonance imaging-based biomarkers and in those of joint metabolism.* Arthritis Res Ther, 2010. **12**(4): p. 212.

31. Zhang, Y., et al., *Lower prevalence of hand osteoarthritis among Chinese subjects in Beijing compared with white subjects in the United States: the Beijing Osteoarthritis Study.* Arthritis Rheum, 2003. **48**(4): p. 1034-40.

32. Nevitt, M.C., et al., *Very low prevalence of hip osteoarthritis among Chinese elderly in Beijing, China, compared with whites in the United States: the Beijing osteoarthritis study.* Arthritis Rheum, 2002. **46**(7): p. 1773-9.

33. Zhang, Y., et al., *Comparison of the prevalence of knee osteoarthritis between the elderly Chinese population in Beijing and whites in the United States: The Beijing Osteoarthritis Study.* Arthritis Rheum, 2001. **44**(9): p. 2065-71.

34. Spector, T.D., et al., *Genetic influences on osteoarthritis in women: a twin study.* BMJ, 1996. **312**(7036): p. 940-3.

35. Miyamoto, Y., et al., *A functional polymorphism in the 5' UTR of GDF5 is associated with susceptibility to osteoarthritis.* Nat Genet, 2007. **39**(4): p. 529-33.

36. Loughlin, J., et al., *Linkage analysis of chromosome 2q in osteoarthritis.* Rheumatology (Oxford), 2000. **39**(4): p. 377-81.

37. Loughlin, J., et al., *Stratification analysis of an osteoarthritis genome screen-suggestive linkage to chromosomes 4, 6, and 16.* Am J Hum Genet, 1999. **65**(6): p. 1795-8.

38. Chapman, K., et al., *Osteoarthritis-susceptibility locus on chromosome 11q, detected by linkage.* Am J Hum Genet, 1999. **65**(1): p. 167-74.

39. arcOGEN, *Identification of new susceptibility loci for osteoarthritis (arcOGEN): a genome-wide association study.* Lancet, 2012. **380**(9844): p. 815-23.

40. Nakajima, M., et al., *New sequence variants in HLA class II/III region associated with susceptibility to knee osteoarthritis identified by genome-wide association study.* PLoS One, 2010. **5**(3): p. e9723.

41.   Kerkhof, H.J., et al., *A genome-wide association study identifies an osteoarthritis susceptibility locus on chromosome 7q22.* Arthritis Rheum, 2010. **62**(2): p. 499-510.

42.   Little, C.B., et al., *Matrix metalloproteinase 13-deficient mice are resistant to osteoarthritic cartilage erosion but not chondrocyte hypertrophy or osteophyte development.* Arthritis Rheum, 2009. **60**(12): p. 3723-33.

43.   Glasson, S.S., et al., *Deletion of active ADAMTS5 prevents cartilage degradation in a murine model of osteoarthritis.* Nature, 2005. **434**(7033): p. 644-8.

44.   Fukui, N., et al., *Zonal gene expression of chondrocytes in osteoarthritic cartilage.* Arthritis Rheum, 2008. **58**(12): p. 3843-53.

45.   Pearle, A.D., R.F. Warren, and S.A. Rodeo, *Basic science of articular cartilage and osteoarthritis.* Clin Sports Med, 2005. **24**(1): p. 1-12.

46.   Aigner, T., et al., *Large-scale gene expression profiling reveals major pathogenetic pathways of cartilage degeneration in osteoarthritis*. 2006. p. 3533-3544.

47.   Blanco, F.J., et al., *Osteoarthritis chondrocytes die by apoptosis. A possible pathway for osteoarthritis pathology.* Arthritis Rheum, 1998. **41**(2): p. 284-9.

48.   Lin, A.C., et al., *Modulating hedgehog signaling can attenuate the severity of osteoarthritis.* Nat Med, 2009. **15**(12): p. 1421-1425.

49.   Shortkroff, S. and K.E. Yates, *Alteration of matrix glycosaminoglycans diminishes articular chondrocytes' response to a canonical Wnt signal.* Osteoarthritis Cartilage, 2007. **15**(2): p. 147-54.

50.   Goldring, M.B., *Are bone morphogenetic proteins effective inducers of cartilage repair? Ex vivo transduction of muscle-derived stem cells.* Arthritis Rheum, 2006. **54**(2): p. 387-9.

51.   Sandell, L.J. and T. Aigner, *Articular cartilage and changes in arthritis. An introduction: cell biology of osteoarthritis.* Arthritis Res, 2001. **3**(2): p. 107-13.

52.   Wang, B., J.F. Fallon, and P.A. Beachy, *Hedgehog-regulated processing of Gli3 produces an anterior/posterior repressor gradient in the developing vertebrate limb.* Cell, 2000. **100**(4): p. 423-34.

53.   Zuzarte-Luis, V. and J.M. Hurle, *Programmed cell death in the embryonic vertebrate limb.* Semin Cell Dev Biol, 2005. **16**(2): p. 261-9.

54.   Dong, Y., et al., *Wnt-mediated regulation of chondrocyte maturation: modulation by TGF-beta.* J Cell Biochem, 2005. **95**(5): p. 1057-68.

55.   Bi, W., et al., *Sox9 is required for cartilage formation.* Nat Genet, 1999. **22**(1): p. 85-9.

56.   Komori, T., *Regulation of skeletal development by the Runx family of transcription factors.* J Cell Biochem, 2005. **95**(3): p. 445-53.

57.   Onyekwelu, I., M.B. Goldring, and C. Hidaka, *Chondrogenesis, joint formation, and articular cartilage regeneration.* J Cell Biochem, 2009. **107**(3): p. 383-92.

58.   Lefebvre, V., R.R. Behringer, and B. de Crombrugghe, *L-Sox5, Sox6 and Sox9 control essential steps of the chondrocyte differentiation pathway.* Osteoarthritis Cartilage, 2001. **9 Suppl A**: p. S69-75.

59.   Mackie, E.J., et al., *Endochondral ossification: How cartilage is converted into bone in the developing skeleton.* The International Journal of Biochemistry & Cell Biology, 2008. **40**(1): p. 46-62.

60.   Minina, E., et al., *BMP and Ihh/PTHrP signaling interact to coordinate chondrocyte proliferation and differentiation.* Development, 2001. **128**(22): p. 4523-34.

61.   Yoon, B.S., et al., *BMPs regulate multiple aspects of growth-plate chondrogenesis through opposing actions on FGF pathways.* Development, 2006. **133**(23): p. 4667-78.

62.   Ducy, P., et al., *Osf2/Cbfa1: a transcriptional activator of osteoblast differentiation.* Cell, 1997. **89**(5): p. 747-54.

63.   Nakashima, K., et al., *The novel zinc finger-containing transcription factor osterix is required for osteoblast differentiation and bone formation.* Cell, 2002. **108**(1): p. 17-29.

64.   Nusslein-Volhard, C. and E. Wieschaus, *Mutations affecting segment number and polarity in Drosophila.* Nature, 1980. **287**(5785): p. 795-801.

65.   Chen, X., et al., *Regulation of articular chondrocyte proliferation and differentiation by indian hedgehog and parathyroid hormone-related protein in mice.* 2008. p. 3788-3797.

66.   Colnot, C., et al., *Indian hedgehog synchronizes skeletal angiogenesis and perichondrial maturation with cartilage development.* 2005. p. 1057-1067.

67.   Pathi, S., et al., *Comparative biological responses to human Sonic, Indian, and Desert hedgehog.* Mech Dev, 2001. **106**(1-2): p. 107-17.

68.   Rohatgi, R., L. Milenkovic, and M.P. Scott, *Patched1 regulates hedgehog signaling at the primary cilium.* Science, 2007. **317**(5836): p. 372-6.

69.   Haycraft, C.J., et al., *Gli2 and Gli3 localize to cilia and require the intraflagellar transport protein polaris for processing and function.* PLoS Genet, 2005. **1**(4): p. e53.

70.   Jia, J., et al., *Shaggy/GSK3 antagonizes Hedgehog signalling by regulating Cubitus interruptus.* Nature, 2002. **416**(6880): p. 548-52.

71.   Incardona, J.P., et al., *Receptor-mediated endocytosis of soluble and membrane-tethered Sonic hedgehog by Patched-1.* Proc Natl Acad Sci U S A, 2000. **97**(22): p. 12044-9.

72.   Chen, Y. and G. Struhl, *Dual roles for patched in sequestering and transducing Hedgehog.* Cell, 1996. **87**(3): p. 553-63.

73.   Sinha, S. and J.K. Chen, *Purmorphamine activates the Hedgehog pathway by targeting Smoothened.* Nat Chem Biol, 2006. **2**(1): p. 29-30.

74.   Chen, J.K., et al., *Small molecule modulation of Smoothened activity.* Proc Natl Acad Sci U S A, 2002. **99**(22): p. 14071-6.

75.   Chung, U.I., et al., *Indian hedgehog couples chondrogenesis to osteogenesis in endochondral bone development.* J Clin Invest, 2001. **107**(3): p. 295-304.

76.   Mak, K.K., et al., *Indian hedgehog signals independently of PTHrP to promote chondrocyte hypertrophy.* Development, 2008. **135**(11): p. 1947-56.

77.   Kronenberg, H.M., *Developmental regulation of the growth plate.* Nature, 2003. **423**(6937): p. 332-6.

78.   Long, F., et al., *Ihh signaling is directly required for the osteoblast lineage in the endochondral skeleton.* Development, 2004. **131**(6): p. 1309-18.

79.   Maeda, Y., et al., *Indian Hedgehog produced by postnatal chondrocytes is essential for maintaining a growth plate and trabecular bone.* Proc Natl Acad Sci U S A, 2007. **104**(15): p. 6382-7.

80.   St-Jacques, B., M. Hammerschmidt, and A.P. McMahon, *Indian hedgehog signaling regulates proliferation and differentiation of chondrocytes and is essential for bone formation.* 1999. p. 2072-2086.

81.   Long, F., et al., *Genetic manipulation of hedgehog signaling in the endochondral skeleton reveals a direct role in the regulation of chondrocyte proliferation.* 2001. p. 5099-5108.

82.   Laurell, T., et al., *A novel 13 base pair insertion in the sonic hedgehog ZRS limb enhancer (ZRS/LMBR1) causes preaxial polydactyly with triphalangeal thumb.* Hum Mutat, 2012. **33**(7): p. 1063-6.

83.   Swingler, T.E., et al., *Degradome expression profiling in human articular cartilage.* Arthritis Res Ther, 2009. **11**(3): p. R96.

84.   Huangfu, D., et al., *Hedgehog signalling in the mouse requires intraflagellar transport proteins.* Nature, 2003. **426**(6962): p. 83-7.

85.   Corbit, K.C., et al., *Vertebrate Smoothened functions at the primary cilium.* Nature, 2005. **437**(7061): p. 1018-21.

86.   Beisson, J. and M. Wright, *Basal body/centriole assembly and continuity.* Curr Opin Cell Biol, 2003. **15**(1): p. 96-104.

87.   Marshall, W.F. and S. Nonaka, *Cilia: tuning in to the cell's antenna.* Curr Biol, 2006. **16**(15): p. R604-14.

88.   Praetorius, H.A. and K.R. Spring, *Bending the MDCK cell primary cilium increases intracellular calcium.* J Membr Biol, 2001. **184**(1): p. 71-9.

89.   Praetorius, H.A. and K.R. Spring, *Removal of the MDCK cell primary cilium abolishes flow sensing.* J Membr Biol, 2003. **191**(1): p. 69-76.

90.   Jensen, C.G., et al., *Ultrastructural, tomographic and confocal imaging of the chondrocyte primary cilium in situ.* Cell Biol Int, 2004. **28**(2): p. 101-10.

91.   Pedersen, L.B., S. Geimer, and J.L. Rosenbaum, *Dissecting the molecular mechanisms of intraflagellar transport in chlamydomonas.* Curr Biol, 2006. **16**(5): p. 450-9.

92.   Scholey, J.M. and K.V. Anderson, *Intraflagellar transport and cilium-based signaling.* Cell, 2006. **125**(3): p. 439-42.

93.   McGlashan, S.R., et al., *Primary cilia in osteoarthritic chondrocytes: from chondrons to clusters.* Dev Dyn, 2008. **237**(8): p. 2013-20.

94.   Wann, A.K. and M.M. Knight, *Primary cilia elongation in response to interleukin-1 mediates the inflammatory response.* Cell Mol Life Sci, 2012. **69**(17): p. 2967-77.

95.   Zhang, Q., et al., *Loss of the Tg737 protein results in skeletal patterning defects.* Dev Dyn, 2003. **227**(1): p. 78-90.

96.   Monroe, D.G., et al., *Update on Wnt signaling in bone cell biology and bone disease.* Gene, 2012. **492**(1): p. 1-18.

97.   Hu, H., et al., *Sequential roles of Hedgehog and Wnt signaling in osteoblast development.* Development, 2005. **132**(1): p. 49-60.

98.   Spater, D., et al., *Wnt9a signaling is required for joint integrity and regulation of Ihh during chondrogenesis.* Development, 2006. **133**(15): p. 3039-49.

99.   Rodda, S.J. and A.P. McMahon, *Distinct roles for Hedgehog and canonical Wnt signaling in specification, differentiation and maintenance of osteoblast progenitors.* Development, 2006. **133**(16): p. 3231-44.

100.   Kim, W., M. Kim, and E.H. Jho, *Wnt/beta-catenin signalling: from plasma membrane to nucleus.* Biochem J, 2013. **450**(1): p. 9-21.

101.   Rao, T.P. and M. Kuhl, *An updated overview on Wnt signaling pathways: a prelude for more.* Circ Res, 2010. **106**(12): p. 1798-806.

102.   Nalesso, G., et al., *WNT-3A modulates articular chondrocyte phenotype by activating both canonical and noncanonical pathways.* J Cell Biol, 2011. **193**(3): p. 551-64.

103.   Gong, Y., et al., *LDL receptor-related protein 5 (LRP5) affects bone accrual and eye development.* Cell, 2001. **107**(4): p. 513-23.

104.   Little, R.D., et al., *A mutation in the LDL receptor-related protein 5 gene results in the autosomal dominant high-bone-mass trait.* Am J Hum Genet, 2002. **70**(1): p. 11-9.

105.   Cui, Y., et al., *Lrp5 functions in bone to regulate bone mass.* Nat Med, 2011. **17**(6): p. 684-91.

106.   Oldridge, M., et al., *Dominant mutations in ROR2, encoding an orphan receptor tyrosine kinase, cause brachydactyly type B.* Nat Genet, 2000. **24**(3): p. 275-8.

107.   Afzal, A.R. and S. Jeffery, *One gene, two phenotypes: ROR2 mutations in autosomal recessive Robinow syndrome and autosomal dominant brachydactyly type B.* Hum Mutat, 2003. **22**(1): p. 1-11.

108.   DeChiara, T.M., et al., *Ror2, encoding a receptor-like tyrosine kinase, is required for cartilage and growth plate development.* Nat Genet, 2000. **24**(3): p. 271-4.

109. Takeuchi, S., et al., *Mouse Ror2 receptor tyrosine kinase is required for the heart development and limb formation.* Genes Cells, 2000. **5**(1): p. 71-8.

110. Bartel, D.P., *MicroRNAs: Genomics, Biogenesis, Mechanism, and Function.* Cell, 2004. **116**(2): p. 281-297.

111. Lee, R.C., R.L. Feinbaum, and V. Ambros, *The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14.* Cell, 1993. **75**(5): p. 843-854.

112. Kozomara, A. and S. Griffiths-Jones, *miRBase: integrating microRNA annotation and deep-sequencing data.* Nucleic Acids Res, 2011. **39**(Database issue): p. D152-7.

113. Lee, Y., et al., *The nuclear RNase III Drosha initiates microRNA processing.* Nature, 2003. **425**(6956): p. 415-9.

114. Yi, R., et al., *Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs.* Genes Dev, 2003. **17**(24): p. 3011-6.

115. Czech, B. and G.J. Hannon, *Small RNA sorting: matchmaking for Argonautes.* Nat Rev Genet, 2010. **12**(1): p. 19-31.

116. Liu, X., K. Fortin, and Z. Mourelatos, *MicroRNAs: biogenesis and molecular functions.* Brain Pathol, 2008. **18**(1): p. 113-21.

117. Griffiths-Jones, S., et al., *MicroRNA evolution by arm switching.* EMBO Rep, 2011. **12**(2): p. 172-7.

118. Su, H., et al., *Essential and overlapping functions for mammalian Argonautes in microRNA silencing.* Genes Dev, 2009. **23**(3): p. 304-17.

119. Grimson, A., et al., *MicroRNA targeting specificity in mammals: determinants beyond seed pairing.* Mol Cell, 2007. **27**(1): p. 91-105.

120. Chang, T.C. and J.T. Mendell, *microRNAs in vertebrate physiology and human disease.* Annu Rev Genomics Hum Genet, 2007. **8**: p. 215-39.

121. Guo, H., et al., *Mammalian microRNAs predominantly act to decrease target mRNA levels.* Nature, 2010. **466**(7308): p. 835-40.

122. Huntzinger, E. and E. Izaurralde, *Gene silencing by microRNAs: contributions of translational repression and mRNA decay.* Nat Rev Genet, 2011. **12**(2): p. 99-110.

123. Tang, R., et al., *Mouse miRNA-709 directly regulates miRNA-15a/16-1 biogenesis at the posttranscriptional level in the nucleus: evidence for a microRNA hierarchy system.* Cell Res, 2012. **22**(3): p. 504-15.

124. Seitz, H., *Redefining microRNA targets.* Curr Biol, 2009. **19**(10): p. 870-3.

125. Salmena, L., et al., *A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language?* Cell, 2011. **146**(3): p. 353-8.

126. Stark, A., et al., *Identification of Drosophila MicroRNA targets.* PLoS Biol, 2003. **1**(3): p. E60.

127. Ferretti, E., et al., *Concerted microRNA control of Hedgehog signalling in cerebellar neuronal progenitor and tumour cells.* Embo J, 2008. **27**(19): p. 2616-27.

128. Lewis, B.P., et al., *Prediction of mammalian microRNA targets.* Cell, 2003. **115**(7): p. 787-98.

129. Lewis, B.P., C.B. Burge, and D.P. Bartel, *Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets.* Cell, 2005. **120**(1): p. 15-20.

130. Friedman, R.C., et al., *Most mammalian mRNAs are conserved targets of microRNAs.* Genome Res, 2009. **19**(1): p. 92-105.

131. Lall, S., et al., *A genome-wide map of conserved microRNA targets in C. elegans.* Curr Biol, 2006. **16**(5): p. 460-71.

132. Kiriakidou, M., et al., *A combined computational-experimental approach predicts human microRNA targets.* Genes Dev, 2004. **18**(10): p. 1165-78.

133. John, B., et al., *Human MicroRNA targets.* PLoS Biol, 2004. **2**(11): p. e363.

134.   Hsu, S.D., et al., *miRTarBase: a database curates experimentally validated microRNA-target interactions.* Nucleic Acids Res, 2011. **39**(Database issue): p. D163-9.

135.   Bernstein, E., et al., *Dicer is essential for mouse development.* Nat Genet, 2003. **35**(3): p. 215-7.

136.   Harfe, B.D., et al., *The RNaseIII enzyme Dicer is required for morphogenesis but not patterning of the vertebrate limb.* Proc Natl Acad Sci U S A, 2005. **102**(31): p. 10898-903.

137.   Kobayashi, T., et al., *Dicer-dependent pathways regulate chondrocyte proliferation and differentiation.* Proc Natl Acad Sci U S A, 2008. **105**(6): p. 1949-54.

138.   Wienholds, E., et al., *MicroRNA expression in zebrafish embryonic development.* Science, 2005. **309**(5732): p. 310-1.

139.   Tuddenham, L., et al., *The cartilage specific microRNA-140 targets histone deacetylase 4 in mouse cells.* FEBS Lett, 2006. **580**(17): p. 4214-7.

140.   Miyaki, S., et al., *MicroRNA-140 is expressed in differentiated human articular chondrocytes and modulates interleukin-1 responses.* Arthritis Rheum, 2009. **60**(9): p. 2723-30.

141.   Yang, J., et al., *MiR-140 is co-expressed with Wwp2-C transcript and activated by Sox9 to target Sp1 in maintaining the chondrocyte proliferation.* FEBS Lett, 2011. **585**(19): p. 2992-7.

142.   Miyaki, S., et al., *MicroRNA-140 plays dual roles in both cartilage development and homeostasis*. 2010. p. 1173-1185.

143.   Song, B., et al., *Mechanism of chemoresistance mediated by miR-140 in human osteosarcoma and colon cancer cells.* Oncogene, 2009. **28**(46): p. 4065-74.

144.   Nicolas, F.E., et al., *Experimental identification of microRNA-140 targets by silencing and overexpressing miR-140.* Rna, 2008. **14**(12): p. 2513-20.

145.   Butz, H., et al., *MicroRNA profile indicates downregulation of the TGFbeta pathway in sporadic non-functioning pituitary adenomas.* Pituitary, 2011. **14**(2): p. 112-24.

146.   Nakamura, Y., et al., *Chondrocyte-specific microRNA-140 regulates endochondral bone development and targets Dnpep to modulate bone morphogenetic protein signaling.* Mol Cell Biol, 2011. **31**(14): p. 3019-28.

147.   Ye, W., et al., *The effect of central loops in miRNA:MRE duplexes on the efficiency of miRNA-mediated gene regulation.* PLoS One, 2008. **3**(3): p. e1719.

148.   Song, J., D. Kim, and E.J. Jin, *MicroRNA-488 suppresses cell migration through modulation of the focal adhesion activity during chondrogenic differentiation of chick limb mesenchymal cells.* Cell Biol Int, 2011. **35**(2): p. 179-85.

149.   Kim, D., et al., *MicroRNA-34a regulates migration of chondroblast and IL-1beta-induced degeneration of chondrocytes by targeting EphA5.* Biochem Biophys Res Commun, 2011. **415**(4): p. 551-7.

150.   Yang, B., et al., *MicroRNA-145 regulates chondrogenic differentiation of mesenchymal stem cells by targeting Sox9.* PLoS One, 2011. **6**(7): p. e21679.

151.   Guan, Y.J., et al., *MiR-365: a mechanosensitive microRNA stimulates chondrocyte differentiation through targeting histone deacetylase 4.* Faseb J, 2011. **25**(12): p. 4457-66.

152.   Dudek, K.A., et al., *Type II collagen expression is regulated by tissue-specific miR-675 in human articular chondrocytes.* J Biol Chem, 2010. **285**(32): p. 24381-7.

153.   Yan, C., et al., *MicroRNA regulation associated chondrogenesis of mouse MSCs grown on polyhydroxyalkanoates.* Biomaterials, 2011. **32**(27): p. 6435-44.

154.   Sumiyoshi, K., et al., *Identification of miR-1 as a micro RNA that supports late-stage differentiation of growth cartilage cells.* Biochem Biophys Res Commun, 2010. **402**(2): p. 286-90.

155. Swingler, T.E., et al., *The expression and function of microRNAs in chondrogenesis and osteoarthritis.* Arthritis Rheum, 2012.

156. Iliopoulos, D., et al., *Integrative MicroRNA and Proteomic Approaches Identify Novel Osteoarthritis Genes and Their Collaborative Metabolic and Inflammatory Networks.* PLoS ONE, 2008. **3**(11): p. e3740.

157. Jones, S.W., et al., *The identification of differentially expressed microRNA in osteoarthritic tissue that modulate the production of TNF-alpha and MMP13.* Osteoarthritis Cartilage, 2009. **17**(4): p. 464-72.

158. Abouheif, M.M., et al., *Silencing microRNA-34a inhibits chondrocyte apoptosis in a rat osteoarthritis model in vitro.* Rheumatology (Oxford), 2010. **49**(11): p. 2054-60.

159. Akhtar, N., et al., *MicroRNA-27b regulates the expression of matrix metalloproteinase 13 in human osteoarthritis chondrocytes.* Arthritis Rheum, 2010. **62**(5): p. 1361-71.

160. Liu, Y., et al., *MicroRNAs modulate the Wnt signaling pathway through targeting its inhibitors.* Biochem Biophys Res Commun, 2011. **408**(2): p. 259-64.

161. Davis, N., E. Mor, and R. Ashery-Padan, *Roles for Dicer1 in the patterning and differentiation of the optic cup neuroepithelium.* Development, 2011. **138**(1): p. 127-38.

162. Andl, T., et al., *The miRNA-processing enzyme dicer is essential for the morphogenesis and maintenance of hair follicles.* Curr Biol, 2006. **16**(10): p. 1041-9.

163. Harris, K.S., et al., *Dicer function is essential for lung epithelium morphogenesis.* Proc Natl Acad Sci U S A, 2006. **103**(7): p. 2208-13.

164. Watanabe, T., et al., *Dnm3os, a non-coding RNA, is required for normal growth and skeletal development in mice.* Dev Dyn, 2008. **237**(12): p. 3738-48.

165. Flynt, A.S., et al., *Zebrafish miR-214 modulates Hedgehog signaling to specify muscle cell fate.* Nat Genet, 2007. **39**(2): p. 259-63.

166. Lin, E.A., et al., *miR-199a, a bone morphogenic protein 2-responsive MicroRNA, regulates chondrogenesis via direct targeting to Smad1.* J Biol Chem, 2009. **284**(17): p. 11326-35.

167. de Pontual, L., et al., *Germline deletion of the miR-17 approximately 92 cluster causes skeletal and growth defects in humans.* Nat Genet, 2011. **43**(10): p. 1026-30.

168. Thatcher, E.J., et al., *MiRNA expression analysis during normal zebrafish development and following inhibition of the Hedgehog and Notch signaling pathways.* Dev Dyn, 2007. **236**(8): p. 2172-80.

169. Mott, J.L., et al., *Transcriptional suppression of mir-29b-1/mir-29a promoter by c-Myc, hedgehog, and NF-kappaB.* J Cell Biochem, 2010. **110**(5): p. 1155-64.

170. Radzikinas, K., et al., *A Shh/miR-206/BDNF cascade coordinates innervation and formation of airway smooth muscle.* J Neurosci, 2011. **31**(43): p. 15407-15.

171. Razumilava, N., et al., *miR-25 targets TNF-related apoptosis inducing ligand (TRAIL) death receptor-4 and promotes apoptosis resistance in cholangiocarcinoma.* Hepatology, 2012. **55**(2): p. 465-75.

172. Friggi-Grelin, F., L. Lavenant-Staccini, and P. Therond, *Control of antagonistic components of the hedgehog signaling pathway by microRNAs in Drosophila.* Genetics, 2008. **179**(1): p. 429-39.

173. Fareh, M., et al., *The miR 302-367 cluster drastically affects self-renewal and infiltration properties of glioma-initiating cells through CXCR4 repression and consequent disruption of the SHH-GLI-NANOG network.* Cell Death Differ, 2012. **19**(2): p. 232-44.

174. Hornstein, E., et al., *The microRNA miR-196 acts upstream of Hoxb8 and Shh in limb development.* Nature, 2005. **438**(7068): p. 671-4.

175. Jopling, C.L., S. Schütz, and P. Sarnow, *Position-Dependent Function for a Tandem MicroRNA miR-122-Binding Site Located in the Hepatitis C Virus RNA Genome.* Cell Host & Microbe, 2008. **4**(1): p. 77-85.

176. Lanford, R.E., et al., *Therapeutic silencing of microRNA-122 in primates with chronic hepatitis C virus infection.* Science, 2010. **327**(5962): p. 198-201.

177. Lanford, R.E., et al., *Therapeutic Silencing of MicroRNA-122 in Primates with Chronic Hepatitis C Virus Infection*. 2010. p. 198-201.

178. Reznikoff, C.A., D.W. Brankow, and C. Heidelberger, *Establishment and characterization of a cloned line of C3H mouse embryo cells sensitive to postconfluence inhibition of division.* Cancer Res, 1973. **33**(12): p. 3231-8.

179. Nakamura, T., et al., *Induction of Osteogenic Differentiation by Hedgehog Proteins.* Biochemical and Biophysical Research Communications, 1997. **237**(2): p. 465-469.

180. Denker, A.E., et al., *Chondrogenic differentiation of murine C3H10T1/2 multipotential mesenchymal cells: I. Stimulation by bone morphogenetic protein-2 in high-density micromass cultures.* Differentiation, 1999. **64**(2): p. 67-76.

181. Zhang, Q., et al., *Differential Toll-like receptor-dependent collagenase expression in chondrocytes*. 2008. p. 1633-1641.

182. Cox, J., et al., *A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics.* Nat Protoc, 2009. **4**(5): p. 698-705.

183. Stanton, H., et al., *ADAMTS5 is the major aggrecanase in mouse cartilage in vivo and in vitro.* Nature, 2005. **434**(7033): p. 648-52.

184. Shigeru, M., et al., *MicroRNA-140 is expressed in differentiated human articular chondrocytes and modulates interleukin-1 responses*. 2009. p. 2723-2730.

185. Abelson, J.F., et al., *Sequence variants in SLITRK1 are associated with Tourette's syndrome.* Science, 2005. **310**(5746): p. 317-20.

186. Clop, A., et al., *A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep.* Nat Genet, 2006. **38**(7): p. 813-8.

187. Wang, G., et al., *Variation in the miRNA-433 binding site of FGF20 confers risk for Parkinson disease by overexpression of alpha-synuclein.* Am J Hum Genet, 2008. **82**(2): p. 283-9.

188. Simon, D., et al., *A mutation in the 3'-UTR of the HDAC6 gene abolishing the post-transcriptional regulation mediated by hsa-miR-433 is linked to a new form of dominant X-linked chondrodysplasia.* Hum Mol Genet, 2010. **19**(10): p. 2015-27.

189. Gong, J., et al., *Genome-wide identification of SNPs in microRNA genes and the SNP effects on microRNA target binding and biogenesis.* Hum Mutat, 2012. **33**(1): p. 254-63.

190. Ling, L., et al., *Single nucleotide polymorphism associated with nonsyndromic cleft palate influences the processing of miR-140*. 2009. p. 856-862.

191. Chen, K. and N. Rajewsky, *Natural selection on human microRNA binding sites inferred from SNP data.* Nat Genet, 2006. **38**(12): p. 1452-6.

192. Saunders, M.A., H. Liang, and W.H. Li, *Human polymorphism at microRNAs and microRNA target sites.* Proc Natl Acad Sci U S A, 2007. **104**(9): p. 3300-5.

193. Eberhart, J.K., et al., *MicroRNA Mirn140 modulates Pdgf signaling during palatogenesis.* Nat Genet, 2008. **40**(3): p. 290-8.

194. Nakamura, Y., et al., *Wwp2 is essential for palatogenesis mediated by the interaction between Sox9 and mediator subunit 25.* Nat Commun, 2011. **2**: p. 251.

195. Forster, T., et al., *Finer linkage mapping of primary osteoarthritis susceptibility loci on chromosomes 4 and 16 in families with affected women.* Arthritis Rheum, 2004. **50**(1): p. 98-102.

196. Xu, Y., et al., *Identification of the pathogenic pathways in osteoarthritic hip cartilage: commonality and discord between hip and knee OA.* Osteoarthritis Cartilage, 2012. **20**(9): p. 1029-38.

197. Loughlin, J., et al., *Functional variants within the secreted frizzled-related protein 3 gene are associated with hip osteoarthritis in females.* Proc Natl Acad Sci U S A, 2004. **101**(26): p. 9757-62.

198. Li, L., et al., *Biological and epidemiological evidence of interaction of infant genotypes at Rs7205289 and maternal passive smoking in cleft palate.* Am J Med Genet A, 2011. **155A**(12): p. 2940-8.

199. Li, L., et al., *Single nucleotide polymorphism associated with nonsyndromic cleft palate influences the processing of miR-140.* Am J Med Genet A, 2010. **152A**(4): p. 856-62.

200. Johnson, A.D., et al., *SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap.* Bioinformatics, 2008. **24**(24): p. 2938-9.

201. Mestdagh, P., et al., *A novel and universal method for microRNA RT-qPCR data normalization.* Genome Biol, 2009. **10**(6): p. R64.

202. Lagos-Quintana, M., et al., *Identification of tissue-specific microRNAs from mouse.* Curr Biol, 2002. **12**(9): p. 735-9.

203. Landgraf, P., et al., *A mammalian microRNA expression atlas based on small RNA library sequencing.* Cell, 2007. **129**(7): p. 1401-14.

204. Weber, M.J., *New human and mouse microRNA genes found by homology search.* Febs J, 2005. **272**(1): p. 59-73.

205. Cai, X., et al., *Kaposi's sarcoma-associated herpesvirus expresses an array of viral microRNAs in latently infected cells.* Proc Natl Acad Sci U S A, 2005. **102**(15): p. 5570-5.

206. Kim, J., et al., *Identification of many microRNAs that copurify with polyribosomes in mammalian neurons.* Proc Natl Acad Sci U S A, 2004. **101**(1): p. 360-5.

207. Poy, M.N., et al., *A pancreatic islet-specific microRNA regulates insulin secretion.* Nature, 2004. **432**(7014): p. 226-30.

208. Ahn, H.W., et al., *MicroRNA transcriptome in the newborn mouse ovaries determined by massive parallel sequencing.* Mol Hum Reprod, 2010. **16**(7): p. 463-71.

209. Chiang, H.R., et al., *Mammalian microRNAs: experimental evaluation of novel and previously annotated genes.* Genes Dev, 2010. **24**(10): p. 992-1009.

210. Sievers, F., et al., *Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega.* Mol Syst Biol, 2011. **7**: p. 539.

211. Huang da, W., B.T. Sherman, and R.A. Lempicki, *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.* Nat Protoc, 2009. **4**(1): p. 44-57.

212. Vortkamp, A., et al., *Regulation of rate of cartilage differentiation by Indian hedgehog and PTH-related protein.* Science, 1996. **273**(5275): p. 613-22.

213. Nakamura, T., et al., *Induction of osteogenic differentiation by hedgehog proteins.* Biochem Biophys Res Commun, 1997. **237**(2): p. 465-9.

214. Kinto, N., et al., *Fibroblasts expressing Sonic hedgehog induce osteoblast differentiation and ectopic bone formation.* FEBS Lett, 1997. **404**(2-3): p. 319-23.

215. Shea, C.M., et al., *BMP treatment of C3H10T1/2 mesenchymal stem cells induces both chondrogenesis and osteogenesis.* J Cell Biochem, 2003. **90**(6): p. 1112-27.

216. Wu, X., et al., *Purmorphamine induces osteogenesis by activation of the hedgehog signaling pathway.* Chem Biol, 2004. **11**(9): p. 1229-38.

217. Kevorkian, L., et al., *Expression profiling of metalloproteinases and their inhibitors in cartilage.* Arthritis Rheum, 2004. **50**(1): p. 131-41.

218. Saetrom, P., et al., *Distance constraints between microRNA target sites dictate efficacy and cooperativity.* Nucleic Acids Res, 2007. **35**(7): p. 2333-42.

219. Hon, L.S. and Z. Zhang, *The roles of binding site arrangement and combinatorial targeting in microRNA repression of gene expression.* Genome Biol, 2007. **8**(8): p. R166.

220. Matsukawa, T., et al., *MicroRNA-125b regulates the expression of aggrecanase-1 (ADAMTS-4) in human osteoarthritic chondrocytes.* Arthritis Res Ther, 2013. **15**(1): p. R28.

221. Xu, N., et al., *MicroRNA-125b down-regulates matrix metallopeptidase 13 and inhibits cutaneous squamous cell carcinoma cell proliferation, migration, and invasion.* J Biol Chem, 2012. **287**(35): p. 29899-908.

222. Stecca, B. and I.A.A. Ruiz, *Context-dependent regulation of the GLI code in cancer by HEDGEHOG and non-HEDGEHOG signals.* J Mol Cell Biol, 2010. **2**(2): p. 84-95.

223. Boissart, C., et al., *miR-125 potentiates early neural specification of human embryonic stem cells.* Development, 2012. **139**(7): p. 1247-57.

224. Vinther, J., et al., *Identification of miRNA targets with stable isotope labeling by amino acids in cell culture.* Nucleic Acids Res, 2006. **34**(16): p. e107.

225. Lim, L.P., et al., *Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs.* Nature, 2005. **433**(7027): p. 769-73.

226. Baigude, H., et al., *miR-TRAP: a benchtop chemical biology strategy to identify microRNA targets.* Angew Chem Int Ed Engl, 2012. **51**(24): p. 5880-3.

227. Yoon, J.H., S. Srikantan, and M. Gorospe, *MS2-TRAP (MS2-tagged RNA affinity purification): tagging RNA to identify associated miRNAs.* Methods, 2012. **58**(2): p. 81-7.

228. Krutzfeldt, J., et al., *Silencing of microRNAs in vivo with 'antagomirs'.* Nature, 2005. **438**(7068): p. 685-9.

229. Xu, G., et al., *Transcriptome and targetome analysis in MIR155 expressing cells using RNA-seq.* Rna, 2010. **16**(8): p. 1610-22.

230. Selbach, M., et al., *Widespread changes in protein synthesis induced by microRNAs.* Nature, 2008. **455**(7209): p. 58-63.

231. Jacobson, E.L., et al., *Effect of nicotinamide analogues on recovery from DNA damage in C3H10T1/2 cells.* Cancer Res, 1984. **44**(6): p. 2485-92.

232. Kaller, M., et al., *Genome-wide characterization of miR-34a induced changes in protein and mRNA expression by a combined pulsed SILAC and microarray analysis.* Mol Cell Proteomics, 2011. **10**(8): p. M111 010462.

233. Baek, D., et al., *The impact of microRNAs on protein output.* Nature, 2008. **455**(7209): p. 64-71.

234. Geiger, T., et al., *Super-SILAC mix for quantitative proteomics of human tumor tissue.* Nat Methods, 2010. **7**(5): p. 383-5.

235. Zanivan, S., M. Krueger, and M. Mann, *In vivo quantitative proteomics: the SILAC mouse.* Methods Mol Biol, 2012. **757**: p. 435-50.

236. Ingolia, N.T., et al., *Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling.* Science, 2009. **324**(5924): p. 218-23.

237. Landthaler, M., et al., *Molecular characterization of human Argonaute-containing ribonucleoprotein complexes and their bound target mRNAs.* Rna, 2008. **14**(12): p. 2580-96.

238. Orom, U.A. and A.H. Lund, *Experimental identification of microRNA targets.* Gene, 2011. **451**(1-2): p. 1-5.

239. Hsu, R.J., H.J. Yang, and H.J. Tsai, *Labeled microRNA pull-down assay system: an experimental approach for high-throughput identification of microRNA-target mRNAs.* Nucleic Acids Res, 2009. **37**(10): p. e77.

240. Beitzinger, M., et al., *Identification of human microRNA targets from isolated argonaute protein complexes.* RNA Biol, 2007. **4**(2): p. 76-84.

241. Tan, L.P., et al., *A high throughput experimental approach to identify miRNA targets in human cells.* Nucleic Acids Res, 2009. **37**(20): p. e137.

242. Chi, S.W., et al., *Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps.* Nature, 2009. **460**(7254): p. 479-86.

243. Hafner, M., et al., *Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP.* Cell, 2010. **141**(1): p. 129-41.

244. Alexiou, P., et al., *Lost in translation: an assessment and perspective for computational microRNA target identification.* Bioinformatics, 2009. **25**(23): p. 3049-55.

245. Xiao, Y., et al., *Prioritizing cancer-related key miRNA-target interactions by integrative genomics.* Nucleic Acids Res, 2012. **40**(16): p. 7653-65.

246. Capurro, M.I., et al., *Glypican-3 Inhibits Hedgehog Signaling during Development by Competing with Patched for Hedgehog Binding.* Developmental Cell, 2008. **14**(5): p. 700-711.

247. Li, F., et al., *Glypican-5 stimulates rhabdomyosarcoma cell proliferation by activating Hedgehog signaling*. 2011. p. 691-704.

248. Djuranovic, S., A. Nahvi, and R. Green, *miRNA-mediated gene silencing by translational repression followed by mRNA deadenylation and decay.* Science, 2012. **336**(6078): p. 237-40.

249. Bazzini, A.A., M.T. Lee, and A.J. Giraldez, *Ribosome profiling shows that miR-430 reduces translation before causing mRNA decay in zebrafish.* Science, 2012. **336**(6078): p. 233-7.

250. Desbordes, S.C. and B. Sanson, *The glypican Dally-like is required for Hedgehog signalling in the embryonic epidermis of Drosophila.* Development, 2003. **130**(25): p. 6245-55.

251. Han, C., et al., *Drosophila glypicans control the cell-to-cell movement of Hedgehog by a dynamin-independent process.* Development, 2004. **131**(3): p. 601-11.

252. Gallet, A., L. Staccini-Lavenant, and P.P. Therond, *Cellular trafficking of the glypican Dally-like is required for full-strength Hedgehog signaling and wingless transcytosis.* Dev Cell, 2008. **14**(5): p. 712-25.

253. Yan, D., et al., *The cell-surface proteins Dally-like and Ihog differentially regulate Hedgehog signaling strength and range during development.* Development, 2010. **137**(12): p. 2033-44.

254. Wang, Y., et al., *Mechanism of microRNA-target interaction: molecular dynamics simulations and thermodynamics analysis.* PLoS Comput Biol, 2010. **6**(7): p. e1000866.

255. Jalvy-Delvaille, S., et al., *Molecular basis of differential target regulation by miR-96 and miR-182: the Glypican-3 as a model.* Nucleic Acids Res, 2012. **40**(3): p. 1356-65.

256. Hu, H.Y., et al., *Evolution of the human-specific microRNA miR-941.* Nat Commun, 2012. **3**: p. 1145.

257. Kedde, M., et al., *RNA-binding protein Dnd1 inhibits microRNA access to target mRNA.* Cell, 2007. **131**(7): p. 1273-86.

258. Rajewsky, N., *microRNA target predictions in animals.* Nat Genet, 2006. **38 Suppl**: p. S8-13.

259. Bethune, J., C.G. Artus-Revel, and W. Filipowicz, *Kinetic analysis reveals successive steps leading to miRNA-mediated silencing in mammalian cells.* EMBO Rep, 2012. **13**(8): p. 716-23.

260. Staton, A.A. and A.J. Giraldez, *Use of target protector morpholinos to analyze the physiological roles of specific miRNA-mRNA pairs in vivo.* Nat Protoc, 2011. **6**(12): p. 2035-49.

261. Place, R.F., et al., *MicroRNA-373 induces expression of genes with complementary promoter sequences.* Proc Natl Acad Sci U S A, 2008. **105**(5): p. 1608-13.

262. Lehmann, S.M., et al., *An unconventional role for miRNA: let-7 activates Toll-like receptor 7 and causes neurodegeneration.* Nat Neurosci, 2012. **15**(6): p. 827-35.

263. Fabbri, M., et al., *MicroRNAs bind to Toll-like receptors to induce prometastatic inflammatory response.* Proc Natl Acad Sci U S A, 2012. **109**(31): p. E2110-6.

264.   Yaddanapudi, K., et al., *Toll-like receptor 3 regulates neural stem cell proliferation by modulating the Sonic Hedgehog pathway.* PLoS One, 2011. **6**(10): p. e26766.

265.   Filmus, J., M. Capurro, and J. Rast, *Glypicans.* Genome Biol, 2008. **9**(5): p. 224.

266.   Traister, A., W. Shi, and J. Filmus, *Mammalian Notum induces the release of glypicans and other GPI-anchored proteins from the cell surface.* Biochem J, 2008. **410**(3): p. 503-11.

267.   Rogers, K.W. and A.F. Schier, *Morphogen gradients: from generation to interpretation.* Annu Rev Cell Dev Biol. **27**: p. 377-407.

268.   McCarthy, R.A., et al., *Megalin functions as an endocytic sonic hedgehog receptor.* J Biol Chem, 2002. **277**(28): p. 25660-7.

269.   Gritli-Linde, A., et al., *The whereabouts of a morphogen: direct evidence for short- and graded long-range activity of hedgehog signaling peptides.* Dev Biol, 2001. **236**(2): p. 364-86.

270.   The, I., Y. Bellaiche, and N. Perrimon, *Hedgehog movement is regulated through tout velu-dependent synthesis of a heparan sulfate proteoglycan.* Mol Cell, 1999. **4**(4): p. 633-9.

271.   Lum, L., et al., *Identification of Hedgehog pathway components by RNAi in Drosophila cultured cells.* Science, 2003. **299**(5615): p. 2039-45.

272.   Yan, D., et al., *The cell-surface proteins Dally-like and Ihog differentially regulate Hedgehog signaling strength and range during development.* Development. **137**(12): p. 2033-44.

273.   Torroja, C., N. Gorfinkiel, and I. Guerrero, *Patched controls the Hedgehog gradient by endocytosis in a dynamin-dependent manner, but this internalization does not play a major role in signal transduction.* Development, 2004. **131**(10): p. 2395-408.

274.   Cui, S., et al., *Evidence from Human and Zebrafish that GPC1 is a Biliary Atresia Susceptibility Gene.* Gastroenterology, 2013.

275.   Bergeron, S.A., et al., *Expression profiling identifies novel Hh/Gli-regulated genes in developing zebrafish embryos.* Genomics, 2008. **91**(2): p. 165-77.

276.   Lin, X., *Functions of heparan sulfate proteoglycans in cell signaling during development.* Development, 2004. **131**(24): p. 6009-21.

277.   Keller, K.M., P.R. Brauer, and J.M. Keller, *Modulation of cell surface heparan sulfate structure by growth of cells in the presence of chlorate.* Biochemistry, 1989. **28**(20): p. 8100-7.

278.   Leong, J.M., et al., *Hemagglutination and proteoglycan binding by the Lyme disease spirochete, Borrelia burgdorferi.* Infect Immun, 1995. **63**(3): p. 874-83.

279.   Ayers, K.L., et al., *Dally and Notum regulate the switch between low and high level Hedgehog pathway signalling.* Development, 2012. **139**(17): p. 3168-79.

280.   Capurro, M.I., et al., *Glypican-3 promotes the growth of hepatocellular carcinoma by stimulating canonical Wnt signaling.* Cancer Res, 2005. **65**(14): p. 6245-54.

281.   Ogden, S.K., et al., *Regulation of Hedgehog signaling: a complex story.* Biochem Pharmacol, 2004. **67**(5): p. 805-14.

282.   Gao, C. and Y.G. Chen, *Dishevelled: The hub of Wnt signaling.* Cell Signal, 2010. **22**(5): p. 717-27.

283.   Nakamura, Y., et al., *Unique roles of microRNA140 and its host gene WWP2 in cartilage biology.* J Musculoskelet Neuronal Interact, 2008. **8**(4): p. 321-2.

284.   Fischer, L., G. Boland, and R.S. Tuan, *Wnt-3A enhances bone morphogenetic protein-2-mediated chondrogenesis of murine C3H10T1/2 mesenchymal cells.* J Biol Chem, 2002. **277**(34): p. 30870-8.

285.   Anton, R., et al., *A systematic screen for micro-RNAs regulating the canonical Wnt pathway.* PLoS One, 2011. **6**(10): p. e26257.

286.   Baeg, G.H., et al., *Heparan sulfate proteoglycans are critical for the organization of the extracellular distribution of Wingless.* Development, 2001. **128**(1): p. 87-94.

287.   Li, G., et al., *MicroRNA-324-3p regulates nasopharyngeal carcinoma radioresistance by directly targeting WNT2B.* Eur J Cancer, 2013.

288.   van der Kraan, P.M., et al., *TGF-beta signaling in chondrocyte terminal differentiation and osteoarthritis: modulation and integration of signaling pathways through receptor-Smads.* Osteoarthritis Cartilage, 2009. **17**(12): p. 1539-45.

289.   Blaney Davidson, E.N., et al., *Increase in ALK1/ALK5 ratio as a cause for elevated MMP-13 expression in osteoarthritis in humans and mice.* J Immunol, 2009. **182**(12): p. 7937-45.

290.   Kim, J., et al., *Functional genomic screen for modulators of ciliogenesis and cilium length.* Nature, 2010. **464**(7291): p. 1048-51.

291.   Blain, E.J., *Involvement of the cytoskeletal elements in articular cartilage homeostasis and pathology.* Int J Exp Pathol, 2009. **90**(1): p. 1-15.

292.   Hayami, T., et al., *Expression of the cartilage derived anti-angiogenic factor chondromodulin-I decreases in the early stage of experimental osteoarthritis.* J Rheumatol, 2003. **30**(10): p. 2207-17.

293.   Suri, S. and D.A. Walsh, *Osteochondral alterations in osteoarthritis.* Bone, 2012. **51**(2): p. 204-11.

294.   Salim, H., et al., *miRNA-214 modulates radiotherapy response of non-small cell lung cancer cells through regulation of p38MAPK, apoptosis and senescence.* Br J Cancer, 2012. **107**(8): p. 1361-73.

295.   Patel, N., *microRNA mediated regulation of APP gene expression*. 2010, Drexel University.

296.   Hooper, N.M., *Glypican-1 facilitates prion conversion in lipid rafts.* J Neurochem, 2011. **116**(5): p. 721-5.

297.   Su, G., et al., *Glypican-1 is frequently overexpressed in human gliomas and enhances FGF-2 signaling in glioma cells.* Am J Pathol, 2006. **168**(6): p. 2014-26.

298.   Stappert, L., et al., *MicroRNA-Based Promotion of Human Neuronal Differentiation and Subtype Specification.* PLoS One, 2013. **8**(3): p. e59011.

299.   Zhu, M., et al., *Activation of beta-catenin signaling in articular chondrocytes leads to osteoarthritis-like phenotype in adult beta-catenin conditional activation mice.* J Bone Miner Res, 2009. **24**(1): p. 12-21.

300.   Zhu, M., et al., *Inhibition of beta-catenin signaling in articular chondrocytes results in articular cartilage destruction.* Arthritis Rheum, 2008. **58**(7): p. 2053-64.

301.   Chen, M., et al., *Inhibition of beta-catenin signaling causes defects in postnatal cartilage development.* J Cell Sci, 2008. **121**(Pt 9): p. 1455-65.

302.   Peterson, R. and J. Turnbull, *Sonic hedgehog is cytoprotective against oxidative challenge in a cellular model of amyotrophic lateral sclerosis.* J Mol Neurosci, 2012. **47**(1): p. 31-41.

303.   Jin, L., et al., *MicroRNA-149*, a p53-responsive microRNA, functions as an oncogenic regulator in human melanoma.* Proc Natl Acad Sci U S A, 2011. **108**(38): p. 15840-5.

304.   Meng, X., et al., *Suppressor of fused negatively regulates beta-catenin signaling.* J Biol Chem, 2001. **276**(43): p. 40113-9.

305.   Price, M.A. and D. Kalderon, *Proteolysis of the Hedgehog signaling effector Cubitus interruptus requires phosphorylation by Glycogen Synthase Kinase 3 and Casein Kinase 1.* Cell, 2002. **108**(6): p. 823-35.

306.   Pallerla, S.R., et al., *Heparan sulfate Ndst1 gene function variably regulates multiple signaling pathways during mouse development.* Dev Dyn, 2007. **236**(2): p. 556-63.

307.   Eugster, C., et al., *Lipoprotein-heparan sulfate interactions in the Hh pathway.* Dev Cell, 2007. **13**(1): p. 57-71.

308.   Glise, B., et al., *Shifted, the Drosophila ortholog of Wnt inhibitory factor-1, controls the distribution and movement of Hedgehog.* Dev Cell, 2005. **8**(2): p. 255-66.

309. McLellan, J.S., et al., *Structure of a heparin-dependent complex of Hedgehog and Ihog.* Proc Natl Acad Sci U S A, 2006. **103**(46): p. 17208-13.

310. Serpe, M., et al., *The BMP-binding protein Crossveinless 2 is a short-range, concentration-dependent, biphasic modulator of BMP signaling in Drosophila.* Dev Cell, 2008. **14**(6): p. 940-53.

311. Chang, S.C., et al., *Two distinct sites in sonic Hedgehog combine for heparan sulfate interactions and cell signaling functions.* J Biol Chem, 2011. **286**(52): p. 44391-402.

312. Chan, J.A., et al., *Proteoglycan interactions with Sonic Hedgehog specify mitogenic responses.* Nat Neurosci, 2009. **12**(4): p. 409-17.

313. Cortes, M., A.T. Baria, and N.B. Schwartz, *Sulfation of chondroitin sulfate proteoglycans is necessary for proper Indian hedgehog signaling in the developing growth plate.* Development, 2009. **136**(10): p. 1697-706.

314. Uziel, T., et al., *The miR-17~92 cluster collaborates with the Sonic Hedgehog pathway in medulloblastoma.* Proc Natl Acad Sci U S A, 2009. **106**(8): p. 2812-7.

315. Ventura, A., et al., *Targeted deletion reveals essential and overlapping functions of the miR-17 through 92 family of miRNA clusters.* Cell, 2008. **132**(5): p. 875-86.

316. Mendell, J.T., *miRiad roles for the miR-17-92 cluster in development and disease.* Cell, 2008. **133**(2): p. 217-22.

317. Guo, L., et al., *Dynamic evolution of mir-17-92 gene cluster and related miRNA gene families in vertebrates.* Mol Biol Rep, 2013. **40**(4): p. 3147-53.

318. Wang, C.L., et al., *Activation of an oncogenic microRNA cistron by provirus integration.* Proc Natl Acad Sci U S A, 2006. **103**(49): p. 18680-4.

319. Dell'Accio, F., C. De Bari, and F.P. Luyten, *Molecular markers predictive of the capacity of expanded human articular chondrocytes to form stable cartilage in vivo.* Arthritis Rheum, 2001. **44**(7): p. 1608-19.

320. Wei, F., et al., *Activation of Indian hedgehog promotes chondrocyte hypertrophy and upregulation of MMP-13 in human osteoarthritic cartilage.* Osteoarthritis Cartilage. **20**(7): p. 755-63.

321. Taschner, M.J., et al., *Ca2+/Calmodulin-dependent kinase II signaling causes skeletal overgrowth and premature chondrocyte maturation.* Dev Biol, 2008. **317**(1): p. 132-46.

322. Koziel, L., et al., *Ext1-dependent heparan sulfate regulates the range of Ihh signaling during endochondral ossification.* Dev Cell, 2004. **6**(6): p. 801-13.

323. Pilia, G., et al., *Mutations in GPC3, a glypican gene, cause the Simpson-Golabi-Behmel overgrowth syndrome.* Nat Genet, 1996. **12**(3): p. 241-7.

324. Milenkovic, L., et al., *Mouse patched1 controls body size determination and limb patterning.* Development, 1999. **126**(20): p. 4431-40.

325. Cook, A., et al., *Genetic heterogeneity in families with hereditary multiple exostoses.* Am J Hum Genet, 1993. **53**(1): p. 71-9.

326. Ahn, J., et al., *Cloning of the putative tumour suppressor gene for hereditary multiple exostoses (EXT1).* Nat Genet, 1995. **11**(2): p. 137-43.

327. Wu, Y.Q., et al., *Assignment of a second locus for multiple exostoses to the pericentromeric region of chromosome 11.* Hum Mol Genet, 1994. **3**(1): p. 167-71.

328. Stickens, D., et al., *The EXT2 multiple exostoses gene defines a family of putative tumour suppressor genes.* Nat Genet, 1996. **14**(1): p. 25-32.

329. Echtermeyer, F., et al., *Syndecan-4 regulates ADAMTS-5 activation and cartilage breakdown in osteoarthritis.* Nat Med, 2009. **15**(9): p. 1072-6.

330. Bertrand, J., et al., *Syndecan 4 supports bone fracture repair, but not fetal skeletal development, in mice.* Arthritis Rheum, 2013. **65**(3): p. 743-52.

331.   Syrrou, M., et al., *Glypican 1 gene: good candidate for brachydactyly type E.* Am J Med Genet, 2002. **108**(4): p. 310-4.

332.   Litwack, E.D., et al., *Expression of the heparan sulfate proteoglycan glypican-1 in the developing rodent.* Dev Dyn, 1998. **211**(1): p. 72-87.