# Dissecting cellular function of fibronectin in osteoarthritic cartilage

Hoolwerff, M. van

# Dissecting cellular function of fibronectin in osteoarthritic cartilage

Marcella van Hoolwerff

**Dissecting cellular function of fibronectin in osteoarthritic cartilage**
M. van Hoolwerff, MSc
ISBN: 978-94-6458-423-3



This research was financially supported by the Reumafonds under agreement DAF-16-1-405 and was performed in the framework of the Medical Delta program Regenerative Medicine 4D: Generating complex tissues with stem cells and printing technology and Improving Mobility with Technology.

Medical Delta and the Nederlandse Vereniging voor Matrix Biologie are gratefully acknowledged for financial support for the printing costs of this thesis.

Bookdesign/lay-out: Selina Abraham, Marcella van Hoolwerff
Printing: Ridderprint

# Dissecting cellular function of fibronectin in osteoarthritic cartilage

**Proefschrift**

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr.ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op dinsdag 6 september 2022
klokke 16.15 uur

door

Marcella van Hoolwerff
geboren te Delft
in 1990

Voor papa

| | |
|---|---|
| Promotor | Prof. dr. I. Meulenbelt |
| Co-promotor | Dr. Y.F.M. Ramos |
| | |
| Commissieleden | Prof. dr. B.T. Heijmans |
| | Prof. dr. P.M. van der Kraan (Radboud Universiteit) |
| | Prof. dr. F. Zaucke (University Hospital Frankfurt) |
| | Prof. dr. J.V.M.G Bovée |

# Contents

# Chapter 1

Introduction

## Osteoarthritis

Osteoarthritis (OA) is an age-related, heterogeneous, degenerative disease of the articular joints, characterized by pathological changes in amongst others cartilage, synovium, and subchondral bone [1]. OA may develop in any joint, but most commonly affects knees, hips, and hands [2]. Risk factors for OA can be divided in systemic factors, including increasing age, female sex, obesity, and genetic predisposition, and local factors, such as joint injury and abnormal loading on the joint [3, 4].

Currently, OA is globally the most prevalent joint disease and affected almost 1.5 million people in the Netherlands in 2019 [4-6]. The societal and economic burden of OA is expected to increase exponentially in the coming years, given the ageing of the population and the increasing prevalence of obesity [4, 5, 7-9]. The social burden is marked by pain and stiffness in the articular joints and decreased mobility, which can result in significant disability in everyday life [8]. The considerable cost of OA lies in the absence from work, social services, and medical costs, where the costs of OA were 1.4% of total health care costs in the Netherlands in 2017 [5, 6, 10]. Despite these detrimental consequences, to date no effective treatment to halt or reverse the progression of the disease is available except pain relief medication and, in a final stage, joint replacement surgery. One major reason for unsuccessful disease modifying OA drug development is inadequate insight into underlying pathophysiological mechanisms.



## Articular cartilage

Articular cartilage covers the end of long bones and is designed to withstand the forces that are exerted upon the joint by providing a low-friction, load-bearing surface between bones [11, 12]. It is aneural and avascular tissue, however, the cellular organisation and composition is complex [13, 14]. To be able to transfer the loads during mechanical loading and joint motion, the tissue consists of mainly water (>70%) and extracellular matrix (ECM) that is synthesized and maintained by the only cell type present: chondrocytes (**Figure 1A**). The major components of cartilage ECM are proteoglycans, mainly aggrecan, various types of collagen, the most abundant being type II collagen, and glycoproteins, such as fibronectin [14, 15]. Proteoglycan consists of a central core protein to which one or multiple negatively charged glycosaminoglycan (GAG) side chains are covalently linked. Generally, proteoglycans exist as aggregates consisting of non-covalently associated proteoglycans with hyaluronic acid and link protein [14]. Type II collagen fibrils form a network that entraps negatively charged proteoglycan aggregates, other small proteoglycans and cartilage matrix proteins. The arrangement of the collagen fibril network differs in the different layers of the cartilage, referred to as superficial layer, middle zone, and deep zone (**Figure 1A**). In the superficial layer the collagen fibrils are parallel to the surface. The distribution becomes more random in the middle zone and is perpendicular to the surface in the deep zone [16], together significantly enhancing tensile strength while the proteoglycan aggregates provide compressive resilience. Upon compression, water molecules

associated with the hydrophilic GAG chains of proteoglycans are dissipated from the cartilage and when compression is released, the proteoglycans attract the water molecules via osmosis back into the matrix, restoring original cartilage composition [15].

Chondrocytes have different morphologies and organization, depending on their localization in the articular cartilage layers. In the superficial layer, chondrocytes are closely together and more flattened, while in the middle zone there are less cells and they are rounder. Furthermore, in the deep zone they become larger and form strings parallel to the collagen fibrils (**Figure 1A**). In healthy adult cartilage the chondrocytes are in a maturational arrested state and there is an equilibrium between anabolic and catabolic processes. Turnover of collagen type II is low, while glycosaminoglycans and other non-collagen matrix proteins are more readily produced [15]. Calcified cartilage is the interface between the articular cartilage and subchondral bone, where there are few chondrocytes with a hypertrophic phenotype. The calcified cartilage is separated from the articular cartilage by a thin line marking the mineralization called the tidemark.

## Pericellular matrix of articular cartilage

One aspect of the complexity of the ECM of articular cartilage lies in the matrix heterogeneity, where interterritorial and pericellular matrices are recognized [17, 18]. The pericellular matrix (PCM) is a narrow matrix region of approximately 2-4 μm directly surrounding the chondrocytes, that is biochemically and biomechanically distinct from the interterritorial matrix (**Figure 1B**) [19]. The PCM is characterized by high concentrations of proteoglycans,

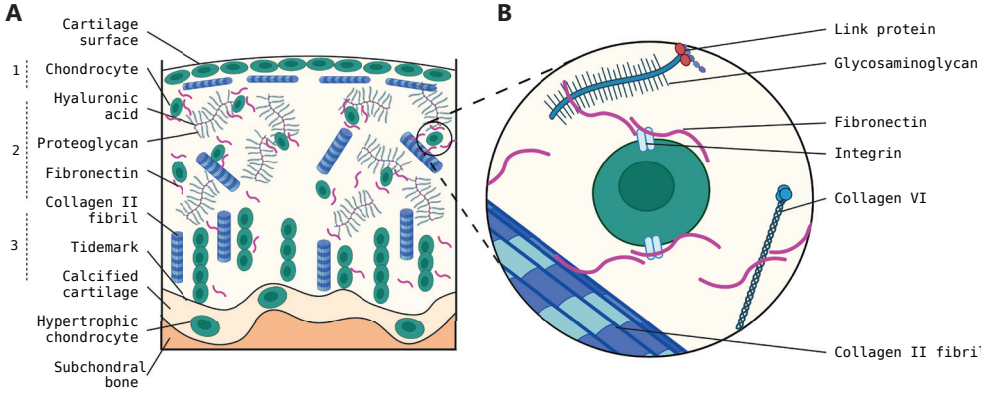



**Figure 1 | Overview of composition articular cartilage and the pericellular matrix directly surrounding the chondrocyte.** (**A**) Schematic representation of structure and morphology of chondrocytes and extracellular matrix proteins in articular cartilage, including collagen type II fibrils, proteoglycan aggregates, and fibronectin. Dotted lines represent the different layers in cartilage: 1) Superficial layer. 2) Middle zone. 3) Deep zone. (**B**) Schematic representation of the pericellular matrix surrounding a chondrocyte. The pericellular matrix is rich in proteoglycans, collagen type II, fibronectin and collagen type VI. (Created with Biorender.com).

collagen type II and fibronectin, as well as collagen type VI, which is exclusively present in the PCM and not in the interterritorial matrix [20]. Although the complete role of the PCM is still unknown, it has been shown to serve as a transducer of biomechanical and biochemical signals to the chondrocytes [17, 21]. The properties of the PCM can regulate mechanical and chemical surroundings of the chondrocyte, thereby influencing cartilage matrix homeostasis [22]. Homozygous inactivation of *COL6A1* in mice showed structurally intact PCM in the articular cartilage, but exhibited significantly reduced mechanical properties compared to wildtype controls. Moreover, the mice also showed accelerated development of OA joint degeneration, giving indirect evidence that alterations in PCM composition and properties can lead to the progression of OA [23].



Chondrocytes interact with the PCM via transmembrane receptors, such as integrins, that mediate cell function, including cell proliferation, cell survival and matrix remodeling. Integrins are heterodimeric receptors, consisting of α and β sununits, that connect the PCM with intracellular cytoskeletal structures and signaling complexes [24]. In humans, there are 18 different α and 8 different β subunits, which can be combined to 24 different integrin heterodimers with different ligand specificities [25]. Chondrocytes express multiple integrins, including α5β1, αVβ1, αVβ3, αVβ5, α3β1, α1β1, and α10β1, which bind to various matrix components [26]. Binding of ligands results in activation of integrins, leading to conformational changes that initiate downstream kinase-mediated intracellular signaling or binding to cytoskeletal proteins via focal adhesion complexes [25]. In this way, integrins play an important role in chondrocyte cell adhesion, ECM remodeling, and mechanotransduction [24]. One key protein involved in these processes linking the cartilage matrix and chondrocytes is fibronectin.

## Fibronectin

In articular cartilage fibronectin is mostly localized in the PCM, where it is organized in a fibrillar network [27]. Fibronectin mediates a wide variety of cellular interactions with the ECM by binding to matrix proteins via multiple binding domains, as well as interactions with chondrocytes via integrins that mediate intracellular signalling [28, 29]. In this way, fibronectin functions as an adhesion molecule facilitating signals from the ECM to the chondrocytes, such as cell adhesion, differentiation and survival [30]. Fibronectin is a glycoprotein, which is usually comprised of a dimer consisting of two high molecular weight subunits (230 – 270 kDa) linked covalently by two C-terminal disulfide bonds [27]. Each monomer consists of three types of repeating units: type I, which are ~40 amino acid residues, type II, which are ~60 amino acids in length, and type III, which are ~90 residues long and the most abundant repeats in the protein [31]. More specifically, fibronectin consists of 12 type I repeats, 2 type II repeats, and 15-17 type III repeats (**Figure 2**).

Fibronectin is encoded by the *FN1* gene, which can give rise to up to 27 transcripts in humans by means of alternative splicing [28]. Alternative splicing occurs at three major sites, called extra domain A (EDA), extra domain B, and variable region (V) [32]. Splicing at the EDA and EDB domain results in inclusion or exclusion of one exon, whereas splicing at the V region can occur at multiple splice sites [31, 33]. Based on solubility two forms of fibronectin can be distinguished, namely the soluble plasma form and the less-soluble cellular form. Plasma fibronectin lacks the EDA and EDB domains, while cellular fibronectin consists of a more heterogenous group of isoforms [28]. This splicing variation provides cells the capacity to generate large numbers of transcripts with different binding properties to precisely alter the ECM composition in a developmental and tissue-specific manner [28]. Fibronectin has been shown to be essential for mesodermal migration, adhesion and proliferation during embryogenesis, as homozygous inactivation of fibronectin caused early embryonic lethality in mice [34].

Fibronectin can bind multiple matrix proteins, including heparin, collagen, and fibrin by distinct structural and functional domains (**Figure 2**). There are two major heparin-binding domains, the strongest is located at the C-terminal side (repeats III 12-14) and the weaker at the N-terminal side (repeats I 1-5) of the protein. The gelatin-binding domain (repeats I 6-9 and II 1-2) can bind to gelatin and collagen. There are two major fibrin-binding domains, the N-terminal domain (repeats I 4-5) and the C-terminal domain (repeats I 10-12) [28]. However, the different specific domains of fibronectin can interact with multiple binding partners. As a result, the variable modular structure results in flexibility of the fibronectin protein and its binding capacities. The main integrin-binding domain contains the arginine-glycine-aspartate (RGD) motif (in repeat III 10), which binds multiple integrin heterodimers, including the classic fibronectin receptor α5β1 [35, 36]. The fibronectin-α5β1 adhesion has been shown to be critical for cartilage regeneration in mice [37], emphasizing the importance

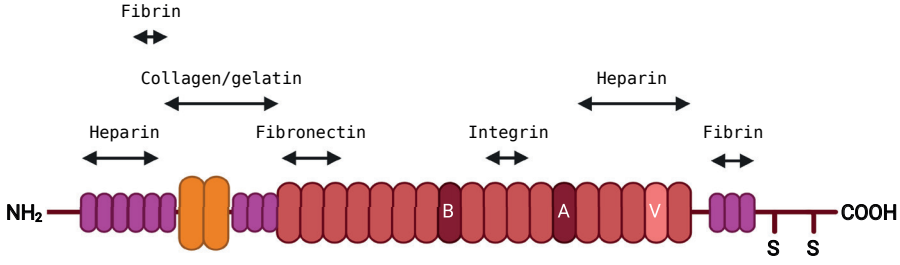



**Figure 2 | Structure of fibronectin.** Linear fibronectin monomer consisting of three types of repeats: type I (purple), type II (orange), and type III (red). Alternative spicing domains EDA (A) and EDB (B) are dark red and V region (V) is light red. Two C-terminal cysteine residues (S) are involved in dimerization of fibronectin via disulfide bonds. Main protein binding domains are indicated as arrow. (Created with Biorender.com).

of fibronectin binding to α5β1 in cartilage. However, in articular cartilage α3β1, αVβ1, αVβ3, and αVβ5 are known to also bind fibronectin [26].

## Pathophysiology osteoarthritic articular cartilage

Currently, OA is considered a disease of the whole joint, characterized by subchondral bone remodeling, formation of osteophytes, and inflammation of the synovium [13]. Moreover, a general hallmark of OA pathophysiology is the degeneration of articular cartilage, which is present to some extend in all OA affected joints (**Figure 3A**) [16]. In OA cartilage, chondrocytes become activated from their quiescent state, characterized by cell hypertrophy, cell clustering, and an increase of both matrix protein and matrix-degrading enzyme synthesis, such as matrix metalloproteinases (MMPs) and a disintegrin and metalloprotease with thrombospondin motif (ADAMTS). The increased deposition of matrix proteins is insufficient to counteract the increased degradation by the matrix-degrading enzymes. Consequently, proteoglycan content is decreased and collagen is cleaved, resulting in disruption of the pericellular matrix and eventually degradation of the cartilage matrix (**Figure 3B**) [38].

Once the collagen network is affected, progression to irreversible cartilage degeneration is inevitable [39]. At the cartilage surface fissures form and chondrocytes become apoptotic, while the calcified zone increases, resulting in tidemark rupture. Moreover, fibronectin content increases in osteoarthritic cartilage due to both enhanced synthesis and enhanced retention [33, 40]. Concomitantly, fibronectin is proteolytically degraded into fragments (FN-fs), that amplify catabolic processes by up-regulating MMPs via the Toll-like receptor signaling pathway [41-43]. Fragmentation results in cryptic binding properties that naïve fibronectin does not share, thereby acquiring catabolic activities [44]. These alterations in the molecular composition and organization of the cartilage PCM and ECM as a result of abnormal matrix turnover lead to changes in the mechanical properties and structural integrity of the articular cartilage [16].

### Molecular pathophysiology of osteoarthritic cartilage

The exact underlying molecular mechanisms of OA pathophysiology are not completely understood. One way to elucidate these mechanisms is to identify genes and pathways involved in the development of OA by performing transcriptomic analyses of cartilage tissue. Comparisons of expression profiles between OA affected and healthy cartilage identified multiple relevant pathways to be differentially expressed. Various genes encoding ECM organisation, such as the collagen-encoding *COL1A1*, *COL2A1*, and *COL6A3*, and genes involved in matrix degradation, such as *MMP-3*, *MMP-13*, and *ADAMTS-5* were found to be up-regulated in OA cartilage (**Table 1**). Moreover, genes involved in skeletal development, such as *ALPL* and *DIO2* were found to be up-regulated, while the chondrogenic transcription

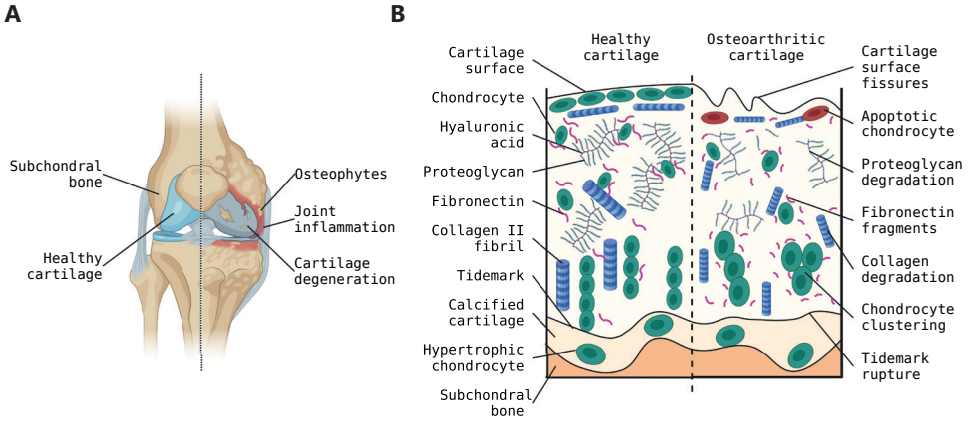



**Figure 3 | Overview of osteoarthritic pathophysiological changes in a joint and articular cartilage.** (**A**) Schematic representation of a knee joint, where the left side represents a normal joint, with cartilage covering the subchondral bone. The right side represents an OA affected joint, with cartilage degeneration, osteophyte formation, and joint inflammation. (**B**) Overview of structural differences in articular cartilage in healthy and osteoarthritic state, including collagen type II fibril and proteoglycan degradation, and fibronectin fragmentation. (Created with Biorender.com).

factor *SOX9* is down-regulated [45-47]. Multiple integrin subunits were found to be up-regulated, including *ITGAV*, *ITGB1*, and *ITGB5*, while *ITGA5* was down-regulated in OA cartilage compared to healthy.

Studies comparing expression profiles of lesioned and preserved cartilage areas from the same joints of patients with OA showed similar pathways to be differentially expressed [48-50]. However, there are numerous examples of genes that were only found differentially expressed in the comparison between lesioned and preserved, suggesting they play a role during the OA pathophysiologic process, but not in the initial phase. One such example is the gene *P3H2* encoding prolyl-3-hydroxylase 2, which is involved in collagen chain assembly and crosslinking [51]. Furthermore, genes such as *MMP-3*, *ITGA5*, and *COL9A1* showed opposite direction of effect in healthy versus OA cartilage compared to lesioned versus preserved cartilage. Therefore, these genes are likely mainly involved in the initial response of chondrocytes to cartilage degeneration and potential causal in the process. In our group, we collect OA relevant tissues, (cartilage, bone, and synovial fluid) and cells (bone-marrow derived mesenchymal stromal cells, primary chondrocytes) from patients undergoing joint replacement surgery due to end-stage OA as part of the Research osteoArthritis and Articular Cartilage (RAAK) study [48]. This study design allows for analyses independent of confounding factors such as sex, age, and joint site, while identifying expression changes specific to OA pathophysiology.

Notably, *FN1* was among the highest expressed genes in OA cartilage from the RAAK study and next to being highly up-regulated between lesioned and preserved OA cartilage, increased

**Table 1 | Pathways and a number of genes in those pathways involved in osteoarthritis pathophysiology, identified by transcriptome- wide profiles of healthy and osteoarthritic cartilage.**

| Pathway | Comparison | Joint | Dir | Gene symbol* | Ref |
|---|---|---|---|---|---|
| ECM organisation | Healthy vs OA | Knee | Up | *COL1A1* , *COL2A1* , *COL6A1* , *COL9A1*, *COL10A1* , *FN1* | [45-47] |
| | | | Down | *ACAN* | |
| | Lesioned vs preserved | Knee | Up | *COL1A1* , *COL6A3* , *FN1*, *P3H2* | [48, 51, 95] |
| | | | Down | *ACAN* , *COL9A1* | |
| Matrix degrading enzymes | Healthy vs OA | Knee | Up | *MMP-3* , *MMP-13*, *ADAMTS-5*, | [46, 47] |
| | Lesioned vs preserved | Knee | Up | *ADAMTS-5* | [51, 95] |
| | | | Down | *MMP-3* | |
| Chondrocyte signaling | Healthy vs OA | Knee | Up | *ITGAV* , *ITGB5*, *ITGB1* | [45, 47] |
| | | | Down | *ITGA5* , *SOX9* | |
| | Lesioned vs preserved | Knee | Up | *ITGA5* , *ITGAV* , *ITGB5* , *ITGB1* | [48, 51, 95] |
| | | | Down | *SOX9* | |
| Skeletal development | Healthy vs OA | Knee | Up | *ALPL* , *DIO2* | [46, 47] |
| | Lesioned vs preserved | Knee & Hip | Up | *TNFRSF11B* | [48, 96] |
| | | | Down | *FRZB* | |

* Gene symbols of genes related to the pathway.
Dir = direction of effect of genes in pathway. Ref = reference.

expression was also found in OA compared to healthy cartilage. (**Table 1**) [45, 48]. Next to investigating the effects of these changes in gene expression, it is of interest to gain better insight into how these gene expression changes are induced and controlled in OA cartilage by epigenetic mechanisms.

## Epigenetic mechanisms in osteoarthritic cartilage

Chondrocytes are maturational arrested cells, but are required to dynamically modify gene expression to maintain cartilage homeostasis upon environmental changes, such as aging, mechanical stress and microtraumas. Therefore, chondrocytes depend on multiple epigenetic mechanisms to regulate gene transcription and translation, including DNA methylation, histone modifications and noncoding RNAs, such as micro-RNAs (miRNAs) and long noncoding RNAs (lncRNAs) (**Figure 4**) [51-55]. DNA methylation is the most studied epigenetic mechanism in OA cartilage, which is a relatively stable epigenetic marker that comprises the addition of a methyl group to a cytosine at a CpG dinucleotide and influences binding of proteins to DNA, resulting in altered transcription [56]. In the RAAK study genome-wide methylation has been shown to characterize the cartilage based on joint site [57]. Furthermore, by combining genome, transcriptome, and methylome data from the same samples, epigenetically regulated genes involved in development and extracellular matrix maintenance pathways were identified in articular cartilage [52]. Therefore, it seems likely that the loss of epigenetic control in OA cartilage causes reactivation of developmental pathways in articular chondrocytes [56].

Previously thought to be transcriptional noise, noncoding RNAs represent the majority of the human genome, namely 80% versus 2% of coding-RNAs [58]. MiRNAs are a widely studied class of noncoding RNAs of < 22 nucleotides in length, which play a role in post-transcriptional regulation of gene expression by binding to complementary sequences in their target messenger RNA (mRNA) molecules, thereby reducing mRNA stability and/or inhibiting translation [53, 59]. In the RAAK study we aimed to uncover the miRNA interactome of the OA pathphysiological process in cartilage, by combining transcriptome-wide mRNA and miRNA data from the same OA cartilage samples. Pathway enrichment analysis revealed that genes involved in nervous system development are likely mediated by miRNA regulatory mechanisms [51]. Targeting dysfunctional miRNA-mRNA interactions shows promise for therapeutic development [60], but this is still ongoing work for OA.

### Long noncoding RNAs

More recently, lncRNAs have received increased attention, which are RNA transcripts of > 200 nucleotides and are similar to mRNAs in structure, but with little to no protein-coding potential. LncRNAs can regulate transcription and translation directly and indirectly by multiple mechanisms, including chromatin remodelling, mRNA processing, stability and

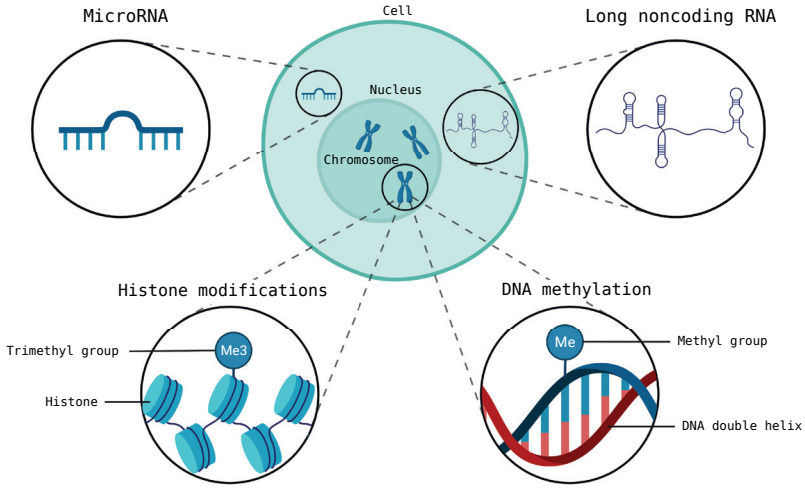



**Figure 4 | Schematic overview of levels of epigenetic regulation in chondrocytes.** Chondrocytes rely on multiple epigenetic mechanisms to regulate their gene expression, including noncoding RNAs, such as microRNAs and long noncoding RNAs, histone modifications, and DNA methylation. (Created with Biorender.com).

splicing, and regulating miRNA function [61, 62]. Based on the genomic location with respect to protein-coding genes, lncRNAs can be classified in biotypes, including antisense RNAs, sense RNAs, pseudogenes, and intergenic lncRNAs. Studies have shown that lncRNAs play an important role in cartilage matrix homeostatis [55, 63-66].

Transcriptome-wide profiles of OA cartilage compared to healthy cartilage resulted in the identification of lncRNAs involved in OA pathophysiology, among others *MEG3*, *MALAT1*, *HOTAIR*, and *GAS5* (**Table 2**) [67-70]. Notably, *MEG3* has been shown to be both up-regulated [68, 71], as well as down-regulated in OA cartilage samples compared to healthy [67, 69]. However, in the studies where *MEG3* was down-regulated, the donors of healthy cartilage were very young and not age-matched to the OA cartilage donors, e.g. 10-51 years, implying that up-regulation of *MEG3* is age-related and down-regulation is specific for the OA process [68, 71]. Moreover, 125 lncRNAs were differentially expressed after IL-1β stimulation of chondrocytes, including *PACER1*, suggesting they play a role in the inflammation-driven cartilage degeneration in OA [55]. These studies have focused primarily on intergenic lncRNAs, even though the proportion of genic and intergenic lncRNAs can be similar depending on the investigated tissue [72]. Therefore, many potentially functional lncRNAs are missing in these studies.

Functional studies of lncRNAs in cartilage showed that overexpression of *MEG3* induced cell proliferation and decreased ECM degradation in IL-1β induced chondrocytes [73].

**Table 2 | Robustly identified long noncoding RNAs in osteoarthritis pathophysiology, identified by transcriptome-wide profiles of healthy and osteoarthritic cartilage and functional studies that have been performed.**

| LncRNA | Biotype | Comparison | Joint | Dir | Ref | Functional studies | Ref |
|---|---|---|---|---|---|---|---|
| *MEG3* | Intergenic | Healthy vs OA | Knee | Up | [68, 71] | *MEG3* induces proliferation and relieves ECM degradation in IL-1β induced chondrocytes via targeting miR-93/TGFBR2 axis. | [73] |
| | | Healthy vs OA | Hip | Down | [67] | *MEG3* targets miR-361-5p/FOXO1 axis and elevates cell proliferation and impairs apoptosis in OA chondrocytes. | [97] |
| | | Healthy vs OA | Knee | Down | [69] | *MEG3* down-regulation leads to OA progression via miR-16/*SMAD7* axis. | [98] |
| *MALAT1* | Intergenic | Healthy vs OA | Hip | Down | [67] | *MALAT1* increases cell proliferation, inhibits apoptosis and inhibits ECM degradation via miR-150-5p/AKT3 axis. | [70] |
| | | Healthy vs OA | Knee | Down | [71] | *MALAT1* /miR-145 axis contributes to ECM degradation in IL-1β induced OA chondrocytes via *ADAMTS-5*. | [99] |
| *HOTAIR* | Intergenic | Healthy vs OA | Knee | Up | [68] | Knockdown of *HOTAIR* in IL-1β induced OA in vitro model reverse IL-1β stimulated expression of matrix-degrading enzymes and significantly reduce apoptosis rate. | [100] |
| | | Healthy vs OA | Knee | Up | [71] | *HOTAIR* promotes chondrocyte apoptosis and ECM degradation via miR-20/PTEN axis in IL-1β induced chondrocytes. | [101] |
| *GAS5* | Intergenic | Healthy vs OA | Knee | Up | [68] | *GAS5* induced apoptosis of chondrocytes through targeting miR-137. | [102] |
| | | Lesioned vs preserved | Knee | Up | [103] | *GAS5* increased expression of matrix-degrading enzymes and *GAS5* was found to act as a negative regulator of miR-21. | [103] |

Dir = direction of effect of genes in pathway. Ref = reference.

Others have shown that *MEG3* targets the TGF-β signaling pathway, which is also active in chondrocytes [74]. Together, these data imply that *MEG3* is involved in the development of OA. An overview of functional studies performed of previously mentioned lncRNAs associated with OA pathophysiology is shown in **Table 2**. These functional studies show that lncRNAs are often involved in miRNA-regulated pathways, however, many of these studies were performed in IL-1β stimulated chondrocytes as OA model, which only provides insight into inflammatory responses. Due to the pleiotropic functions of lncRNAs, functional studies with other OA models are necessary to elucidate the exact molecular mechanisms of OA-relevant lncRNAs.



## Insight into etiology of osteoarthritis

The identification of genes and their epigenetic regulation as markers in OA pathophysiology can help find potential therapeutic targets. However, these studies do not provide insight in the etiology of the disease. Causal pathways can be identified by unravelling the substantial genetic component of OA. The genetic background of common OA is highly polygenetic, where many genetic variants have been identified with small effects conferring risk to OA [75-78]. On the other hand, family-based linkage studies with more severe OA phenotypes result in the identification of high-impact mutations [79]. These rare variants are often located in the protein coding region of the gene and induce an amino acid change, thereby affecting protein function directly. To identify pathogenic variants a prioritization scheme has to be applied, where intergenic, intronic, synonymous, common variants, and tolerated missense mutations are filtered. Tolerated missense mutations are filtered based on sequence homology and physical properties of amino acids, thereby likely not affecting protein function. Filtering can result in up to 150 missense, damaging variants in one affected person, after which linkage analysis has to be performed, followed by genotyping in multiple affected and non-affected individuals to identify the likely causal mutation.

For that matter, whole exome sequencing combined with linkage in early-onset OA families with OA associated phenotypes has been highly successful in identifying likely causal mutations in genes in familial patients in mostly matrix protein encoding genes, including *TNFRSF11B* [80], *COL2A1* [81], *GDF5* [82], *COL11A1* and *COL11A2* [83], *COMP* [84], and *SMAD3* [85] (**Table 3**). Strikingly, functional follow-up of these mutations is mostly not performed. Nonetheless, because of their strong effect, these high-impact mutations can help in the characterization of causal underlying pathways, which can likely be extrapolated to confer risk to common OA. Moreover, they are more actionable to express a disease state in experimental human in vitro tissue models, in which they likely elucidate underlying pathways of OA [86]. Therefore, to understand how causality is explained by these high-impact mutations, functional disease modeling is necessary to increase translation to potential drug development.

**Table 3 | Previously identified high-impact mutations in early-onset osteoarthritis linkage studies and functional follow-up performed.**

| Gene | Protein function | Variant | Functional validation | Ref |
|---|---|---|---|---|
| *COL2A1* | Main collagen protein in articular cartilage | p.Gly204Ala | No functional validation. | [81] |
| *COL11A1*<br>*COL11A2* | Minor collagen protein in articular cartilage | p.Pro446Gln<br>p.Arg53Trp | No functional validation. | [83] |
| *COMP* | Glycoprotein in articular cartilage | p.Arg718Trp | No functional validation. | [84] |
| *GDF5* | Cytokine essential for cartilage development and homeostasis | p.Leu441Pro<br>p.Arg438Leu | Expressed mutant proteins in chicken limb bud micromass culture and treated ATDC5 and C2C12 cells with recombinant GDF5. Leu441Pro mutant was inactive, Arg438Leu mutant showed increased biological activity. | [82] |
| *SMAD3* | Intracellular signal transducer protein in TGF-β signaling | p.Cys89Tyr | No functional validation. | [85] |
| *TNFRSF11B* | Decoy receptor inhibiting osteoclastogenesis | p.Stop402Leu | Bone resorption assay showed mutant OPG had enhanced capacity to inhibit osteoclastogenesis. | [80] |

Ref = reference.

## In vitro OA disease modeling

After the identification of OA related genes, human 3D in vitro cell models can be applied to investigate the role of the gene in underlying pathophysiology. OA relevant cells include bone-marrow derived mesenchymal stromal cells and primary articular chondrocytes, which can subsequently be used in 3D chondrogenesis models. Lentiviral up- or down-regulation of a specific gene in these cells can mimick abberant gene function, after which the effect on cartilage deposition can be assessed. For example, *SOX9* overexpression resulted in increased *COL2A1* expression in chondrocytes in alginate bead culture, showing *SOX9* overexpressing could counteract the loss of chondropotential in 2D culture [87]. Moreover, *DIO2* up-regulation in a human 3D in vitro chondrogenesis model of bone marrow-derived mesenchymal stromal cells resulted in a significant reduced capacity to deposit ECM [88]. More recently, up-regulation of *TNFRSF11B* was shown to be detrimental for neo-cartilage production in a human 3D in vitro chondrogenesis model of primary articular chondrocytes [89]. Together, these studies show that abberant gene expression in vitro can reveal functional consequences of candidate genes. However, the drawback of these cell models is that the cells have limited availability and are generally obtained from OA affected tissue from donors receiving joint replacement surgery. Consequently, the material represents end-stage disease and effects in healthy and young chondrocytes remain mostly unknown.

To circumvent aforementioned issues, human induced pluripotent stem cell (hiPSC) technology has been groundbreaking, which allows reprogramming of readily available primary cells into a sustainable pluripotent cell source [90]. Subsequently, the hiPSCs can be used for directed differentiation into specialized cells of interest [91, 92]. To characterize genetic variation, accessible human donor cells can be reprogrammed into hiPSCs, which contain the unmodified genome of the donor. Functional studies after chondrogenic differentiation of healthy and OA-predisposed patient donors can elucidate underlying pathways of disease causing variants. The drawback of this method is that there is high variability of chondrogenic differentiation potential as a result of different genetic backgrounds among cell lines, which can not be corrected for. However, because of relative novel advancements in CRISPR/Cas9, precise genome editing of cells is easier than ever before. The main CRISPR system utilizes a guide RNA (gRNA) in complex with Cas9, an endonuclease. The gRNA contains the desired target sequence as well as the Cas9 binding sequence, so that Cas9 cleaves the DNA at the target site [93]. The resulting double stranded breaks are repaired by either the non-homologous end joining pathway or the homology directed repair pathway, if a double stranded DNA template is provided. In this way, either gene knockouts or specific mutations can be introduced in the cell's genome. The process of genome editing using CRISPR/Cas9 is infeasible in human primary articular chondrocytes, as they lose chondrogenic potential with prolonged 2D cell culturing [94]. Therefore, the creation of isogenic hiPSC lines followed by chondrogenic differentiation can give insight in the functional characterization of genetic variation. Taken together, the combination of hiPSCs and CRISPR/Cas9 holds immense potential for translational human OA disease modeling [86].



## Outline of this thesis

In this thesis, we aim to obtain insight into causal underlying pathways in OA pathophysiology by combining transcriptomics, genetics and OA disease modeling. In **chapter 2** we applied an improved detection strategy to detect robustly differentially expressed lncRNAs in OA cartilage and identify lncRNAs associated with the OA process in samples obtained in the RAAK study. Integration of these lncRNAs with differential expression levels in the same samples provided insight into their regulatory networks, which we functionally characterized. In **chapter 3**, a high-impact mutation in *FN1* was identified, after which disease modeling was applied to elucidate the underlying mechanism of the mutation. To this end, the mutation was introduced in hiPSCs using CRISPR/Cas9, after which functional studies were performed in our in vitro organoid cartilage model. In **chapter 4**, we sought out to identify *FN1* transcripts associated with OA pathophysiology and investigate downstream effect of modulation of *FN1* expression and relative transcript ratio in our 3D in vitro chondrogenesis model of primary articular chondrocytes obtained from the RAAK study. As such, studies performed in this thesis contributed to the translation from genetics to mechanisms of disease.

# References


1. Goldring, M.B. and S.R. Goldring, Articular cartilage and subchondral bone in the pathogenesis of osteoarthritis. Ann N Y Acad Sci, 2010. 1192: p. 230-7.
2. Allen, K.D., L.M. Thoma, and Y.M. Golightly, Epidemiology of osteoarthritis. Osteoarthritis Cartilage, 2021.
3. Palazzo, C., et al., Risk factors and burden of osteoarthritis. Ann Phys Rehabil Med, 2016. 59(3): p. 134-138.
4. Johnson, V.L. and D.J. Hunter, The epidemiology of osteoarthritis. Best Pract Res Clin Rheumatol, 2014. 28(1): p. 5-15.
5. Conaghan, P.G., et al., Osteoarthritis research priorities: a report from a EULAR ad hoc expert committee. Ann Rheum Dis, 2014. 73(8): p. 1442-5.
6. www.volksgezondheidzorg.info/onderwerp/artrose. Artrose.
7. van der Kraan, P.M., et al., Translation of clinical problems in osteoarthritis into pathophysiological research goals. RMD Open, 2016. 2(1): p. e000224.
8. Litwic, A., et al., Epidemiology and burden of osteoarthritis. Br Med Bull, 2013. 105: p. 185-99.
9. Loeser, R.F., J.A. Collins, and B.O. Diekman, Ageing and the pathogenesis of osteoarthritis. Nat Rev Rheumatol, 2016. 12(7): p. 412-20.
10. Rabenda, V., et al., Direct and indirect costs attributable to osteoarthritis in active subjects. J Rheumatol, 2006. 33(6): p. 1152-8.
11. Drissi, H., et al., Transcriptional regulation of chondrocyte maturation: potential involvement of transcription factors in OA pathogenesis. Mol Aspects Med, 2005. 26(3): p. 169-79.
12. Sophia Fox, A.J., A. Bedi, and S.A. Rodeo, The basic science of articular cartilage: structure, composition, and function. Sports Health, 2009. 1(6): p. 461-8.
13. Loeser, R.F., et al., Osteoarthritis: a disease of the joint as an organ. Arthritis Rheum, 2012. 64(6): p. 1697-707.
14. Kiani, C., et al., Structure and function of aggrecan. Cell Res, 2002. 12(1): p. 19-32.
15. Goldring, S.R. and M.B. Goldring, Changes in the osteochondral unit during osteoarthritis: structure, function and cartilage-bone crosstalk. Nat Rev Rheumatol, 2016. 12(11): p. 632-644.
16. Goldring, M.B., Chondrogenesis, chondrocyte differentiation, and articular cartilage metabolism in health and osteoarthritis. Ther Adv Musculoskelet Dis, 2012. 4(4): p. 269-85.
17. Poole, C.A., Articular cartilage chondrons: form, function and failure. J Anat, 1997. 191 ( Pt 1): p. 1-13.
18. Poole, C.A., M.H. Flint, and B.W. Beaumont, Chondrons in cartilage: ultrastructural analysis of the pericellular microenvironment in adult human articular cartilages. J Orthop Res, 1987. 5(4): p. 509-22.
19. Wilusz, R.E., J. Sanchez-Adams, and F. Guilak, The structure and function of the pericellular matrix of articular cartilage. Matrix Biol, 2014. 39: p. 25-32.
20. Chang, J., H. Nakajima, and C.A. Poole, Structural colocalisation of type VI collagen and fibronectin in agarose cultured chondrocytes and isolated chondrons extracted from adult canine tibial cartilage. J Anat, 1997. 190 ( Pt 4): p. 523-32.
21. Guilak, F., et al., The pericellular matrix as a transducer of biomechanical and biochemical signals in articular cartilage. Ann N Y Acad Sci, 2006. 1068: p. 498-512.
22. Guilak, F., et al., Osteoarthritis as a disease of the cartilage pericellular matrix. Matrix Biol, 2018. 71-72: p. 40-50.
23. Alexopoulos, L.G., et al., Developmental and osteoarthritic changes in Col6a1-knockout mice: biomechanics of type VI collagen in the cartilage pericellular matrix. Arthritis Rheum, 2009. 60(3): p. 771-9.v
24. Tian, J., F.J. Zhang, and G.H. Lei, Role of integrins and their ligands in osteoarthritic cartilage. Rheumatol Int, 2015. 35(5): p. 787-98.
25. Hynes, R.O., Integrins: bidirectional, allosteric signaling machines. Cell, 2002. 110(6): p. 673-87.
26. Loeser, R.F., Integrins and chondrocyte-matrix interactions in articular cartilage. Matrix Biol, 2014. 39: p. 11-6.
27. Singh, P., C. Carraher, and J.E. Schwarzbauer, Assembly of fibronectin extracellular matrix. Annu Rev Cell Dev Biol, 2010. 26: p. 397-419.

28. Pankov, R. and K.M. Yamada, Fibronectin at a glance. J Cell Sci, 2002. 115(Pt 20): p. 3861-3.
29. Stoffels, J.M., C. Zhao, and W. Baron, Fibronectin in tissue regeneration: timely disassembly of the scaffold is necessary to complete the build. Cell Mol Life Sci, 2013. 70(22): p. 4243-53.
30. Schwartz, M.A. and R.K. Assoian, Integrins and cell proliferation: regulation of cyclin-dependent kinases via cytoplasmic signaling pathways. J Cell Sci, 2001. 114(Pt 14): p. 2553-60.
31. Schwarzbauer, J.E. and D.W. DeSimone, Fibronectins, their fibrillogenesis, and in vivo functions. Cold Spring Harb Perspect Biol, 2011. 3(7).
32. Schwarzbauer, J.E., Fibronectin: from gene to protein. Curr Opin Cell Biol, 1991. 3(5): p. 786-91.
33. Chevalier, X., Fibronectin, cartilage, and osteoarthritis. Semin Arthritis Rheum, 1993. 22(5): p. 307-18.
34. George, E.L., et al., Defects in mesoderm, neural tube and vascular development in mouse embryos lacking fibronectin. Development, 1993. 119(4): p. 1079-91.
35. Pierschbacher, M.D. and E. Ruoslahti, Variants of the cell recognition site of fibronectin that retain attachment-promoting activity. Proc Natl Acad Sci U S A, 1984. 81(19): p. 5985-8.
36. Pytela, R., M.D. Pierschbacher, and E. Ruoslahti, Identification and isolation of a 140 kd cell surface glycoprotein with properties expected of a fibronectin receptor. Cell, 1985. 40(1): p. 191-8.
37. Almonte-Becerril, M., et al., Genetic abrogation of the fibronectin-alpha5beta1 integrin interaction in articular cartilage aggravates osteoarthritis in mice. PLoS One, 2018. 13(6): p. e0198559.
38. Goldring, M.B. and K.B. Marcu, Cartilage homeostasis in health and rheumatic diseases. Arthritis Res Ther, 2009. 11(3): p. 224.
39. Karsdal, M.A., et al., Cartilage degradation is fully reversible in the presence of aggrecanase but not matrix metalloproteinase activity. Arthritis Res Ther, 2008. 10(3): p. R63.
40. Chevalier, X., N. Groult, and J. Labat-Robert, Biosynthesis and distribution of fibronectin in normal and osteoarthritic human cartilage. Clin Physiol Biochem, 1992. 9(1): p. 1-6.
41. Homandberg, G.A., R. Meyers, and J.M. Williams, Intraarticular injection of fibronectin fragments causes severe depletion of cartilage proteoglycans in vivo. J Rheumatol, 1993. 20(8): p. 1378-82.
42. Hwang, H.S., et al., Fibronectin fragment-induced expression of matrix metalloproteinases is mediated by MyD88-dependent TLR-2 signaling pathway in human chondrocytes. Arthritis Res Ther, 2015. 17: p. 320.
43. Stanton, H., L. Ung, and A.J. Fosang, The 45 kDa collagen-binding fragment of fibronectin induces matrix metalloproteinase-13 synthesis by chondrocytes and aggrecan degradation by aggrecanases. Biochem J, 2002. 364(Pt 1): p. 181-90.
44. Homandberg, G.A., C. Wen, and F. Hui, Cartilage damaging activities of fibronectin fragments derived from cartilage and synovial fluid. Osteoarthritis Cartilage, 1998. 6(4): p. 231-44.
45. Aigner, T., et al., Large-scale gene expression profiling reveals major pathogenetic pathways of cartilage degeneration in osteoarthritis. Arthritis Rheum, 2006. 54(11): p. 3533-44.
46. Karlsson, C., et al., Genome-wide expression profiling reveals new candidate genes associated with osteoarthritis. Osteoarthritis Cartilage, 2010. 18(4): p. 581-92.
47. Fisch, K.M., et al., Identification of transcription factors responsible for dysregulated networks in human osteoarthritis cartilage by global gene expression analysis. Osteoarthritis Cartilage, 2018. 26(11): p. 1531-1538.
48. Ramos, Y.F., et al., Genes involved in the osteoarthritis process identified through genome wide expression analysis in articular cartilage; the RAAK study. PLoS One, 2014. 9(7): p. e103056.
49. Geyer, M., et al., Differential transcriptome analysis of intraarticular lesional vs intact cartilage reveals new candidate genes in osteoarthritis pathophysiology. Osteoarthritis Cartilage, 2009. 17(3): p. 328-35.
50. Sato, T., et al., Comparative analysis of gene expression profiles in intact and damaged regions of human osteoarthritic cartilage. Arthritis Rheum, 2006. 54(3): p. 808-17.
51. Coutinho de Almeida, R., et al., RNA sequencing data integration reveals an miRNA interactome of osteoarthritis cartilage. Ann Rheum Dis, 2019. 78(2): p. 270-277.
52. den Hollander, W., et al., Transcriptional associations of osteoarthritis-mediated loss of epigenetic control in articular cartilage. Arthritis Rheumatol, 2015. 67(8): p. 2108-16.

53. Coutinho De Almeida, R., Y.F.M. Ramos, and I. Meulenbelt, Involvement of epigenetics in osteoarthritis. Best Practice & Research Clinical Rheumatology, 2017. 31(5): p. 634-648.
54. Wan, C., et al., Histone Modifications and Chondrocyte Fate: Regulation and Therapeutic Implications. Front Cell Dev Biol, 2021. 9: p. 626708.
55. Pearson, M.J., et al., Long Intergenic Noncoding RNAs Mediate the Human Chondrocyte Inflammatory Response and Are Differentially Expressed in Osteoarthritis Cartilage. Arthritis Rheumatol, 2016. 68(4): p. 845-56.
56. den Hollander, W. and I. Meulenbelt, DNA Methylation in Osteoarthritis. Curr Genomics, 2015. 16(6): p. 419-26.
57. den Hollander, W., et al., Knee and hip articular cartilage have distinct epigenomic landscapes: implications for future cartilage regeneration approaches. Ann Rheum Dis, 2014. 73(12): p. 2208-12.
58. Mercer, T.R., M.E. Dinger, and J.S. Mattick, Long non-coding RNAs: insights into functions. Nat Rev Genet, 2009. 10(3): p. 155-9.
59. Fabian, M.R., N. Sonenberg, and W. Filipowicz, Regulation of mRNA translation and stability by microRNAs. Annu Rev Biochem, 2010. 79: p. 351-79.
60. Winkle, M., et al., Noncoding RNA therapeutics - challenges and potential solutions. Nat Rev Drug Discov, 2021. 20(8): p. 629-651.
61. Jarroux, J., A. Morillon, and M. Pinskaya, History, Discovery, and Classification of lncRNAs. Adv Exp Med Biol, 2017. 1008: p. 1-46.
62. Geisler, S. and J. Coller, RNA in unexpected places: long non-coding RNA functions in diverse cellular contexts. Nat Rev Mol Cell Biol, 2013. 14(11): p. 699-712.
63. Jiang, S.D., et al., Long noncoding RNAs in osteoarthritis. Joint Bone Spine, 2017. 84(5): p. 553-556.
64. Liu, Q., et al., Long Noncoding RNA Related to Cartilage Injury Promotes Chondrocyte Extracellular Matrix Degradation in Osteoarthritis. Arthritis & Rheumatology, 2014. 66(4): p. 969-978.
65. Huynh, N.P., et al., Emerging roles for long noncoding RNAs in skeletal biology and disease. Connect Tissue Res, 2017. 58(1): p. 116-141.
66. Sun, H., et al., Emerging roles of long noncoding RNA in chondrogenesis, osteogenesis, and osteoarthritis. Am J Transl Res, 2019. 11(1): p. 16-30.
67. Ajekigbe, B., et al., Identification of long non-coding RNAs expressed in knee and hip osteoarthritic cartilage. Osteoarthritis Cartilage, 2019. 27(4): p. 694-702.
68. Xing, D., et al., Identification of long noncoding RNA associated with osteoarthritis in humans. Orthop Surg, 2014. 6(4): p. 288-93.
69. Su, W., et al., The Long Noncoding RNA MEG3 Is Downregulated and Inversely Associated with VEGF Levels in Osteoarthritis. Biomed Res Int, 2015. 2015: p. 356893.
70. Zhang, Y., et al., LncRNA MALAT1 promotes osteoarthritis by modulating miR-150-5p/AKT3 axis. Cell & Bioscience, 2019. 9(1).
71. Fu, M., et al., Expression profile of long noncoding RNAs in cartilage from knee osteoarthritis patients. Osteoarthritis and Cartilage, 2015. 23(3): p. 423-432.
72. Uszczynska-Ratajczak, B., et al., Towards a complete map of the human long non-coding RNA transcriptome. Nature Reviews Genetics, 2018. 19(9): p. 535-548.
73. Chen, K., et al., LncRNA MEG3 Inhibits the Degradation of the Extracellular Matrix of Chondrocytes in Osteoarthritis via Targeting miR-93/TGFBR2 Axis. Cartilage, 2019: p. 1947603519855759.
74. Mondal, T., et al., MEG3 long noncoding RNA regulates the TGF-beta pathway genes through formation of RNA-DNA triplex structures. Nat Commun, 2015. 6: p. 7743.
75. Boer, C.G., et al., Deciphering osteoarthritis genetics across 826,690 individuals from 9 populations. Cell, 2021. 184(18): p. 4784-4818 e17.
76. Tachmazidou, I., et al., Identification of new therapeutic targets for osteoarthritis through genome-wide analyses of UK Biobank data. Nat Genet, 2019. 51(2): p. 230-236.
77. arc, O.C., et al., Identification of new susceptibility loci for osteoarthritis (arcOGEN): a genome-wide association study. Lancet, 2012. 380(9844): p. 815-23.
78. Meulenbelt, I., et al., Identification of DIO2 as a new susceptibility locus for symptomatic osteoarthritis. Hum Mol Genet, 2008. 17(12): p. 1867-75.
79. Manolio, T.A., et al., Finding the missing heritability of complex diseases. Nature, 2009. 461(7265): p. 747-53.

80. Ramos, Y.F., et al., A gain of function mutation in TNFRSF11B encoding osteoprotegerin causes osteoarthritis with chondrocalcinosis. Ann Rheum Dis, 2015. 74(9): p. 1756-62.
81. Kannu, P., et al., Premature arthritis is a distinct type II collagen phenotype. Arthritis Rheum, 2010. 62(5): p. 1421-30.
82. Seemann, P., et al., Activating and deactivating mutations in the receptor interaction site of GDF5 cause symphalangism or brachydactyly type A2. J Clin Invest, 2005. 115(9): p. 2373-81.
83. Jakkula, E., et al., The role of sequence variations within the genes encoding collagen II, IX and XI in non-syndromic, early-onset osteoarthritis. Osteoarthritis Cartilage, 2005. 13(6): p. 497-507.
84. Mu, S.C., et al., A mutation in cartilage oligomeric matrix protein (COMP) causes early-onset osteoarthritis in a large kindred study. Ann Hum Genet, 2011. 75(5): p. 575-83.
85. van de Laar, I.M., et al., Mutations in SMAD3 cause a syndromic form of aortic aneurysms and dissections with early-onset osteoarthritis. Nat Genet, 2011. 43(2): p. 121-6.
86. Adkar, S.S., et al., Genome Engineering for Personalized Arthritis Therapeutics. Trends Mol Med, 2017. 23(10): p. 917-931.
87. Li, Y., et al., Transduction of passaged human articular chondrocytes with adenoviral, retroviral, and lentiviral vectors and the effects of enhanced expression of SOX9. Tissue Eng, 2004. 10(3-4): p. 575-84.
88. Bomer, N., et al., Underlying molecular mechanisms of DIO2 susceptibility in symptomatic osteoarthritis. Ann Rheum Dis, 2015. 74(8): p. 1571-9.
89. Ruiz, A.R., et al., The role of TNFRSF11B in development of osteoarthritic cartilage. Rheumatology (Oxford), 2021.
90. Takahashi, K., et al., Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell, 2007. 131(5): p. 861-72.
91. Adkar, S.S., et al., Step-Wise Chondrogenesis of Human Induced Pluripotent Stem Cells and Purification Via a Reporter Allele Generated by CRISPR-Cas9 Genome Editing. Stem Cells, 2019. 37(1): p. 65-76.
92. Rodriguez Ruiz, A., et al., Cartilage from human-induced pluripotent stem cells: comparison with neo-cartilage from chondrocytes and bone marrow mesenchymal stromal cells. Cell Tissue Res, 2021.
93. Ran, F.A., et al., Genome engineering using the CRISPR-Cas9 system. Nat Protoc, 2013. 8(11): p. 2281-2308.
94. Benya, P.D., S.R. Padilla, and M.E. Nimni, Independent regulation of collagen types by chondrocytes during the loss of differentiated function in culture. Cell, 1978. 15(4): p. 1313-21.
95. Dunn, S.L., et al., Gene expression changes in damaged osteoarthritic cartilage identify a signature of non-chondrogenic and mechanical responses. Osteoarthritis Cartilage, 2016. 24(8): p. 1431-40.
96. Snelling, S., et al., A gene expression study of normal and damaged cartilage in anteromedial gonarthrosis, a phenotype of osteoarthritis. Osteoarthritis Cartilage, 2014. 22(2): p. 334-43.
97. Wang, A., et al., MEG3 promotes proliferation and inhibits apoptosis in osteoarthritis chondrocytes by miR-361-5p/FOXO1 axis. BMC Med Genomics, 2019. 12(1): p. 201.
98. Xu, J. and Y. Xu, The lncRNA MEG3 downregulation leads to osteoarthritis progression via miR-16/SMAD7 axis. Cell Biosci, 2017. 7: p. 69.
99. Liu, C., et al., LncRNA MALAT1/MiR-145 Adjusts IL-1beta-Induced Chondrocytes Viability and Cartilage Matrix Degradation by Regulating ADAMTS5 in Human Osteoarthritis. Yonsei Med J, 2019. 60(11): p. 1081-1092.
100. Zhang, C., et al., Upregulation of lncRNA HOTAIR contributes to IL-1beta-induced MMP overexpression and chondrocytes apoptosis in temporomandibular joint osteoarthritis. Gene, 2016. 586(2): p. 248-53.
101. Chen, Y., et al., Long-chain non-coding RNA HOTAIR promotes the progression of osteoarthritis via sponging miR-20b/PTEN axis. Life Sci, 2020. 253: p. 117685.
102. Gao, S.T., et al., LncRNA GAS5 induces chondrocyte apoptosis by down-regulating miR-137. Eur Rev Med Pharmacol Sci, 2020. 24(21): p. 10984-10991.
103. Song, J., et al., A long non-coding RNA, GAS5, plays a critical role in the regulation of miR-21 during osteoarthritis. J Orthop Res, 2014. 32(12): p. 1628-35.

# Chapter 2

## Elucidating epigenetic regulation by identifying functional *cis*-acting long noncoding RNAs and their targets in osteoarthritic articular cartilage

Marcella van Hoolwerff[1], Paula I. Metselaar[1], Margo Tuerlings[1], H. Eka D. Suchiman[1], Nico Lakenberg[1], Yolande F.M. Ramos[1], Davy Cats[2], Rob G.H.H. Nelissen[3], Demiën Broekhuis[3], Hailiang Mei[2], Rodrigo Coutinho de Almeida[1], Ingrid Meulenbelt[1]

[1] Dept. of Biomedical Data Sciences, Section Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands
[2] Sequence Analysis Support Core, Leiden University Medical Center, Leiden, The Netherlands
[3] Dept. of Orthopaedics, Leiden University Medical Center, Leiden, The Netherlands

Arthritis Rheumatol. 2020 Nov;72(11):1845-1854
DOI 10.1002/art.41396

## Abstract

**Objective.** To identify robustly differentially expressed long noncoding RNAs (lncRNAs) with osteoarthritis (OA) pathophysiology in cartilage and to explore potential target messenger RNAs (mRNAs) by establishing coexpression networks, followed by functional validation.

**Methods.** RNA sequencing was performed on macroscopically lesioned and preserved OA cartilage of patients who underwent joint replacement surgery due to OA (N=98). Differential expression analysis was performed on lncRNAs that were annotated in GENCODE and Ensembl databases. To identify potential interactions, correlations were calculated between the identified differentially expressed lncRNAs and previously reported differentially expressed protein-coding genes in the same samples. Modulation of chondrocyte lncRNA expression was achieved using locked nucleic acid GapmeRs.

**Results.** By applying our in-house pipeline we identified 5,053 lncRNAs to be robustly expressed, of which 191 were significantly differentially expressed (according to false discovery rate) between lesioned and preserved OA cartilage. Upon integrating mRNA sequencing data, we showed that intergenic and antisense differentially expressed lncRNAs demonstrate high, positive correlations with their respective flanking or sense genes. To functionally validate this observation, we selected *P3H2-AS1*, which was down-regulated in primary chondrocytes, resulting in down-regulation of *P3H2* gene expression levels. As such, we can confirm that *P3H2-AS1* regulates its sense gene *P3H2*.

**Conclusion.** By applying an improved detection strategy, robustly differentially expressed lncRNAs in OA cartilage were detected. Integration of these lncRNAs with differential mRNA expression levels in the same samples provided insight into their regulatory networks. Our data indicate that intergenic and antisense lncRNAs play an important role in regulating the pathophysiology of OA.

## Introduction

Osteoarthritis (OA) is an age-related, heterogenous, degenerative disease of the articular joints, characterized in part by cartilage degeneration and remodeling of subchondral bone, which results in stiff and painful joints and decreased mobility [1]. Despite the fact that OA is the most globally prevalent joint disease, no effective treatment is currently available [2]. It has been demonstrated that OA pathophysiology in cartilage is marked by altered gene expression regulation in chondrocytes [3, 4]. This alteration of gene expression regulation could be triggered by adaptation processes occurring due to aging, genetic predisposition, or environmental stimuli, and is in part caused by aberrant epigenetic mechanisms. These mechanisms include DNA methylation, histone modifications and expression of microRNAs (< 22 nucleotides) [4-6]. More recently, long noncoding RNAs (lncRNAs; > 200 nucleotides) have been shown to play an important role in the homeostasis of the extracellular matrix of cartilage [5, 7-10].



LncRNAs are defined as RNA transcripts with little or no protein-coding potential and are known to regulate transcription and translation by numerous mechanisms, such as chromatin remodeling, messenger RNA (mRNA) stabilization, microRNA modulation, and recruitment of scaffolding proteins. One classification type of lncRNAs is based on the genomic location with respect to protein-coding genes, so-called biotypes, including antisense RNAs, sense RNAs, pseudogenes, and long intergenic noncoding RNAs (lincRNAs). Another type of classification is based on the location at which the lncRNA functions relative to its transcription site, which can be in *trans* or *cis* [11-13]. *Cis*-acting lncRNAs comprise a considerable portion of known lncRNAs and can be positioned at various distances and orientations relative to their target genes, such as lincRNAs around transcription factor start sites, as well as sense and antisense lncRNAs that overlap with their sense gene [13, 14]. Potentially, lncRNAs could be candidate targets in OA treatment, since their expression can be highly tissue specific [9].

RNA sequencing (RNA-seq) has improved the ability to detect lncRNAs, but mapping and annotating lncRNAs remains challenging. These challenges arise from the fact that they are usually expressed at very low levels and their sequence-function relationship is still poorly understood. Moreover, recent findings from studies on ribosome profiling and bioinformatics suggested that a large proportion of transcripts has unknown coding potential [15]. Recent studies on OA have focused on intergenic lncRNAs, even though the proportion of genic and intergenic lncRNAs can be similar depending on the investigated tissue [15, 16]. To determine the complete lncRNA transcriptome, we used an in-house pipeline to robustly capture lncRNAs in a previously assessed RNA-seq dataset of lesioned and preserved OA cartilage samples [4]. Subsequently, lncRNAs associated with OA pathophysiology were identified, and potential interactions with OA-specific mRNAs were investigated.

# Materials and Methods

## Sample collection

Macroscopically lesioned and preserved articular cartilage samples were obtained from participants in the Research osteoArthritis and Articular Cartilage (RAAK) study described by Ramos *et al.* [3]. In the present study, a total of 98 samples was used (65 knees, 33 hips) (**Supplementary Table 1**). Ethical approval was obtained from the medical ethics committee of the Leiden University Medical Center (no. P08.239/P19.013) and informed consent was obtained from all participants.

## RNA sequencing

Total RNA from articular cartilage was isolated using a Qiagen RNeasy Mini Kit (Qiagen, GmbH, Hilden, Germany). Paired-end 2×100 bp RNA sequencing (Illumina TruSeq RNA Library Prep Kit, Illumina HiSeq2000 and Illumina HiSeq4000) was performed. Strand-specific RNA-seq libraries were generated, which yielded a mean of 20 million reads per sample. Quality control was performed as previously described [4], and reads were subsequently aligned to the GRCh38 reference genome with the RNA-seq aligner STAR (version 2.6.0) [17]. Thereafter, aligned reads were processed into individual transcripts using StringTie (version 1.3.4) [18]. LncRNAs were identified by mapping the transcripts to GENCODE (version 29) [11] and Ensembl (version 94) [19].

In order to filter transcripts with unknown protein-coding potential, we integrated two sources of evidence: (1) predictions from the alignment-free Coding Potential Assessment Tool (CPAT, version 1.2.2) [20], and (2) predictions from the LncFinder R package (version 1.1.3) [21]. CPAT is a machine learning-based method that analyzes the sequence features of transcript open-reading frames (ORFs) using a logistic regression model built from ORF size, Fickett TESTCODE statistic and hexamer usage bias. In CPAT, a transcript with a coding probability ≥0.364 was considered to be a coding sequence. LncFinder predicts lncRNAs using heterologous features and machine learning model [21]. Transcripts with protein-coding potential predicted by both tools were removed from the dataset.

## Differential expression analysis and replication

Differential expression analysis was performed on 32 paired samples (25 knees and 7 hips) (**Supplementary Table 1A**) using the DESeq2 R package (version 1.24) [22]. A general linear model assuming a negative binomial distribution was applied, followed by a paired Wald's test comparing lesioned OA cartilage samples and preserved OA cartilage samples, with the preserved samples as the referent. P values less than 0.05 (after Benjamini-Hochberg correction) were considered significant and are reported as the false discovery rate (FDR). This analysis was repeated for knee and hip samples separately.

Furthermore, to validate the results, five significant differentially expressed lncRNAs were selected and measured by reverse transcription-quantitative polymerase chain

reaction (RT-qPCR) in ten paired cartilage samples overlapping with the RNA-seq samples (**Supplementary Table 1B**), and replication was performed in an independent cohort of ten paired cartilage samples (**Supplementary Table 1C**). Total RNA was isolated using an RNeasy Mini Kit, followed by complementary DNA (cDNA) synthesis using 100 ng RNA with a First Strand cDNA synthesis kit according to the instructions of the manufacturer (Roche Applied Science). Expression levels of *AC025370.1*, *AC090877.2*, *MEG3*, *P3H2-AS1*, *TBILA*, and *GAPDH* were determined using FastStart SYBR Green Master reaction mix (Roche Applied Science). Primer sequences are shown in **Supplementary Table 2**. Relative gene expression levels were calculated with the $2^{-\Delta\Delta Ct}$ method, using *GAPDH* as internal control. A paired t-test was performed on the -ΔCt values, and P less than 0.05 were considered significant.



## LncRNA-mRNA interactions

To identify potential interactions, correlations were calculated between the identified differentially expressed lncRNAs and previously reported differentially expressed protein-coding genes in the same samples. LncRNA expression data were normalized and variance stabilizing transformed using the DESeq2 R package (version 1.24) [22], and batch effect was removed using the limma R package (version 3.40.6) [23]. Our previously published mRNA data [4] were equally normalized and transformed, and batch effect was removed. Subsequently, Spearman's correlations were calculated between the significantly differentially expressed lncRNAs identified in the combined analysis of knee and hip samples and the differentially expressed protein-coding genes previously published [4], using the Hmisc R package (version 4.2.0) for OA cartilage samples (**Supplementary Table 1D**). Correlations with P values less than 0.05 were considered significant. Network visualization was performed using the RedeR package (version 3.10) [24].

## In vitro down-regulation lncRNA using locked nucleic acid GapmeRs

Primary chondrocytes were isolated from three independent donors and passaged twice or thrice, as previously described [25]. Chondrocytes were transfected in duplo with antisense locked nucleic acid (LNA) GapmeR (Qiagen) targeting *P3H2-AS1* (TGAGCAACTAGGTGTA) or GapmeR negative control (AACACGTCTATACGC) at 10 nM final concentration using Lipofectamine RNAiMax Transfection Reagent according to instructions of the manufacturer (Invitrogen). Cells were lysed 30 hours posttransfection with TRIzol Reagent (Thermo Fisher Scientific) for RNA isolation, which was done using an RNeasy Mini Kit. Synthesis of cDNA was performed with 150 ng of total RNA using a First Strand cDNA Synthesis kit according to the instructions of the manufacturer. Expression levels of *P3H2-AS1*, *P3H2*, and *GAPDH* were determined using FastStart SYBR Green Master reaction mix. Primer sequences are shown in **Supplementary Table 2**. Relative gene expression levels were calculated with the $2^{-\Delta\Delta Ct}$ method, using *GAPDH* as internal control. A paired t-test was performed on the -ΔCt values, P values less than 0.05 were considered significant.

## Data availability

FASTQ files are available on ArrayExpress E-MTAB-7313.

# Results

## Characterization of lncRNAs in OA cartilage

To characterize lncRNAs in OA cartilage, we used our previously assessed RNA-seq data of 32 paired samples (25 knees, 7 hips) of lesioned and preserved OA cartilage [4] (**Supplementary Table 1A**). Our in-house pipeline was applied to capture lncRNAs from 2 databases (GENCODE and Ensembl). As shown in **Figure 1**, 30,354 lncRNAs were initially detected

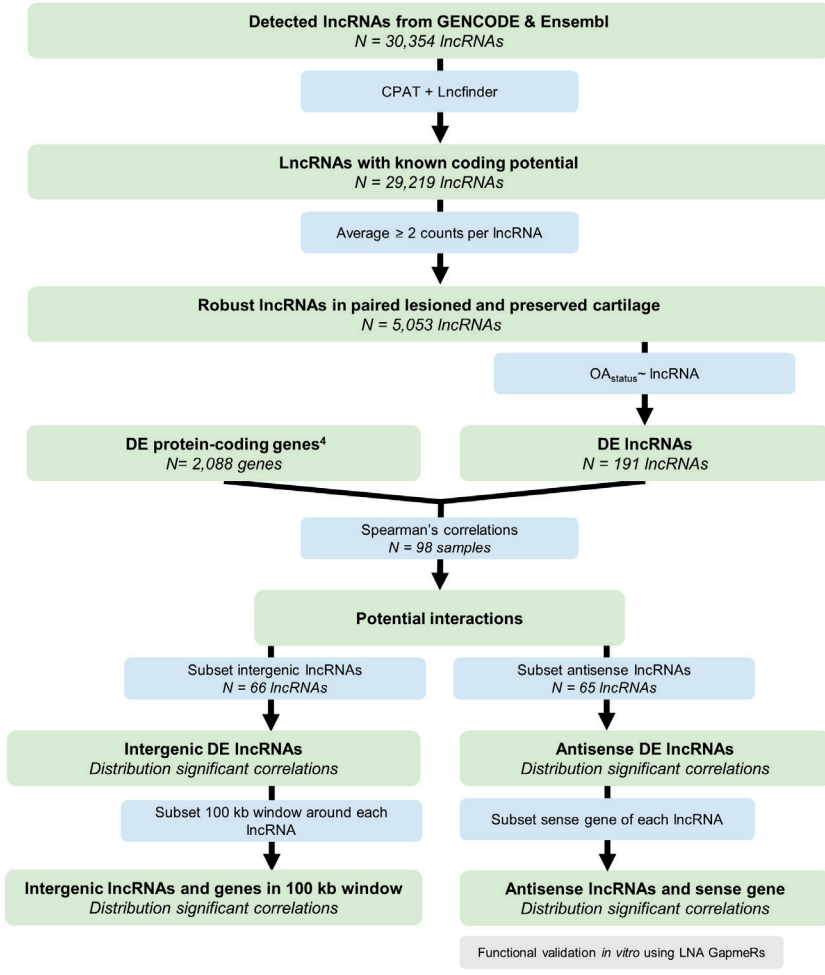



**Figure 1 | Overview of applied strategy |** Numbers of genes or long noncoding RNAs (lncRNAs) represent significantly differentially expressed (DE) genes or lncRNAs (according to false discovery rate). OA = osteoarthritis; LNA = locked nucleic acid.

in our dataset. To filter out possible transcripts of unknown coding potential, we integrated results from two machine learning approaches (CPAT [20] and Lncfinder, [21]). After removing these transcripts, 29,219 lncRNAs remained in the dataset and were considered for further analyses. To robustly detect lncRNAs expressed in OA cartilage, a cutoff of ≥2 counts per lncRNA was applied, resulting in a total of 5,053 lncRNAs expressed in cartilage (**Figure 1**). Classification of these lncRNAs based on biotype showed that 1,989 were antisense RNAs (39.4%), 249 sense RNAs (4.9%), 1,532 were pseudogenes (30.3%), and 900 were lincRNAs (17.8%) (**Figure 2**).

## Differential expression of lncRNAs between lesioned and preserved OA cartilage

To identify lncRNAs associated with the OA process, differential expression analysis was performed on paired lesioned and preserved OA cartilage samples, resulting in 191 significantly differentially expressed lncRNAs (FDR < 0.05, **Figure 1**). Of these, 65 were antisense RNAs (34.0%), 10 were sense RNAs (5.2%), 33 were pseudogenes (17.3%), and 66 were lincRNAs (34.6%) (**Figure 2**, **Supplementary Table 3**). When comparing the biotypes of the total expressed lncRNAs to the biotypes of the differentially expressed lncRNAs (**Figure 2**), we observed an increase of lincRNAs and a decrease of pseudogenes. The most significantly differentially expressed lncRNA was lincRNA *AL139220.2* (fold change 2.2, FDR 2.0 x $10^{-10}$). As depicted in **Figure 3**, 114 lncRNAs were down-regulated and 77 were up-regulated, with a fold change ranging from 0.3 (*AC100782.1*, FDR 6.5 x $10^{-4}$) to 4.5 (*LINC01411*, FDR 2.6 x $10^{-6}$).

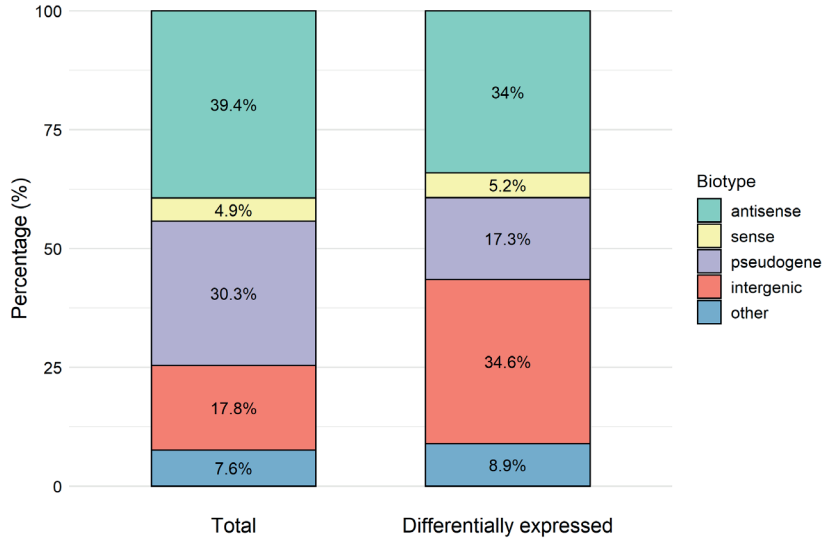



**Figure 2 | Distribution of biotypes of long noncoding RNAs (lncRNAs) |** Distribution of biotypes of lncRNAs expressed in cartilage compared to lncRNAs that were significantly differentially expressed (according to the false discovery rate) between lesioned osteoarthritis (OA) cartilage and preserved OA cartilage.

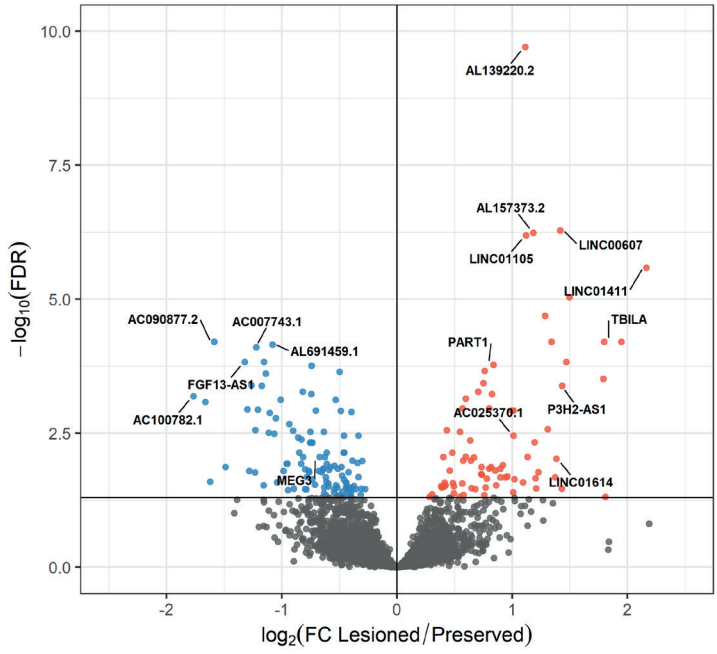



**Figure 3 | Differential expression analysis of long noncoding RNAs (lncRNAs) between lesioned osteoarthritis (OA) and preserved OA cartilage.** Volcano plot shows differentially expressed lncRNAs, with down-regulated lncRNAs represented by blue circles and up-regulated lncRNAs represented by red circles. Top differentially expressed lncRNAs are labeled, as are known and novel OA-associated lncRNAs. FDR = false discovery rate; FC = fold change.

The 191 identified lncRNAs in this study included several previously found to be associated to OA, such as *MEG3* (fold change 0.6, FDR 8.8 x $10^{-3}$), *PART1* (fold change 1.8, FDR 1.7 x $10^{-4}$), and *LINC01614* (fold change 2.6, FDR 9.5 x $10^{-3}$) [16, 26], as well as novel OA-associated lncRNAs, including *P3H2-AS1* (fold change 2.7, FDR 4.1 x $10^{-4}$) and *AC090877.2* (fold change 0.3, FDR 6.2 x $10^{-5}$). Notably, previously identified lncRNAs such as *MALAT1* (fold change 1.3, FDR 0.4) [27], *TUG1* (fold change 1.1, FDR 0.7) [28], *HOTAIR* (fold change 0.8, FDR 0.5), and *GAS5* (fold change 1.1, FDR 0.8) [29] were not found to be significantly differentially expressed in the present study.

To validate the differential expression results, we selected five lncRNAs (*AC025370.1*, *AC090877.2*, *MEG3*, *P3H2-AS1*, and *TBILA)* based on highest absolute fold change and genomic location, using RT-qPCR in a cohort consisting of 10 paired samples (**Supplementary Table 1B**) overlapping with the RNA-seq samples. All five lncRNAs were detected using RT-qPCR with equal direction of effect as those found in the RNA-seq analysis (**Supplementary Table 4**). Furthermore, replication was performed in an independent cohort of 10 paired cartilage samples (**Supplementary Table 1C**), which also showed comparable effect sizes and directions (**Supplementary Table 4**).

To explore whether joint-specific lncRNAs could be detected, stratified analyses were performed for knee samples (25 pairs) and hip samples (7 pairs). Upon performing differential expression analysis on the knee samples, 90 significantly differential expressed lncRNAs were identified (**Supplementary Table 5A**), of which 12 were not found in the previous combined analysis and therefore were unique to knee cartilage (**Supplementary Table 6A**). In the hip samples, 31 lncRNAs were significantly differentially expressed (**Supplementary Table 5B)**, of which 13 were unique to hip cartilage **(Supplementary Table 6B**). The most significantly differentially expressed lncRNA unique to the knee was *MSL3P1* (fold change 1.5, FDR 1.49 x $10^{-2}$), while one of the most significantly differentially expressed lncRNAs unique to the hip was *PAPPA-AS1* (fold change 9.4, FDR 2.77 x $10^{-4}$). Notably, the most up-regulated lncRNA in the hip, *AP001515.1* (fold change 21.5, FDR 2.8 x $10^{-4}$), was also unique to the hip, while the most up-regulated lncRNA in the knee, *LINC01411* (fold change 5.8, FDR 6.1 x $10^{-6}$), was not unique to the knee.



## Potential interactions between lncRNAs and mRNAs relevant in OA pathophysiology

We next aimed to investigate whether mRNAs associated with the OA process are regulated by differentially expressed lncRNAs. Based on the assumption that interactions between lncRNAs and mRNAs likely show coexpression [30] among lesioned and preserved OA cartilage samples, correlations were calculated between our previously reported differentially expressed protein-coding genes [4] and differentially expressed lncRNAs (**Supplementary Table 1D**), as shown in **Figure 1**. This resulted in 343 significant correlations (r>0.8) (**Supplementary Table 7**), comprising 47 unique lncRNAs, of which 17 were antisense (36%) and 14 were intergenic (30%) (**Supplementary Table 8**). This distribution is comparable to that found among all differentially expressed lncRNAs (**Figure 2**), supporting the notion that lncRNAs regulate mRNAs, independent of biotype. Notably, the most significantly differentially expressed lncRNA, *AL139220.2* (fold change 2.2, FDR 2.0 x $10^{-10}$), showed one of the highest correlations with *COL6A3* (r = 0.8, P = 2.2 x $10^{-16}$), encoding a collagen type VI chain.

To visualize these interactions, an OA-specific lncRNA-mRNA coexpression network was generated. As shown in **Figure 4**, three relative large clusters of interacting lncRNAs and mRNAs were observed. One cluster was characterized by being highly interlinked with a cluster of the same genes (e.g., *ITGB1BP1* correlated to the six lncRNAs *IER-AS1, AL355075.3*, *AC234917.1*, *AC091564.4*, *AC108449.3*, and *AL450306.1*), whereas the other two were characterized by lncRNAs interlinked with mostly unique genes (e.g. *LNCSRLR* with 18 genes). In addition to the clusters, there are a number of singular interlinked lncRNAs, such as *AC090877.2* (fold change 0.3, FDR 6.2 x $10^{-5}$) with *GREM1* (r = 0.9, P= 2.2 x $10^{-16}$), which encodes a cytokine of the BMP antagonist family (**Figure 4**). Interestingly, *GREM1* is the gene located closest to *AC090877.2*, suggesting that this lncRNA *cis*-regulates this gene.

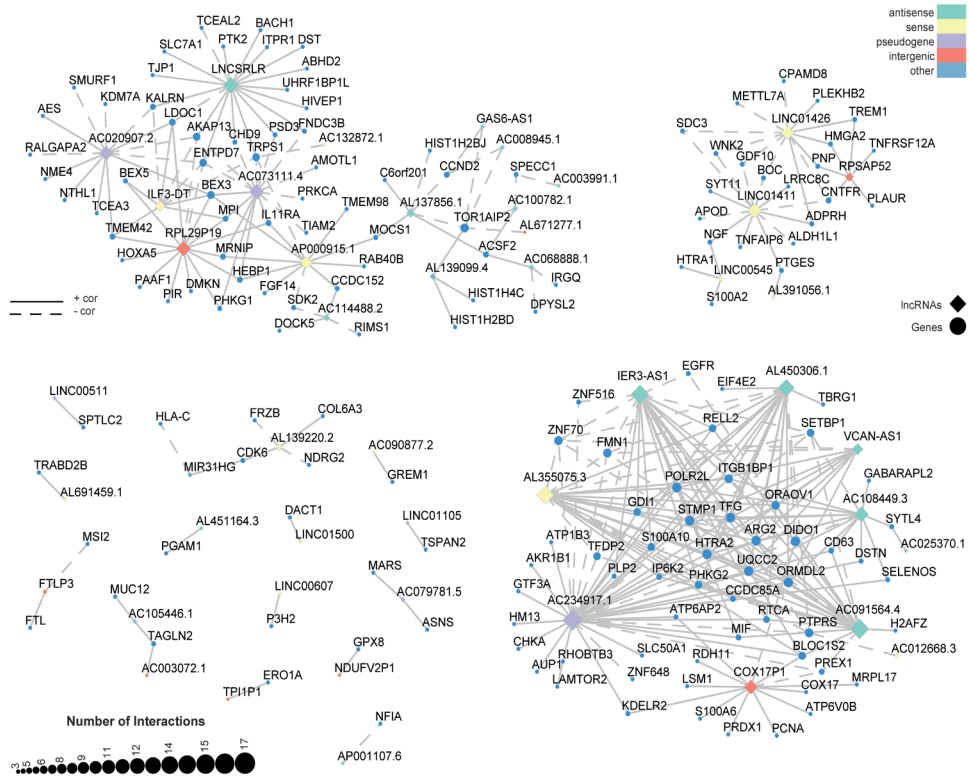



**Figure 4 | Osteoarthritis (OA)-specific long noncoding RNA (lncRNA)-mRNA coexpression network.** Network of differentially expressed lncRNAs and mRNAs with a correlation (cor) of >0.8 between lesioned OA and preserved OA cartilage is shown.

One of our objectives in the present study was to generalize the identification of potential *cis*-regulation of differentially expressed lincRNAs (**Figure 1**). As shown in **Figure 5A**, we compared the distribution of significant correlations between differentially expressed lincRNAs and all genes and between differentially expressed lincRNAs and genes that lie within a 100 kb window of the transcription start site. The proportion of significant correlations >0.5 with all differentially expressed genes was 11%, but this increased to 44% when we only considered the 100 kb window. Since the percentage differentially expressed antisense lncRNAs (34%) was comparable to that of intergenic lncRNAs (34.6%), we also aimed to identify potential *cis*-regulation of antisense lncRNAs. To this end, we compared the distribution of correlations between differentially expressed antisense lncRNAs and all protein-coding mRNAs and between differentially expressed antisense lncRNAs and their sense genes (**Figure 5B**). The percentage of correlations >0.5 was 10% with all genes and 61% with only the sense genes, showing that there is an enrichment for higher, positive correlations between antisense lncRNAs and their sense gene. Taken together, these data suggest that both intergenic and antisense lncRNAs are prone to regulate mRNAs in *cis* in OA cartilage.

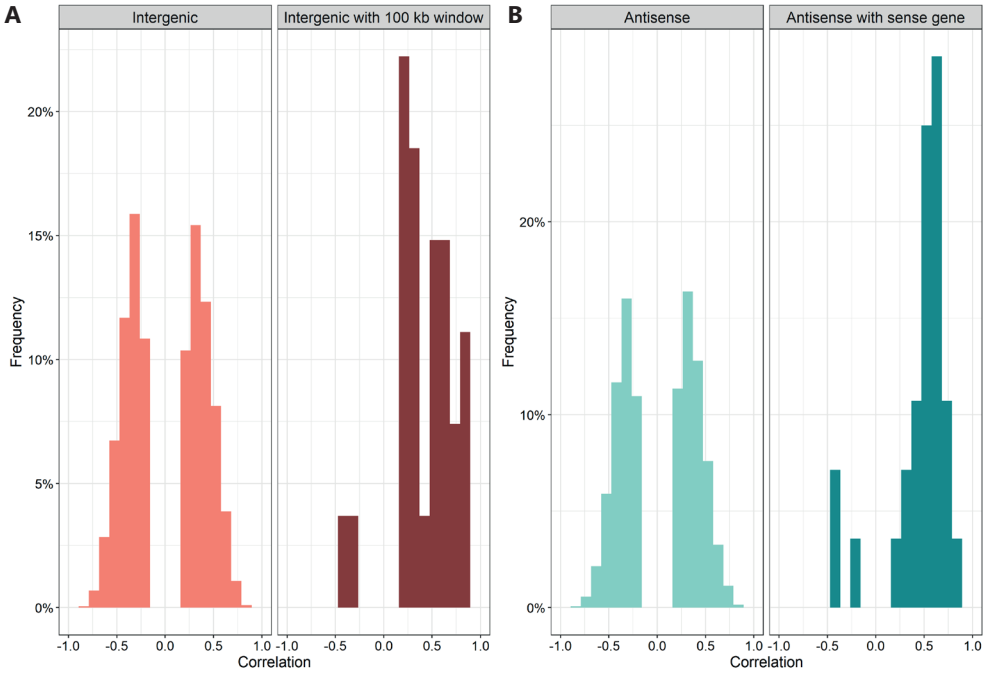



**Figure 5 | Distribution of significant correlations between intergenic differentially expressed long noncoding (lncRNAs) and previously identified differentially expressed protein-coding genes or protein-coding genes in a 100 kb window (A), and between antisense differentially expressed lncRNAs and differentially expressed protein-coding genes or their sense genes (B).** Correlations between lncRNA and mRNA data were calculated from the same osteoarthritis cartilage samples (n=98).

## Down-regulation of lncRNA expression using LNA GapmeRs

To validate whether the previously identified *cis*-regulation between lncRNAs and their surrounding genes is caused by a direct effect, *P3H2-AS1* was selected as a proof of concept for functional validation. *P3H2-AS1* is an antisense lncRNA, which was found to be highly up-regulated in lesioned OA cartilage (fold change 2.7, FDR 4.1 x $10^{-4}$) [4] and the highest correlation was with its sense gene *P3H2* (r = 0.63, P = 1.0 x $10^{-13}$) (**Supplementary Figure 1**). To this end, primary chondrocytes were transfected with a *P3H2-AS1* targeting LNA GapmeR. As shown in **Figure 6A**, this resulted in a significant down-regulation of *P3H2-AS1* compared to a non-targeting LNA GapmeR (fold change 0.28, P = 0.0035). Subsequently, *P3H2* expression levels were measured, which showed that *P3H2* expression was significantly down-regulated compared to cells transfected with non-targeting control LNA GapmeRs (fold change 0.36, P = 0.001) (**Figure 6B**).

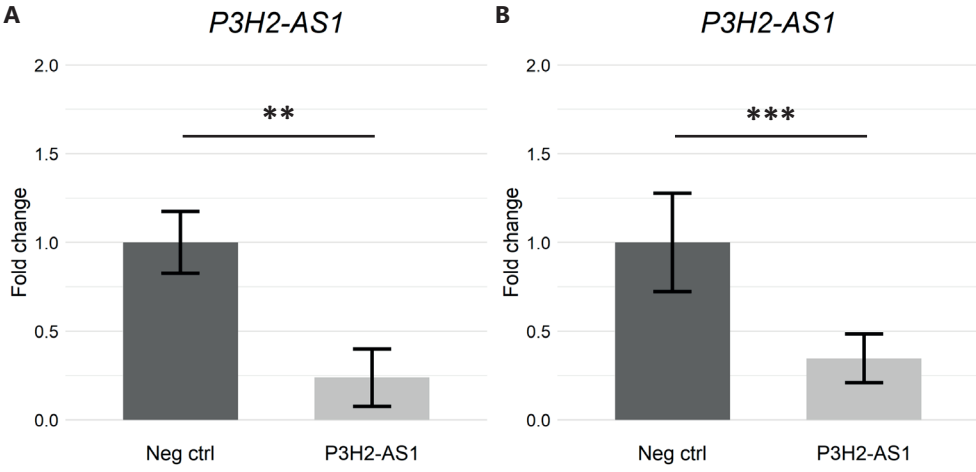



**Figure 6 | Expression of long noncoding RNA (lncRNA) *P3H2-AS1* and gene *P3H2* in primary chondrocytes transfected with *P3H2-AS1* targeting antisense locked nucleic acid (LNA) GapmeRs compared to non-targeting LNA GapmeRs.** (**A**) *P3H2-AS1* expression was significantly down-regulated by the *P3H2-AS1* targeting LNA GapmeRs. (**B**) *P3H2* expression was significantly down-regulated in chondrocytes transfected with *P3H2-AS1* targeting LNA GapmeRs. Bars show the mean ± SD, ** P < 0.01 *** P < 0.001 by paired t-test (N = 3 donors).

## Discussion

To our knowledge, we are the first to report on robust differential expression of lncRNAs as related to OA pathophysiology, while integrating them with data on differential mRNA expression levels of the same samples using RNA sequencing. As a result, our new in-house pipeline identified 5,053 lncRNAs that were robustly expressed, 191 of which were significantly differentially expressed (according to FDR) between lesioned and preserved OA cartilage. Notably, we observed an increase in the percentage of lincRNAs, highlighting their general involvement in the OA pathophysiology process. The directions of effect of *AC025370.1* (fold change 2.0, FDR 3.5 x $10^{-3}$), *AC090877.2* (fold change 0.3, FDR 6.2 x $10^{-5}$), *MEG3* (fold change 0.63, FDR 8.8 x $10^{-3}$), *P3H2-AS1* (fold change 2.7, FDR 4.1 x $10^{-4}$), and *TBILA* (fold change 3.5, FDR 1.1 x $10^{-7}$) was validated and replicated by RT-qPCR, indicating robustness of our lncRNA mapping strategy. Correlations were calculated to identify potential interactions between expression levels of differentially expressed lncRNAs and differentially expressed protein-coding genes [4] in the same OA cartilage samples. As a result, both intergenic and antisense differentially expressed lncRNAs showed an enrichment for higher, positive correlations with their respective flanking or sense genes compared to the total dataset. Validating this *cis*-regulation in vitro, *P3H2-AS1* levels were down-regulated in primary chondrocytes, which resulted in down-regulation of the sense gene *P3H2* expression levels, thereby confirming that *P3H2-AS1* regulates its sense gene *P3H2*.

We identified 29,219 lncRNAs that were expressed in OA cartilage. However, after applying a filter with a cutoff of an average of ≥2 counts per lncRNA, the detected lncRNAs were reduced by ~83% to 5,053. Since lncRNAs are known to be expressed at very low levels, this was to be expected, yet lncRNAs expressed at low levels can still be functional [12]. To allow exploratory analyses with lncRNAs expressed at such low levels, deeper sequencing would be necessary, with a read-dept of ~50 million reads per sample. Additionally, to be able to report on valid lncRNAs in OA articular cartilage and their potential target mRNAs, we prioritized reporting known lncRNAs with a predicted non-protein-coding potential. Nonetheless, by focusing on these known lncRNAs, we may have disregarded compelling novel OA-relevant lncRNAs.



Given that we had a (within patient) paired lesioned cartilage – preserved cartilage study design, with pairs sequenced on the same batch, we applied a paired Wald's test as implemented in the DESeq2 R package. Since our dataset also included lncRNAs expressed at low levels, the addition of a random effect to compensate for technical errors may have been a better, yet more conservative, approach. As such, the lncRNAs in our dataset, particularly those with low read counts, could be subject to false positive results and therefore require replication and verification.

We observed a particular enrichment of lincRNAs in the differential expression analysis compared to the total dataset (34.6% versus 17.8%) (**Figure 2**), showing that lincRNAs indeed play an important role in OA pathophysiology, as seen in previous studies [8, 16, 30]. Nonetheless, in comparison to the fraction of significantly differentially expressed lncRNAs reported by Pearson *et al.* [8], this proportion is still relative small. However, Pearson and colleagues performed RNA-seq on samples from isolated chondrocytes in contrast to the RNA isolated from cartilage in our study, and focused on profiling lncRNAs up-regulated by interleukin-1β. The activation of chondrocyte proliferation in tissue culture will likely induce expression of RNAs involved in transcriptional regulation, compared to the transcriptome of maturational arrested chondrocytes residing in cartilage.

Of the 191 lncRNAs that were significantly differentially expressed between lesioned and preserved OA cartilage (**Figure 3**), multiple lncRNAs have been previously identified, including *MEG3*, *LINC01614*, and *PART1* [16, 26]. However, there were also examples of lncRNAs previously associated with OA, which were not significantly differentially expressed, such as *MALAT1*, *HOTAIR*, *GAS5*, and *TUG1* [27-29]. A possible explanation could be that they were found to be differentially expressed between preserved OA and healthy cartilage, as opposed to our comparison between lesioned OA cartilage and preserved OA cartilage [7]. The cross-sectional study design comparing OA cartilage and healthy cartilage provides insight into which lncRNAs are involved in the early phase of OA pathophysiology and are therefore potentially causal in the process and which lncRNAs are specific for healthy cartilage; this

was not possible with our study design. Nonetheless, the paired analysis allowed for detection of lncRNA expression changes specific to the OA pathophysiological process, independent of confounding factors such as sex and age. At least 35 differentially expressed lncRNAs in our dataset were previously found to be associated with OA [10, 16, 30], but the most significantly differentially expressed lncRNA, *AL139220.2,* and the most up- and down-regulated differentially expressed lncRNAs, *LINC01411* and *AC100782.1*, respectively, have not previously been associated with OA, showing that a paired study design allows for the detection of many more lncRNAs involved in the OA pathophysiological process [3].

Previous studies have demonstrated differences in dysregulated pathways between knee and hip OA cartilage and epigenetic differences based on DNA methylation [8, 16, 31, 32]; thus, we aimed to identify joint-specific lncRNAs. Differential expression analysis in knee samples resulted in a higher number of significantly differentially expressed lncRNAs (N = 90) than in hip samples (N = 31), which could be due to the smaller sample size of the hip samples (25 knee pairs versus 7 hip pairs). However, the number of unique lncRNAs per joint site was similar: 12 unique knee lncRNAs and 13 unique hip lncRNAs. This suggests that there is more heterogeneity in the processes in the knee, which could be due in part to anatomical joint site-specific differences. This is also supported by the fact that the average fold change of the up-regulated lncRNAs unique to the hip is 8.3, while it was 1.5 for knee. The unique lncRNAs with the highest fold change in the knee (*AC068768.1*, fold change 1.6, FDR $2.3 \times 10^{-2}$) and hip (*AP001615.1*, fold change 21.5, FDR $2.8 \times 10^{-4}$) were not previously found to be associated with OA. The identification of these joint specific lncRNAs are interesting for follow-up studies to determine potential joint specific therapeutic targets.

Unlike conserved microRNAs, it is difficult to predict the function of lncRNAs based solely on nucleotide sequence, due to their lack of conservation of the primary sequence [15]. To explore potential regulatory interactions between lncRNAs and mRNAs in cartilage, correlations were calculated between differentially expressed lncRNAs and differentially expressed protein-coding mRNAs (**Figure 4**). At the transcriptional level, lncRNAs can exert their function in *trans* or *cis* [13], both of which we observed in this study. The most significantly differentially expressed lncRNA, *AL139220.2*, showed one of the highest correlations with *COL6A3* ($r = 0.8$, $P = 2.2 \times 10^{-16}$), encoding one of the collagen type VI chains as part of the complete collagen type VI molecule, which is mostly present in the pericellular matrix of cartilage. *AL139220.2* is located on chromosome 1 and, at present, little is known about its function. Since *COL6A3* is located on chromosome 2, it seems likely that *AL139220.2* regulates *COL6A3* expression in *trans*. Notably, *AC090877.2* showed the highest correlation with its sense gene *GREM1* ($r = 0.9$, $P=2.2 \times 10^{-16}$), suggesting that this lncRNA *cis*-regulates this gene. In previous studies, it has been shown that lincRNAs often regulate flanking mRNAs in *cis* in OA, in which a positive correlation was found between the expression of mRNA-flanking lincRNAs and their nearest

coding mRNA [8, 30]. This observation was confirmed by our findings, as the percentage of higher, positive correlations (r > 0.5) was considerably larger between lincRNAs and the differentially expressed genes that lie within a 100 kb window (44%) than with all differentially expressed genes (11%) (**Figure 5A**).

Furthermore, it is known that antisense lncRNAs can regulate their overlapping sense genes in *cis* [14], which has not previously been investigated in OA. We found an enrichment for higher, positive correlations between antisense differentially expressed lncRNAs and their sense gene (r > 0.5 in 61%) compared to correlations between antisense differentially expressed lncRNAs and all differentially expressed genes (r > 0.5 in 10%), suggesting that indeed antisense lncRNAs often regulate their sense gene in *cis* (**Figure 5B**). Therefore, to completely understand the transcriptional regulation of lncRNAs in the OA process, the total lncRNA transcriptome should be considered, not solely the lincRNAs. Of importance is the notion that these correlations are not yet proof of a (direct) downstream effect of lncRNAs on the mRNAs.



Given these observations, we selected the antisense lncRNA *P3H2-AS1* as proof of principle to establish whether it regulates its sense gene. Down-regulation of *P3H2-AS1* resulted in a significant down-regulation of *P3H2* expression levels (**Figure 6**), thereby confirming that *P3H2-AS1* regulates its sense gene in *cis*. *P3H2* encodes an enzyme that catalyzes post-translational 3-hydroxylation of proline residues and plays a critical role in collagen chain assembly, stability, and cross-linking and was recently found to be highly up-regulated in lesioned OA cartilage, and therefore likely involved in the OA process [4]. Antisense lncRNAs can affect biogenesis or mobilization of target RNA on multiple levels, such as transcription, splicing, and translation [14]. To elucidate the exact mechanism of *P3H2-AS1* regulating *P3H2* and investigate whether *P3H2-AS1* can be used as a potential preclinical target by modulating *P3H2* expression levels via *P3H2-AS1*, complementary functional studies employing e.g. CRISPR/Cas9, RNA fluorescence in situ hybridization, or crosslinked immunoprecipitation are necessary [33].

In conclusion, our improved detection strategy resulted in the characterization of lncRNAs robustly expressed in OA cartilage. Our data signify that intergenic, as well as antisense lncRNAs play an important role in regulating the pathophysiology of OA. Moreover, we observed that in addition to the previous finding that intergenic lncRNAs function in *cis*, antisense lncRNAs can exert their function in *cis,* which we confirmed in vitro. Future studies regarding lncRNAs and OA should be complemented by functional validation, e.g., by modulating lncRNA expression levels using antisense LNA GapmeRs, in order to confirm whether a correlation signifies a biological relation between lncRNA and mRNA.

## Acknowlegdements

We thank all the participants of the RAAK study (supported by the Leiden University Medical Center). We thank Evelyn Houtman, Enrike van der Linden, Robert van der Wal, Peter van Schie, Shaho Hasan, Maartje Meijer, Daisy Latijnhouwers, and Geert Spierenburg for their contribution to the collection of the joint tissue. Data were generated within the scope of the Medical Delta programs Regenerative Medicine 4D: Generating complex tissues with stem cells and printing technology and Improving Mobility with Technology.

**Funding:**

The study was funded by the Dutch Scientific Research council NWO /ZonMW VICI scheme (nr 91816631/528), the Dutch Arthritis Society (DAA-10-1-402, DAF-16-1-405), and TreatOA (grant 200800 from the European Commission Seventh Framework Program).

## References

1. Goldring, S.R. and M.B. Goldring, *Changes in the osteochondral unit during osteoarthritis: structure, function and cartilage-bone crosstalk*. Nat Rev Rheumatol, 2016. **12**(11): p. 632-644.
2. Tonge, D.P., M.J. Pearson, and S.W. Jones, *The hallmarks of osteoarthritis and the potential to develop personalised disease-modifying pharmacological therapeutics*. Osteoarthritis and Cartilage, 2014. **22**(5): p. 609-621.
3. Ramos, Y.F., et al., *Genes involved in the osteoarthritis process identified through genome wide expression analysis in articular cartilage; the RAAK study*. PLoS One, 2014. **9**(7): p. e103056.
4. Coutinho de Almeida, R., et al., *RNA sequencing data integration reveals an miRNA interactome of osteoarthritis cartilage*. Ann Rheum Dis, 2019. **78**(2): p. 270-277.
5. Coutinho De Almeida, R., Y.F.M. Ramos, and I. Meulenbelt, *Involvement of epigenetics in osteoarthritis*. Best Practice & Research Clinical Rheumatology, 2017. **31**(5): p. 634-648.
6. den Hollander, W., et al., *Transcriptional associations of osteoarthritis-mediated loss of epigenetic control in articular cartilage*. Arthritis Rheumatol, 2015. **67**(8): p. 2108-16.
7. Jiang, S.D., et al., *Long noncoding RNAs in osteoarthritis*. Joint Bone Spine, 2017. **84**(5): p. 553-556.
8. Pearson, M.J., et al., *Long Intergenic Noncoding RNAs Mediate the Human Chondrocyte Inflammatory Response and Are Differentially Expressed in Osteoarthritis Cartilage*. Arthritis Rheumatol, 2016. **68**(4): p. 845-56.
9. Sun, H., et al., *Emerging roles of long noncoding RNA in chondrogenesis, osteogenesis, and osteoarthritis*. Am J Transl Res, 2019. **11**(1): p. 16-30.
10. Liu, Q., et al., *Long Noncoding RNA Related to Cartilage Injury Promotes Chondrocyte Extracellular Matrix Degradation in Osteoarthritis*. Arthritis & Rheumatology, 2014. **66**(4): p. 969-978.
11. Frankish, A., et al., *GENCODE reference annotation for the human and mouse genomes*. Nucleic Acids Research, 2019. **47**(D1): p. D766-D773.
12. Jarroux, J., A. Morillon, and M. Pinskaya, *History, Discovery, and Classification of lncRNAs*. Adv Exp Med Biol, 2017. **1008**: p. 1-46.
13. Gil, N. and I. Ulitsky, *Regulation of gene expression by cis-acting long noncoding RNAs*. Nat Rev Genet, 2020. **21**(2): p. 102-117.
14. Villegas, V.E. and P.G. Zaphiropoulos, *Neighboring gene regulation by antisense long non-coding RNAs*. Int J Mol Sci, 2015. **16**(2): p. 3251-66.

15. Uszczynska-Ratajczak, B., et al., *Towards a complete map of the human long non-coding RNA transcriptome*. Nature Reviews Genetics, 2018. **19**(9): p. 535-548.
16. Ajekigbe, B., et al., *Identification of long non-coding RNAs expressed in knee and hip osteoarthritic cartilage*. Osteoarthritis Cartilage, 2019. **27**(4): p. 694-702.
17. Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner*. Bioinformatics, 2013. **29**(1): p. 15-21.
18. Pertea, M., et al., *StringTie enables improved reconstruction of a transcriptome from RNA-seq reads*. Nature Biotechnology, 2015. **33**(3): p. 290-295.
19. Cunningham, F., et al., *Ensembl 2019*. Nucleic acids research, 2019. **47**(D1): p. D745-D751.20. Wang, L., et al., *CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model*. Nucleic Acids Res, 2013. **41**(6): p. e74.
21. Han, S., et al., *LncFinder: an integrated platform for long non-coding RNA identification utilizing sequence intrinsic composition, structural information and physicochemical property*. Briefings in Bioinformatics, 2018.
22. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2*. Genome Biology, 2014. **15**(12).
23. Ritchie, M.E., et al., *limma powers differential expression analyses for RNA-sequencing and microarray studies*. Nucleic Acids Research, 2015. **43**(7): p. e47-e47.
24. Castro, M.A., et al., *RedeR: R/Bioconductor package for representing modular structures, nested networks and multiple levels of hierarchical associations*. Genome Biol, 2012. **13**(4): p. R29.
25. Bomer, N., et al., *Neo-cartilage engineered from primary chondrocytes is epigenetically similar to autologous cartilage, in contrast to using mesenchymal stem cells*. Osteoarthritis Cartilage, 2016. **24**(8): p. 1423-30.
26. Chen, K., et al., *LncRNA MEG3 Inhibits the Degradation of the Extracellular Matrix of Chondrocytes in Osteoarthritis via Targeting miR-93/TGFBR2 Axis*. Cartilage, 2019: p. 1947603519855759.
27. Zhang, Y., et al., *LncRNA MALAT1 promotes osteoarthritis by modulating miR-150-5p/AKT3 axis*. Cell & Bioscience, 2019. **9**(1).
28. Tang, L.P., et al., *LncRNA TUG1 promotes osteoarthritis-induced degradation of chondrocyte extracellular matrix via miR-195/MMP-13 axis*. Eur Rev Med Pharmacol Sci, 2018. **22**(24): p. 8574-8581.
29. Xing, D., et al., *Identification of long noncoding RNA associated with osteoarthritis in humans*. Orthop Surg, 2014. **6**(4): p. 288-93.
30. Fu, M., et al., *Expression profile of long noncoding RNAs in cartilage from knee osteoarthritis patients*. Osteoarthritis and Cartilage, 2015. **23**(3): p. 423-432.
31. den Hollander, W., et al., *Knee and hip articular cartilage have distinct epigenomic landscapes: implications for future cartilage regeneration approaches*. Ann Rheum Dis, 2014. **73**(12): p. 2208-12.
32. Loughlin, J. and L.N. Reynard, *Osteoarthritis: Epigenetics of articular cartilage in knee and hip OA*. Nat Rev Rheumatol, 2015. **11**(1): p. 6-7.
33. Charles Richard, J.L. and P.J.A. Eichhorn, *Platforms for Investigating LncRNA Functions*. SLAS TECHNOLOGY: Translating Life Sciences Innovation, 2018. **23**(6): p. 493-506.

# Supplementary Materials

## Supplementary Tables

**Supplementary Table 1 | Sample characteristics included in the analysis in the (A) discovery, (B) validation (C) replication, and (D) correlation analyses.**

**A**

| Discovery | | | | | | |
|---|---|---|---|---|---|---|
| **Tissue type** | **Mean age (range)** | **# Men** | **# Women** | **# Preserved** | **# Lesioned** | **# Total** |
| Knee | 68.8 (48 -79) | 6 | 19 | 25 | 25 | 50 |
| Hip | 64.4 (48 -79) | 2 | 12 | 7 | 7 | 14 |
| All | 67.8 (48 -79) | 7 | 25 | 32 | 32 | 64 |

**B**

| Validation | | | | | | |
|---|---|---|---|---|---|---|
| **Tissue type** | **Mean age (range)** | **# Men** | **# Women** | **# Preserved** | **# Lesioned** | **# Total** |
| Knee | 69.7 (60 - 77) | 2 | 5 | 7 | 7 | 14 |
| Hip | 69.7 (65 - 79) | 1 | 2 | 3 | 3 | 6 |
| All | 69.7 (60 - 79) | 3 | 7 | 10 | 10 | 20 |

**C**

| Replication | | | | | | |
|---|---|---|---|---|---|---|
| **Tissue type** | **Mean age (range)** | **# Men** | **# Women** | **# Preserved** | **# Lesioned** | **# Total** |
| Knee | 64.6 (57 - 75) | 3 | 2 | 5 | 5 | 10 |
| Hip | 68.2 (59 - 75) | 2 | 3 | 5 | 5 | 10 |
| All | 66.4 (57 - 75) | 5 | 5 | 10 | 10 | 20 |

**D**

| Correlations | | | | | | |
|---|---|---|---|---|---|---|
| **Tissue type** | **Mean age (range)** | **# Men** | **# Women** | **# Preserved** | **# Lesioned** | **# Total** |
| Knee | 68.5 (46 - 79) | 54 | 11 | 35 | 30 | 65 |
| Hip | 66.2 (48 - 82) | 27 | 6 | 22 | 11 | 33 |
| All | 67.7 (46 - 82) | 81 | 17 | 57 | 41 | 98 |

**Supplementary Table 2 | Primer sequences to measure mRNA and lncRNA expression levels.**

| Gene | Forward primer (5' --> 3') | Reverse primer (5' --> 3') |
|---|---|---|
| *AC025370.1* | AGCCAGCTTTTAAGTGAACCTG | GTGCTATAACTCTCCTGCCCA |
| *AC090877.2* | AAGCACATGGGACCCTCTCA | TGAATTGTGAAGAACCATCGCG |
| *GAPDH* | TGCCATGTAGACCCCTTGAAG | ATGGTACATGACAAGGTGCGG |
| *MEG3* | CCACCCCTCTTGCTTGTCTT | CCTGGAGTGCTGTTGGAGAA |
| *P3H2-AS1* | CACTGCCTGATGGGTACTAGC | TTGAGACTTGGAGAGGCCTTG |
| *P3H2* | AGAGAAGCCAAGCCACACAT | GCTTGTTCGAAGTGCCTGAT |
| *TBILA* | CGGGACAGGAATCATGGATTTT | ACAGATGAGTGACCAAACTGGA |

**Supplementary Table 3 | Significant (according to false discovery rate) differentially expressed long noncoding RNAs in lesioned osteoarthritic cartilage compared to preserved osteoarthritic cartilage samples.**

| GeneID | GeneName | baseMean | log2FoldChange | FoldChange | Pvalue | Padj |
|---|---|---|---|---|---|---|
| ENSG00000230615 | *AL139220.2* | 52.47 | 1.11 | 2.16 | 3.90E-14 | 1.97E-10 |
| ENSG00000235770 | *LINC00607* | 58.39 | 1.42 | 2.67 | 2.06E-10 | 5.20E-07 |
| ENSG00000229896 | *AL157373.2* | 51.56 | 1.18 | 2.27 | 3.44E-10 | 5.80E-07 |
| ENSG00000232044 | *LINC01105* | 88.54 | 1.12 | 2.17 | 5.06E-10 | 6.39E-07 |
| ENSG00000249306 | *LINC01411* | 50.68 | 2.16 | 4.48 | 2.55E-09 | 2.58E-06 |
| ENSG00000236094 | *LINC00545* | 18.23 | 1.50 | 2.82 | 1.09E-08 | 9.21E-06 |
| ENSG00000258137 | *AC079313.2* | 6.31 | 1.28 | 2.44 | 2.83E-08 | 2.05E-05 |
| ENSG00000184669 | *OR7E14P* | 4.94 | 1.95 | 3.86 | 1.35E-07 | 6.22E-05 |
| ENSG00000261488 | *TBILA* | 6.10 | 1.80 | 3.48 | 1.14E-07 | 6.22E-05 |
| ENSG00000258520 | *AL359317.1* | 20.67 | 1.34 | 2.53 | 1.34E-07 | 6.22E-05 |
| ENSG00000259721 | *AC090877.2* | 22.03 | -1.59 | 0.33 | 1.01E-07 | 6.22E-05 |
| ENSG00000223814 | *AL691459.1* | 20.86 | -1.08 | 0.47 | 1.67E-07 | 7.03E-05 |
| ENSG00000233251 | *AC007743.1* | 86.02 | -1.22 | 0.43 | 2.03E-07 | 7.88E-05 |
| ENSG00000237525 | *AC012668.3* | 7.18 | 1.47 | 2.77 | 4.65E-07 | 1.48E-04 |
| ENSG00000256995 | *AC084816.1* | 60.00 | -1.16 | 0.45 | 4.70E-07 | 1.48E-04 |
| ENSG00000226031 | *FGF13-AS1* | 3.29 | -1.32 | 0.40 | 4.25E-07 | 1.48E-04 |
| ENSG00000152931 | *PART1* | 267.61 | 0.84 | 1.78 | 5.62E-07 | 1.67E-04 |
| ENSG00000269609 | *RPARP-AS1* | 24.73 | -0.74 | 0.60 | 6.18E-07 | 1.73E-04 |
| ENSG00000225511 | *LINC00475* | 28.26 | 0.76 | 1.69 | 8.13E-07 | 2.16E-04 |
| ENSG00000257337 | *AC068888.1* | 268.12 | -0.50 | 0.71 | 8.90E-07 | 2.25E-04 |
| ENSG00000258583 | *LINC01500* | 20.02 | -1.14 | 0.45 | 1.02E-06 | 2.45E-04 |
| ENSG00000241749 | *RPSAP52* | 9.65 | 1.79 | 3.46 | 1.33E-06 | 3.05E-04 |
| ENSG00000226453 | *LINC02542* | 11.13 | 0.75 | 1.68 | 1.66E-06 | 3.65E-04 |
| ENSG00000228113 | *AC003991.1* | 78.53 | -1.26 | 0.42 | 1.90E-06 | 4.00E-04 |
| ENSG00000225764 | *P3H2-AS1* | 3.09 | 1.43 | 2.70 | 2.11E-06 | 4.12E-04 |
| ENSG00000278869 | *BX539320.1* | 5.91 | -1.17 | 0.44 | 2.12E-06 | 4.12E-04 |
| ENSG00000231690 | *LINC00574* | 26.22 | 0.71 | 1.63 | 2.95E-06 | 5.32E-04 |
| ENSG00000280007 | *AC008079.1* | 20.64 | -0.82 | 0.57 | 2.93E-06 | 5.32E-04 |
| ENSG00000234380 | *LINC01426* | 14.78 | 0.82 | 1.77 | 3.44E-06 | 5.91E-04 |
| ENSG00000131797 | *CLUHP3* | 20.79 | -0.74 | 0.60 | 3.51E-06 | 5.91E-04 |
| ENSG00000254238 | *AC100782.1* | 196.11 | -1.77 | 0.29 | 3.98E-06 | 6.49E-04 |
| ENSG00000278419 | *AL451164.3* | 291.29 | 0.59 | 1.51 | 4.55E-06 | 7.18E-04 |
| ENSG00000220785 | *MTMR9LP* | 65.31 | -0.53 | 0.69 | 5.07E-06 | 7.53E-04 |
| ENSG00000255471 | *AP001528.2* | 7.82 | -1.01 | 0.50 | 4.94E-06 | 7.53E-04 |
| ENSG00000266968 | *AC023421.1* | 3.96 | -1.66 | 0.32 | 5.75E-06 | 8.30E-04 |
| ENSG00000253210 | *AC040970.1* | 14.98 | 0.80 | 1.74 | 8.01E-06 | 1.09E-03 |
| ENSG00000255343 | *AC234917.1* | 18.06 | 0.57 | 1.48 | 7.96E-06 | 1.09E-03 |
| ENSG00000225194 | *LINC00092* | 15.69 | -1.30 | 0.41 | 8.55E-06 | 1.14E-03 |
| ENSG00000260244 | *AC104083.1* | 16.23 | -1.21 | 0.43 | 8.93E-06 | 1.16E-03 |
| ENSG00000283029 | *AL139099.4* | 37836.17 | 1.01 | 2.02 | 1.00E-05 | 1.20E-03 |
| ENSG00000270412 | *AL136084.3* | 11.50 | 0.99 | 1.99 | 9.65E-06 | 1.20E-03 |
| ENSG00000232677 | *LINC00665* | 14.12 | -0.70 | 0.61 | 9.78E-06 | 1.20E-03 |
| ENSG00000183935 | *HTR7P1* | 31.02 | -0.49 | 0.71 | 1.03E-05 | 1.21E-03 |
| ENSG00000188242 | *PP7080* | 82.64 | -0.40 | 0.76 | 1.11E-05 | 1.28E-03 |
| ENSG00000223561 | *AC005165.1* | 47.47 | -1.10 | 0.47 | 1.18E-05 | 1.33E-03 |
| ENSG00000271474 | *AC106881.1* | 8.29 | -1.05 | 0.48 | 1.51E-05 | 1.65E-03 |
| ENSG00000267248 | *AC025048.2* | 3.61 | -0.94 | 0.52 | 1.99E-05 | 2.14E-03 |
| ENSG00000182366 | *FAM87A* | 3.25 | 1.31 | 2.48 | 2.52E-05 | 2.66E-03 |
| ENSG00000229214 | *LINC00242* | 18.90 | 0.43 | 1.35 | 2.69E-05 | 2.75E-03 |

| GeneID | GeneName | baseMean | log2FoldChange | FoldChange | Pvalue | Padj |
|---|---|---|---|---|---|---|
| ENSG00000186466 | *AQP7P1* | 8.77 | -1.23 | 0.43 | 2.72E-05 | 2.75E-03 |
| ENSG00000225978 | *HAR1A* | 12.31 | -0.90 | 0.54 | 2.85E-05 | 2.83E-03 |
| ENSG00000227036 | *LINC00511* | 137.66 | 0.54 | 1.46 | 3.19E-05 | 2.99E-03 |
| ENSG00000232368 | *FTLP2* | 12.62 | -0.64 | 0.64 | 3.16E-05 | 2.99E-03 |
| ENSG00000255153 | *TOLLIP-AS1* | 8.05 | -0.75 | 0.59 | 3.14E-05 | 2.99E-03 |
| ENSG00000197320 | *AC060834.1* | 122.01 | -1.11 | 0.46 | 3.38E-05 | 3.10E-03 |
| ENSG00000256321 | *AC087258.1* | 14.94 | -1.07 | 0.48 | 3.61E-05 | 3.26E-03 |
| ENSG00000253851 | *AC025370.1* | 58.50 | 1.01 | 2.02 | 4.08E-05 | 3.50E-03 |
| ENSG00000232699 | *BDH2P1* | 37.99 | -0.33 | 0.79 | 3.97E-05 | 3.50E-03 |
| ENSG00000267100 | *ILF3-DT* | 143.23 | -0.46 | 0.73 | 4.03E-05 | 3.50E-03 |
| ENSG00000234840 | *LINC01239* | 9.84 | -0.86 | 0.55 | 4.50E-05 | 3.79E-03 |
| ENSG00000233725 | *LINC00284* | 9.25 | -0.83 | 0.56 | 5.01E-05 | 4.15E-03 |
| ENSG00000266954 | *AP001010.1* | 29.87 | 0.63 | 1.55 | 5.30E-05 | 4.32E-03 |
| ENSG00000280693 | *SH3PXD2A-AS1* | 5.72 | 1.20 | 2.29 | 5.90E-05 | 4.67E-03 |
| ENSG00000223652 | *AC106786.1* | 6.95 | -0.75 | 0.60 | 5.91E-05 | 4.67E-03 |
| ENSG00000251661 | *AC136475.1* | 10.57 | -0.73 | 0.60 | 6.18E-05 | 4.77E-03 |
| ENSG00000267414 | *AC120049.1* | 10.20 | -0.76 | 0.59 | 6.23E-05 | 4.77E-03 |
| ENSG00000257122 | *RRN3P3* | 53.03 | 0.48 | 1.39 | 1.00E-04 | 7.34E-03 |
| ENSG00000229044 | *AL451070.1* | 25.71 | -0.46 | 0.72 | 1.02E-04 | 7.34E-03 |
| ENSG00000224729 | *PCOLCE-AS1* | 670.82 | -0.61 | 0.66 | 9.90E-05 | 7.34E-03 |
| ENSG00000162913 | *OBSCN-AS1* | 9.15 | -0.85 | 0.56 | 9.96E-05 | 7.34E-03 |
| ENSG00000266904 | *LINC00663* | 23.15 | -0.46 | 0.73 | 1.03E-04 | 7.34E-03 |
| ENSG00000260549 | *MT1L* | 80.60 | 0.59 | 1.51 | 1.23E-04 | 8.65E-03 |
| ENSG00000268941 | *LINC01711* | 9.50 | 1.13 | 2.19 | 1.29E-04 | 8.74E-03 |
| ENSG00000282057 | *AC092807.3* | 34.23 | 0.40 | 1.32 | 1.27E-04 | 8.74E-03 |
| ENSG00000272668 | *AL590560.1* | 69.60 | -0.63 | 0.65 | 1.31E-04 | 8.74E-03 |
| ENSG00000261959 | *AC015909.3* | 30.43 | -0.82 | 0.57 | 1.30E-04 | 8.74E-03 |
| ENSG00000214548 | *MEG3* | 775.47 | -0.67 | 0.63 | 1.34E-04 | 8.81E-03 |
| ENSG00000205105 | *COX17P1* | 20.18 | 0.67 | 1.59 | 1.40E-04 | 9.05E-03 |
| ENSG00000230838 | *LINC01614* | 21.19 | 1.38 | 2.60 | 1.48E-04 | 9.46E-03 |
| ENSG00000273328 | *AC099329.2* | 20.31 | 0.64 | 1.56 | 1.64E-04 | 1.02E-02 |
| ENSG00000231711 | *LINC00899* | 28.16 | -0.39 | 0.76 | 1.62E-04 | 1.02E-02 |
| ENSG00000250318 | *AC003072.1* | 13.23 | 0.58 | 1.49 | 1.68E-04 | 1.03E-02 |
| ENSG00000225791 | *TRAM2-AS1* | 64.91 | -0.30 | 0.81 | 1.72E-04 | 1.05E-02 |
| ENSG00000204622 | *HLA-J* | 76.75 | -0.34 | 0.79 | 1.88E-04 | 1.13E-02 |
| ENSG00000237238 | *BMS1P10* | 5.89 | -0.83 | 0.56 | 2.02E-04 | 1.17E-02 |
| ENSG00000231394 | *AC099681.2* | 5.25 | -0.95 | 0.52 | 2.02E-04 | 1.17E-02 |
| ENSG00000269416 | *LINC01224* | 4.42 | -0.96 | 0.52 | 2.02E-04 | 1.17E-02 |
| ENSG00000260563 | *AC132872.1* | 32.58 | -0.61 | 0.66 | 2.06E-04 | 1.18E-02 |
| ENSG00000171889 | *MIR31HG* | 13.35 | 0.92 | 1.89 | 2.21E-04 | 1.26E-02 |
| ENSG00000262580 | *AC087741.1* | 23.27 | -0.54 | 0.69 | 2.28E-04 | 1.28E-02 |
| ENSG00000273186 | *AL359091.5* | 5.26 | 0.73 | 1.66 | 2.48E-04 | 1.35E-02 |
| ENSG00000259728 | *LINC00933* | 12.58 | -0.48 | 0.72 | 2.46E-04 | 1.35E-02 |
| ENSG00000271239 | *AC007423.1* | 3.65 | -1.49 | 0.36 | 2.46E-04 | 1.35E-02 |
| ENSG00000272273 | *IER3-AS1* | 226.80 | 0.81 | 1.76 | 2.53E-04 | 1.36E-02 |
| ENSG00000271858 | *CYB561D2* | 16.03 | -0.58 | 0.67 | 2.56E-04 | 1.36E-02 |
| ENSG00000184385 | *UMODL1-AS1* | 3.95 | 0.90 | 1.86 | 2.82E-04 | 1.44E-02 |
| ENSG00000226816 | *AC005082.1* | 4.92 | 0.80 | 1.74 | 2.77E-04 | 1.44E-02 |
| ENSG00000257261 | *AC008014.1* | 9.40 | -0.51 | 0.70 | 2.74E-04 | 1.44E-02 |
| ENSG00000254693 | *AC010768.1* | 18.42 | -0.64 | 0.64 | 2.81E-04 | 1.44E-02 |

| GeneID | GeneName | baseMean | log2FoldChange | FoldChange | Pvalue | Padj |
|---|---|---|---|---|---|---|
| ENSG00000276791 | *AC092117.1* | 20.75 | -0.80 | 0.58 | 2.92E-04 | 1.48E-02 |
| ENSG00000221857 | *AC020907.2* | 252.95 | -0.64 | 0.64 | 3.02E-04 | 1.51E-02 |
| ENSG00000225138 | *SLC9A3-AS1* | 56.43 | -0.59 | 0.66 | 3.08E-04 | 1.52E-02 |
| ENSG00000224126 | *UBE2SP2* | 4.58 | 0.85 | 1.81 | 3.24E-04 | 1.59E-02 |
| ENSG00000259607 | *AC108449.3* | 148.06 | 0.46 | 1.37 | 3.28E-04 | 1.59E-02 |
| ENSG00000231128 | *AL137856.1* | 10.38 | -0.68 | 0.62 | 3.38E-04 | 1.60E-02 |
| ENSG00000285802 | *AL450043.1* | 3.83 | -0.99 | 0.50 | 3.37E-04 | 1.60E-02 |
| ENSG00000231324 | *AP000696.1* | 3.16 | -1.28 | 0.41 | 3.37E-04 | 1.60E-02 |
| ENSG00000277142 | *LINC00235* | 13.89 | -0.77 | 0.59 | 3.44E-04 | 1.61E-02 |
| ENSG00000285686 | *AL353613.2* | 3.72 | 1.22 | 2.34 | 3.65E-04 | 1.69E-02 |
| ENSG00000272501 | *AL662844.4* | 27.72 | -0.42 | 0.75 | 3.72E-04 | 1.70E-02 |
| ENSG00000278058 | *AC009159.3* | 6.36 | -0.66 | 0.64 | 3.77E-04 | 1.70E-02 |
| ENSG00000259891 | *AC107375.1* | 13.49 | -0.76 | 0.59 | 3.76E-04 | 1.70E-02 |
| ENSG00000225492 | *GBP1P1* | 4.68 | -1.23 | 0.43 | 3.82E-04 | 1.71E-02 |
| ENSG00000260391 | *AC022336.2* | 8.21 | 0.73 | 1.66 | 4.14E-04 | 1.83E-02 |
| ENSG00000260496 | *AC009041.1* | 8.79 | 0.73 | 1.66 | 4.26E-04 | 1.87E-02 |
| ENSG00000233695 | *GAS6-AS1* | 205.80 | -0.48 | 0.72 | 4.49E-04 | 1.96E-02 |
| ENSG00000249835 | *VCAN-AS1* | 102.26 | 0.96 | 1.94 | 4.71E-04 | 2.03E-02 |
| ENSG00000236213 | *AC006369.1* | 12.15 | -0.78 | 0.58 | 4.89E-04 | 2.09E-02 |
| ENSG00000263072 | *ZNF213-AS1* | 64.76 | -0.33 | 0.79 | 4.94E-04 | 2.10E-02 |
| ENSG00000227619 | *AL391056.1* | 6.94 | 1.37 | 2.59 | 5.07E-04 | 2.11E-02 |
| ENSG00000272746 | *AP005131.7* | 5.60 | 0.94 | 1.92 | 5.11E-04 | 2.11E-02 |
| ENSG00000236173 | *AL049612.1* | 3.97 | 0.90 | 1.86 | 5.13E-04 | 2.11E-02 |
| ENSG00000244041 | *LINC01011* | 17.42 | -0.62 | 0.65 | 5.09E-04 | 2.11E-02 |
| ENSG00000230498 | *AL035409.1* | 5.15 | 1.20 | 2.30 | 5.36E-04 | 2.18E-02 |
| ENSG00000285650 | *AL157827.2* | 8.88 | 0.78 | 1.72 | 5.49E-04 | 2.22E-02 |
| ENSG00000255121 | *AP003392.4* | 5.91 | -0.73 | 0.60 | 5.70E-04 | 2.28E-02 |
| ENSG00000225913 | *AL138767.3* | 5.09 | 1.02 | 2.03 | 5.77E-04 | 2.30E-02 |
| ENSG00000259943 | *AL050341.2* | 25.56 | -0.46 | 0.73 | 5.87E-04 | 2.32E-02 |
| ENSG00000235790 | *AC114488.2* | 104.36 | -0.61 | 0.65 | 6.49E-04 | 2.54E-02 |
| ENSG00000260433 | *LINC01917* | 3.73 | -1.62 | 0.32 | 6.54E-04 | 2.54E-02 |
| ENSG00000246022 | *ALDH1L1-AS2* | 3.36 | -1.04 | 0.49 | 6.76E-04 | 2.61E-02 |
| ENSG00000233785 | *AC131011.1* | 3.33 | 1.09 | 2.13 | 6.84E-04 | 2.62E-02 |
| ENSG00000230091 | *TMEM254-AS1* | 11.23 | -0.44 | 0.73 | 6.91E-04 | 2.62E-02 |
| ENSG00000261553 | *AL137782.1* | 70.69 | 0.41 | 1.33 | 7.02E-04 | 2.65E-02 |
| ENSG00000273038 | *AL365203.2* | 18.02 | 0.48 | 1.40 | 7.21E-04 | 2.70E-02 |
| ENSG00000227766 | *AL671277.1* | 385.83 | -0.44 | 0.74 | 7.30E-04 | 2.71E-02 |
| ENSG00000254409 | *AC087521.3* | 138.82 | 0.57 | 1.48 | 7.44E-04 | 2.75E-02 |
| ENSG00000226696 | *LENG8-AS1* | 17.80 | -0.60 | 0.66 | 7.53E-04 | 2.76E-02 |
| ENSG00000274925 | *ZKSCAN2-DT* | 12.47 | -0.71 | 0.61 | 7.92E-04 | 2.88E-02 |
| ENSG00000227053 | *AC105446.1* | 7.36 | 0.86 | 1.81 | 8.36E-04 | 2.97E-02 |
| ENSG00000267809 | *NDUFV2P1* | 55.15 | 0.42 | 1.34 | 8.34E-04 | 2.97E-02 |
| ENSG00000186056 | *MATN1-AS1* | 18.48 | -0.53 | 0.69 | 8.36E-04 | 2.97E-02 |
| ENSG00000267573 | *KRT8P5* | 8.50 | -1.16 | 0.45 | 8.42E-04 | 2.97E-02 |
| ENSG00000237424 | *FOXD2-AS1* | 38.23 | -0.37 | 0.77 | 8.70E-04 | 3.05E-02 |
| ENSG00000213383 | *AC104297.1* | 10.03 | 0.49 | 1.41 | 8.90E-04 | 3.08E-02 |
| ENSG00000284707 | *AC079781.5* | 86.37 | 0.38 | 1.30 | 8.88E-04 | 3.08E-02 |
| ENSG00000217648 | *AL136116.3* | 12.62 | 0.41 | 1.33 | 9.23E-04 | 3.14E-02 |
| ENSG00000271614 | *ATP2B1-AS1* | 18.88 | -0.44 | 0.74 | 9.27E-04 | 3.14E-02 |
| ENSG00000264514 | *AP000915.1* | 32.45 | -0.60 | 0.66 | 9.17E-04 | 3.14E-02 |

| GeneID | GeneName | baseMean | log2FoldChange | FoldChange | Pvalue | Padj |
|---|---|---|---|---|---|---|
| ENSG00000226380 | *LINC-PINT* | 23.89 | 0.77 | 1.70 | 9.83E-04 | 3.24E-02 |
| ENSG00000267321 | *LINC02001* | 26.49 | -0.40 | 0.76 | 9.82E-04 | 3.24E-02 |
| ENSG00000274828 | *AC068473.5* | 20.73 | -0.44 | 0.74 | 9.73E-04 | 3.24E-02 |
| ENSG00000260236 | *AC099778.1* | 10.41 | -0.52 | 0.70 | 9.88E-04 | 3.24E-02 |
| ENSG00000206344 | *HCG27* | 10.90 | -0.63 | 0.64 | 9.87E-04 | 3.24E-02 |
| ENSG00000255390 | *AC091564.4* | 185.46 | 0.39 | 1.31 | 9.95E-04 | 3.25E-02 |
| ENSG00000240032 | *LNCSR1R* | 19.66 | 0.65 | 1.57 | 1.04E-03 | 3.36E-02 |
| ENSG00000230695 | *AC012462.3* | 2.74 | 1.21 | 2.31 | 1.05E-03 | 3.39E-02 |
| ENSG00000204054 | *LINC00963* | 440.69 | -0.31 | 0.81 | 1.07E-03 | 3.41E-02 |
| ENSG00000225472 | *AL136366.1* | 3.34 | -0.90 | 0.54 | 1.07E-03 | 3.41E-02 |
| ENSG00000250697 | *AC010343.3* | 3.01 | 1.43 | 2.69 | 1.10E-03 | 3.44E-02 |
| ENSG00000272037 | *AP002907.1* | 19.67 | -0.42 | 0.75 | 1.09E-03 | 3.44E-02 |
| ENSG00000284048 | *AC073111.4* | 29.55 | -0.36 | 0.78 | 1.11E-03 | 3.48E-02 |
| ENSG00000254756 | *AP001107.6* | 8.37 | 0.68 | 1.60 | 1.13E-03 | 3.49E-02 |
| ENSG00000243431 | *RPL5P30* | 3.88 | -0.80 | 0.58 | 1.13E-03 | 3.49E-02 |
| ENSG00000235257 | *ITGA9-AS1* | 35.98 | -0.28 | 0.83 | 1.14E-03 | 3.50E-02 |
| ENSG00000226608 | *FTLP3* | 631.57 | -0.42 | 0.75 | 1.16E-03 | 3.52E-02 |
| ENSG00000272970 | *AC107294.2* | 12.07 | -0.54 | 0.69 | 1.17E-03 | 3.52E-02 |
| ENSG00000224769 | *MUC20P1* | 8.28 | -0.79 | 0.58 | 1.17E-03 | 3.52E-02 |
| ENSG00000275120 | *AC048382.5* | 30.50 | -0.39 | 0.76 | 1.20E-03 | 3.60E-02 |
| ENSG00000272234 | *AC008945.1* | 26.19 | -0.95 | 0.52 | 1.24E-03 | 3.68E-02 |
| ENSG00000277449 | *CEBPB-AS1* | 27.96 | -0.43 | 0.74 | 1.33E-03 | 3.92E-02 |
| ENSG00000220517 | *ASS1P1* | 30.67 | -0.48 | 0.72 | 1.33E-03 | 3.92E-02 |
| ENSG00000282121 | *AL592430.2* | 10.99 | 1.01 | 2.01 | 1.37E-03 | 4.00E-02 |
| ENSG00000227258 | *SMIM2-AS1* | 6.85 | -0.55 | 0.68 | 1.38E-03 | 4.02E-02 |
| ENSG00000215386 | *MIR99AHG* | 145.01 | -0.40 | 0.76 | 1.40E-03 | 4.05E-02 |
| ENSG00000275549 | *STPG3-AS1* | 11.64 | -0.63 | 0.64 | 1.45E-03 | 4.17E-02 |
| ENSG00000228106 | *AL392172.1* | 23.57 | 0.49 | 1.41 | 1.46E-03 | 4.18E-02 |
| ENSG00000226415 | *TPI1P1* | 378.41 | 0.31 | 1.24 | 1.53E-03 | 4.34E-02 |
| ENSG00000261051 | *AC107021.2* | 40.06 | 0.58 | 1.49 | 1.59E-03 | 4.46E-02 |
| ENSG00000226711 | *FAM66C* | 29.68 | -0.31 | 0.80 | 1.58E-03 | 4.46E-02 |
| ENSG00000258908 | *AL355075.3* | 48.27 | 0.76 | 1.69 | 1.61E-03 | 4.50E-02 |
| ENSG00000227199 | *ST7-AS1* | 15.17 | -0.41 | 0.75 | 1.67E-03 | 4.61E-02 |
| ENSG00000277283 | *AC004812.2* | 9.49 | -0.50 | 0.71 | 1.68E-03 | 4.61E-02 |
| ENSG00000271894 | *AC007744.1* | 104.72 | -0.51 | 0.70 | 1.67E-03 | 4.61E-02 |
| ENSG00000251548 | *AC106760.2* | 9.07 | -0.61 | 0.65 | 1.71E-03 | 4.67E-02 |
| ENSG00000215241 | *LINC02449* | 17.55 | -0.40 | 0.76 | 1.73E-03 | 4.71E-02 |
| ENSG00000226812 | *AL117382.1* | 8.40 | 1.81 | 3.51 | 1.81E-03 | 4.88E-02 |
| ENSG00000237276 | *ANO7L1* | 7.07 | 0.54 | 1.46 | 1.82E-03 | 4.88E-02 |
| ENSG00000176998 | *HCG4* | 32.10 | -0.46 | 0.73 | 1.82E-03 | 4.88E-02 |
| ENSG00000228748 | *AL450306.1* | 53.91 | 0.28 | 1.22 | 1.84E-03 | 4.90E-02 |
| ENSG00000224594 | *RPL29P19* | 40.38 | -0.46 | 0.73 | 1.88E-03 | 4.96E-02 |

baseMean = mean of normalized counts of all samples normalized for transcript length and sequencing depth, FoldChange = fold change between lesioned and preserved OA cartilage samples, Pvalue = nominal P value, Padj = P value according to false discovery rate.

**Supplementary Table 4 | Results of the validation and replication of differentially expressed long noncoding RNAs between lesioned osteoarthritic cartilage and preserved osteoarthritic cartilage, paired t-test was performed on the -ΔCt values in the validation and replication analyses.**

| | Discovery | | Validation | | Replication | |
|---|---|---|---|---|---|---|
| **LncRNA** | **FC** | **FDR** | **FC** | **P value** | **FC** | **P value** |
| *AC025370.1* | 2.02 | 3.50E-03 | 2.33 | 2.16E-02 | 19.8 | 1.70E-02 |
| *AC090877.2* | 0.33 | 6.22E-05 | 0.28 | 1.82E-05 | 0.07 | 3.56E-04 |
| *MEG3* | 0.63 | 8.81E-03 | 0.53 | 6.16E-03 | 0.77 | 6.02E-02 |
| *P3H2-AS1* | 2.70 | 4.12E-04 | 2.68 | 6.58E-03 | 1.19 | 9.31E-01 |
| *TBILA* | 3.48 | 6.22E-05 | 2.50 | 9.00E-02 | 3.54 | 2.22E-03 |

FC = fold change, FDR = false discovery rate

**Supplementary Table 5 | Significant (according to false discovery rate) differentially expressed lncRNAs in lesioned versus preserved OA cartilage in knee (A) and hip (B) samples.**

**A** **Knee**

| GeneID | GeneName | baseMean | log2FoldChange | FoldChange | Pval | Padj |
|---|---|---|---|---|---|---|
| ENSG00000230615 | *AL139220.2* | 38.13 | 1.17 | 2.25 | 1.96E-10 | 9.66E-07 |
| ENSG00000232044 | *LINC01105* | 82.83 | 1.24 | 2.37 | 7.43E-10 | 1.83E-06 |
| ENSG00000152931 | *PART1* | 249.85 | 1.08 | 2.11 | 1.62E-09 | 2.67E-06 |
| ENSG00000249306 | *LINC01411* | 37.45 | 2.53 | 5.78 | 4.91E-09 | 6.07E-06 |
| ENSG00000235770 | *LINC00607* | 47.65 | 1.32 | 2.50 | 7.28E-09 | 7.19E-06 |
| ENSG00000229896 | *AL157373.2* | 45.42 | 1.13 | 2.19 | 1.16E-07 | 9.55E-05 |
| ENSG00000261488 | *TBILA* | 4.66 | 1.96 | 3.90 | 3.41E-07 | 2.41E-04 |
| ENSG00000258520 | *AL359317.1* | 20.24 | 1.44 | 2.71 | 1.03E-06 | 6.34E-04 |
| ENSG00000226031 | *FGF13-AS1* | 3.43 | -1.45 | 0.37 | 1.22E-06 | 6.68E-04 |
| ENSG00000258583 | *LINC01500* | 23.21 | -1.20 | 0.44 | 3.14E-06 | 1.55E-03 |
| ENSG00000225511 | *LINC00475* | 24.26 | 0.77 | 1.71 | 4.07E-06 | 1.83E-03 |
| ENSG00000225194 | *LINC00092* | 16.83 | -1.36 | 0.39 | 5.09E-06 | 2.02E-03 |
| ENSG00000230838 | *LINC01614* | 19.89 | 1.83 | 3.56 | 5.31E-06 | 2.02E-03 |
| ENSG00000226453 | *LINC02542* | 9.79 | 0.76 | 1.69 | 7.04E-06 | 2.48E-03 |
| ENSG00000220785 | *MTMR9LP* | 62.84 | -0.57 | 0.67 | 9.45E-06 | 3.11E-03 |
| ENSG00000267809 | *NDUFV2P1* | 51.40 | 0.51 | 1.42 | 1.09E-05 | 3.35E-03 |
| ENSG00000188242 | *PP7080* | 82.86 | -0.46 | 0.73 | 1.41E-05 | 3.86E-03 |
| ENSG00000258137 | *AC079313.2* | 4.86 | 1.16 | 2.24 | 1.40E-05 | 3.86E-03 |
| ENSG00000269609 | *RPARP-AS1* | 26.79 | -0.70 | 0.62 | 2.20E-05 | 5.72E-03 |
| ENSG00000232677 | *LINC00665* | 14.09 | -0.81 | 0.57 | 3.08E-05 | 6.67E-03 |
| ENSG00000255471 | *AP001528.2* | 9.24 | -0.99 | 0.50 | 2.71E-05 | 6.67E-03 |
| ENSG00000261051 | *AC107021.2* | 35.12 | 0.69 | 1.61 | 3.01E-05 | 6.67E-03 |
| ENSG00000283029 | *AL139099.4* | 37199.75 | 1.07 | 2.10 | 3.11E-05 | 6.67E-03 |
| ENSG00000131797 | *CLUHP3* | 21.87 | -0.78 | 0.58 | 4.55E-05 | 8.32E-03 |
| ENSG00000254238 | *AC100782.1* | 207.35 | -1.45 | 0.36 | 4.05E-05 | 8.32E-03 |
| ENSG00000266968 | *AC023421.1* | 4.58 | -1.67 | 0.32 | 4.53E-05 | 8.32E-03 |
| ENSG00000278419 | *AL451164.3* | 240.86 | 0.63 | 1.55 | 4.31E-05 | 8.32E-03 |
| ENSG00000255343 | *AC234917.1* | 15.44 | 0.62 | 1.53 | 6.14E-05 | 1.05E-02 |
| ENSG00000259721 | *AC090877.2* | 24.12 | -1.30 | 0.41 | 6.11E-05 | 1.05E-02 |
| ENSG00000223814 | *AL691459.1* | 22.59 | -1.00 | 0.50 | 6.71E-05 | 1.10E-02 |
| ENSG00000236094 | *LINC00545* | 9.01 | 1.22 | 2.34 | 7.30E-05 | 1.16E-02 |
| ENSG00000280007 | *AC008079.1* | 19.30 | -0.80 | 0.57 | 8.21E-05 | 1.27E-02 |
| ENSG00000267100 | *ILF3-DT* | 146.81 | -0.51 | 0.70 | 9.08E-05 | 1.28E-02 |
| ENSG00000267414 | *AC120049.1* | 10.19 | -0.77 | 0.59 | 8.84E-05 | 1.28E-02 |
| ENSG00000276791 | *AC092117.1* | 22.31 | -0.96 | 0.51 | 8.96E-05 | 1.28E-02 |
| ENSG00000261553 | *AL137782.1* | 65.06 | 0.51 | 1.42 | 1.01E-04 | 1.39E-02 |
| ENSG00000229214 | *LINC00242* | 18.82 | 0.48 | 1.39 | 1.07E-04 | 1.43E-02 |
| ENSG00000234380 | *LINC01426* | 10.59 | 0.75 | 1.68 | 1.11E-04 | 1.44E-02 |
| ENSG00000251661 | *AC136475.1* | 11.21 | -0.75 | 0.59 | 1.17E-04 | 1.48E-02 |
| ENSG00000224287 | *MSL3P1* | 12.55 | 0.59 | 1.50 | 1.27E-04 | 1.49E-02 |
| ENSG00000226415 | *TPI1P1* | 323.68 | 0.39 | 1.31 | 1.29E-04 | 1.49E-02 |
| ENSG00000260244 | *AC104083.1* | 16.76 | -1.06 | 0.48 | 1.28E-04 | 1.49E-02 |
| ENSG00000267248 | *AC025048.2* | 3.64 | -0.99 | 0.50 | 1.28E-04 | 1.49E-02 |
| ENSG00000228113 | *AC003991.1* | 78.55 | -1.15 | 0.45 | 1.38E-04 | 1.55E-02 |
| ENSG00000225138 | *SLC9A3-AS1* | 58.20 | -0.67 | 0.63 | 1.41E-04 | 1.55E-02 |
| ENSG00000224729 | *PCOLCE-AS1* | 560.41 | -0.61 | 0.66 | 1.58E-04 | 1.66E-02 |
| ENSG00000233251 | *AC007743.1* | 100.74 | -1.04 | 0.48 | 1.55E-04 | 1.66E-02 |
| ENSG00000225913 | *AL138767.3* | 4.68 | 1.22 | 2.33 | 1.70E-04 | 1.71E-02 |
| ENSG00000237424 | *FOXD2-AS1* | 39.92 | -0.46 | 0.73 | 1.68E-04 | 1.71E-02 |
| ENSG00000255153 | *TOLLIP-AS1* | 8.61 | -0.77 | 0.59 | 1.73E-04 | 1.71E-02 |
| ENSG00000257337 | *AC068888.1* | 279.17 | -0.45 | 0.73 | 1.91E-04 | 1.81E-02 |
| ENSG00000271239 | *AC007423.1* | 4.29 | -1.59 | 0.33 | 1.91E-04 | 1.81E-02 |
| ENSG00000271474 | *AC106881.1* | 9.40 | -1.00 | 0.50 | 1.94E-04 | 1.81E-02 |
| ENSG00000182366 | *FAM87A* | 2.90 | 1.56 | 2.95 | 2.09E-04 | 1.84E-02 |
| ENSG00000237238 | *BMS1P10* | 5.86 | -0.99 | 0.50 | 2.08E-04 | 1.84E-02 |
| ENSG00000269416 | *LINC01224* | 5.06 | -1.08 | 0.47 | 2.03E-04 | 1.84E-02 |

**Knee**

| GeneID | GeneName | baseMean | log2FoldChange | FoldChange | Pval | Padj |
|---|---|---|---|---|---|---|
| ENSG00000183935 | *HTR7P1* | 32.22 | -0.46 | 0.73 | 2.28E-04 | 1.97E-02 |
| ENSG00000226816 | *AC005082.1* | 4.18 | 0.91 | 1.87 | 2.33E-04 | 1.99E-02 |
| ENSG00000182397 | *DNM1P46* | 8.81 | -0.76 | 0.59 | 2.46E-04 | 2.03E-02 |
| ENSG00000251095 | *AC097478.1* | 3.60 | -1.46 | 0.36 | 2.44E-04 | 2.03E-02 |
| ENSG00000184669 | *OR7E14P* | 2.75 | 1.46 | 2.76 | 2.56E-04 | 2.06E-02 |
| ENSG00000278869 | *BX539320.1* | 6.19 | -1.01 | 0.50 | 2.59E-04 | 2.06E-02 |
| ENSG00000235423 | *AC068768.1* | 628.38 | 0.64 | 1.56 | 2.94E-04 | 2.31E-02 |
| ENSG00000256995 | *AC084816.1* | 68.75 | -0.89 | 0.54 | 3.06E-04 | 2.36E-02 |
| ENSG00000272970 | *AC107294.2* | 12.86 | -0.68 | 0.62 | 3.18E-04 | 2.42E-02 |
| ENSG00000237525 | *AC012668.3* | 5.45 | 1.24 | 2.36 | 4.15E-04 | 2.98E-02 |
| ENSG00000250218 | *ALDH1L1-AS1* | 10.22 | -0.79 | 0.58 | 4.17E-04 | 2.98E-02 |
| ENSG00000261959 | *AC015909.3* | 31.90 | -0.85 | 0.55 | 4.11E-04 | 2.98E-02 |
| ENSG00000270412 | *AL136084.3* | 10.68 | 0.96 | 1.94 | 4.03E-04 | 2.98E-02 |
| ENSG00000267573 | *KRT8P5* | 9.01 | -1.39 | 0.38 | 4.56E-04 | 3.21E-02 |
| ENSG00000267272 | *LINC01140* | 119.60 | 0.46 | 1.37 | 4.79E-04 | 3.33E-02 |
| ENSG00000231394 | *AC099681.2* | 5.74 | -0.98 | 0.51 | 5.25E-04 | 3.50E-02 |
| ENSG00000260822 | *AC004656.1* | 425.41 | 0.52 | 1.43 | 5.21E-04 | 3.50E-02 |
| ENSG00000273038 | *AL365203.2* | 15.66 | 0.64 | 1.56 | 5.13E-04 | 3.50E-02 |
| ENSG00000227258 | *SMIM2-AS1* | 7.05 | -0.65 | 0.64 | 5.39E-04 | 3.55E-02 |
| ENSG00000230487 | *PSMG3-AS1* | 59.40 | -0.47 | 0.72 | 6.02E-04 | 3.91E-02 |
| ENSG00000230498 | *AL035409.1* | 4.83 | 1.35 | 2.56 | 6.30E-04 | 4.04E-02 |
| ENSG00000232368 | *FTLP2* | 10.92 | -0.63 | 0.65 | 6.39E-04 | 4.04E-02 |
| ENSG00000234840 | *LINC01239* | 11.21 | -0.75 | 0.59 | 6.46E-04 | 4.04E-02 |
| ENSG00000240032 | *LNCSRLR* | 18.09 | 0.68 | 1.60 | 6.56E-04 | 4.05E-02 |
| ENSG00000225472 | *AL136366.1* | 3.63 | -1.04 | 0.49 | 6.70E-04 | 4.08E-02 |
| ENSG00000225492 | *GBP1P1* | 5.56 | -1.22 | 0.43 | 6.97E-04 | 4.19E-02 |
| ENSG00000256210 | *AC005255.1* | 34.11 | 0.68 | 1.60 | 7.26E-04 | 4.32E-02 |
| ENSG00000231690 | *LINC00574* | 24.37 | 0.54 | 1.46 | 7.52E-04 | 4.37E-02 |
| ENSG00000231711 | *LINC00899* | 28.78 | -0.41 | 0.75 | 7.50E-04 | 4.37E-02 |
| ENSG00000221990 | *EXOC3-AS1* | 16.33 | -0.48 | 0.72 | 7.69E-04 | 4.41E-02 |
| ENSG00000171889 | *MIR31HG* | 10.43 | 0.91 | 1.88 | 7.87E-04 | 4.47E-02 |
| ENSG00000282057 | *AC092807.3* | 36.31 | 0.41 | 1.32 | 8.41E-04 | 4.72E-02 |
| ENSG00000227855 | *DPY19L2P3* | 8.18 | -0.81 | 0.57 | 8.79E-04 | 4.88E-02 |
| ENSG00000260139 | *CSPG4P13* | 40.87 | -0.81 | 0.57 | 9.08E-04 | 4.98E-02 |

**B** **Hip**

| GeneID | GeneName | baseMean | log2FoldChange | FoldChange | Pval | Padj |
|---|---|---|---|---|---|---|
| ENSG00000236094 | *LINC00545* | 52.87 | 2.23 | 4.68 | 7.03E-08 | 2.71E-04 |
| ENSG00000233251 | *AC007743.1* | 24.98 | -1.86 | 0.28 | 2.25E-07 | 2.77E-04 |
| ENSG00000236883 | *AP001615.1* | 6.95 | 4.43 | 21.54 | 3.19E-07 | 2.77E-04 |
| ENSG00000256040 | *PAPPA-AS1* | 115.96 | 3.23 | 9.40 | 3.59E-07 | 2.77E-04 |
| ENSG00000269275 | *AC020922.3* | 11.20 | 3.79 | 13.87 | 2.54E-07 | 2.77E-04 |
| ENSG00000281371 | *INE2* | 136.78 | 1.12 | 2.18 | 9.48E-07 | 6.10E-04 |
| ENSG00000258525 | *AL049830.3* | 128.19 | 2.70 | 6.51 | 1.59E-06 | 8.78E-04 |
| ENSG00000236886 | *AC007563.2* | 63.43 | 1.92 | 3.78 | 2.33E-06 | 1.12E-03 |
| ENSG00000278192 | *AL118505.1* | 17.38 | 3.61 | 12.19 | 2.61E-06 | 1.12E-03 |
| ENSG00000241749 | *RPSAP52* | 36.76 | 2.38 | 5.19 | 4.98E-06 | 1.92E-03 |
| ENSG00000226812 | *AL117382.1* | 18.28 | 3.08 | 8.43 | 7.07E-06 | 2.48E-03 |
| ENSG00000184669 | *OR7E14P* | 13.19 | 3.46 | 10.97 | 9.94E-06 | 3.06E-03 |
| ENSG00000260231 | *KDM7A-DT* | 116.45 | 1.34 | 2.53 | 1.03E-05 | 3.06E-03 |
| ENSG00000223561 | *AC005165.1* | 15.46 | -3.03 | 0.12 | 1.64E-05 | 4.52E-03 |
| ENSG00000256995 | *AC084816.1* | 23.25 | -2.20 | 0.22 | 1.79E-05 | 4.60E-03 |
| ENSG00000259721 | *AC090877.2* | 12.62 | -2.73 | 0.15 | 3.73E-05 | 9.00E-03 |
| ENSG00000237525 | *AC012668.3* | 13.48 | 2.17 | 4.49 | 4.42E-05 | 1.00E-02 |
| ENSG00000230615 | *AL139220.2* | 109.63 | 0.96 | 1.94 | 5.52E-05 | 1.12E-02 |
| ENSG00000254409 | *AC087521.3* | 316.72 | 1.11 | 2.16 | 5.42E-05 | 1.12E-02 |
| ENSG00000272273 | *IER3-AS1* | 580.15 | 1.62 | 3.08 | 7.70E-05 | 1.49E-02 |
| ENSG00000231690 | *LINC00574* | 33.09 | 1.29 | 2.44 | 1.06E-04 | 1.95E-02 |
| ENSG00000224743 | *TEX26-AS1* | 32.18 | 1.92 | 3.78 | 1.17E-04 | 2.06E-02 |
| ENSG00000258908 | *AL355075.3* | 113.22 | 1.34 | 2.54 | 1.65E-04 | 2.77E-02 |
| ENSG00000271894 | *AC007744.1* | 68.60 | -1.14 | 0.45 | 1.85E-04 | 2.97E-02 |
| ENSG00000229896 | *AL157373.2* | 73.33 | 1.37 | 2.59 | 2.45E-04 | 3.78E-02 |
| ENSG00000223855 | *HRAT92* | 7.66 | 2.57 | 5.95 | 2.77E-04 | 4.07E-02 |
| ENSG00000270547 | *LINC01235* | 31.37 | 2.68 | 6.39 | 2.85E-04 | 4.07E-02 |
| ENSG00000234964 | *FABP5P7* | 51.19 | -1.55 | 0.34 | 3.35E-04 | 4.62E-02 |
| ENSG00000228113 | *AC003991.1* | 74.28 | -1.64 | 0.32 | 3.66E-04 | 4.77E-02 |
| ENSG00000232530 | *LIF-AS1* | 4.75 | 3.57 | 11.86 | 3.71E-04 | 4.77E-02 |
| ENSG00000223814 | *AL691459.1* | 13.09 | -1.38 | 0.38 | 3.84E-04 | 4.78E-02 |

baseMean = mean of normalized counts of all samples normalized for transcript length and sequencing depth, FoldChange = fold change between lesioned and preserved OA cartilage samples, Pval = nominal P value, Padj = P value according to false discovery rate.

**Supplementary Table 6 | Significant (according to false discovery rate) differentially expressed long noncoding RNAs in lesioned osteoarthritic cartilage compared to preserved osteoarthritic cartilage unique for knee (A) and hip (B) samples.**

**A** **Knee**

| GeneID | GeneName | baseMean | log2FoldChange | FoldChange | pvalue | padj |
|---|---|---|---|---|---|---|
| ENSG00000224287 | *MSL3P1* | 12.55 | 0.59 | 1.50 | 1.27E-04 | 1.49E-02 |
| ENSG00000182397 | *DNM1P46* | 8.81 | -0.76 | 0.59 | 2.46E-04 | 2.03E-02 |
| ENSG00000251095 | *AC097478.1* | 3.60 | -1.46 | 0.36 | 2.44E-04 | 2.03E-02 |
| ENSG00000235423 | *AC068768.1* | 628.38 | 0.64 | 1.56 | 2.94E-04 | 2.31E-02 |
| ENSG00000250218 | *ALDH1L1-AS1* | 10.22 | -0.79 | 0.58 | 4.17E-04 | 2.98E-02 |
| ENSG00000267272 | *LINC01140* | 119.60 | 0.46 | 1.37 | 4.79E-04 | 3.33E-02 |
| ENSG00000260822 | *AC004656.1* | 425.41 | 0.52 | 1.43 | 5.21E-04 | 3.50E-02 |
| ENSG00000230487 | *PSMG3-AS1* | 59.40 | -0.47 | 0.72 | 6.02E-04 | 3.91E-02 |
| ENSG00000256210 | *AC005255.1* | 34.11 | 0.68 | 1.60 | 7.26E-04 | 4.32E-02 |
| ENSG00000221990 | *EXOC3-AS1* | 16.33 | -0.48 | 0.72 | 7.69E-04 | 4.41E-02 |
| ENSG00000227855 | *DPY19L2P3* | 8.18 | -0.81 | 0.57 | 8.79E-04 | 4.88E-02 |
| ENSG00000260139 | *CSPG4P13* | 40.87 | -0.81 | 0.57 | 9.08E-04 | 4.98E-02 |

**B** **Hip**

| GeneID | GeneName | baseMean | log2FoldChange | FoldChange | pvalue | padj |
|---|---|---|---|---|---|---|
| ENSG00000236883 | *AP001615.1* | 6.95 | 4.43 | 21.54 | 3.19E-07 | 2.77E-04 |
| ENSG00000256040 | *PAPPA-AS1* | 115.96 | 3.23 | 9.40 | 3.59E-07 | 2.77E-04 |
| ENSG00000269275 | *AC020922.3* | 11.20 | 3.79 | 13.87 | 2.54E-07 | 2.77E-04 |
| ENSG00000281371 | *INE2* | 136.78 | 1.12 | 2.18 | 9.48E-07 | 6.10E-04 |
| ENSG00000258525 | *AL049830.3* | 128.19 | 2.70 | 6.51 | 1.59E-06 | 8.78E-04 |
| ENSG00000236886 | *AC007563.2* | 63.43 | 1.92 | 3.78 | 2.33E-06 | 1.12E-03 |
| ENSG00000278192 | *AL118505.1* | 17.38 | 3.61 | 12.19 | 2.61E-06 | 1.12E-03 |
| ENSG00000260231 | *KDM7A-DT* | 116.45 | 1.34 | 2.53 | 1.03E-05 | 3.06E-03 |
| ENSG00000224743 | *TEX26-AS1* | 32.18 | 1.92 | 3.78 | 1.17E-04 | 2.06E-02 |
| ENSG00000223855 | *HRAT92* | 7.66 | 2.57 | 5.95 | 2.77E-04 | 4.07E-02 |
| ENSG00000270547 | *LINC01235* | 31.37 | 2.68 | 6.39 | 2.85E-04 | 4.07E-02 |
| ENSG00000234964 | *FABP5P7* | 51.19 | -1.55 | 0.34 | 3.35E-04 | 4.62E-02 |
| ENSG00000232530 | *LIF-AS1* | 4.75 | 3.57 | 11.86 | 3.71E-04 | 4.77E-02 |

baseMean = mean of normalized counts of all samples normalized for transcript length and sequencing depth, FoldChange = fold change between lesioned and preserved OA cartilage samples, pvalue = nominal P value, Padj = P value according to false discovery rate.

**Supplementary Table 7 | Nominal significant correlations > |0.8| between significant differentially expressed long noncoding RNAs and significant (according to false discovery rate) differentially expressed protein-coding genes in the same osteoarthritis cartilage samples.**

| lncRNA ID | lncRNA name | Biotype | mRNA ID | mRNA name | Cor | Pval |
|---|---|---|---|---|---|---|
| ENSG00000171889 | *MIR31HG* | sense | ENSG00000105810 | *CDK6* | 0.808 | 2.22E-16 |
| ENSG00000171889 | *MIR31HG* | sense | ENSG00000204525 | *HLA-C* | -0.808 | 2.22E-16 |
| ENSG00000205105 | *COX17P1* | pseudogene | ENSG00000158042 | *MRPL17* | 0.806 | 2.22E-16 |
| ENSG00000205105 | *COX17P1* | pseudogene | ENSG00000196072 | *BLOC1S2* | 0.830 | 2.22E-16 |
| ENSG00000205105 | *COX17P1* | pseudogene | ENSG00000117410 | *ATP6V0B* | 0.807 | 2.22E-16 |
| ENSG00000205105 | *COX17P1* | pseudogene | ENSG00000124126 | *PREX1* | -0.815 | 2.22E-16 |
| ENSG00000205105 | *COX17P1* | pseudogene | ENSG00000175324 | *LSM1* | 0.807 | 2.22E-16 |
| ENSG00000205105 | *COX17P1* | pseudogene | ENSG00000182220 | *ATP6AP2* | 0.812 | 2.22E-16 |
| ENSG00000205105 | *COX17P1* | pseudogene | ENSG00000101191 | *DIDO1* | -0.803 | 2.22E-16 |
| ENSG00000205105 | *COX17P1* | pseudogene | ENSG00000117450 | *PRDX1* | 0.814 | 2.22E-16 |
| ENSG00000205105 | *COX17P1* | pseudogene | ENSG00000132646 | *PCNA* | 0.805 | 2.22E-16 |
| ENSG00000205105 | *COX17P1* | pseudogene | ENSG00000197956 | *S100A6* | 0.813 | 2.22E-16 |
| ENSG00000205105 | *COX17P1* | pseudogene | ENSG00000138495 | *COX17* | 0.852 | 2.22E-16 |
| ENSG00000205105 | *COX17P1* | pseudogene | ENSG00000072042 | *RDH11* | 0.860 | 2.22E-16 |
| ENSG00000205105 | *COX17P1* | pseudogene | ENSG00000136240 | *KDELR2* | 0.801 | 2.22E-16 |
| ENSG00000221857 | *AC020907.2* | other | ENSG00000103202 | *NME4* | 0.819 | 2.22E-16 |
| ENSG00000221857 | *AC020907.2* | other | ENSG00000188559 | *RALGAPA2* | -0.831 | 2.22E-16 |
| ENSG00000221857 | *AC020907.2* | other | ENSG00000160145 | *KALRN* | -0.837 | 2.22E-16 |
| ENSG00000221857 | *AC020907.2* | other | ENSG00000198018 | *ENTPD7* | -0.872 | 2.22E-16 |
| ENSG00000221857 | *AC020907.2* | other | ENSG00000170776 | *AKAP13* | -0.804 | 2.22E-16 |
| ENSG00000221857 | *AC020907.2* | other | ENSG00000204219 | *TCEA3* | 0.820 | 2.22E-16 |
| ENSG00000221857 | *AC020907.2* | other | ENSG00000065057 | *NTHL1* | 0.802 | 2.22E-16 |
| ENSG00000221857 | *AC020907.2* | other | ENSG00000169964 | *TMEM42* | 0.809 | 2.22E-16 |
| ENSG00000221857 | *AC020907.2* | other | ENSG00000182195 | *LDOC1* | 0.833 | 2.22E-16 |
| ENSG00000221857 | *AC020907.2* | other | ENSG00000166681 | *BEX3* | 0.826 | 2.22E-16 |
| ENSG00000221857 | *AC020907.2* | other | ENSG00000104964 | *AES* | 0.801 | 2.22E-16 |
| ENSG00000221857 | *AC020907.2* | other | ENSG00000006459 | *KDM7A* | -0.850 | 2.22E-16 |
| ENSG00000221857 | *AC020907.2* | other | ENSG00000198742 | *SMURF1* | -0.831 | 2.22E-16 |
| ENSG00000223814 | *AL691459.1* | intergenic | ENSG00000269113 | *TRABD2B* | 0.826 | 2.22E-16 |
| ENSG00000224594 | *RPL29P19* | pseudogene | ENSG00000164776 | *PHKG1* | 0.840 | 2.22E-16 |
| ENSG00000224594 | *RPL29P19* | pseudogene | ENSG00000166681 | *BEX3* | 0.804 | 2.22E-16 |
| ENSG00000224594 | *RPL29P19* | pseudogene | ENSG00000137070 | *IL11RA* | 0.806 | 2.22E-16 |
| ENSG00000224594 | *RPL29P19* | pseudogene | ENSG00000161010 | *MRNIP* | 0.855 | 2.22E-16 |
| ENSG00000224594 | *RPL29P19* | pseudogene | ENSG00000170776 | *AKAP13* | -0.810 | 2.22E-16 |
| ENSG00000224594 | *RPL29P19* | pseudogene | ENSG00000087842 | *PIR* | 0.809 | 2.22E-16 |
| ENSG00000224594 | *RPL29P19* | pseudogene | ENSG00000013583 | *HEBP1* | 0.810 | 2.22E-16 |
| ENSG00000224594 | *RPL29P19* | pseudogene | ENSG00000178802 | *MPI* | 0.886 | 2.22E-16 |
| ENSG00000224594 | *RPL29P19* | pseudogene | ENSG00000161249 | *DMKN* | 0.849 | 2.22E-16 |
| ENSG00000224594 | *RPL29P19* | pseudogene | ENSG00000182195 | *LDOC1* | 0.829 | 2.22E-16 |
| ENSG00000224594 | *RPL29P19* | pseudogene | ENSG00000198018 | *ENTPD7* | -0.814 | 2.22E-16 |
| ENSG00000224594 | *RPL29P19* | pseudogene | ENSG00000169964 | *TMEM42* | 0.874 | 2.22E-16 |
| ENSG00000224594 | *RPL29P19* | pseudogene | ENSG00000106004 | *HOXA5* | 0.802 | 2.22E-16 |
| ENSG00000224594 | *RPL29P19* | pseudogene | ENSG00000184515 | *BEX5* | 0.837 | 2.22E-16 |
| ENSG00000224594 | *RPL29P19* | pseudogene | ENSG00000175575 | *PAAF1* | 0.805 | 2.22E-16 |
| ENSG00000226415 | *TPI1P1* | pseudogene | ENSG00000197930 | *ERO1A* | 0.814 | 2.22E-16 |
| ENSG00000226608 | *FTLP3* | pseudogene | ENSG00000087086 | *FTL* | 0.855 | 2.22E-16 |
| ENSG00000226608 | *FTLP3* | pseudogene | ENSG00000153944 | *MSI2* | -0.801 | 2.22E-16 |
| ENSG00000227036 | *LINC00511* | other | ENSG00000100596 | *SPTLC2* | 0.821 | 2.22E-16 |
| ENSG00000227053 | *AC105446.1* | antisense | ENSG00000205277 | *MUC12* | 0.808 | 2.22E-16 |
| ENSG00000227053 | *AC105446.1* | antisense | ENSG00000158710 | *TAGLN2* | 0.839 | 2.22E-16 |
| ENSG00000227619 | *AL391056.1* | intergenic | ENSG00000148344 | *PTGES* | 0.804 | 2.22E-16 |
| ENSG00000227766 | *AL671277.1* | pseudogene | ENSG00000169905 | *TOR1AIP2* | -0.813 | 2.22E-16 |
| ENSG00000228113 | *AC003991.1* | antisense | ENSG00000128487 | *SPECC1* | -0.818 | 2.22E-16 |
| ENSG00000228748 | *AL450306.1* | antisense | ENSG00000119185 | *ITGB1BP1* | 0.809 | 2.22E-16 |
| ENSG00000228748 | *AL450306.1* | antisense | ENSG00000137288 | *UQCC2* | 0.873 | 2.22E-16 |

| lncRNA ID | lncRNA name | Biotype | mRNA ID | mRNA name | Cor | Pval |
|---|---|---|---|---|---|---|
| ENSG00000228748 | *AL450306.1* | antisense | ENSG00000114354 | *TFG* | 0.837 | 2.22E-16 |
| ENSG00000228748 | *AL450306.1* | antisense | ENSG00000156873 | *PHKG2* | 0.813 | 2.22E-16 |
| ENSG00000228748 | *AL450306.1* | antisense | ENSG00000203879 | *GDI1* | 0.837 | 2.22E-16 |
| ENSG00000228748 | *AL450306.1* | antisense | ENSG00000105426 | *PTPRS* | -0.806 | 2.22E-16 |
| ENSG00000228748 | *AL450306.1* | antisense | ENSG00000135930 | *EIF4E2* | 0.828 | 2.22E-16 |
| ENSG00000228748 | *AL450306.1* | antisense | ENSG00000114126 | *TFDP2* | -0.830 | 2.22E-16 |
| ENSG00000228748 | *AL450306.1* | antisense | ENSG00000154144 | *TBRG1* | 0.805 | 2.22E-16 |
| ENSG00000228748 | *AL450306.1* | antisense | ENSG00000137996 | *RTCA* | 0.819 | 2.22E-16 |
| ENSG00000228748 | *AL450306.1* | antisense | ENSG00000177700 | *POLR2L* | 0.844 | 2.22E-16 |
| ENSG00000228748 | *AL450306.1* | antisense | ENSG00000101191 | *DIDO1* | -0.810 | 2.22E-16 |
| ENSG00000228748 | *AL450306.1* | antisense | ENSG00000115317 | *HTRA2* | 0.825 | 2.22E-16 |
| ENSG00000228748 | *AL450306.1* | antisense | ENSG00000123353 | *ORMDL2* | 0.846 | 2.22E-16 |
| ENSG00000228748 | *AL450306.1* | antisense | ENSG00000248905 | *FMN1* | -0.835 | 2.22E-16 |
| ENSG00000228748 | *AL450306.1* | antisense | ENSG00000243317 | *STMP1* | 0.833 | 2.22E-16 |
| ENSG00000228748 | *AL450306.1* | antisense | ENSG00000164620 | *RELL2* | 0.837 | 2.22E-16 |
| ENSG00000228748 | *AL450306.1* | antisense | ENSG00000187792 | *ZNF70* | -0.834 | 2.22E-16 |
| ENSG00000228748 | *AL450306.1* | antisense | ENSG00000149716 | *ORAOV1* | 0.892 | 2.22E-16 |
| ENSG00000228748 | *AL450306.1* | antisense | ENSG00000197747 | *S100A10* | 0.825 | 2.22E-16 |
| ENSG00000228748 | *AL450306.1* | antisense | ENSG00000081181 | *ARG2* | 0.851 | 2.22E-16 |
| ENSG00000230615 | *AL139220.2* | intergenic | ENSG00000105810 | *CDK6* | 0.818 | 2.22E-16 |
| ENSG00000230615 | *AL139220.2* | intergenic | ENSG00000163359 | *COL6A3* | 0.806 | 2.22E-16 |
| ENSG00000230615 | *AL139220.2* | intergenic | ENSG00000162998 | *FRZB* | -0.847 | 2.22E-16 |
| ENSG00000230615 | *AL139220.2* | intergenic | ENSG00000165795 | *NDRG2* | -0.856 | 2.22E-16 |
| ENSG00000231128 | *AL137856.1* | antisense | ENSG00000185689 | *C6orf201* | 0.840 | 2.22E-16 |
| ENSG00000231128 | *AL137856.1* | antisense | ENSG00000118971 | *CCND2* | -0.843 | 2.22E-16 |
| ENSG00000231128 | *AL137856.1* | antisense | ENSG00000167107 | *ACSF2* | 0.830 | 2.22E-16 |
| ENSG00000231128 | *AL137856.1* | antisense | ENSG00000124635 | *HIST1H2BJ* | -0.821 | 2.22E-16 |
| ENSG00000231128 | *AL137856.1* | antisense | ENSG00000169905 | *TOR1AIP2* | -0.835 | 2.22E-16 |
| ENSG00000231128 | *AL137856.1* | antisense | ENSG00000124615 | *MOCS1* | 0.808 | 2.22E-16 |
| ENSG00000232044 | *LINC01105* | other | ENSG00000134198 | *TSPAN2* | 0.841 | 2.22E-16 |
| ENSG00000233695 | *GAS6-AS1* | antisense | ENSG00000118971 | *CCND2* | -0.802 | 2.22E-16 |
| ENSG00000233695 | *GAS6-AS1* | antisense | ENSG00000169905 | *TOR1AIP2* | -0.813 | 2.22E-16 |
| ENSG00000234380 | *LINC01426* | intergenic | ENSG00000149948 | *HMGA2* | 0.819 | 2.22E-16 |
| ENSG00000234380 | *LINC01426* | intergenic | ENSG00000162512 | *SDC3* | -0.804 | 2.22E-16 |
| ENSG00000234380 | *LINC01426* | intergenic | ENSG00000122756 | *CNTFR* | -0.839 | 2.22E-16 |
| ENSG00000234380 | *LINC01426* | intergenic | ENSG00000144843 | *ADPRH* | 0.848 | 2.22E-16 |
| ENSG00000234380 | *LINC01426* | intergenic | ENSG00000144857 | *BOC* | -0.816 | 2.22E-16 |
| ENSG00000234380 | *LINC01426* | intergenic | ENSG00000198805 | *PNP* | 0.824 | 2.22E-16 |
| ENSG00000234380 | *LINC01426* | intergenic | ENSG00000115762 | *PLEKHB2* | 0.804 | 2.22E-16 |
| ENSG00000234380 | *LINC01426* | intergenic | ENSG00000160111 | *CPAMD8* | -0.808 | 2.22E-16 |
| ENSG00000234380 | *LINC01426* | intergenic | ENSG00000266524 | *GDF10* | -0.835 | 2.22E-16 |
| ENSG00000234380 | *LINC01426* | intergenic | ENSG00000165238 | *WNK2* | -0.813 | 2.22E-16 |
| ENSG00000234380 | *LINC01426* | intergenic | ENSG00000124731 | *TREM1* | 0.819 | 2.22E-16 |
| ENSG00000234380 | *LINC01426* | intergenic | ENSG00000185432 | *METTL7A* | -0.803 | 2.22E-16 |
| ENSG00000235770 | *LINC00607* | intergenic | ENSG00000090530 | *P3H2* | 0.805 | 2.22E-16 |
| ENSG00000235790 | *AC114488.2* | antisense | ENSG00000198865 | *CCDC152* | 0.831 | 2.22E-16 |
| ENSG00000235790 | *AC114488.2* | antisense | ENSG00000147459 | *DOCK5* | -0.801 | 2.22E-16 |
| ENSG00000235790 | *AC114488.2* | antisense | ENSG00000079841 | *RIMS1* | -0.845 | 2.22E-16 |
| ENSG00000235790 | *AC114488.2* | antisense | ENSG00000069188 | *SDK2* | -0.802 | 2.22E-16 |
| ENSG00000236094 | *LINC00545* | intergenic | ENSG00000166033 | *HTRA1* | 0.822 | 2.22E-16 |
| ENSG00000236094 | *LINC00545* | intergenic | ENSG00000196754 | *S100A2* | 0.812 | 2.22E-16 |
| ENSG00000236094 | *LINC00545* | intergenic | ENSG00000134259 | *NGF* | 0.810 | 2.22E-16 |
| ENSG00000237525 | *AC012668.3* | intergenic | ENSG00000105426 | *PTPRS* | -0.804 | 2.22E-16 |
| ENSG00000240032 | *LNCSRLR* | antisense | ENSG00000160145 | *KALRN* | 0.846 | 2.22E-16 |
| ENSG00000240032 | *LNCSRLR* | antisense | ENSG00000156273 | *BACH1* | 0.832 | 2.22E-16 |
| ENSG00000240032 | *LNCSRLR* | antisense | ENSG00000095951 | *HIVEP1* | 0.812 | 2.22E-16 |

| lncRNA ID | lncRNA name | Biotype | mRNA ID | mRNA name | Cor | Pval |
|---|---|---|---|---|---|---|
| ENSG00000240032 | *LNCSRLR* | antisense | ENSG00000104067 | *TJP1* | 0.804 | 2.22E-16 |
| ENSG00000240032 | *LNCSRLR* | antisense | ENSG00000169398 | *PTK2* | 0.842 | 2.22E-16 |
| ENSG00000240032 | *LNCSRLR* | antisense | ENSG00000150995 | *ITPR1* | 0.816 | 2.22E-16 |
| ENSG00000240032 | *LNCSRLR* | antisense | ENSG00000177200 | *CHD9* | 0.806 | 2.22E-16 |
| ENSG00000240032 | *LNCSRLR* | antisense | ENSG00000075420 | *FNDC3B* | 0.852 | 2.22E-16 |
| ENSG00000240032 | *LNCSRLR* | antisense | ENSG00000156011 | *PSD3* | 0.825 | 2.22E-16 |
| ENSG00000240032 | *LNCSRLR* | antisense | ENSG00000170776 | *AKAP13* | 0.802 | 2.22E-16 |
| ENSG00000240032 | *LNCSRLR* | antisense | ENSG00000151914 | *DST* | 0.849 | 2.22E-16 |
| ENSG00000240032 | *LNCSRLR* | antisense | ENSG00000104447 | *TRPS1* | 0.839 | 2.22E-16 |
| ENSG00000240032 | *LNCSRLR* | antisense | ENSG00000182195 | *LDOC1* | -0.817 | 2.22E-16 |
| ENSG00000240032 | *LNCSRLR* | antisense | ENSG00000198018 | *ENTPD7* | 0.869 | 2.22E-16 |
| ENSG00000240032 | *LNCSRLR* | antisense | ENSG00000111647 | *UHRF1BP1L* | 0.800 | 2.22E-16 |
| ENSG00000240032 | *LNCSRLR* | antisense | ENSG00000140526 | *ABHD2* | 0.805 | 2.22E-16 |
| ENSG00000240032 | *LNCSRLR* | antisense | ENSG00000184905 | *TCEAL2* | -0.815 | 2.22E-16 |
| ENSG00000240032 | *LNCSRLR* | antisense | ENSG00000139514 | *SLC7A1* | 0.844 | 2.22E-16 |
| ENSG00000241749 | *RPSAP52* | pseudogene | ENSG00000149948 | *HMGA2* | 0.845 | 2.22E-16 |
| ENSG00000241749 | *RPSAP52* | pseudogene | ENSG00000198805 | *PNP* | 0.832 | 2.22E-16 |
| ENSG00000241749 | *RPSAP52* | pseudogene | ENSG00000011422 | *PLAUR* | 0.844 | 2.22E-16 |
| ENSG00000241749 | *RPSAP52* | pseudogene | ENSG00000006327 | *TNFRSF12A* | 0.819 | 2.22E-16 |
| ENSG00000241749 | *RPSAP52* | pseudogene | ENSG00000122756 | *CNTFR* | -0.813 | 2.22E-16 |
| ENSG00000241749 | *RPSAP52* | pseudogene | ENSG00000124731 | *TREM1* | 0.803 | 2.22E-16 |
| ENSG00000249306 | *LINC01411* | intergenic | ENSG00000122756 | *CNTFR* | -0.805 | 2.22E-16 |
| ENSG00000249306 | *LINC01411* | intergenic | ENSG00000266524 | *GDF10* | -0.804 | 2.22E-16 |
| ENSG00000249306 | *LINC01411* | intergenic | ENSG00000189058 | *APOD* | -0.837 | 2.22E-16 |
| ENSG00000249306 | *LINC01411* | intergenic | ENSG00000171488 | *LRRC8C* | 0.836 | 2.22E-16 |
| ENSG00000249306 | *LINC01411* | intergenic | ENSG00000123610 | *TNFAIP6* | 0.814 | 2.22E-16 |
| ENSG00000249306 | *LINC01411* | intergenic | ENSG00000132718 | *SYT11* | 0.819 | 2.22E-16 |
| ENSG00000249306 | *LINC01411* | intergenic | ENSG00000148344 | *PTGES* | 0.823 | 2.22E-16 |
| ENSG00000249306 | *LINC01411* | intergenic | ENSG00000165238 | *WNK2* | -0.814 | 2.22E-16 |
| ENSG00000249306 | *LINC01411* | intergenic | ENSG00000144908 | *ALDH1L1* | -0.815 | 2.22E-16 |
| ENSG00000249306 | *LINC01411* | intergenic | ENSG00000162512 | *SDC3* | -0.810 | 2.22E-16 |
| ENSG00000249306 | *LINC01411* | intergenic | ENSG00000134259 | *NGF* | 0.836 | 2.22E-16 |
| ENSG00000249306 | *LINC01411* | intergenic | ENSG00000144843 | *ADPRH* | 0.808 | 2.22E-16 |
| ENSG00000249306 | *LINC01411* | intergenic | ENSG00000144857 | *BOC* | -0.804 | 2.22E-16 |
| ENSG00000249835 | *VCAN-AS1* | antisense | ENSG00000149716 | *ORAOV1* | 0.843 | 2.22E-16 |
| ENSG00000249835 | *VCAN-AS1* | antisense | ENSG00000114354 | *TFG* | 0.825 | 2.22E-16 |
| ENSG00000249835 | *VCAN-AS1* | antisense | ENSG00000248905 | *FMN1* | -0.806 | 2.22E-16 |
| ENSG00000249835 | *VCAN-AS1* | antisense | ENSG00000177700 | *POLR2L* | 0.822 | 2.22E-16 |
| ENSG00000249835 | *VCAN-AS1* | antisense | ENSG00000081181 | *ARG2* | 0.823 | 2.22E-16 |
| ENSG00000249835 | *VCAN-AS1* | antisense | ENSG00000137288 | *UQCC2* | 0.827 | 2.22E-16 |
| ENSG00000249835 | *VCAN-AS1* | antisense | ENSG00000123353 | *ORMDL2* | 0.818 | 2.22E-16 |
| ENSG00000249835 | *VCAN-AS1* | antisense | ENSG00000115317 | *HTRA2* | 0.806 | 2.22E-16 |
| ENSG00000249835 | *VCAN-AS1* | antisense | ENSG00000187792 | *ZNF70* | -0.837 | 2.22E-16 |
| ENSG00000249835 | *VCAN-AS1* | antisense | ENSG00000243317 | *STMP1* | 0.824 | 2.22E-16 |
| ENSG00000250318 | *AC003072.1* | pseudogene | ENSG00000158710 | *TAGLN2* | 0.805 | 2.22E-16 |
| ENSG00000253851 | *AC025370.1* | antisense | ENSG00000102362 | *SYTL4* | 0.825 | 2.22E-16 |
| ENSG00000254238 | *AC100782.1* | antisense | ENSG00000128487 | *SPECC1* | -0.810 | 2.22E-16 |
| ENSG00000254238 | *AC100782.1* | antisense | ENSG00000169905 | *TOR1AIP2* | -0.828 | 2.22E-16 |
| ENSG00000254238 | *AC100782.1* | antisense | ENSG00000167107 | *ACSF2* | 0.820 | 2.22E-16 |
| ENSG00000254756 | *AP001107.6* | antisense | ENSG00000162599 | *NFIA* | -0.804 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000135404 | *CD63* | 0.845 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000119185 | *ITGB1BP1* | 0.835 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000114126 | *TFDP2* | -0.831 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000149716 | *ORAOV1* | 0.842 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000203879 | *GDI1* | 0.853 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000182220 | *ATP6AP2* | 0.827 | 2.22E-16 |

| lncRNA ID | lncRNA name | Biotype | mRNA ID | mRNA name | Cor | Pval |
|---|---|---|---|---|---|---|
| ENSG00000255343 | *AC234917.1* | other | ENSG00000114354 | *TFG* | 0.874 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000116586 | *LAMTOR2* | 0.812 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000137996 | *RTCA* | 0.820 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000248905 | *FMN1* | -0.850 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000115307 | *AUP1* | 0.836 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000115317 | *HTRA2* | 0.862 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000055813 | *CCDC85A* | -0.854 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000125868 | *DSTN* | 0.809 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000240972 | *MIF* | 0.811 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000101294 | *HM13* | 0.804 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000105426 | *PTPRS* | -0.841 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000169241 | *SLC50A1* | 0.803 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000101191 | *DIDO1* | -0.852 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000101493 | *ZNF516* | -0.809 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000131871 | *SELENOS* | 0.814 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000196072 | *BLOC1S2* | 0.834 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000187792 | *ZNF70* | -0.824 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000152217 | *SETBP1* | -0.851 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000197747 | *S100A10* | 0.860 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000179930 | *ZNF648* | -0.822 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000068745 | *IP6K2* | 0.838 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000081181 | *ARG2* | 0.879 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000122034 | *GTF3A* | 0.821 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000124126 | *PREX1* | -0.832 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000137288 | *UQCC2* | 0.884 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000102007 | *PLP2* | 0.811 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000156873 | *PHKG2* | 0.855 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000177700 | *POLR2L* | 0.868 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000243317 | *STMP1* | 0.869 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000164292 | *RHOBTB3* | -0.809 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000085662 | *AKR1B1* | 0.828 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000110721 | *CHKA* | 0.821 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000136240 | *KDELR2* | 0.809 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000164620 | *RELL2* | 0.855 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000123353 | *ORMDL2* | 0.884 | 2.22E-16 |
| ENSG00000255343 | *AC234917.1* | other | ENSG00000069849 | *ATP1B3* | 0.824 | 2.22E-16 |
| ENSG00000255390 | *AC091564.4* | antisense | ENSG00000119185 | *ITGB1BP1* | 0.844 | 2.22E-16 |
| ENSG00000255390 | *AC091564.4* | antisense | ENSG00000164032 | *H2AFZ* | 0.801 | 2.22E-16 |
| ENSG00000255390 | *AC091564.4* | antisense | ENSG00000055813 | *CCDC85A* | -0.820 | 2.22E-16 |
| ENSG00000255390 | *AC091564.4* | antisense | ENSG00000101191 | *DIDO1* | -0.846 | 2.22E-16 |
| ENSG00000255390 | *AC091564.4* | antisense | ENSG00000137996 | *RTCA* | 0.838 | 2.22E-16 |
| ENSG00000255390 | *AC091564.4* | antisense | ENSG00000124126 | *PREX1* | -0.808 | 2.22E-16 |
| ENSG00000255390 | *AC091564.4* | antisense | ENSG00000115317 | *HTRA2* | 0.884 | 2.22E-16 |
| ENSG00000255390 | *AC091564.4* | antisense | ENSG00000114126 | *TFDP2* | -0.834 | 2.22E-16 |
| ENSG00000255390 | *AC091564.4* | antisense | ENSG00000203879 | *GDI1* | 0.877 | 2.22E-16 |
| ENSG00000255390 | *AC091564.4* | antisense | ENSG00000135404 | *CD63* | 0.822 | 2.22E-16 |
| ENSG00000255390 | *AC091564.4* | antisense | ENSG00000114354 | *TFG* | 0.860 | 2.22E-16 |
| ENSG00000255390 | *AC091564.4* | antisense | ENSG00000196072 | *BLOC1S2* | 0.826 | 2.22E-16 |
| ENSG00000255390 | *AC091564.4* | antisense | ENSG00000156873 | *PHKG2* | 0.854 | 2.22E-16 |
| ENSG00000255390 | *AC091564.4* | antisense | ENSG00000068745 | *IP6K2* | 0.825 | 2.22E-16 |
| ENSG00000255390 | *AC091564.4* | antisense | ENSG00000187792 | *ZNF70* | -0.832 | 2.22E-16 |
| ENSG00000255390 | *AC091564.4* | antisense | ENSG00000248905 | *FMN1* | -0.865 | 2.22E-16 |
| ENSG00000255390 | *AC091564.4* | antisense | ENSG00000197747 | *S100A10* | 0.843 | 2.22E-16 |
| ENSG00000255390 | *AC091564.4* | antisense | ENSG00000240972 | *MIF* | 0.820 | 2.22E-16 |
| ENSG00000255390 | *AC091564.4* | antisense | ENSG00000152217 | *SETBP1* | -0.833 | 2.22E-16 |
| ENSG00000255390 | *AC091564.4* | antisense | ENSG00000081181 | *ARG2* | 0.890 | 2.22E-16 |

| lncRNA ID | lncRNA name | Biotype | mRNA ID | mRNA name | Cor | Pval |
|---|---|---|---|---|---|---|
| ENSG00000255390 | *AC091564.4* | antisense | ENSG00000164620 | *RELL2* | 0.844 | 2.22E-16 |
| ENSG00000255390 | *AC091564.4* | antisense | ENSG00000243317 | *STMP1* | 0.871 | 2.22E-16 |
| ENSG00000255390 | *AC091564.4* | antisense | ENSG00000177700 | *POLR2L* | 0.884 | 2.22E-16 |
| ENSG00000255390 | *AC091564.4* | antisense | ENSG00000137288 | *UQCC2* | 0.911 | 2.22E-16 |
| ENSG00000255390 | *AC091564.4* | antisense | ENSG00000123353 | *ORMDL2* | 0.877 | 2.22E-16 |
| ENSG00000255390 | *AC091564.4* | antisense | ENSG00000149716 | *ORAOV1* | 0.892 | 2.22E-16 |
| ENSG00000255390 | *AC091564.4* | antisense | ENSG00000105426 | *PTPRS* | -0.835 | 2.22E-16 |
| ENSG00000255390 | *AC091564.4* | antisense | ENSG00000146648 | *EGFR* | -0.805 | 2.22E-16 |
| ENSG00000257337 | *AC068888.1* | antisense | ENSG00000167378 | *IRGQ* | -0.807 | 2.22E-16 |
| ENSG00000257337 | *AC068888.1* | antisense | ENSG00000092964 | *DPYSL2* | -0.863 | 2.22E-16 |
| ENSG00000257337 | *AC068888.1* | antisense | ENSG00000167107 | *ACSF2* | 0.813 | 2.22E-16 |
| ENSG00000258583 | *LINC01500* | intergenic | ENSG00000165617 | *DACT1* | 0.825 | 2.22E-16 |
| ENSG00000258908 | *AL355075.3* | intergenic | ENSG00000203879 | *GDI1* | 0.841 | 2.22E-16 |
| ENSG00000258908 | *AL355075.3* | intergenic | ENSG00000177700 | *POLR2L* | 0.858 | 2.22E-16 |
| ENSG00000258908 | *AL355075.3* | intergenic | ENSG00000119185 | *ITGB1BP1* | 0.808 | 2.22E-16 |
| ENSG00000258908 | *AL355075.3* | intergenic | ENSG00000114126 | *TFDP2* | -0.840 | 2.22E-16 |
| ENSG00000258908 | *AL355075.3* | intergenic | ENSG00000055813 | *CCDC85A* | -0.804 | 2.22E-16 |
| ENSG00000258908 | *AL355075.3* | intergenic | ENSG00000105426 | *PTPRS* | -0.832 | 2.22E-16 |
| ENSG00000258908 | *AL355075.3* | intergenic | ENSG00000114354 | *TFG* | 0.832 | 2.22E-16 |
| ENSG00000258908 | *AL355075.3* | intergenic | ENSG00000248905 | *FMN1* | -0.827 | 2.22E-16 |
| ENSG00000258908 | *AL355075.3* | intergenic | ENSG00000187792 | *ZNF70* | -0.846 | 2.22E-16 |
| ENSG00000258908 | *AL355075.3* | intergenic | ENSG00000137288 | *UQCC2* | 0.890 | 2.22E-16 |
| ENSG00000258908 | *AL355075.3* | intergenic | ENSG00000152217 | *SETBP1* | -0.809 | 2.22E-16 |
| ENSG00000258908 | *AL355075.3* | intergenic | ENSG00000149716 | *ORAOV1* | 0.899 | 2.22E-16 |
| ENSG00000258908 | *AL355075.3* | intergenic | ENSG00000081181 | *ARG2* | 0.863 | 2.22E-16 |
| ENSG00000258908 | *AL355075.3* | intergenic | ENSG00000197747 | *S100A10* | 0.833 | 2.22E-16 |
| ENSG00000258908 | *AL355075.3* | intergenic | ENSG00000156873 | *PHKG2* | 0.806 | 2.22E-16 |
| ENSG00000258908 | *AL355075.3* | intergenic | ENSG00000123353 | *ORMDL2* | 0.864 | 2.22E-16 |
| ENSG00000258908 | *AL355075.3* | intergenic | ENSG00000243317 | *STMP1* | 0.841 | 2.22E-16 |
| ENSG00000258908 | *AL355075.3* | intergenic | ENSG00000196072 | *BLOC1S2* | 0.808 | 2.22E-16 |
| ENSG00000258908 | *AL355075.3* | intergenic | ENSG00000115317 | *HTRA2* | 0.834 | 2.22E-16 |
| ENSG00000258908 | *AL355075.3* | intergenic | ENSG00000164620 | *RELL2* | 0.818 | 2.22E-16 |
| ENSG00000258908 | *AL355075.3* | intergenic | ENSG00000101191 | *DIDO1* | -0.832 | 2.22E-16 |
| ENSG00000258908 | *AL355075.3* | intergenic | ENSG00000068745 | *IP6K2* | 0.807 | 2.22E-16 |
| ENSG00000259607 | *AC108449.3* | antisense | ENSG00000125868 | *DSTN* | 0.860 | 2.22E-16 |
| ENSG00000259607 | *AC108449.3* | antisense | ENSG00000135404 | *CD63* | 0.803 | 2.22E-16 |
| ENSG00000259607 | *AC108449.3* | antisense | ENSG00000177700 | *POLR2L* | 0.829 | 2.22E-16 |
| ENSG00000259607 | *AC108449.3* | antisense | ENSG00000114354 | *TFG* | 0.867 | 2.22E-16 |
| ENSG00000259607 | *AC108449.3* | antisense | ENSG00000137288 | *UQCC2* | 0.833 | 2.22E-16 |
| ENSG00000259607 | *AC108449.3* | antisense | ENSG00000102362 | *SYTL4* | 0.867 | 2.22E-16 |
| ENSG00000259607 | *AC108449.3* | antisense | ENSG00000243317 | *STMP1* | 0.847 | 2.22E-16 |
| ENSG00000259607 | *AC108449.3* | antisense | ENSG00000119185 | *ITGB1BP1* | 0.866 | 2.22E-16 |
| ENSG00000259607 | *AC108449.3* | antisense | ENSG00000152217 | *SETBP1* | -0.835 | 2.22E-16 |
| ENSG00000259607 | *AC108449.3* | antisense | ENSG00000034713 | *GABARAPL2* | 0.801 | 2.22E-16 |
| ENSG00000259607 | *AC108449.3* | antisense | ENSG00000123353 | *ORMDL2* | 0.830 | 2.22E-16 |
| ENSG00000259607 | *AC108449.3* | antisense | ENSG00000131871 | *SELENOS* | 0.811 | 2.22E-16 |
| ENSG00000259607 | *AC108449.3* | antisense | ENSG00000115317 | *HTRA2* | 0.883 | 2.22E-16 |
| ENSG00000259607 | *AC108449.3* | antisense | ENSG00000101191 | *DIDO1* | -0.806 | 2.22E-16 |
| ENSG00000259607 | *AC108449.3* | antisense | ENSG00000081181 | *ARG2* | 0.838 | 2.22E-16 |
| ENSG00000259721 | *AC090877.2* | intergenic | ENSG00000166923 | *GREM1* | 0.873 | 2.22E-16 |
| ENSG00000260563 | *AC132872.1* | intergenic | ENSG00000104447 | *TRPS1* | -0.836 | 2.22E-16 |
| ENSG00000264514 | *AP000915.1* | intergenic | ENSG00000006042 | *TMEM98* | 0.835 | 2.22E-16 |
| ENSG00000264514 | *AP000915.1* | intergenic | ENSG00000198865 | *CCDC152* | 0.867 | 2.22E-16 |
| ENSG00000264514 | *AP000915.1* | intergenic | ENSG00000137070 | *IL11RA* | 0.801 | 2.22E-16 |
| ENSG00000264514 | *AP000915.1* | intergenic | ENSG00000141542 | *RAB40B* | 0.805 | 2.22E-16 |
| ENSG00000264514 | *AP000915.1* | intergenic | ENSG00000166681 | *BEX3* | 0.809 | 2.22E-16 |

| lncRNA ID | lncRNA name | Biotype | mRNA ID | mRNA name | Cor | Pval |
|---|---|---|---|---|---|---|
| ENSG00000264514 | *AP000915.1* | intergenic | ENSG00000013583 | *HEBP1* | 0.810 | 2.22E-16 |
| ENSG00000264514 | *AP000915.1* | intergenic | ENSG00000102466 | *FGF14* | 0.814 | 2.22E-16 |
| ENSG00000264514 | *AP000915.1* | intergenic | ENSG00000069188 | *SDK2* | -0.834 | 2.22E-16 |
| ENSG00000264514 | *AP000915.1* | intergenic | ENSG00000104447 | *TRPS1* | -0.807 | 2.22E-16 |
| ENSG00000264514 | *AP000915.1* | intergenic | ENSG00000161010 | *MRNIP* | 0.854 | 2.22E-16 |
| ENSG00000264514 | *AP000915.1* | intergenic | ENSG00000124615 | *MOCS1* | 0.842 | 2.22E-16 |
| ENSG00000267100 | *ILF3-DT* | intergenic | ENSG00000104447 | *TRPS1* | -0.832 | 2.22E-16 |
| ENSG00000267100 | *ILF3-DT* | intergenic | ENSG00000170776 | *AKAP13* | -0.827 | 2.22E-16 |
| ENSG00000267100 | *ILF3-DT* | intergenic | ENSG00000160145 | *KALRN* | -0.800 | 2.22E-16 |
| ENSG00000267100 | *ILF3-DT* | intergenic | ENSG00000184515 | *BEX5* | 0.846 | 2.22E-16 |
| ENSG00000267100 | *ILF3-DT* | intergenic | ENSG00000169964 | *TMEM42* | 0.809 | 2.22E-16 |
| ENSG00000267100 | *ILF3-DT* | intergenic | ENSG00000182195 | *LDOC1* | 0.843 | 2.22E-16 |
| ENSG00000267100 | *ILF3-DT* | intergenic | ENSG00000166681 | *BEX3* | 0.810 | 2.22E-16 |
| ENSG00000267100 | *ILF3-DT* | intergenic | ENSG00000198018 | *ENTPD7* | -0.820 | 2.22E-16 |
| ENSG00000267100 | *ILF3-DT* | intergenic | ENSG00000178802 | *MPI* | 0.801 | 2.22E-16 |
| ENSG00000267809 | *NDUFV2P1* | pseudogene | ENSG00000164294 | *GPX8* | 0.808 | 2.22E-16 |
| ENSG00000272234 | *AC008945.1* | antisense | ENSG00000169905 | *TOR1AIP2* | -0.812 | 2.22E-16 |
| ENSG00000272234 | *AC008945.1* | antisense | ENSG00000118971 | *CCND2* | -0.836 | 2.22E-16 |
| ENSG00000272273 | *IER3-AS1* | antisense | ENSG00000146648 | *EGFR* | -0.810 | 2.22E-16 |
| ENSG00000272273 | *IER3-AS1* | antisense | ENSG00000101493 | *ZNF516* | -0.808 | 2.22E-16 |
| ENSG00000272273 | *IER3-AS1* | antisense | ENSG00000105426 | *PTPRS* | -0.862 | 2.22E-16 |
| ENSG00000272273 | *IER3-AS1* | antisense | ENSG00000102007 | *PLP2* | 0.809 | 2.22E-16 |
| ENSG00000272273 | *IER3-AS1* | antisense | ENSG00000149716 | *ORAOV1* | 0.873 | 2.22E-16 |
| ENSG00000272273 | *IER3-AS1* | antisense | ENSG00000248905 | *FMN1* | -0.833 | 2.22E-16 |
| ENSG00000272273 | *IER3-AS1* | antisense | ENSG00000115317 | *HTRA2* | 0.846 | 2.22E-16 |
| ENSG00000272273 | *IER3-AS1* | antisense | ENSG00000123353 | *ORMDL2* | 0.865 | 2.22E-16 |
| ENSG00000272273 | *IER3-AS1* | antisense | ENSG00000114354 | *TFG* | 0.824 | 2.22E-16 |
| ENSG00000272273 | *IER3-AS1* | antisense | ENSG00000137996 | *RTCA* | 0.825 | 2.22E-16 |
| ENSG00000272273 | *IER3-AS1* | antisense | ENSG00000196072 | *BLOC1S2* | 0.835 | 2.22E-16 |
| ENSG00000272273 | *IER3-AS1* | antisense | ENSG00000187792 | *ZNF70* | -0.833 | 2.22E-16 |
| ENSG00000272273 | *IER3-AS1* | antisense | ENSG00000156873 | *PHKG2* | 0.820 | 2.22E-16 |
| ENSG00000272273 | *IER3-AS1* | antisense | ENSG00000137288 | *UQCC2* | 0.891 | 2.22E-16 |
| ENSG00000272273 | *IER3-AS1* | antisense | ENSG00000177700 | *POLR2L* | 0.857 | 2.22E-16 |
| ENSG00000272273 | *IER3-AS1* | antisense | ENSG00000164620 | *RELL2* | 0.808 | 2.22E-16 |
| ENSG00000272273 | *IER3-AS1* | antisense | ENSG00000081181 | *ARG2* | 0.862 | 2.22E-16 |
| ENSG00000272273 | *IER3-AS1* | antisense | ENSG00000101191 | *DIDO1* | -0.842 | 2.22E-16 |
| ENSG00000272273 | *IER3-AS1* | antisense | ENSG00000114126 | *TFDP2* | -0.820 | 2.22E-16 |
| ENSG00000272273 | *IER3-AS1* | antisense | ENSG00000152217 | *SETBP1* | -0.823 | 2.22E-16 |
| ENSG00000272273 | *IER3-AS1* | antisense | ENSG00000119185 | *ITGB1BP1* | 0.802 | 2.22E-16 |
| ENSG00000272273 | *IER3-AS1* | antisense | ENSG00000203879 | *GDI1* | 0.832 | 2.22E-16 |
| ENSG00000272273 | *IER3-AS1* | antisense | ENSG00000197747 | *S100A10* | 0.828 | 2.22E-16 |
| ENSG00000272273 | *IER3-AS1* | antisense | ENSG00000240972 | *MIF* | 0.820 | 2.22E-16 |
| ENSG00000272273 | *IER3-AS1* | antisense | ENSG00000243317 | *STMP1* | 0.844 | 2.22E-16 |
| ENSG00000278419 | *AL451164.3* | antisense | ENSG00000171314 | *PGAM1* | 0.834 | 2.22E-16 |
| ENSG00000283029 | *AL139099.4* | sense | ENSG00000158373 | *HIST1H2BD* | 0.801 | 2.22E-16 |
| ENSG00000283029 | *AL139099.4* | sense | ENSG00000197061 | *HIST1H4C* | 0.808 | 2.22E-16 |
| ENSG00000283029 | *AL139099.4* | sense | ENSG00000169905 | *TOR1AIP2* | 0.809 | 2.22E-16 |
| ENSG00000284048 | *AC073111.4* | other | ENSG00000146426 | *TIAM2* | -0.836 | 2.22E-16 |
| ENSG00000284048 | *AC073111.4* | other | ENSG00000156011 | *PSD3* | -0.820 | 2.22E-16 |
| ENSG00000284048 | *AC073111.4* | other | ENSG00000161010 | *MRNIP* | 0.830 | 2.22E-16 |
| ENSG00000284048 | *AC073111.4* | other | ENSG00000170776 | *AKAP13* | -0.837 | 2.22E-16 |
| ENSG00000284048 | *AC073111.4* | other | ENSG00000137070 | *IL11RA* | 0.822 | 2.22E-16 |
| ENSG00000284048 | *AC073111.4* | other | ENSG00000075420 | *FNDC3B* | -0.835 | 2.22E-16 |
| ENSG00000284048 | *AC073111.4* | other | ENSG00000104447 | *TRPS1* | -0.814 | 2.22E-16 |
| ENSG00000284048 | *AC073111.4* | other | ENSG00000154229 | *PRKCA* | -0.811 | 2.22E-16 |
| ENSG00000284048 | *AC073111.4* | other | ENSG00000177200 | *CHD9* | -0.800 | 2.22E-16 |

| lncRNA ID | lncRNA name | Biotype | mRNA ID | mRNA name | Cor | Pval |
|---|---|---|---|---|---|---|
| ENSG00000284048 | *AC073111.4* | other | ENSG00000013583 | *HEBP1* | 0.822 | 2.22E-16 |
| ENSG00000284048 | *AC073111.4* | other | ENSG00000198018 | *ENTPD7* | -0.821 | 2.22E-16 |
| ENSG00000284048 | *AC073111.4* | other | ENSG00000166681 | *BEX3* | 0.811 | 2.22E-16 |
| ENSG00000284048 | *AC073111.4* | other | ENSG00000178802 | *MPI* | 0.801 | 2.22E-16 |
| ENSG00000284048 | *AC073111.4* | other | ENSG00000164776 | *PHKG1* | 0.815 | 2.22E-16 |
| ENSG00000284048 | *AC073111.4* | other | ENSG00000166025 | *AMOTL1* | -0.825 | 2.22E-16 |
| ENSG00000284707 | *AC079781.5* | other | ENSG00000166986 | *MARS* | 0.841 | 2.22E-16 |
| ENSG00000284707 | *AC079781.5* | other | ENSG00000070669 | *ASNS* | 0.871 | 2.22E-16 |

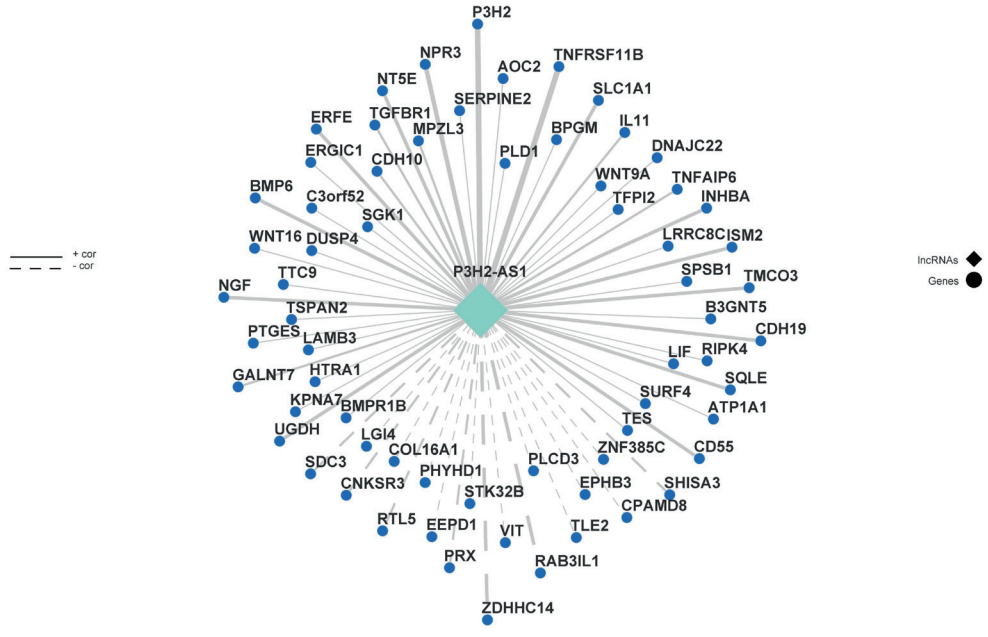
Cor = correlation between lncRNA and mRNA, Pval = nominal P value.

**Supplementary Table 8 | Count and percentage of unique long noncoding RNAs that have correlations > 0.8 with differentially expressed genes from the same samples.**

| Biotype | Number | Percentage (%) |
|---|---|---|
| Antisense | 17 | 36 |
| Sense | 2 | 4 |
| Pseudogene | 8 | 17 |
| Intergenic | 14 | 30 |
| Other | 6 | 13 |



## Supplementary Figures




**Supplementary Figure 1 | *P3H2-AS1*-mRNA coexpression network**

# Chapter 3

## High-impact *FN1* mutation decreases chondrogenic potential and affects cartilage deposition via decreased binding to collagen type II

Marcella van Hoolwerff[1†], Alejandro Rodríguez Ruiz[1†], Marga Bouma[2,3], H. Eka D. Suchiman[1], Roman I. Koning[4], Carolina R. Jost[4], Aat A. Mulder[4], Christian Freund[2,3], Farshid Guilak[5], Yolande F.M. Ramos[1], Ingrid Meulenbelt[1]

[1] Dept. of Biomedical Data Sciences, Section Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands.
[2] LUMC hiPSC Hotel, Leiden University Medical Center, Leiden, The Netherlands.
[3] Dept. of Anatomy and Embryology, Leiden University Medical Center, Leiden, The Netherlands.
[4] Section Electron Microscopy, Department of Cell and Chemical Biology, Leiden University Medical Center, Leiden, The Netherlands.
[5] Dept. of Orthopedic Surgery, Washington University and Shriners Hospitals for Children, St. Louis, USA.
[†] These authors contributed equally to this work.

Sci Adv. 2021 Nov 5;7(45):eabg8583
doi: 10.1126/sciadv.abg8583

## Abstract

Osteoarthritis is the most prevalent joint disease worldwide, yet progress in development of effective disease-modifying treatments is slow because of lack of insight into the underlying disease pathways. Therefore, we aimed to identify the causal pathogenic mutation in an early-onset osteoarthritis family, followed by functional studies in human induced pluripotent stem cells (hiPSCs) in an in vitro organoid cartilage model. We demonstrated that the identified causal missense mutation in the gelatin-binding domain of the extracellular matrix protein fibronectin resulted in significant decreased binding capacity to collagen type II. Further analyses of formed hiPSC-derived neo-cartilage tissue highlighted that mutated fibronectin affected chondrogenic capacity and propensity to a procatabolic osteoarthritic state. Together, we demonstrate that binding of fibronectin to collagen type II is crucial for fibronectin downstream gene expression of chondrocytes. We advocate that effective treatment development should focus on restoring or maintaining proper binding between fibronectin and collagen type II.

## Introduction

Osteoarthritis (OA) is an age-related, degenerative, heterogeneous disease of the articular joints, characterized by pathological changes of the articular cartilage, synovium, and subchondral bone [1]. Currently, OA is the most prevalent joint disease worldwide and affects more than 40 million people in Europe alone [2, 3]. Patients with OA suffer pain and stiffness in the articular joints, resulting in significant disability in everyday life [4]. Despite these detrimental consequences, to date, no effective treatment is available except pain relief medication and, in the final stage, joint replacement surgery. Lacking insight into underlying OA pathophysiology has considerably contributed to failures in disease modifying OA drug development.

For that matter, next generation exome sequencing has been highly successful in identifying likely causal mutations in patients with familial, more severe phenotypes [5] or Mendelian disorders [6]. Because of their strong effect, these mutations can provide direct clues to genotype-phenotype relations and are more actionable to express a particular disease state in experimental human in vitro tissue models. Since the pathways in which these genes operate are likely extrapolated to confer risk to common OA, these mutations are destined to elucidate OA-specific aspects of the disease. Using cells carrying high-impact mutations in tailored human in vitro tissue models may thus elucidate previously unknown, likely causal underlying pathways of OA [7, 8]. Moreover, human induced pluripotent stem cell (hiPSC) technology allows reprogramming of primary cells into a sustainable pluripotent cell source, which is subsequently used for the directed differentiation into specialized cells of interest [9]. Even more, because of novel advancements in genomic engineering with CRISPR/Cas9 technology, cells are now easily edited or modulated. As a result, the creation of isogenic lines can give insight in the functional characterization of genetic variation. Since OA-relevant cells, such as chondrocytes, are not readily available, the combination of hiPSCs and CRISPR/Cas9 holds immense potential for translational human OA disease modeling to uncover the role of high-impact mutations [8].

By applying a genome-wide linkage scan in seven extended early-onset (EO)-OA families without dysplasia, we previously identified significant linkage on chromosome 2q33.3, with one family (family 2) that substantially contributed to this linkage area [10]. However, Sanger sequencing of 20 positional candidate genes located in the confined linkage region of these seven families failed to definitively identify the causal mutation in these families. More recently, an exome sequencing dataset was established in these OA families, resulting in the successful identification of a high-impact mutation in one of the families [11]. By applying a pathogenic prioritization scheme to the exome sequencing dataset of family 2, we identified a previously unidentified mutation and used hiPSCs and CRISPR/Cas9 technology to develop an in vitro organoid cartilage model to functionally characterize the effects of the causal mutation in this family.

# Materials and Methods

## Experimental design

The objective of the current study was to identify a likely causal pathogenic mutation in an EO-OA family and underlying disease pathways by functional analyses of isogenic hiPSCs. Exome sequencing was applied to identify a pathogenic mutation, whereas linkage analysis was applied to confirm co-segregation of the mutation and the OA phenotype. To study underlying mechanisms, the mutation was introduced in hiPSCs using CRISPR/Cas9 genome engineering, after which they were used in an in vitro organoid cartilage model, followed by functional analyses. A solid-phase binding assay was performed to quantify binding capacity of mutant and wild-type fibronectin to collagen type II.

## Study populations

### *EO-OA family members*

Previously, we reported on an EO family with primary generalized OA without dysplasia (**Supplementary Figure 1**) [10, 12]. The proband was collected through a Rheumatology outpatient clinic; family members were recruited via the proband. The age of onset of OA was between 23 and 45 and the phenotype is characterized by progressive OA with symptoms at multiple joints simultaneously. Two family members reported OA-related complaints, however, no radiographic evidence was available to confirm the diagnosis. An overview of the phenotypes is shown in **Supplementary Table 1**. The study was approved by the medical ethical committee and written and informed consent was obtained from all participants.

### *Controls*

Occurrence of the *FN1* missense mutation was tested in the general population by de novo genotyping in healthy participants and OA cases. Middle-aged partners (N=744) of the offspring of nonagenarian siblings from the Leiden Longevity Study (LLS) [13] were considered as controls for OA cases were named 'random controls'. OA cases included were selected from the Genetics Osteoarthritis And Progression (GARP) study (N = 177) [14], the Patients Prospectively Recruited in Knee and hip Arthroplasty (PAPRIKA) study (N = 1,137) [15], and the ongoing Research osteoArthritis and Articular Cartilage (RAAK) study (N = 153) [16]. For more details on ascertainment of RAAK study biospecimen see Ramos *et al.* [11]. Ethical permission for the described studies was obtained from the appropriate medical ethical committee under protocol numbers P08.239 and P19.013. Written informed consent was obtained from all participants.

## Exome sequencing

Exome sequencing of an affected EO-OA family member was performed by Illumina HiSeq2000 technology (Bejing Genome Institute). The sequences were generated as 100 base pair paired-end reads, after enrichment of 44 MB exonic sequences by NimbleGen EZ

(Roche NimbleGen). Raw imaging files were processed by Illumina base calling software version 1.7 with default parameters. SOAPaligner/SOAP2.21 was used to align reads to the GRCh37 reference genome at the UCSC Genome Browser website (http://genome.ucsc.edu/). Approximately, 70% of bases originated from the targeted exome, resulting in a mean coverage of 55.7-fold for the proband. More than 83% of the targeted exons were covered more than 10 times. Single-nucleotide variants were subsequently called by the SOAPsnps. The variant-filtering scheme, which resulted in 122 pathogenic mutations, is shown in **Supplementary Table 2**.

## Genotyping

Single-nucleotide polymorphisms were genotyped by mass spectrometry (MassARRAY system, Sequenom, San Diego, CA) using standard conditions as described by Meulenbelt *et al.* [17]. Variants werWWe genotyped across available EO-OA family members (see **Supplementary Table 3**), OA case studies (GARP, PAPRIKA, and RAAK study), and random controls (LLS).

## hiPSC line and cell culture

The hiPSC line LUMC0004iCTRL10 ("LUMC0004" in this manuscript) was generated from skin fibroblasts of a male donor without known genetic diseases ("wild type") as described previously [18], and registered at the Human pluripotent stem cell registry. The generation of the hiPSC line was approved by the Leiden University Medical Center ethical committee under P13.080. The hiPSCs were maintained under standard conditions (37 °C, 5% $CO_2$) on Vitronectin-XF (Stemcell Technologies)-coated plates and refreshed daily with TeSR-E8 medium (Stemcell Technologies). Cells were passaged in aggregates using Gentle Cell Dissociation Reagent (Stemcell Technologies) upon reaching approximately 80% confluency.

## Genome editing of hiPSCs and clonal isolation

hiPSCs were dissociated with Gentle Cell Dissociation Reagent (Stemcell Technologies) and 1 x $10^5$ cells were transfected with the Alt-R Cas9 RNP complex and the single-stranded oligodeoxynucleotide (ssODN; both IDT) using the NEON transfection System (Invitrogen) at 1200V/30 ms/1 pulse (for guide RNA and ssODN sequence see **Supplementary Table 4**). Cells were plated in a Synthemax II-SC (Corning)-coated plate using TeSR-E8 supplemented with CloneR (Stemcell Technologies). After recovery, 1000 cells were plated in a Synthemax II-SC-coated 10 cm dish in TeSR-E8 supplemented with CloneR and refreshed daily. After ~14 days, single cell-derived hiPSC colonies were manually picked, and one half was used for DNA isolation with QuickExtract Solution (Lucigen). Positive clones were identified by polymerase chain reaction (PCR) screening of the region of interest, followed by successful digestion by restriction enzyme HincII (New England Biolabs), according to manufacturer's instructions. Primer sequences are shown in **Supplementary Table 5**. The introduced mutation was confirmed by Sanger sequencing, performed by the Leiden Genome Technology Center.

## Characterization of gene-edited hiPSCs

In silico analysis with the Cas-OFFinder tool [19] was used to identify most likely sites for off-target mutations. When allowing two mismatches, one potential off-target was identified, which was sequenced as described above (for primer sequences see **Supplementary Table 5**). G-banding analysis was conducted at the Laboratory of Clinical Genetics Leiden according to standard procedures. A total of 18 metaphases were analyzed for one heterozygous clone and 20 metaphases for one homozygous clone.

## hiPSC differentiation to induced chondroprogenitor cells

Generation of human induced chondroprogenitor cells (hiCPCs) was based on a protocol previously described [20], which was shown to produce similar neo-cartilage to that produced by human primary articular chondrocytes [21]. When hiPSCs reached 60% confluence, the culture medium was switched to mesodermal differentiation (MD) medium, composed of IMDM GlutaMAX (IMDM; Thermofisher Scientific) and Ham's F12 Nutrient Mix (F12; Sigma-Aldrich) with 1% chemically defined lipid concentrate (Gibco), 1% insulin/human transferrin/selenous (ITS+; Corning), 0.5% penicillin/streptomycin (P/S; Gibco), and 450 μM 1-thioglycerol (Sigma-Aldrich). Before induction of anterior primitive streak (day 0) hiPSCs were washed with wash medium (IMDM/F12, 0.5% P/S) and then fed with MD medium supplemented with 30 ng/ml Activin A (Stemgent), 4 μM CHIR99021 (CHIR; Stemgent), and 20 ng/ml human fibroblast growth factor (FGF-2; R&D Systems) for 24 hours. Subsequently, the cells were washed again with wash medium and paraxial mesoderm was induced on day 1, by MD medium supplemented with 2 μM SB-505124 (Tocris), 3 μM CHIR, 20 ng/ml FGF-2, and 4 μM dorsomorphin (Tocris) for 24 hours. Before induction of early somite (day 2), cells were washed with wash medium, and then cells were fed with MD medium supplemented with 2 μM SB-505124, 4 μM dorsomorphin, 1 μM C59 (Cellagen Technology), and 500 nM PD173074 (Tocris) for 24 hours. Subsequently, cells were washed with wash medium, and for induction of sclerotome, cells (day 3-5) were fed daily with MD medium supplemented with 2 μM puromorphamine (Stemgent), and 1 μM C59. To induce chondroprogenitor cells (day 6-14), cells were washed briefly with wash medium and fed daily with MD medium supplemented with 20 ng/ml human bone morphogenetic protein 4 (BMP-4; Miltenyibiotec). Five independent differentiations were done per clone.

## Chondrogenic differentiation

Monolayer cultured hiCPC aggregates present at day 14 of the differentiation were washed with MD medium, manually picked, and centrifuged for 4 minutes at 1200 rpm. Cell aggregates were subsequently maintained in Dulbecco's modified Eagle's medium/F12 (Gibco), supplemented with 1% ITS+, 55 μM 2-mercaptoethanol (Gibco), 1% non-essential amino acids (NEAA, Gibco), 0.5% P/S, 50 μg/ml L-ascorbate-2-phosphate (Sigma-Aldrich), 40 μg/ml L-proline (Sigma-Aldrich), and 10 ng/ml transforming growth factor-β1 (Prepotech) for 35 days while

refreshing medium every 3 to 4 days. The remaining hiCPC aggregates in the well were pooled and lysed with 500 µl TRIzol reagent (Thermo Fisher Scientific) and stored at -80 °C until further processing.

### Sulfated glycosaminoglycan measurement

Sulfated glycosaminoglycan (sGAG) concentration in the chondrogenic pellets was measured with the 1,9-dimethylmethylene blue (DMMB) assay [22]. Pellets were digested with papain from papaya (Sigma-Aldrich) at 60 °C overnight. Shark chondroitin sulfate (Sigma-Aldrich) was used as a reference standard. The absorbance was measured at 525 and 595 nm in a microplate reader (Spectramax iD3; Molecular Devices).

### Histology and immunohistochemistry

Chondrogenic pellets were fixed in 4% formaldehyde overnight and stored in 70% ethanol at 4 °C. They were then embedded in paraffin and sectioned at 5 µm. After sectioning, slides were deparaffinized and rehydrated. Overall cellular and tissue structure was visualized with hematoxylin and eosin staining. Sections were stained with 1% Alcian blue 8-GX (Sigma-Aldrich) and Nuclear Fast Red (Sigma-Aldrich) to visualize glycosaminoglycan deposition. To detect fibronectin (1:400; ab2413, Abcam), collagen type I (1:1000; ab34710, Abcam), and collagen type II (1:100; ab34712, Abcam), immunohistochemistry was performed with 3-diaminobenzidine solution (Sigma-Aldrich) and hematoxylin (Klinipath) as described before [23]. Alcian blue staining was quantified by loading the images in Fiji and splitting the color channels. Subsequently, grey values were measured of three to five separate squares per pellet and corrected for the grey value of the background.

### RNA isolation and reverse transcription-quantitative PCR

For RNA isolation, two pellets were pooled and lysed in 200 µl of TRIzol reagent (Thermo Fisher Scientific) and homogenized using micro pestles. RNA was extracted with chloroform and purified from the supernatant using an RNeasy Mini kit (Qiagen). Synthesis of complementary DNA (cDNA) was performed with 150 ng of total RNA using a First Strand cDNA Synthesis kit according to manufacturer's protocol (Roche Applied Science). cDNA was diluted five times and preamplification with Taqman preamp master mix (Thermo Fisher Scientific) was performed for eight genes of interest. Primer sequences are shown in **Supplementary Table 5**. Gene expression was measured with the Fluidigm Biomark HD machine using a 96.96 IFC chip or with QuantStudio 6 Real-Time PCR system (Applied Biosystems) using FastStart SYBR Green Master reaction mix (Roche Applied Science). Quality control of the data was performed and non-detected values were removed. Relative gene expression levels were calculated with the $2^{-\Delta\Delta Ct}$ method as follows; the average of triplicate Ct values of the genes were averaged. Ct values of *GAPDH* and *SDHA* were averaged to function as reference genes (RG). Next, ΔCt values of the gene of interest (GOI) were calculated as $Ct_{GOI} - Ct_{RG}$.

Subsequently, we averaged the ΔCt values of *COL1A1* and *COL2A1* as chondrogenic controls (CC) to calculate final ΔCt values of the gene of interest, as followed: $\Delta Ct_{GOI\ new} = \Delta Ct_{GOI} - \Delta Ct_{CC}$. Fold changes were calculated with $2^{-\Delta\Delta Ct}$, with the wild type as reference.

## Solid-phase binding assay

Conditioned medium of wild-type and *FN1* mutant pellets was collected and concentrated in preparation for the binding assay. To this end, 450 µl medium was collected in 100 K molecular weight cut-off Pierce Protein Concentrators (Thermo Scientific) and centrifuged for 10 min at 12,000 *g*. Subsequently, fibronectin concentration was determined using the Human Fibronectin enzyme-linked immunosorbent assay (ELISA) kit (Invitrogen) according to manufacturer's protocol.

Clear multiwell plates (R&D Systems) were coated overnight with 100 µl of purified human collagen type II (10 µg/ml; Merck) in phosphate-buffered saline (PBS) at 4 °C, followed by four wash steps with wash buffer (0.05% Tween-20 in PBS). Nonspecific binding was blocked for 1 hour with 3% (w/v) bovine serum albumin (BSA) in PBS. After washing with wash buffer, the plates were incubated with 100 µl concentrated medium samples at 40 ng/ml FN1 concentration in assay buffer (0.05% Tween-20, 0.5% BSA in PBS) for 2 hours. Plates were then washed four times with wash buffer and incubated with rabbit anti-fibronectin biotin conjugated antibody (Rockland) at 0.2 µg/ml in assay buffer for 1 hour. Plates were washed, after which the plates were incubated with streptavidin-horseradish peroxidase (Thermo Scientific) at 0.1 µg/ml in assay buffer for 1 hour. After washing, color development was performed with 100 µl tetramethyl-benzidine substrate (Thermo Fisher) for 10 min, reaction was stopped with 100 µl 1 M HCl and absorbance was measured at 450 nm. Assays were performed in triplicate. Error bars are SDs.

## Transmission electron microscopy

Chondrogenic pellets that where embedded in paraffin were melted from the paraffin for 30 minutes at 62 °C. Subsequently, the samples were deparaffinized and rehydrated with a series of ethanol. The rehydrated samples were fixed again in 1.5% glutaraldehyde/0.1 M cacodylate buffer and after three times rinsing with 0.1 M cacodylate buffer, postfixed in 1% OsO4/0.1 M cacodylate buffer. A dehydration with a series of ethanol and series of propylene oxide mixed with EPON (LX112, Ladd research industries) was performed. After a final step with EPON the samples were each placed in a BEEM capsule; these were filled with EPON and polymerized at 70°C for 48 hours. Ultrathin sections (90 nm) were made on a Reichert Ultracut S (Leica Microsystems). After the poststaining of the sections with uranyl acetate and lead citrate, the electron microscopy images were taken with a twin transmission electron microscope (Thermo Fisher Scientific, formerly Fei) operating at 120 kV, using a Gatan Oneview camera (Gatan, Pleasanton) on binning 2. Overlapping images were collected and stitched together into a big image as previously described [24].

### Statistical analysis

Statistical analyses were performed using SPSS version 25 (IBM). Data are shown as mean ± SD. With respect to reverse transcription-quantitative PCR (RT-qPCR) gene expression data, individual -ΔCt values are shown, as well as box plots per genotype. The box plots represent 25th, 50th and 75th percentiles, and whiskers represent the lowest and largest data point lying within 1.5 times the interquartile range. Individual samples are depicted by dots in each graph. The reported P values were determined by applying a generalized estimating equations (GEE) to the experimental readout i.e. -ΔCt values, relative Alcian blue intensity, and FN1 concentration, to effectively adjust for dependencies of the independent differentiations per genotype. In general we followed a linear GEE model, with the readout data as dependent variable, genotype as factor, and exchangeable working matrix: Read-out ~ Genotype + (1| Differentiation). To determine a dose-response effect, we applied a GEE to the experimental read-out, i.e. absorbance, and sGAG/DNA, with the read-out as dependent variable, genotype as covariate, and exchangeable working matrix. For the analysis of the pellet morphology, we followed a binary logistic GEE model, with normal and not-normal pellets as readout, genotype as factor, and adjusted for independent differentiations. P values < 0.05 were considered statistically significant. * P < 0.05, ** P < 0.01, *** P < 0.005.



## Results

### Identification of heterozygous missense mutation in *FN1*

Whole-exome sequencing was applied to an affected individual of an EO generalized OA family (**Figure 1A, subject 10**) with an average age of onset at 37 years and without dysplasia (**Supplementary Table 1**). This resulted in the identification of 73,407 variants after quality control, after which a prioritization scheme was applied to identify pathogenic variants (**Supplementary Table 2**). Hence, intergenic variants, intron variants, synonymous variants, and tolerated missense variants as determined by PolyPhen or SIFT (Sorting Intolerant From Tolerant) were excluded, as well as variants present in dbSNP build 124 (including 1000 Genomes Pilot project data). In addition, variants were excluded that were previously detected by whole-genome sequencing projects, in-house (N = 221), and by means of the BBMRI-Genome of The Netherlands project (N = 473) [25]. As a consequence, the number of variants was reduced to 122 missense variants (**Supplementary Table 6**) that were predicted to have a functional impact on the gene.

Previously, our group showed strong linkage on 2q33.3 with seven extended EO-OA families, including the family described in this study [10]. This family showed substantial contribution to this linkage area (logarithm of odds score 2.93); hence, we investigated the chromosomal location of the previously mentioned 122 missense variants. Of these variants, three were located around the linkage area on chromosome 2, namely, *ALS2*, *FN1*, and *ABCB6*. To explore

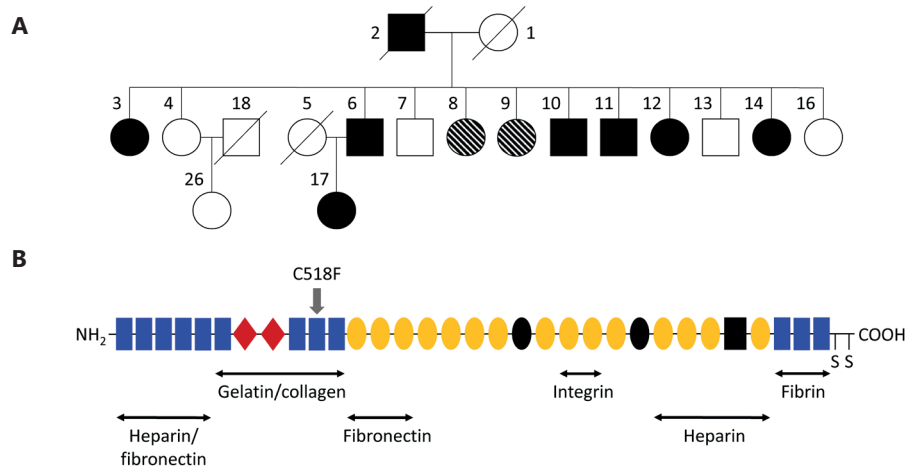



**Figure 1 | High-impact mutation identified in *FN1* in early-onset osteoarthritis family.** (**A**) Pedigree of investigated early-onset osteoarthritis family. Individual 10 was used for exome sequencing. Circles represent females, squares represent males. Closed and open shapes represent affected and healthy individuals, respectively, and striped shapes represent uncertain phenotypes, i.e., some complaints, but not confirmed by radiography. The diagonal line indicates deceased status. (**B**) Schematic diagram of a fibronectin subunit. Fibronectin consists of three types of repeats: type I (rectangle), type II (diamond), and type III (sphere). The C-terminal cysteines are involved in dimerization. The alternatively spliced sites are black shapes. Main binding sites for heparin, collagen, fibronectin and integrin are shown. Position of the identified C518F mutation is indicated by the grey arrow.

functionality of these genes in OA-relevant tissue, gene expression levels in our previously published RNA sequencing data of lesioned and preserved cartilage and bone samples of patients undergoing joint replacement surgery due to end-stage OA, from the RAAK study, were examined [26, 27]. *FN1* showed the highest gene expression in both cartilage and bone samples and showed a high, significant up-regulation (fold change, 2.12; false discovery rate = $1.17 \times 10^{-3}$) in lesioned compared with preserved OA cartilage (**Table 1**). Furthermore, *ALS2* was expressed at low levels in both cartilage and bone samples, while *ABCB6* expression was even lower in cartilage and not detected in bone. Therefore, de novo genotyping was performed for the *ALS2* and *FN1* variant, showing that the *ALS2* variant was not detected, yet confirming that the *FN1* variant was carried in seven affected individuals, but not in six unaffected family

**Table 1 | Reproduced gene expression analysis of detected previously unidentified variants located around linkage area 2q33.3 in osteoarthritic cartilage and bone.**

| | | | | | Cartilage† | | | Bone† | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Position** | **Gene** | **Codon** | **AA sub.** | **Base** | **Quartile** | **FC** | **FDR** | **Quartile** | **FC** | **FDR** |
| 202590176 | *ALS2* | TAT-gAT | Y1084D | A/C | Q3 | 1.1 | $4.95 \times 10^{-1}$ | Q3 | 1.0 | $8.14 \times 10^{-1}$ |
| 216285518 | *FN1* | TGC-TtC | C518F | C/A | Q4 | 2.1 | $1.17 \times 10^{-3}$ | Q4 | 1.2 | $6.04 \times 10^{-1}$ |
| 220075737 | *ABCB6* | CGT-gGT | R688G | G/C | Q1 | 1.1 | $8.72 \times 10^{-1}$ | ND | ND | ND |

AA sub.* = amino acid substitution; Base = base change; FC = fold change; FDR = false discovery rate of differential expression; ND = not detected. * Variants are novel, missense damaging, nonsense, readthrough, frame error or splice-sites not present in dbSNPv124, as well as not detected in in-house genome sequencing projects (LLS, N = 221) and analyzed human genomes by means of the Genome of The Netherlands (N = 473)[25]. † Expression levels in quartiles, with Q1 being the lowest and Q4 being the highest, fold changes and false discovery rate of differential expression levels of identified variants reproduced from RNA sequencing dataset of articular cartilage [26] and subchondral bone [27].

members. These data indicate complete linkage to robustly affected and unaffected subjects in the family (**Supplementary Table 3**). The *FN1* variant was neither present in the additional 1467 OA cases (GARP, RAAK, and PAPRIKA study) and 744 random controls (LLS), nor in the Genome Aggregation Database (version 2.1.1), and is therefore likely specific for this EO-OA family. Taken together, these data indicate that the G-to-T nucleotide change (c1,819), resulting in an amino acid change from cysteine to phenylalanine (C518F; **Figure 1B**) in the *FN1* gene is likely causal to the EO phenotype in this family.

*FN1* encodes fibronectin, which is a dimeric glycoprotein, highly abundant in the extracellular matrix (ECM) of articular cartilage and more specifically localized in the pericellular matrix surrounding the chondrocyte. Fibronectin can bind to matrix proteins by multiple binding domains, such as the gelatin-binding domain which binds to collagen type II, as well as to cell surface integrins, thereby functioning as an adhesion molecule facilitating signals from the extracellular matrix to chondrocytes [28, 29]. Next, we aimed to investigate the molecular implications of the identified *FN1* mutation for the pathophysiology of OA.



## Targeted gene editing of *FN1* mutation using CRISPR/Cas9 in hiPSCs

To elucidate the potential pathogenic mechanism of the identified *FN1* mutation, we sought to introduce it into wild-type hiPSCs, resulting in an isogenic pair of mutated and non-targeted hiPSCs. Gene editing was achieved by using CRISPR/Cas9, a guide RNA binding in close to proximity to the target site, as well as a 141-mer ssODN containing the specific *FN1* mutation (**Supplementary Table 4**). The creation of a HincII target site by introducing the mutation was used for screening of successfully targeted clones (**Figure 2A**). The presence of the introduced mutation was confirmed by Sanger sequencing (**Figure 2B**). A total of 274 hiPSC clones were screened, of which 14 (5.1%) were heterozygous and 3 (1.1%) were homozygous for the *FN1* mutation. Last, one heterozygous and one homozygous *FN1* mutant clone were selected for further experiments. Both clones had a normal karyotype, and the predicted off-target site allowing two mismatches was not altered (**Supplementary Figure 2**).

## Effect of missense mutation *FN1* on chondrogenesis

The wild-type, heterozygous and homozygous *FN1* hiPSC clones were differentiated stepwise to the mesodermal lineage and further towards hiCPCs. Subsequently, hiCPC aggregates were manually picked, and chondrogenesis was initiated in our organoid cartilage model. After 5 weeks of chondrogenesis, the chondrogenic pellets were collected, where we observed different morphologies of the pellets. We categorized the pellets in groups based on shape (round or protrusive) and opaqueness of the matrix (**Figure 3A**). The abundance of round, dense pellets (**Figure 3A, group 1**) was 53% of the pellets in the wild-type group, while this decreased substantially to 15% in the heterozygous group, and to 18% in the homozygous group (**Figure 3B**). Furthermore, the abundance of protrusive pellets with more transparent

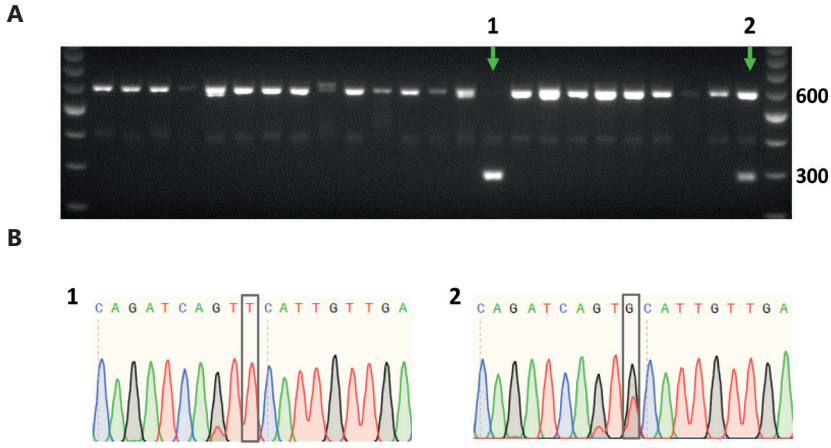



**Figure 2 | Introducing identified *FN1* mutation in human induced pluripotent stem cells (hiPSCs).** Clonal screen for *FN1* mutation in hiPSCs by genomic povlymerase chain reaction (PCR) followed by HincII digestion (**A**), confirmed by Sanger sequencing (**B**). The PCR product is 600 base pairs (bp) in size. Upon integration of the provided single-stranded oligodeoxynucleotide, HincII digestion results in 303- and 297-bp products. (**A**) Representative gel image, where *FN1* mutant clones are indicated by a green arrow, showing a homozygous clone (lane 1) and a heterozygous clone (lane 2). (**B**) Sanger sequencing results of homozygous (1) and heterozygous (2) clone, respectively. Grey box represents *FN1* mutation.

matrix (**Figure 3A, group 3**) in the heterozygous (60%), and homozygous (54%) groups was increased compared to the wild-type group (25%). When considering group 1 as 'normal' pellets and groups 2-4 as 'not-normal' pellets, we found a significant association between the occurrence of not-normal pellets and the presence of the *FN1* mutation, for both the heterozygous group (Beta = 1.84, 95% CI 2.8 – 14.1, P value = $9 \times 10^{-6}$) and the homozygous group (Beta = 1.63, 95% CI 2.4 – 11.1, P value = $3.1 \times 10^{-5}$). These data show that there is a dominant effect of the mutation on the formation of dense, round pellets, thus indicating an inhibited chondrogenic potential of the mutant hiCPCs.

To investigate whether the chondrogenic potential of the mutant hiCPCs was also inhibited before the chondrogenesis, we measured expression of the previously identified chondroprogenitor and chondrogenic markers *MCAM, PDGFRβ,* and *COL2A1* of hiCPCs at day 0 of chondrogenesis (**Figure 3C**) [20, 30-32]. *MCAM* was not differentially expressed between the groups, yet *PDGFRβ* and *COL2A1* expression was significantly down-regulated in the homozygous group, compared with wild type. Together, these data show that the chondrogenic potential is decreased as a result of the *FN1* mutation, with the effect being the most pronounced in the homozygous hiCPCs.

To characterize the differences between the normal and the not-normal chondrogenic pellets, we performed histological and immunohistochemical analyses (**Figure 4A**). Hematoxylin and eosin staining of chondrogenic pellets from group 3 showed that there was a clear difference

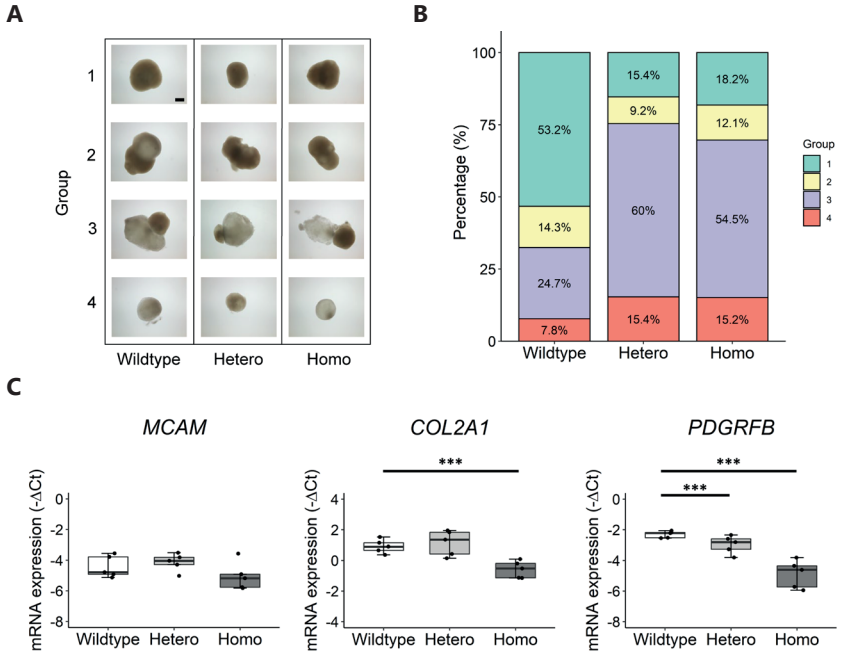



**Figure 3 | Decreased chondrogenic potential of human induced chondroprogenitor cells (hiCPCs) as a result of the *FN1* mutation.** (**A**) Representative images of the observed morphology of wild-type, *FN1* C518F heterozygous (Hetero), and C518F homozygous (Homo) hiCPC-derived chondrogenic pellets after 5 weeks of chondrogenesis in five independent differentiations. Pellets were categorized in four groups based on shape and opaqueness of the matrix: (1) Round and dense, (2) Protrusive and dense, (3) Protrusive and transparent, (4) Round and transparent. Scale bar, 400 μm. (**B**) Distribution of pellets in groups 1 to 4 of wild-type (N=77), *FN1* heterozygous (N=65), and *FN1* homozygous pellets (N=66). (N = 5-20 pellets per differentiation). (**C**) Gene expression levels depicted by -ΔCt values of hiCPC markers *MCAM*, *COL2A1*, and *PDGFRβ* [20] at day 0 of chondrogenesis of pooled wild-type and *FN1* mutant hiCPCs (N=5). The box plots represent 25th, 50th and 75th percentiles, and whiskers extend up to 1.5 times the interquartile range. Individual samples are depicted by black dots in each graph. P values were determined by generalized estimation equations, with gene expression levels (-ΔCt) as dependent variable and genotype as factor. *** P < 0.001.

in structure and composition between the dense and transparent matrix. Moreover, Alcian blue, fibronectin, collagen type I and II staining revealed that the transparent matrix was non-cartilaginous tissue. When only considering normal pellets (group 1), visually there was not a clear difference in matrix intensity staining of collagen type II and fibronectin (**Figure 4B**). However, quantification of the Alcian blue staining in the chondrogenic parts of pellets from groups 1 to 3 revealed a significant decrease in intensity in the *FN1* mutant chondrogenic pellets (**Figure 4C**). Quantification of the neo-cartilage sGAG deposition normalized to DNA content showed a significant linear reduction (Beta = -1.44, 95% CI 0.07 – 0.86, P = 2.8 x $10^{-2}$) with the presence of the *FN1* mutation (**Figure 4D**). Moreover, fibronectin concentration was decreased in the medium of the *FN1* mutant pellets, albeit not significantly in the homozygous group (**Figure 4E**). Together, our data show that the *FN1* mutation results in a negative effect on chondrogenic potential and neo-cartilage deposition.

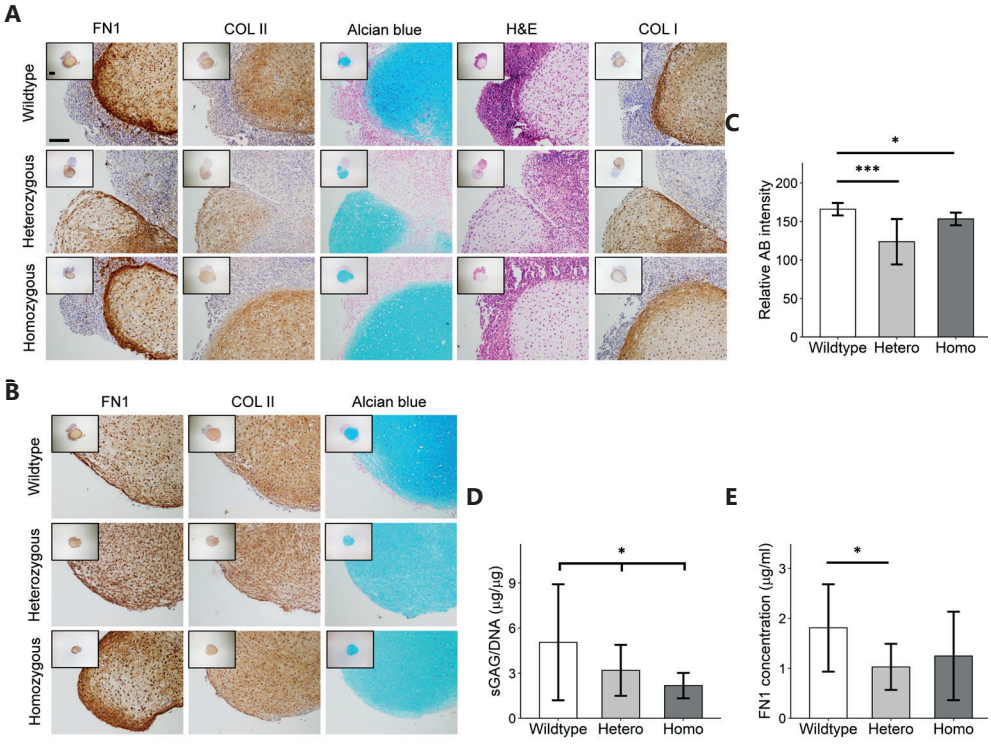



**Figure 4 | Decreased neo-cartilage deposition in C518F *FN1* mutant chondrogenic pellets.** (**A**) Representative images of wild-type and *FN1* mutant human induced chondroprogenitor cells (hiCPC)-derived chondrogenic pellets categorized in groups 2 to 4 analyzed by hematoxylin and eosin (H&E), Alcian blue, fibronectin (FN1), collagen type I (COL I), and collagen type II (COL II). Scale bar, 200 µm. (**B**) Representative images of wild-type and *FN1* mutant hiCPC-derived chondrogenic pellets categorized in group 1 analyzed by Alcian blue, fibronectin (FN1), and collagen type II (COL II). Scale bar, 200 µm. (**C**) Quantification of Alcian blue (AB) staining of Alcian blue-positive areas in wild-type and *FN1* mutant hiCPC-derived chondrogenic pellets (N=3). (**D**) Dose-response reduction (Beta = -1.44, P = 2.8 x $10^{-2}$) of sulfated glycosaminoglycan (sGAG) normalized to DNA in wild-type, *FN1* heterozygous (Hetero) and homozygous (Homo) chondrogenic pellets analyzed by dimethylmethylene blue (N=8). (**E**) Fibronectin concentration in conditioned medium from *FN1* mutant pellets is decreased compared with wild-type, as determined by ELISA (N=3) . Data are means ± SD. P values were determined by generalized estimation equations, with experimental readout (relative Alcian Blue intensity, sGAG/DNA, and fibronectin concentration) as dependent variable and genotype as covariate (relative Alcian Blue intensity, sGAG/DNA) or factor (fibronectin concentration). * P < 0.05, *** P < 0.005.

## Mechanism of reduced chondrogenesis in mutant chondrogenic pellets

The identified C518F mutation is located in the gelatin-binding domain of fibronectin, which is particularly involved in binding to collagen type II [29]. The cysteine is involved in intramolecular disulfide bond formation and was shown to be highly conserved among multiple species, highlighting the importance of this amino acid (**Supplementary Figure 3A**). The change from a polar cysteine to a nonpolar phenylalanine resulted in a predicted conformational change of the protein as determined by RaptorX [33] (**Supplementary**

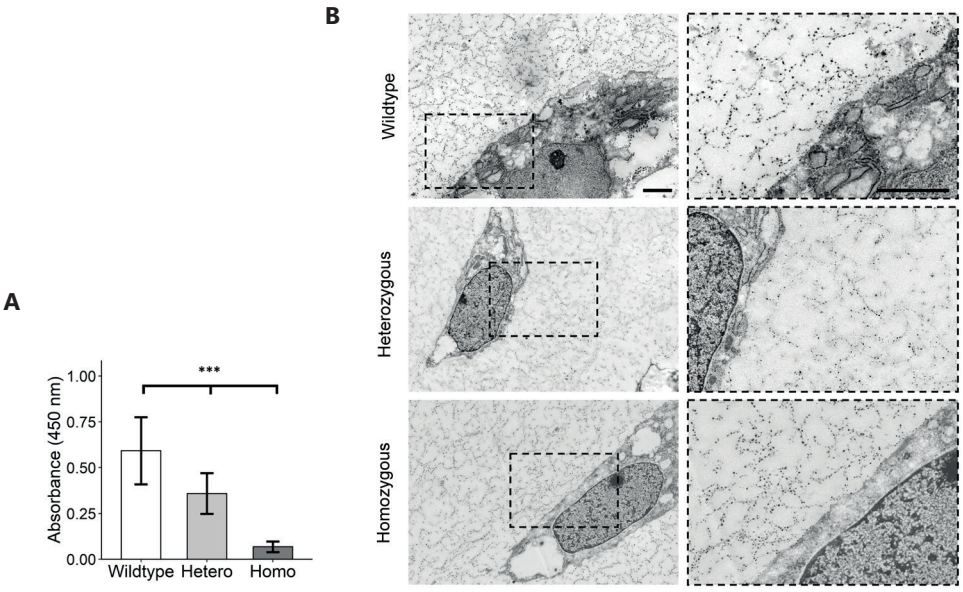



**Figure 5 | *FN1* mutant fibronectin showed a dose-response decrease of binding capacity to collagen type II.** (**A**) Solid-phase binding assay with collagen type II-coated wells and wild-type and C518F mutant fibronectin isolated from chondrogenic pellet cultures. Heterozygous (Hetero) and homozygous (Homo) mutant fibronectin showed linear reduction (Beta = -0.26 , P = 2.26 x $10^{-9}$) of binding to collagen type II compared to wild-type fibronectin. Data are means ± SD (N = 3). P value was determined by generalized estimation equations, with absorbance as dependent variable and genotype as covariate. *** P < 0.005. (**B**) Representative transmission electron microscopy images showing a cell surrounded by cartilaginous matrix of wild-type and *FN1* mutant human induced chondroprogenitor cell-derived chondrogenic pellets. The fibrillar network surrounding the cells is suggestive of hyaluronic acid aggregates, and the dense particles are proteoglycans containing granules. Scale bars, 2 µm.

**Figure 3B**). When replaced by the phenylalanine, the disulfide bond in this part of the domain cannot be formed. Therefore we hypothesized that the mutation resulted in a conformational change in the gelatin-binding domain, with concurrent decreased binding of fibronectin to collagen type II. To explore this, we performed a collagen type II binding assay with wild-type, heterozygous, and homozygous mutant fibronectin. As a result, we observed a linear reduction (Beta = -0.26, 95% CI 0.71 – 0.84, P = 2.3x$10^{-9}$) in binding of mutant fibronectin to collagen type II relative to wild-type fibronectin, thereby confirming our hypothesis (**Figure 5A**). To further investigate whether we could identify structural differences in the neo-cartilage as a result of this decreased binding between fibronectin and collagen type II, we visualized cartilage by transmission electron microscopy. We observed electron dense particles that were present both attached to the plasma membrane and to a branching fibrillar network obviously devoid of striations emanating from the plasma membrane. The fibrillar network suggests a network of hyaluronic acid aggregates [34], and the electron dense particles are proteoglycans containing granules (**Figure 5B**) [35]. However, we could not identify any clear differences of the abundance and structure of observed structures between the wild-type and *FN1* mutant.

To study the effect of the mutation on markers reflecting chondrocytes in health or disease state, we next explored chondrocyte gene expression in the pellets as a result of the mutation by performing RT-qPCR for 10 relevant genes. Relative gene expression levels were first calculated by normalizing to expression levels of *GAPDH* and *SDHA* (**Supplementary Figure 4**). The observed difference in formation of neo-cartilage between the genotypes (**Figure 4A**) was potentially a confounding factor in the gene expression analysis. Therefore, we aimed to correct for the differences in efficiency of the neo-cartilage formation between the genotypes. Hence, we selected cartilage-specific markers that were highly expressed in our lesioned and preserved osteoarthritic cartilage samples (RAAK study), but without significant differential expression [26], and that we showed to be specific markers for the neo-cartilage (**Figure 4A**). As a result, *COL1A1* and *COL2A1* were selected as reference genes. The anabolic cartilage markers *ACAN* and *FN1* were significantly down-regulated in the *FN1* mutant chondrocytes as compared with wild type, while the catabolic and hypertrophic markers *ADAMTS-5*, *ALPL*, and *RUNX2* were significantly up-regulated in the *FN1* homozygous group (**Figure 6** and **Table 2**). These data indicate that the mutant chondrocytes were in a more OA disease state, where the effect was stronger in the homozygous mutant group. To investigate the downstream effects of the fibronectin mutation on the fibronectin-binding chondrocyte cell surface integrin receptors, we investigated the gene expression levels of *ITGA3, ITGB1,* and *ITGA5*. Integrin α5β1 is the main fibronectin-binding integrin, while α3β1 and αVβ1 can also bind fibronectin [36, 37]. *ITGA3* and *ITGB1* were significantly up-regulated in *FN1* homozygous chondrocytes (**Figure 6**). Notably, *ITGA5* was down-regulated in the heterozygous group, while its expression was not changed in the homozygous group.

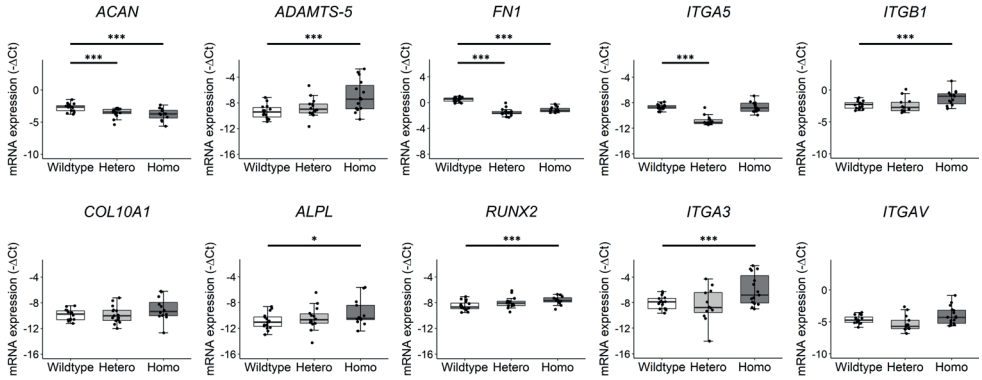



**Figure 6 | *FN1* mutation results in aberrant chondrocyte gene expression.** Box plots of –ΔCt values of *ACAN*, *ADAMTS-5*, *FN1*, *ITGA5*, *ITGB1*, *COL10A1*, *ALPL*, *RUNX2*, *ITGA3*, *ITGAV* in wild-type, C518F *FN1* heterozygous (Hetero) and homozygous (Homo) chondrogenic pellets. –ΔCt values shown were corrected for *COL1A1* and *COL2A1* expression levels to correct for the difference in efficiency of chondrogenic differentiation between genotypes. The box plots represent 25th, 50th and 75th percentiles, and whiskers extend to 1.5 times the interquartile range. Individual samples are depicted by black dots in each graph. P values were determined by generalized estimation equations, with gene expression levels (-ΔCt) as dependent variable and genotype as factor. * P < 0.05, ** P < 0.01, *** P < 0.005.

**Table 2 | Differences in gene expression between C518F *FN1* homozygous and heterozygous human induced chondroprogenitor-derived chondrogenic pellets compared to wild-type after 5 weeks of chondrogenesis.**

| | Hetero vs Wt* | | | | Homo vs Wt* | | | |
|---|---|---|---|---|---|---|---|---|
| **Gene** | **FC** | **Beta** | **SE** | **P value** | **FC** | **Beta** | **SE** | **P value** |
| *ACAN* | 0.57 | -0.81 | 0.23 | $4.52 \times 10^{-4}$ | 0.57 | -1.1 | 0.24 | $2.62 \times 10^{-4}$ |
| *ADAMTS-5* | 1.55 | 0.63 | 0.48 | $1.93 \times 10^{-1}$ | 5.14 | 2.36 | 0.68 | $5.45 \times 10^{-4}$ |
| *ALPL* | 1.41 | 0.5 | 0.56 | $3.79 \times 10^{-1}$ | 2.45 | 1.3 | 0.63 | $4.03 \times 10^{-2}$ |
| *COL10A1* | 1.03 | 0.05 | 0.41 | $9.11 \times 10^{-1}$ | 1.86 | 0.9 | 0.56 | $1.09 \times 10^{-1}$ |
| *FN1* | 0.25 | -1.98 | 0.16 | $1.00 \times 10^{-4}$ | 0.33 | -1.59 | 0.13 | $1.00 \times 10^{-4}$ |
| *ITGA3* | 1.18 | -0.41 | 0.77 | $5.93 \times 10^{-1}$ | 3.91 | 1.97 | 0.65 | $2.51 \times 10^{-3}$ |
| *ITGA5* | 0.23 | -2.13 | 0.23 | $1.00 \times 10^{-4}$ | 1.01 | 0.02 | 0.24 | $9.34 \times 10^{-1}$ |
| *ITGAV* | 0.66 | -0.6 | 0.37 | $1.03 \times 10^{-1}$ | 1.53 | 0.61 | 0.37 | $1.03 \times 10^{-1}$ |
| *ITB1* | 0.92 | -0.12 | 0.33 | $7.20 \times 10^{-1}$ | 2.09 | 1.07 | 0.33 | $1.11 \times 10^{-3}$ |
| *RUNX2* | 1.42 | 0.5 | 0.28 | $6.75 \times 10^{-2}$ | 1.79 | 0.84 | 0.26 | $9.48 \times 10^{-4}$ |

*Fold change (FC), Beta value, standard error (SE), and P value for expression levels (-ΔCt) of 10 genes as determined by generalized estimation equations, with gene expression levels (-ΔCt) as dependent variable and genotype as factor, when comparing *FN1* heterozygous (Hetero) and *FN1* homozygous (Homo) pellets to wild-type (Wt) pellets.



## Discussion

In the current study, we identified a heterozygous missense mutation (c1,819 G> T; C518F) in *FN1* that is likely causal to the EO phenotype in an extended EO-OA family (**Figure 1**). By introducing this high-impact mutation in hiPSCs using CRISPR/Cas9 gene editing and using them in an established in vitro organoid cartilage model, we showed that the chondrogenic potential (**Figure 3**) and deposition of neo-cartilage (**Figure 4**) of the *FN1* mutant chondrocytes was decreased. Moreover, we demonstrated that the underlying pathogenic mechanism of the mutation was caused by a decreased binding of fibronectin to the surrounding ECM protein collagen type II (**Figure 5**). As a downstream effect, the formed neo-cartilage was prone to a hypertrophic and procatabolic (OA) state, as reflected by the changed chondrocyte gene expression pattern, particularly of *ADAMTS-5* and *RUNX2* (**Figure 6**). This is further confirmed by the up-regulation of *ITGA3* and *ITGB1*, which is comparable to an OA disease state of chondrocytes [37].

Since fibronectin is an important adhesion protein between ECM and chondrocytes via binding to the cell surface integrins, we advocate that the identified C518F mutation in fibronectin hampers essential interaction between chondrocytes and the ECM; hence, plasticity of chondrocytes to adequately respond to cues from the ECM is decreased. In osteoarthritic joints, it has been shown that fibronectin is degraded proteolytically into fragments (FN-fs), which amplify catabolic processes in articular cartilage [38-41]. These FN-fs have distinct activities due to different binding capacities and thereby lead to changes in communication

between the ECM and the chondrocyte. Therefore, it is tempting to speculate that the phenomenon of changed binding could be directly contributing to the onset of age-related OA in the population.

One of the findings in this study is that the *FN1* mutant hiCPCs compared to wild type had both reduced chondrogenic potential (**Figure 3**). These results are in line with previous findings showing that the presence of fibronectin matrix is essential for mesenchymal stromal cell condensation and chondrogenic differentiation [42, 43]. The results of our study therefore suggest that decreased binding between fibronectin and collagen type II may negatively affect these processes. To elucidate the exact mechanism of during differentiation, longitudinal analysis of the differentiation would have to be performed in future studies.

The less efficient formation of dense, round pellets as a result of the mutation (**Figure 3**), hence variable quantity of neo-cartilage formed in the *FN1* mutant pellets, was a confounding factor in the gene expression data analyses. We observed a lower amount of cartilage-producing cells relative to the total amount of cells in *FN1* mutated pellets compared with wild type, as well as a decrease in the deposition of neo-cartilage (**Figure 4**). As a result, an overall stepwise decrease in cartilage-specific gene expression was observed in the *FN1* mutant pellets. To adjust for the relatively lower amount of cartilage formed in the *FN1* mutant chondrogenic pellets, we used the cartilage-specific markers *COL2A1* and *COL1A1* as reference genes. We chose *COL2A1* and *COL1A1* for their typically high expression in cartilage, as well as their stability in ongoing OA pathophysiology as reflected by a previously published large RNA sequencing dataset of autologous cartilage [26], and that we showed to be specific markers for the neo-cartilage (**Figure 4A**). Although we are confident that the observed significant increase in *ADAMTS-5*, *ALPL*, and *RUNX2* and the significant decrease in *ACAN* and *FN1* expression reflected true biological effects (**Figure 6**), the accuracy of the fold changes provided may be affected. There are different possibilities to overcome this problem in the future. Recently, Adkar *et al.* [20] engineered a knock-in *COL2A1-GFP* reporter system in hiPSCs to purify chondroprogenitors, consequently improving chondrogenic capacity during differentiation. To be able to further investigate the effect of the mutation on chondrogenesis without the decrease in chondrogenic potential, the *COL2A1* reporter system could be used for our *FN1* OA disease model to circumvent this effect. On the other hand, single-cell RNA sequencing could identify the chondroprogenitor subpopulation in the heterogeneous pellets, so that gene expression analysis of only that subpopulation would be possible [31].

Notably, even after correcting for *COL2A1* and *COL1A1* as cartilage specific markers, *FN1* and *ACAN* expression was significantly decreased in *FN1* mutant pellets, likely affecting quality of deposited cartilage matrix. Moreover, *ITGA3* and *ITGB1* were significantly up-regulated in the *FN1* homozygous pellets. Since α3β1 integrin was shown to be up-regulated in OA chondrocytes,

the here observed switch from α5β1 to α3β1 integrins in *FN1* homozygous chondrocytes may reflect an unbeneficial disease state of the *FN1* homozygous chondrocytes [37]. Conversely, *ITGA5* was down-regulated only in the *FN1* heterozygous pellets, which could have been a direct result of the formation of wild type-mutant fibronectin dimers, thereby also potentially being responsible for the decreased *FN1* and *ACAN* expression. Recently, it has been shown that the fibronectin-α5β1 adhesion is essential for tissue regeneration in mice [44]. These data indicate that the mutation results in aberrant chondrocyte gene expression, potentially via a change in integrin subunit expression levels toward a more hypertrophic, catabolic state, possibly due to decreased binding between fibronectin and collagen type II. However, this should be confirmed by more investigative studies on protein level, e.g., by Western blot to quantify protein levels of the integrin subunits.

Previous studies regarding familial, high-impact *FN1* mutations have been associated with glomerulopathy with fibronectin deposits (GDND), which is a hereditary kidney disease with proteinuria, microscopic hematuria, and hypertension, as well as spondylometaphyseal dysplasia with "corner factures" [45, 46]. These mutations were located in multiple heparin-binding domains, which play an essential role in regulating fibronectin assembly into organized fibrils in ECM. When comparing the phenotypes associated to the *FN1* mutations, it is notable that the phenotype in our family is less severe since glomerulopathy results in renal failure and the spine and growth plates are affected in spondylometaphyseal dysplasia, resulting in, amongst others, shortened trunk and scoliotic posture. Apparently, a mutation in the gelatin-binding domain results in a more specific pathology involving only the joints. The previously identified mutations also often involved cysteine residues, confirming that proper folding of fibronectin, e.g., via disulfide bonds, is crucial for the binding of fibronectin to other matrix components. Together, these data highlight the importance of proper binding of fibronectin to collagen in articular cartilage.

Together, by combining exome sequencing and linkage analysis we demonstrated that the identified *FN1* mutation is causal to human primary OA at an age of onset around 20 to 40 years. Moreover, the mutation does not cause severe developmental aberrations; yet, it has sufficient impact on cartilage tissues to allow measurable functional in vitro effects. Furthermore, by precise genetic engineering of a registered hiPSC line without off-target effects (**Supplementary Figure 2**), we established both a hetero- and homozygous isogenic clone with the OA causing *FN1* mutation (**Figure 2**). By subsequently applying an established differentiation protocol able to produce biomimetic human neo-cartilage, we generated multiple reliable biological replicates on how aberrant function of fibronectin in cartilage is causal to OA onset. Hence, we are confident that our approach was able to create reliable data highly translating to the human in vivo situation while contributing to the societal need to reduce animal studies.

In conclusion, our approach highlights the immense potential of combining exome sequencing, hiPSCs, CRISPR/Cas9, and organoid disease modeling to uncover previously unknown underlying disease mechanisms that are readily extrapolated to novel therapeutic avenues in common, complex human genetic diseases. Particularly, we here show that the underlying pathogenic mechanism of the mutation was caused by a decreased binding of fibronectin to the surrounding ECM protein collagen type II. Moreover, as reflected by the changes in chondrocyte gene expression, particularly of *ADAMTS-5* and *RUNX2,* such impaired binding results in the cartilage becoming prone to enter an OA disease state. In addition, we observed a switch in expression levels of the main fibronectin-binding receptor integrin α5β1 to integrin α3β1, reflecting a more OA disease state of the chondrocytes. Together our work merits further exploration of fibronectin as potential target for therapeutic interventions. We advocate that restoring or maintaining proper binding between fibronectin and collagen type II should be the focus of such a quest.

## Acknowledgements

We thank all the members of the MolEpi Osteoarthritis group for valuable discussion and feedback. We thank specifically Rodrigo Coutinho de Almeida and Margo Tuerlings for providing the RNA sequencing data of lesioned and preserved OA cartilage and bone. We thank prof. dr. Christine Mummery for the initial help with the hiPSC work and dr. Manuel Gonçalves for the help with the design of the CRISPR/Cas9 components. Data were generated within the scope of the Medical Delta Programs Regenerative Medicine 4D: Generating Complex Tissues With Stem Cells and Printing Technology and improving Mobility with Technology.

**Funding:**

Dutch Scientific Research council NWO /ZonMW VICI scheme grant nr 91816631/528 (IM)
Dutch Arthritis Society grant nr DAF-16-1-405 & DAF-16-1-406 (IM)
US National Institutes of Health grant nr AG15768 (FG)

## References

1. Goldring, S.R. and M.B. Goldring, *Changes in the osteochondral unit during osteoarthritis: structure, function and cartilage-bone crosstalk*. Nat Rev Rheumatol, 2016. **12**(11): p. 632-644.
2. Conaghan, P.G., et al., *Osteoarthritis research priorities: a report from a EULAR ad hoc expert committee*. Ann Rheum Dis, 2014. **73**(8): p. 1442-5.
3. van der Kraan, P.M., et al., *Translation of clinical problems in osteoarthritis into pathophysiological research goals*. RMD Open, 2016. **2**(1): p. e000224.
4. Litwic, A., et al., *Epidemiology and burden of osteoarthritis*. Br Med Bull, 2013. **105**: p. 185-99.

5. Cirulli, E.T. and D.B. Goldstein, *Uncovering the roles of rare variants in common disease through whole-genome sequencing*. Nat Rev Genet, 2010. **11**(6): p. 415-25.
6. Rabbani, B., et al., *Next-generation sequencing: impact of exome sequencing in characterizing Mendelian disorders*. J Hum Genet, 2012. **57**(10): p. 621-32.
7. Goldstein, D.B., et al., *Sequencing studies in human genetics: design and interpretation*. Nat Rev Genet, 2013. **14**(7): p. 460-70.
8. Adkar, S.S., et al., *Genome Engineering for Personalized Arthritis Therapeutics*. Trends Mol Med, 2017. **23**(10): p. 917-931.
9. Takahashi, K., et al., *Induction of pluripotent stem cells from adult human fibroblasts by defined factors*. Cell, 2007. **131**(5): p. 861-72.
10. Meulenbelt, I., et al., *Strong linkage on 2q33.3 to familial early-onset generalized osteoarthritis and a consideration of two positional candidate genes*. Eur J Hum Genet, 2006. **14**(12): p. 1280-7.
11. Ramos, Y.F., et al., *A gain of function mutation in TNFRSF11B encoding osteoprotegerin causes osteoarthritis with chondrocalcinosis*. Ann Rheum Dis, 2015. **74**(9): p. 1756-62.
12. Min, J.L., et al., *Mutation analysis of candidate genes within the 2q33.3 linkage area for familial early-onset generalised osteoarthritis*. Eur J Hum Genet, 2007. **15**(7): p. 791-9.
13. Deelen, J., et al., *Genome-wide association study identifies a single major locus contributing to survival into old age; the APOE locus revisited*. Aging Cell, 2011. **10**(4): p. 686-98.
14. Riyazi, N., et al., *Evidence for familial aggregation of hand, hip, and spine but not knee osteoarthritis in siblings with multiple joint involvement: the GARP study*. Ann Rheum Dis, 2005. **64**(3): p. 438-43.
15. Keurentjes, J.C., et al., *Patients with severe radiographic osteoarthritis have a better prognosis in physical functioning after hip and knee replacement: a cohort-study*. PLoS One, 2013. **8**(4): p. e59500.
16. Ramos, Y.F., et al., *Genes involved in the osteoarthritis process identified through genome wide expression analysis in articular cartilage; the RAAK study*. PLoS One, 2014. **9**(7): p. e103056.
17. Meulenbelt, I., et al., *Meta-analyses of genes modulating intracellular T3 bio-availability reveal a possible role for the DIO3 gene in osteoarthritis susceptibility*. Ann Rheum Dis, 2011. **70**(1): p. 164-7.
18. Dambrot, C., et al., *Polycistronic lentivirus induced pluripotent stem cells from skin biopsies after long term storage, blood outgrowth endothelial cells and cells from milk teeth*. Differentiation, 2013. **85**(3): p. 101-9.
19. Bae, S., J. Park, and J.S. Kim, *Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases*. Bioinformatics, 2014. **30**(10): p. 1473-5.
20. Adkar, S.S., et al., *Step-Wise Chondrogenesis of Human Induced Pluripotent Stem Cells and Purification Via a Reporter Allele Generated by CRISPR-Cas9 Genome Editing*. Stem Cells, 2019. **37**(1): p. 65-76.
21. Rodriguez Ruiz, A., et al., *Cartilage from human-induced pluripotent stem cells: comparison with neo-cartilage from chondrocytes and bone marrow mesenchymal stromal cells*. Cell Tissue Res, 2021.
22. Farndale, R.W., D.J. Buttle, and A.J. Barrett, *Improved quantitation and discrimination of sulphated glycosaminoglycans by use of dimethylmethylene blue*. Biochim Biophys Acta, 1986. **883**(2): p. 173-7.
23. Bomer, N., et al., *Underlying molecular mechanisms of DIO2 susceptibility in symptomatic osteoarthritis*. Ann Rheum Dis, 2015. **74**(8): p. 1571-9.
24. Faas, F.G., et al., *Virtual nanoscopy: generation of ultra-large high resolution electron microscopy maps*. J Cell Biol, 2012. **198**(3): p. 457-69.
25. Kiezun, A., et al., *Deleterious alleles in the human genome are on average younger than neutral alleles of the same frequency*. PLoS Genet, 2013. **9**(2): p. e1003301.
26. Coutinho de Almeida, R., et al., *RNA sequencing data integration reveals an miRNA interactome of osteoarthritis cartilage*. Ann Rheum Dis, 2019. **78**(2): p. 270-277.
27. Tuerlings, M., et al., *RNA sequencing reveals interacting key determinants of osteoarthritis acting in subchondral bone and articular cartilage*. Arthritis Rheumatol, 2020.
28. Stoffels, J.M., C. Zhao, and W. Baron, *Fibronectin in tissue regeneration: timely disassembly of the scaffold is necessary to complete the build*. Cell Mol Life Sci, 2013. **70**(22): p. 4243-53.

29. Pankov, R. and K.M. Yamada, *Fibronectin at a glance*. J Cell Sci, 2002. **115**(Pt 20): p. 3861-3.
30. Vinod, E., P. Boopalan, and S. Sathishkumar, *Reserve or Resident Progenitors in Cartilage? Comparative Analysis of Chondrocytes versus Chondroprogenitors and Their Role in Cartilage Repair*. Cartilage, 2018. **9**(2): p. 171-182.
31. Dicks, A., et al., *Prospective isolation of chondroprogenitors from human iPSCs based on cell surface markers identified using a CRISPR-Cas9-generated reporter*. Stem Cell Res Ther, 2020. **11**(1): p. 66.
32. Su, X., et al., *CD146 as a new marker for an increased chondroprogenitor cell sub-population in the later stages of osteoarthritis*. J Orthop Res, 2015. **33**(1): p. 84-91.
33. Wang, S., et al., *RaptorX-Property: a web server for protein structure property prediction*. Nucleic Acids Res, 2016. **44**(W1): p. W430-5.
34. Hadler, N.M., et al., *Ultrastructure of a hyaluronic acid matrix*. Proc Natl Acad Sci U S A, 1982. **79**(2): p. 307-9.
35. Thyberg, J., *Electron microscopy of cartilage proteoglycans*. Histochem J, 1977. **9**(3): p. 259-66.
36. Loeser, R.F., *Integrins and chondrocyte-matrix interactions in articular cartilage*. Matrix Biol, 2014. **39**: p. 11-6.
37. Loeser, R.F., C.S. Carlson, and M.P. McGee, *Expression of beta 1 integrins by cultured articular chondrocytes and in osteoarthritic cartilage*. Exp Cell Res, 1995. **217**(2): p. 248-57.
38. Zack, M.D., et al., *Identification of fibronectin neoepitopes present in human osteoarthritic cartilage*. Arthritis Rheum, 2006. **54**(9): p. 2912-22.
39. Homandberg, G.A., R. Meyers, and D.L. Xie, *Fibronectin fragments cause chondrolysis of bovine articular cartilage slices in culture*. J Biol Chem, 1992. **267**(6): p. 3597-604.
40. Reed, K.S.M., et al., *Transcriptional response of human articular chondrocytes treated with fibronectin fragments: an in vitro model of the osteoarthritis phenotype*. Osteoarthritis Cartilage, 2020.
41. Homandberg, G.A., R. Meyers, and J.M. Williams, *Intraarticular injection of fibronectin fragments causes severe depletion of cartilage proteoglycans in vivo*. J Rheumatol, 1993. **20**(8): p. 1378-82.
42. Singh, P. and J.E. Schwarzbauer, *Fibronectin matrix assembly is essential for cell condensation during chondrogenesis*. J Cell Sci, 2014. **127**(Pt 20): p. 4420-8.
43. Singh, P. and J.E. Schwarzbauer, *Fibronectin and stem cell differentiation - lessons from chondrogenesis*. J Cell Sci, 2012. **125**(Pt 16): p. 3703-12.
44. Almonte-Becerril, M., et al., *Genetic abrogation of the fibronectin-alpha5beta1 integrin interaction in articular cartilage aggravates osteoarthritis in mice*. PLoS One, 2018. **13**(6): p. e0198559.
45. Lee, C.S., et al., *Mutations in Fibronectin Cause a Subtype of Spondylometaphyseal Dysplasia with "Corner Fractures"*. Am J Hum Genet, 2017. **101**(5): p. 815-823.
46. Castelletti, F., et al., *Mutations in FN1 cause glomerulopathy with fibronectin deposits*. Proc Natl Acad Sci U S A, 2008. **105**(7): p. 2538-43.

**Supplementary Table 1 | Radiographic and clinical abnormalities in members of a family with early-onset osteoarthritis.**

| Patient # | Sex | Age of onset | DIJ | PIJ | MCJ | CMC | Elbow | Shoulder | Hip | Knee | Foot | Ankle | CS | LS | Confirmed ROA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | F | 45 | + | + | | OA | | | | OA | | OA | | | Yes |
| 4 | F | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 6 | M | 40 | + | + | | | | | | OA | + | + | | | Yes |
| 7 | M | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 8 | F | 50 | + | + | | | | | | + | | | + | | No |
| 9 | F | 40 | + | + | | | | | | + | + | + | | | No |
| 10 | M | 43 | OA | OA | OA | | + | + | | + | | OA | | | Yes |
| 11 | M | 40 | + | + | | | | | | | + | OA | OA | OA | Yes |
| 12 | F | 39 | OA | OA | OA | | | | OA | + | + | + | | | Yes |
| 13 | M | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 14 | F | 28 | OA | OA | OA | | | | | OA | | OA | | OA | Yes |
| 15* | F | 30 | + | + | + | | | | | + | | | | | No |
| 16 | F | - | - | - | - | - | - | - | - | - | - | - | - | - | - |
| 17 | F | 23 | H | B | | OA | | - | - | OA | - | - | | | Yes |

F = female, M = male, OA = Radiographic osteoarthritis, Kellgren Lawrence grade 2 or higher in one or both joints, + = joint complaints, - = no x-ray available or no joint complaints, H = Heberden nodes, B = Bouchard nodes, DIJ = distal interphalangeal joints, PIJ = proximal interphalangeal joints, MCJ = metacarpophalangeal joint, CMC = carpometacarpal joint, CS = cervical spine, LS = lumbar spine, confirmed ROA = confirmed radiological OA.
* patient 15 is a monozygotic twin with patient 14, since there was no confirmed ROA, patient 15 was removed from the linkage analysis

**Supplementary Table 2 | Variant prioritization scheme applied after exome sequencing.**

| | Subject 10 |
|---|---|
| Total number of variants detected | 73,407 |
| Exonic including 5'UTR and 3'UTR | 20,062 |
| Missense, nonsense, readthrough, frame error, splice site | 7,907 |
| Novel | 1,032 |
| Missense damaging, nonsense, readthrough, frame error, splice site | 122 |

**Supplementary Table 3 | Genotyping of the detected *FN1* variant in familial osteoarthritis family members.**

| Patient # | Status | FN1 - C518F |
|---|---|---|
| 3 | Affected | C A |
| 4 | Unaffected | C C |
| 6 | Affected | C A |
| 7 | Unaffected | C C |
| 8 | Diagnostic uncertain | C C |
| 9 | Diagnostic uncertain | C C |
| 10 | Affected | C A |
| 11 | Affected | C A |
| 12 | Affected | C A |
| 13 | Unaffected | C C |
| 14 | Affected | C A |
| 16 | Unaffected | C C |
| 17 | Affected | C A |

**Supplementary Table 4 | Design of CRISPR/Cas9 system.** In the single-strand oligonucleotide template the introduced silent mutation to give a HincII restriction site is depicted in italic font and the restriction site is shown in grey, the wanted *FN1* mutation is depicted in bold font, and the mutated PAM site is depicted underlined.

| | |
|---|---|
| guide RNA | 5'- AACAAUGCACUGAUCUGUUU-3' |
| single-strand oligonucleotide template | 5'-CAG CAT GTG CCC CTC TTC ATG ACG CTT GTG GAA TGT GTC GTT CAC ATT GTA AGT GAT GTC *G*TC AAC AAT G**A**A CTG ATC TGT TTA CGA AAC AGG TGG GTG AGT GAG AAA CTT TTT AAA GTT CCA TTG ATG AAA GAA AAA AGG-3' |

**Supplementary Table 5 | Primer sequences to measure mRNA expression levels, amplify region of interest around *FN1* mutation (exon 11), and amplify off-target region.**

| Region of interest | Forward primer (5' --> 3') | Reverse primer (5' --> 3') |
|---|---|---|
| *ACAN* | AGAGACTCACACAGTCGAAACAGC | CTATGTTACAGTGCTCGCCAGTG |
| *ADAMTS-5* | GTGGTGAAGGTGGTGGTGCT | CTCATGGTCATCTCCCAGCTG |
| *ALPL* | CAAAGGCTTCTTCTTGCTGGTG | CCTGCTTGGCTTTTCCTTCA |
| *COL1A1* | GTGCTAAAGGTGCCAATGGT | ACCAGGTTCACCGCTGTTAC |
| *COL2A1* | CTACCCCAATCCAGCAAACGT | AGGTGATGTTCTGGGAGCCTT |
| *COL10A1* | GGCAACAGCATTATGACCCA | TGAGATCGATGATGGCACTCC |
| *FN1* | CCGACCAGAAGTTTGGGTTC | CACGACCATTCCCAACACAC |
| *GAPDH* | TGCCATGTAGACCCCTTGAAG | ATGGTACATGACAAGGTGCGG |
| *ITGA5* | GTCGGGGGCTTCAACTTAGAC | AGCACACTGACCCCGTCTG |
| *ITGA3* | TGCCTACAACTGGAAAGGAAAC | CTGCCTACCTGCATCGTGTA |
| *ITGAV* | TCTCTCGGGACTCCTGCTAC | AAGAAACATCCGGGAAGACGC |
| *ITGB1* | AGCGAAGGCATCCCTGAAAG | AATGTCTACCAACACGCCCT |
| *MCAM* | GGAAGTCACCGTCCCTGTTT | ATGCTGAAGTGTGGTGGAGG |
| *PDGRFβ* | GGACATACCCCCGCAAAGAA | CTCTCCGTCACATTGCAGGT |
| *SDHA* | TGGGAACAAGAGGGCATCTG | GCCTACCACCACTGCATCAA |
| *RUNX2* | CTGTGGTTACTGTCATGGCG | AGGTAGCTACTTGGGGAGGA |
| *FN1* (exon 11) | TGGCTTCTCACAGGAAAGTGT | CCAGACGACTACCAATATCGAG |
| gRNA off-target | AGTGAAGGAGATGTGGCTGC | AAGAGCAACGGGGTGAACAA |

**Supplementary Table 6 | List of damaging missense variants in an affected family member of an early-onset OA family after our applied prioritization scheme.** Base change is base of the sense strand of the DNA.

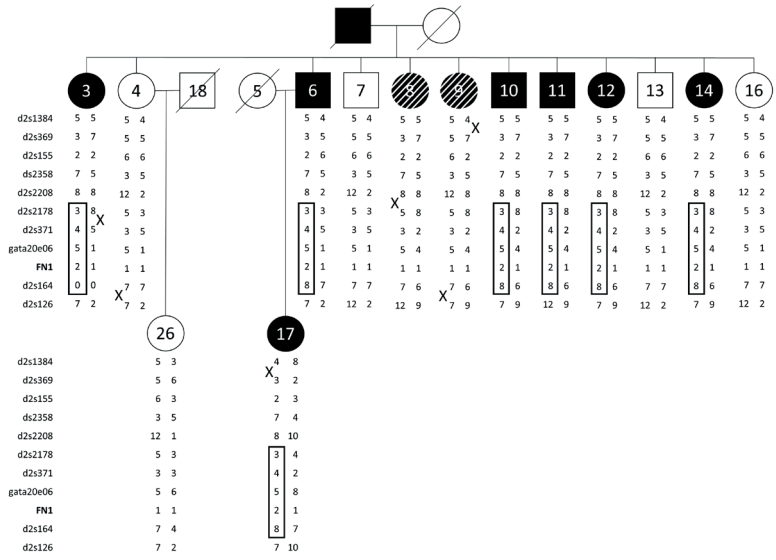
| Chrom. | Gene Name | Ensembl ID | Base change | Base position | Codon change | Substitution |
|---|---|---|---|---|---|---|
| 1 | CPSF3L | ENSG00000127054 | A/G | 1250997 | GTA-GcA | V46A |
| 1 | RBMXL1 | ENSG00000213516 | T/G | 89448812 | GAT-GcT | D233A |
| 1 | ITGA10 | ENSG00000143127 | C/T | 145530939 | ACG-AtG | T224M |
| 1 | NBPF14 | ENSG00000122497 | C/A | 148004625 | GTG-tTG | V897L |
| 1 | ADAMTSL4 | ENSG00000143382 | G/A | 150531177 | GAC-aAC | D871N |
| 1 | LINGO4 | ENSG00000213171 | G/A | 151774330 | TCC-TtC | S284F |
| 1 | SELE | ENSG00000007908 | A/T | 169697065 | GTG-GaG | V428E |
| 1 | C1orf129 | ENSG00000117501 | G/A | 170961360 | GTG-aTG | V362M |
| 1 | HEATR1 | ENSG00000119285 | A/G | 236722383 | GTG-GcG | V1527A |
| 1 | ZP4 | ENSG00000116996 | G/A | 238050846 | ACC-AtC | T190I |
| 2 | GPR113 | ENSG00000173567 | G/A | 26534245 | GCG-GtG | A784V |
| 2 | MOBKL1B | ENSG00000114978 | A/G | 74399827 | TCT-cCT | S23P |
| 2 | DNAH6 | ENSG00000115423 | A/T | 84899498 | ATG-tTG | M2168L |
| 2 | MMADHC | ENSG00000168288 | C/T | 150436081 | GGT-GaT | G79D |
| 2 | ALS2 | ENSG00000003393 | A/C | 202590176 | TAT-gAT | Y1084D |
| 2 | FN1 | ENSG00000115414 | C/A | 216285518 | TGC-TtC | C518F |
| 2 | ABCB6 | ENSG00000115657 | G/C | 220075737 | CGT-gGT | R688G |
| 3 | DNAH12 | ENSG00000174844 | G/C | 57401178 | GCA-GgA | A1924G |
| 3 | NEK11 | ENSG00000114670 | G/C | 130748679 | GTC-cTC | V43L |
| 3 | PAK2 | ENSG00000180370 | G/C | 196529902 | CAG-CAc | Q101H |
| 4 | MFSD10 | ENSG00000109736 | A/C | 2935381 | TAT-gAT | Y61D |
| 4 | MAN2B2 | ENSG00000013288 | T/A | 6596307 | ATG-AaG | M302K |
| 4 | TMPRSS11 | ENSG00000198092 | G/A | 68930465 | TCT-TtT | S318F |
| 4 | FAM13A | ENSG00000138640 | A/C | 89660254 | GTG-GgG | V830G |
| 4 | GPRIN3 | ENSG00000185477 | T/G | 90170862 | ACC-cCC | T134P |
| 4 | NHEDC1 | ENSG00000164037 | A/C | 103832643 | CTG-CgG | L294R |
| 4 | SEC24D | ENSG00000150961 | G/A | 119678880 | CCA-tCA | P473S |
| 4 | FRG1 | ENSG00000109536 | C/T | 190874234 | CCT-tCT | P91S |
| 4 | FRG1 | ENSG00000109536 | C/A | 190878563 | GCT-GaT | A148D |
| 5 | CDC20B | ENSG00000164287 | G/A | 54423155 | CGG-tGG | R307W |
| 5 | PPIC | ENSG00000168938 | C/T | 122361517 | GGC-aGC | G158S |
| 5 | LECT2 | ENSG00000145826 | C/T | 135287013 | GGA-GaA | G63E |
| 5 | PCDHB8 | ENSG00000120322 | G/A | 140558064 | GGG-GaG | G150E |
| 5 | ZNF354C | ENSG00000177932 | A/G | 178507004 | TAT-TgT | Y524C |
| 6 | ZNF318 | ENSG00000171467 | G/A | 43307399 | CCA-CtA | P1446L |
| 7 | SDK1 | ENSG00000146555 | G/A | 4002375 | GGA-aGA | G441R |
| 7 | HBP1 | ENSG00000105856 | G/C | 106836335 | CGC-CcC | R375P |
| 7 | SLC26A3 | ENSG00000091138 | T/G | 107427841 | AAA-AAc | K283N |
| 7 | CFTR | ENSG00000001626 | C/T | 117188750 | TCT-TtT | S422F |
| 7 | OR9A2 | ENSG00000179468 | A/C | 142723621 | GTT-GgT | V200G |
| 7 | GALNT11 | ENSG00000178234 | T/C | 151810377 | GTA-GcA | V376A |
| 7 | MLL3 | ENSG00000055609 | A/G | 151919751 | TGT-cGT | C1114R |
| 8 | PABPC1 | ENSG00000070756 | G/A | 101721812 | CGC-tGC | R374C |
| 8 | PABPC1 | ENSG00000070756 | T/C | 101721817 | GAA-GgA | E372G |
| 8 | PABPC1 | ENSG00000070756 | C/A | 101721839 | GTA-tTA | V365L |
| 8 | FER1L6 | ENSG00000214814 | C/T | 124989741 | CGG-tGG | R319W |
| 8 | PLEC | ENSG00000178209 | A/G | 145003292 | CTG-CcG | L1217P |
| 8 | ZNF251 | ENSG00000198169 | T/C | 145979657 | GAG-GgG | E28G |
| 9 | EXOSC3 | ENSG00000107371 | G/C | 37782053 | CGA-gGA | R186G |

| Chrom. | Gene Name | Ensembl ID | Base change | Base position | Codon change | Substitution |
|---|---|---|---|---|---|---|
| 9 | NOTCH1 | ENSG00000148400 | T/C | 139390627 | AGC-gGC | S2522G |
| 10 | IL15RA | ENSG00000134470 | G/A | 6008296 | ACG-AtG | T32M |
| 10 | ARMC3 | ENSG00000165309 | A/G | 23250814 | GAT-GgT | D180G |
| 10 | PLAU | ENSG00000122861 | C/T | 75673437 | CGG-tGG | R165W |
| 10 | LIPM | ENSG00000173239 | C/A | 90579993 | ACT-AaT | T336N |
| 10 | TM9SF3 | ENSG00000077147 | C/A | 98292928 | TGT-TtT | C402F |
| 10 | CPXM2 | ENSG00000121898 | C/G | 125521390 | GGA-GcA | G592A |
| 11 | OR51S1 | ENSG00000176922 | C/A | 4870162 | GCC-tCC | A93S |
| 11 | DNHD1 | ENSG00000179532 | C/T | 6585157 | CGC-tGC | R3363C |
| 11 | OLFML1 | ENSG00000183801 | C/T | 7509386 | ACG-AtG | T53M |
| 11 | SAA4 | ENSG00000148965 | T/A | 18254071 | GAC-GtC | D34V |
| 11 | LGR4 | ENSG00000205213 | A/T | 27390683 | TTT-TTa | F529L |
| 11 | BTBD18 | ENSG00000233436 | A/G | 57513573 | TTC-cTC | F58L |
| 11 | ATM | ENSG00000149311 | T/C | 108214007 | ATT-AcT | I2776T |
| 11 | CRYAB | ENSG00000109846 | C/A | 111781098 | GTG-tTG | V93L |
| 12 | FOXM1 | ENSG00000111206 | T/C | 2967952 | AAT-AgT | N753S |
| 12 | PZP | ENSG00000126838 | G/A | 9307415 | CGC-tGC | R1191C |
| 12 | SLC11A2 | ENSG00000110911 | G/C | 51384724 | CGG-gGG | R477G |
| 12 | NUAK1 | ENSG00000074590 | C/A | 106460782 | CGC-CtC | R595L |
| 13 | CCNA1 | ENSG00000133101 | G/A | 37014187 | CGA-CaA | R322Q |
| 13 | KIAA1704 | ENSG00000133114 | G/T | 45589633 | GGT-tGT | G33C |
| 14 | HEATR5A | ENSG00000129493 | G/A | 31762650 | TCT-TtT | S1995F |
| 14 | RALGAPA1 | ENSG00000174373 | A/G | 36103797 | CTC-CcC | L1534P |
| 14 | MTHFD1 | ENSG00000100714 | A/C | 64884707 | ACC-cCC | T194P |
| 14 | PLEKHG3 | ENSG00000126822 | C/T | 65198849 | CAC-tAC | H332Y |
| 14 | DIO2 | ENSG00000211448 | T/C | 80669078 | GAG-GgG | E259G |
| 14 | DDX24 | ENSG00000089737 | A/C | 94521529 | GTC-GgC | V664G |
| 14 | CDC42BPB | ENSG00000198752 | C/G | 103410810 | GCT-cCT | A1276P |
| 14 | PACS2 | ENSG00000179364 | A/G | 105821407 | AGG-gGG | R106G |
| 15 | HERC2 | ENSG00000128731 | G/A | 28483865 | CGC-tGC | R1211C |
| 15 | ARHGAP11 | ENSG00000198826 | G/C | 32921855 | GAT-cAT | D333H |
| 15 | RYR3 | ENSG00000198838 | G/T | 34152828 | GAC-tAC | D4774Y |
| 15 | KIF23 | ENSG00000137807 | C/A | 69709740 | CCA-aCA | P34T |
| 15 | AKAP13 | ENSG00000170776 | G/T | 86270411 | GTC-tTC | V2317F |
| 16 | RGS11 | ENSG00000076344 | T/C | 323545 | CAG-CgG | Q152R |
| 16 | TBL3 | ENSG00000183751 | A/G | 2024410 | AGC-gGC | S75G |
| 16 | TSC2 | ENSG00000103197 | T/C | 2130361 | CTG-CcG | L1198P |
| 16 | ACSM5 | ENSG00000183549 | C/T | 20429546 | CGG-tGG | R124W |
| 16 | ABCC11 | ENSG00000121270 | C/T | 48218359 | GTG-aTG | V1084M |
| 16 | SALL1 | ENSG00000103449 | C/T | 51171034 | GTC-aTC | V1322I |
| 16 | GLG1 | ENSG00000090863 | T/C | 74514287 | GAC-GgC | D560G |
| 16 | KIAA1609 | ENSG00000140950 | A/G | 84531559 | TTC-TcC | F45S |
| 17 | FAM57A | ENSG00000167695 | G/A | 644611 | CGG-CaG | R160Q |
| 17 | ALOX12B | ENSG00000179477 | T/C | 7977033 | TAT-TgT | Y566C |
| 17 | CCDC144A | ENSG00000170160 | G/A | 16664922 | GAA-aAA | E1186K |
| 17 | LGALS9C | ENSG00000171916 | G/A | 18397608 | CGC-CaC | R333H |
| 17 | TRAF4 | ENSG00000076604 | C/T | 27076578 | CGG-tGG | R466W |
| 17 | SLFN11 | ENSG00000172716 | A/G | 33680081 | ATT-AcT | I667T |
| 17 | CDC27 | ENSG00000004897 | A/C | 45216160 | GTT-GgT | V550G |

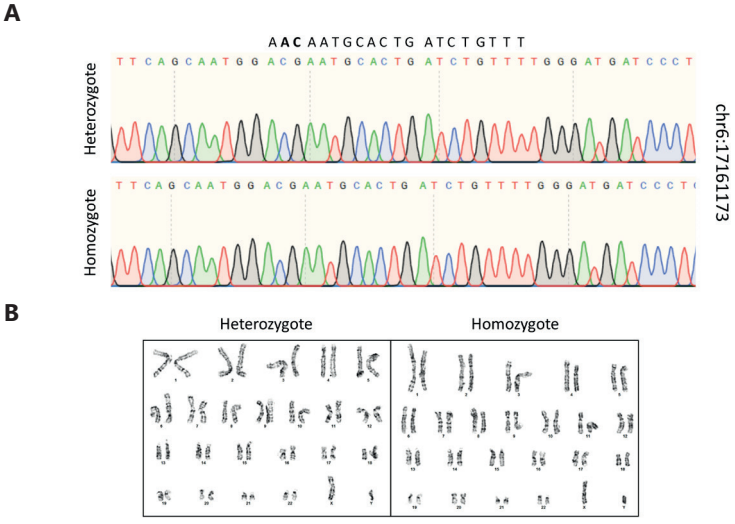
| Chrom. | Gene Name | Ensembl ID | Base change | Base position | Codon change | Substitution |
|---|---|---|---|---|---|---|
| 17 | CDC27 | ENSG00000004897 | G/A | 45216182 | CTT-tTT | L543F |
| 17 | ABCC3 | ENSG00000108846 | G/A | 48762223 | GGG-aGG | G1423R |
| 17 | PYCR1 | ENSG00000183010 | G/A | 79893008 | CGG-tGG | R112W |
| 18 | DSG1 | ENSG00000134760 | C/T | 28934644 | CCT-tCT | P829S |
| 19 | MATK | ENSG00000007264 | C/G | 3779073 | GGC-cGC | G372R |
| 19 | SH2D3A | ENSG00000125731 | T/C | 6755228 | AGG-gGG | R199G |
| 19 | XAB2 | ENSG00000076924 | A/G | 7685456 | TGC-cGC | C691R |
| 19 | HAUS8 | ENSG00000131351 | C/T | 17170885 | GGA-aGA | G83R |
| 19 | ISYNA1 | ENSG00000105655 | C/T | 18546943 | GCG-aCG | A90T |
| 19 | ZNF98 | ENSG00000197360 | T/C | 22574988 | TAC-TgC | Y350C |
| 19 | MEGF8 | ENSG00000105429 | T/C | 42855627 | TGG-cGG | W968R |
| 19 | ZNF285 | ENSG00000062370 | C/T | 44890654 | GAG-aAG | E585K |
| 19 | ZNF285 | ENSG00000062370 | G/C | 44891010 | GCG-GgG | A466G |
| 19 | PPP1R13L | ENSG00000104881 | G/A | 45885802 | CCG-tCG | P385S |
| 19 | DMWD | ENSG00000185800 | T/C | 46290117 | AAG-gAG | K213E |
| 19 | ZNF845 | ENSG00000213799 | T/C | 53856761 | TGT-cGT | C861R |
| 19 | ZNF813 | ENSG00000198346 | T/C | 53995143 | TGT-cGT | C553R |
| 19 | ZNF543 | ENSG00000178229 | C/T | 57839675 | CCT-CtT | P282L |
| 20 | XKR7 | ENSG00000101321 | C/G | 30556211 | GCG-GgG | A78G |
| 21 | TSPEAR | ENSG00000175894 | T/C | 45941975 | AGT-gGT | S453G |
| 22 | SOX10 | ENSG00000100146 | T/C | 38374107 | GAG-GgG | E155G |
| 22 | L3MBTL2 | ENSG00000100395 | G/C | 41615519 | GCC-cCC | A233P |
| 22 | SCUBE1 | ENSG00000159307 | C/T | 43604155 | CGC-CaC | R886H |
| X | MAGEA6 | ENSG00000197172 | A/C | 151870145 | ATT-cTT | I279L |

Chrom. = chromosome.

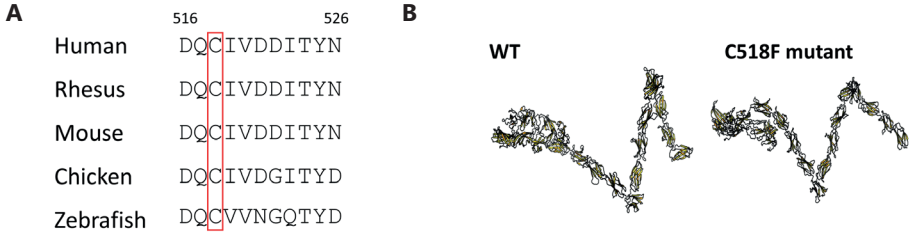
## Supplementary Figures



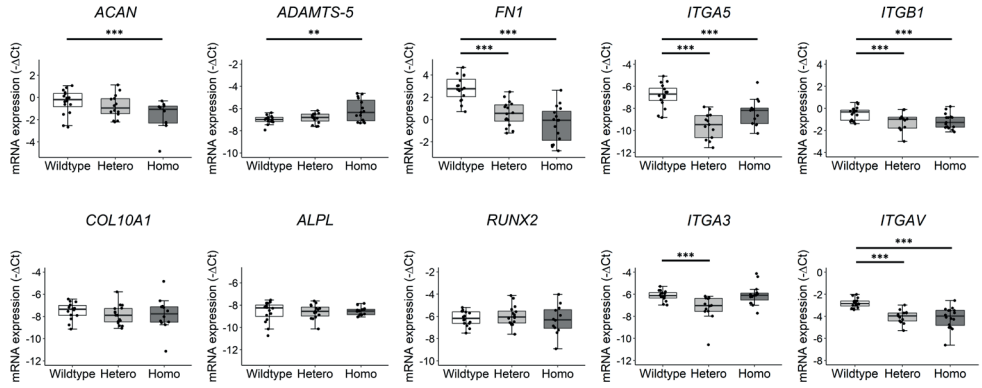

**Supplementary Figure 1 | Pedigree of investigated early-onset osteoarthritis family.** Co-segregating haplotypes among affected individuals are depicted as squared boxes and recombination as crosses. Individual 10 was used for exome sequencing. Circles represent females, squares represent males. Closed and open shapes represent affected and healthy individuals, respectively; striped shapes represent uncertain phenotypes.




**Supplementary Figure 2 | Quality control *FN1* mutant iPSC clones.** (**A**) Sanger sequencing results of in silico predicted off-target site at chr6:17161173 allowing 2 mismatches for the heterozygote and homozygote FN1 clone. Sequence of guide RNA is depicted above, where mismatches are depicted in bold. (**B**) Karyotypes of isogenic iPSC lines heterozygous and homozygous for FN1 mutation, showing absence of chromosomal abnormalities.




**Supplementary Figure 3 | Investigation of C518F fibronectin mutation.** (**A**) Alignment of the amino acid sequence around C518 of *FN1*, where the identified mutation is located. Mutated residue is outlined in red. (**B**) Structure prediction of wild-type (WT) and C518F mutant fibronectin by RaptorX of the 1,980 N-terminal amino acid residues.




**Supplementary Figure 4 | Box plots of –ΔCt values of *ACAN*, *ADAMTS-5*, *FN1*, *ITGA5*, *ITGB1*, *COL10A1*, *ALPL*, *RUNX2*, *ITGA3*, *ITGAV* in wild-type, C518F *FN1* hetero- (Hetero) and homozygous (Homo) chondrogenic pellets.** –ΔCt values shown were corrected for *GAPDH* and *SDHA* expression levels. The box plots represent 25th, 50th and 75th percentiles, and whiskers extend to 1.5 times the interquartile range. Individual samples are depicted by black dots in each graph. P values were determined by generalized estimation equations, with gene expression levels (-ΔCt) as dependent variable and genotype as factor. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.005$.

# Chapter 4

## Identification and functional characterization of imbalanced osteoarthritis-associated fibronectin splice variants

Marcella van Hoolwerff[1], Margo Tuerlings[1], Imke J.L. Wijnen[1], H. Eka D. Suchiman[1], Davy Cats[2], Hailiang Mei[2], Rob G.H.H. Nelissen[3], Henrike M.J. van der Linden – van der Zwaag[3], Yolande F.M. Ramos[1], Rodrigo Coutinho de Almeida[1], Ingrid Meulenbelt[1]

[1] Dept. of Biomedical Data Sciences, Section Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands
[2] Sequence Analysis Support Core, Leiden University Medical Center, Leiden, The Netherlands
[3] Dept. of Orthopaedics, Leiden University Medical Center, Leiden, The Netherlands

Rheumatology (Oxford). 2022 May 9:keac272
DOI: 10.1093/rheumatology/keac272

## Abstract

**Objective.** To identify *FN1* transcripts associated with osteoarthritis pathophysiology and investigate the downstream effects of modulating *FN1* expression and relative transcript ratio.

**Methods.** *FN1* transcriptomic data was obtained from our previously assessed RNA sequencing dataset of lesioned and preserved osteoarthritic cartilage samples from the Research osteoArthritis Articular Cartilage (RAAK) study. Differential transcript expression analysis was performed on all 27 *FN1* transcripts annotated in the Ensembl database. Human primary chondrocytes were transduced with lentiviral particles containing short hairpin RNA (shRNA) targeting full-length *FN1* transcripts or non-targeting shRNA. Subsequently, matrix deposition was induced in our 3D in vitro neo-cartilage model. Effects of changes in the *FN1* transcript ratio on sulfated glycosaminoglycan (sGAG) deposition were investigated by Alcian blue staining and dimethylmethylene blue assay. Moreover, gene expression levels of 17 cartilage-relevant markers were determined by reverse transcription quantitative polymerase chain reaction.

**Results.** We identified 16 *FN1* transcripts differentially expressed between lesioned and preserved cartilage. *FN1-208*, encoding migration-stimulating factor, was the most significantly differentially expressed protein coding transcript. Down-regulation of full-length *FN1* and a concomitant increased *FN1-208* ratio resulted in decreased sGAG deposition as well as decreased *ACAN* and *COL2A1* and increased *ADAMTS-5*, *ITGB1*, and *ITGB5* gene expression levels.

**Conclusion.** We show that full-length *FN1* down-regulation and concomitant relative *FN1-208* up-regulation was unbeneficial for deposition of cartilage matrix, likely due to decreased availability of the classical arginine-glycine-aspartate (RGD) integrin-binding site of fibronectin.

## Introduction

Currently, osteoarthritis (OA) is the most prevalent degenerative joint disease worldwide, associated with a high societal and economic burden [1]. A general hallmark of OA is the degeneration of articular cartilage [2, 3]. To date, no effective treatment to reverse or slow down the disease is available, except pain relief medication and joint replacement surgery. Therefore, more insight into the underlying pathophysiology of OA is necessary for the development of druggable targets.

In this regard, transcriptome-wide analyses of cartilage have been performed to identify underlying biological mechanisms driving OA [4-6]. Specifically, among the highest expressed and significantly up-regulated genes in affected OA cartilage is *fibronectin* (*FN1*) [4, 6]. *FN1* encodes a high molecular weight, dimeric glycoprotein that in articular cartilage is deposited by chondrocytes and mostly localized in the pericellular matrix directly surrounding the chondrocytes [7]. Fibronectin mediates a wide variety of cellular interactions with the extracellular matrix (ECM) by binding to matrix proteins via multiple binding domains, as well as interactions with chondrocytes via integrins that mediate intracellular signaling. The main integrin-binding domain of fibronectin is the arginine-glycine-aspartate (RGD) motif, which binds multiple integrin heterodimers, including the classic fibronectin receptor integrin α5β1 [8]. Recently, it has been shown that fibronectin-α5β1 adhesion is critical for cartilage regeneration in mice [9]. More recently, our group identified a high-impact mutation in the gelatin-binding domain of *FN1* in an early-onset OA family, resulting in decreased binding capacity of fibronectin to collagen type II [10]. Furthermore, fibronectin can be degraded by proteases and the resulting fibronectin fragments are known to amplify catabolic processes in the articular cartilage [11, 12]. Taken together, these studies show that proper binding of fibronectin to both ECM components and integrins is important for cartilage homeostasis.

The fibronectin gene is known to give rise to 20 different protein coding transcripts by virtue of alternative splicing as well as multiple non-protein coding transcripts [13]. Alternative splicing occurs at three major sites, called extra domain A (EDA), extra domain B (EDB), and variable region (V) [14, 15]. Splicing at the EDA and EDB domains results in inclusion or exclusion of one exon, whereas splicing at the V region can occur at multiple splice sites [14]. This splicing variation provides cells with the capacity to generate large numbers of protein isoforms with different binding properties to precisely alter the ECM composition in a developmental and tissue-specific manner. As a result, each isoform has a unique function in cell-ECM interactions [8]. Among *FN1* splice variants is the intact 70 kDa N-terminus of the full-length protein, also known as migration-stimulating factor (MSF) [16]. MSF includes the heparin- and gelatin-binding domains, but does not have the classical RGD integrin-binding domain. Previously, $EDB^+$, $EDB^-$, $EDA^-$ and $V^+$ transcript variants were shown to be present in multiple joint tissues, while $EDA^+$ transcript variants were rarely detected [17].

It is still unknown which specific *FN1* transcripts may be involved in the response to a healthy or disease state of cartilage, as well as what the effect of changed *FN1* expression is in OA cartilage. Therefore, we aimed to identify *FN1* transcripts associated with the OA process by assessing our previously published RNA sequencing dataset with paired lesioned and preserved OA cartilage samples [4]. Subsequently, the downstream effects of modulating *FN1* expression and the transcript ratio was investigated in our established human 3D in vitro OA cartilage model.

# Materials and Methods

## Sample description

Macroscopically lesioned and preserved articular cartilage samples were obtained from patients who underwent joint replacement surgery due to OA in the Research Osteoarthritis and Articular Cartilage (RAAK) study, as described previously [18]. The RAAK study was approved by the medical ethics committee of the Leiden University Medical Center (P08.239/P19.013), and written informed consent was obtained from all participants. In the current study, previously assessed RNA-seq data of 101 cartilage samples were used, of which 35 paired samples between lesioned and preserved (25 knees, 7 hips) [4]. The replication cohort consisted of an additional 10 paired cartilage samples (5 knees, 5 hips). Primary articular chondrocytes obtained from knee replacement surgeries of six participants of the RAAK study were isolated and cultured to perform lentiviral transduction. For all sample characteristics, see **Supplementary Table 1**.

## RNA sequencing

Total RNA isolation from articular cartilage, sequencing, and quality control was performed as previously described [4]. Detailed information on the alignment, mapping and filtering is available in the **Supplementary Methods**.

## Differential expression analysis and replication

Differential expression analysis was performed between lesioned and preserved OA cartilage samples. Results were validated by means of visualizing exon count data and replicated by means of reverse-transcription quantitative polymerase chain reaction (RT-qPCR). More detailed information is available in the **Supplementary Methods**.

## Lentiviral production and cell culture

For knockdown experiments, the pLKO.1-puro vector from the Sigma-Aldrich Mission short hairpin RNA (shRNA) library targeting all full-length protein coding transcripts of *FN1* (TRCN0000286357) and non-targeting control virus particles (SHC002) were kindly provided by Martijn Rabelink (Dept. of Cell & Chemical Biology, Leiden University Medical

Center, Leiden, The Netherlands). Detailed information on the lentiviral production and chondrocyte cell culture and transduction is available in the **Supplementary Methods**. In vitro chondrogenesis was induced as previously described [18]. Neo-cartilage pellets and medium were collected following 3 days of chondrogenesis.

### Enzyme-linked immunosorbent assay

Culture medium of neo-cartilage pellets of primary chondrocytes transduced with non-targeting shRNA (control) and *FN1* targeting shRNA was collected following 3 days of chondrogenesis. Fibronectin concentration was determined using the Human Fibronectin Enzyme-linked Immunosorbent assay (ELISA) kit (Thermo Fisher Scientific, Vienna, Austria) according to manufacturer's protocol. The absorbance was measured at 450 nm in a microplate reader (Spectramax iD3, Molecular Devices, San Jose, CAL, USA).

### Histology and immunohistochemistry

Neo-cartilage pellets were fixed in 4% formaldehyde overnight and stored in 70% ethanol at 4 °C. Detailed information on histology and immunohistochemistry is available in the **Supplementary Methods**.

### RNA isolation and relative gene expression analysis

Two neo-cartilage pellets were pooled in 200 µl Trizol reagent (Thermo Fisher Scientific, Carlsbad, CA, USA) and homogenized using micro pestles. Further details on RNA isolation and gene expression analysis is available in the **Supplementary Methods**. Primer sequences are shown in **Supplementary Table 2**.

### Sulfated glycosaminoglycan measurement

The sulfated glycosaminoglycan (sGAG) content in the neo-cartilage pellets was measured with the 1,9-dimethylmethylene blue (DMMB) assay [19]. Pellets were digested with 200 µl papain from papaya (Sigma-Aldrich, Zwijndrecht, The Netherlands) at 60 °C overnight. Shark chondroitin sulfate (Sigma-Aldrich, Zwijndrecht, The Netherlands) was used as a reference standard. The absorbance was measured at 525 and 595 nm in a microplate reader (Spectramax iD3, Molecular Devices, San Jose, CA, USA).

### *FN1* downstream interactions

To identify potential new *FN1* downstream interactions, Pearson correlations were calculated between *FN1* and the previously reported significantly differentially expressed genes (N = 2,387) [4] in the same lesioned (N = 44) and preserved (N = 57) OA cartilage samples (**Supplementary Table 1D**). Genes with $|r| \geq 0.8$ were analyzed for enrichment in protein-protein interactions with the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING, version 11.0) [20].

### Statistical analysis

Statistical analyses were performed using SPSS version 25 (IBM). Data are either shown as mean ± SD or box plots, representing the 25$^{th}$, 50$^{th}$ and 75$^{th}$ percentiles and whiskers representing the lowest and highest data point lying within 1.5 times the interquartile range. Individual samples are depicted by dots in each box plot. The reported P values of the lentiviral experiments were determined by applying a generalized estimating equations (GEE) to the experimental read-out to adjust for dependencies of the different donors [21]. We followed a linear GEE model, with the read-out data as dependent variable, and group and donor as covariate and exchangeable working matrix: Read-out ~ Group + (1|Donor). P values < 0.05 were considered statistically significant. * P < 0.05, ** P < 0.01, *** P < 0.005.

### Data availability

The processed dataset generated and the code to reproduce the differential expression analysis is available on https://git.lumc.nl/mvanhoolwerff/fn1-transcripts.

## Results

### Characterization of *FN1* transcripts in OA cartilage

To characterize the *FN1* transcripts in cartilage, we used our previously assessed RNA-seq data on 35 paired samples (28 knees, 7 hips) of lesioned and preserved OA cartilage (**Supplementary Table 1A**) [4]. Our in-house pipeline was applied to obtain transcriptomic *FN1* data from the Ensembl database, which has 27 *FN1* transcripts annotated (**Supplementary Figure 1**). To robustly detect the *FN1* transcripts expressed in OA cartilage, a cut-off was applied in the lesioned and preserved cartilage samples separately of an average of ≥ 10 counts per three samples. As a result, 22 transcripts were found to be expressed in cartilage, represented as the mean of normalized counts corrected for transcript length and sequencing depth, as shown in **Table 1**. Notably, the highest expressed protein coding transcripts were full-length transcripts *FN1-211* (base mean counts = 430,585.6; quartile 4), which is EDB$^-$, EDA$^-$, and V$^+$, and *FN1-209* (base mean counts = 120,949.2; quartile 4), which is EDB$^+$, EDA$^-$, and V$^+$. Expression levels of *FN1-211* and *FN1-209* represent 39.4% and 21.4%, respectively of total base mean counts, indicating that these were the main protein coding *FN1* transcripts transcribed in OA cartilage. Moreover, among the top 10 highest expressed transcripts, 2 are classified as retained introns and therefore not protein coding, namely *FN1-227* (base mean counts = 50,673.0; quartile 4), and *FN1-225* (base mean counts = 33,536.1; quartile 3), suggesting that these non-protein coding transcripts may be functional in OA cartilage.

**Table 1 | Expression levels of the 22 *FN1* transcripts that were robustly expressed in lesioned and preserved osteoarthritic cartilage samples.**

| Ensembl ID | Name | Biotype | baseMean | Quartile | EDA | EDB | V |
|---|---|---|---|---|---|---|---|
| ENST00000443816.5 | FN1-211 | Protein coding | 430585.6 | 4 | No | No | Yes |
| ENST00000432072.6 | FN1-209 | Protein coding | 234116.5 | 4 | No | Yes | No |
| ENST00000456923.5 | FN1-213 | Protein coding | 113350 | 4 | No | Yes | Yes |
| ENST00000356005.8 | FN1-204 | Protein coding | 112854.7 | 4 | No | No | Yes |
| ENST00000498719.1 | FN1-227 | Retained intron | 50673 | 4 | | | |
| ENST00000446046.5 | FN1-212 | Protein coding | 37427.2 | 3 | Yes | No | Yes |
| ENST00000494446.1 | FN1-225 | Retained intron | 33536.1 | 3 | | | |
| ENST00000421182.5 | FN1-207 | Protein coding | 30469 | 3 | No | No | Yes |
| ENST00000357867.8 | FN1-205 | Protein coding | 23405.6 | 3 | No | No | No |
| ENST00000426059.1 | FN1-208 | Protein coding | 16523.2 | 3 | No | No | No |
| ENST00000438981.1 | FN1-210 | Protein coding | 3472.7 | 2 | No | No | Yes |
| ENST00000461974.1 | FN1-215 | Retained intron | 2163.2 | 2 | | | |
| ENST00000492816.6 | FN1-224 | Retained intron | 2028 | 2 | | | |
| ENST00000471193.1 | FN1-217 | Retained intron | 903.3 | 2 | | | |
| ENST00000460217.1 | FN1-214 | Retained intron | 447.5 | 2 | | | |
| ENST00000480024.1 | FN1-220 | Retained intron | 277.2 | 2 | | | |
| ENST00000473614.1 | FN1-218 | Retained intron | 155.2 | 1 | | | |
| ENST00000496542.1 | FN1-226 | Retained intron | 87.6 | 1 | | | |
| ENST00000474036.1 | FN1-219 | Retained intron | 68 | 1 | | | |
| ENST00000469569.1 | FN1-216 | Retained intron | 41.5 | 1 | | | |
| ENST00000480737.1 | FN1-221 | Retained intron | 38.8 | 1 | | | |
| ENST00000490833.5 | FN1-223 | Processed transcript | 8.8 | 1 | | | |

The splice variant of the protein coding transcripts is indicated. baseMean = mean of normalized counts of all samples normalized for transcript length and sequencing depth; Quartile = expression in quartiles, with 1 being the lowest en 4 being the highest; EDA = presence EDA domain; EDB = presence EDB domain; V = presence V region.

## Differential expression of *FN1* transcripts between lesioned and preserved OA cartilage

To identify specific *FN1* transcripts associated with the OA process, differential expression analysis was performed on the 22 previously identified expressed *FN1* transcripts in paired lesioned and preserved OA cartilage samples, resulting in 16 significantly up-regulated *FN1* transcripts (false discovery rate (FDR) < 0.05, **Table 2**). The most significantly up-regulated transcript was *FN1-220* (fold change 2.8, FDR 5.8 x $10^{-13}$), which is classified as a retained intron, but had relatively low expression levels (base mean counts = 277.2; quartile 2). Notably, the most significant up-regulated protein coding transcript was *FN1-208* (fold change 2.3, FDR 4.9 x $10^{-6}$, **Supplementary Figure 2**), which encodes migration-stimulating factor. This truncated fibronectin protein contains the heparin- and gelatin-binding domain of fibronectin, but not the integrin-binding domain, and has been identified as a potent motogenic factor yet has not been previously identified in OA cartilage. To validate the differential expression, we analyzed *FN1* exon count data using DEXSeq, which separates the abundance of exons and parts of exons that are not the same for all transcripts in counting bins. We observed

**Table 2 | False discovery rate significantly differentially expressed *FN1* transcripts between lesioned and preserved osteoarthritic cartilage samples.**

| Ensembl ID | Name | Biotype | Quartile | FC | FDR | EDA | EDB | V |
|---|---|---|---|---|---|---|---|---|
| ENST00000480024.1 | FN1-220 | Retained intron | 2 | 2.8 | $5.8 \times 10^{-13}$ | | | |
| ENST00000494446.1 | FN1-225 | Retained intron | 3 | 2.4 | $4.4 \times 10^{-8}$ | | | |
| ENST00000473614.1 | FN1-218 | Retained intron | 1 | 2.3 | $1.7 \times 10^{-7}$ | | | |
| ENST00000492816.6 | FN1-224 | Retained intron | 2 | 2.4 | $3.3 \times 10^{-6}$ | | | |
| ENST00000426059.1 | FN1-208 | Protein coding | 3 | 2.3 | $4.9 \times 10^{-6}$ | No | No | No |
| ENST00000496542.1 | FN1-226 | Retained intron | 1 | 2.9 | $5.5 \times 10^{-6}$ | | | |
| ENST00000421182.5 | FN1-207 | Protein coding | 3 | 2.6 | $5.4 \times 10^{-4}$ | No | No | Yes |
| ENST00000490833.5 | FN1-223 | Processed transcript | 1 | 2.3 | $1.1 \times 10^{-3}$ | | | |
| ENST00000460217.1 | FN1-214 | Retained intron | 2 | 2.2 | $2.1 \times 10^{-3}$ | | | |
| ENST00000471193.1 | FN1-217 | Retained intron | 2 | 1.8 | $2.1 \times 10^{-3}$ | | | |
| ENST00000498719.1 | FN1-227 | Retained intron | 4 | 2.3 | $5.2 \times 10^{-3}$ | | | |
| ENST00000469569.1 | FN1-216 | Retained intron | 1 | 1.7 | $1.5 \times 10^{-2}$ | | | |
| ENST00000461974.1 | FN1-215 | Retained intron | 2 | 1.8 | $1.7 \times 10^{-2}$ | | | |
| ENST00000446046.5 | FN1-212 | Protein coding | 3 | 2.1 | $3.0 \times 10^{-2}$ | Yes | No | Yes |
| ENST00000356005.8 | FN1-204 | Protein coding | 4 | 2.0 | $5.0 \times 10^{-2}$ | No | No | Yes |
| ENST00000432072.6 | FN1-209 | Protein coding | 4 | 2.1 | $5.0 \times 10^{-2}$ | No | Yes | No |

The splice variant of the protein coding transcripts is indicated. Quartile = expression in quartiles, with 1 being the lowest en 4 being the highest; FC = fold change; FDR = false discovery rate; EDA = presence EDA domain; EDB = presence EDB domain; V = presence V region.

that the exon specific for *FN1-208* was higher expressed in lesioned OA cartilage compared to preserved, consequently validating its differential expression (**Supplementary Figure 3**). To replicate the identified up-regulation of *FN1-208* in lesioned cartilage, we performed RT-qPCR for *FN1-208* in an independent cohort consisting of 10 paired cartilage samples (**Supplementary Table 1B**). The transcript was detected in all samples and showed significant up-regulation (fold change 2.0, $P = 9.2 \times 10^{-3}$, **Supplementary Figure 4**).

To explore whether joint-specific *FN1* transcripts could be identified, we performed stratified analyses for knee (28 pairs) and hip samples (7 pairs). After filtering, we identified 22 transcripts to be robustly expressed in knee OA cartilage samples, while 19 transcripts were robustly expressed in hip samples. This difference could be partly explained by the lower number of hip samples in the analysis. Upon performing differential expression analysis on the knee samples, we identified 16 significantly differentially expressed *FN1* transcripts (**Supplementary Table 3**). In the hip samples, we also identified 16 significantly differentially expressed *FN1* transcripts (**Supplementary Table 4**). Notably, we identified one differentially expressed *FN1* transcript (ENST00000490833.5, *FN1-223*) only present in the knee stratified analysis. However, its expression was very low (basemean 8.8; quartile 1); as a result, it was likely removed in the hip analysis due to the filtering. Furthermore, we identified one transcript (ENST00000357867.8, *FN1-205*) that was only significantly up-regulated in the hip stratified analysis. However, upon closer inspection this transcript was

up-regulated in the knee samples, but with a P value slight more than the cut off value of 0.05, namely 0.051. Consequently, we could not identify joint-specific *FN1* transcripts in OA cartilage.

## Effect of modulation of *FN1* and *FN1-208* ratio levels on matrix deposition

Next, we aimed to study downstream effects of observed *FN1* transcript changes, including *FN1-208*, in our established human 3D in vitro neo-cartilage model. Since overall *FN1* expression is inherently high in cartilage, we could not obtain big changes by overexpressing *FN1-208* (**Supplementary Figure 5**); therefore, we aimed to down-regulate full-length *FN1*. To this end, primary chondrocytes were lentivirally transduced with *FN1* targeting shRNA, as depicted in **Figure 1A**. The shRNA targets all full-length protein coding *FN1* transcript but not *FN1-208*. Human primary chondrocytes of six donors were transduced (**Supplementary Table 1C**), after which in vitro chondrogenesis was induced in 3D pellet culture for 3 days. Consequently, we observed a down-regulation of the full-length transcripts (*FN1*¸ fold change 0.3, P = 7.6 x $10^{-7}$), but since *FN1-208* is not targeted, the *FN1-208* expression relative to full-length transcripts was increased (fold change 3.5, P = 6.0 x $10^{-6}$) (**Figure 1B**), thereby mimicking lesioned OA cartilage status. The overall down-regulation of fibronectin was also observed at protein level, both in the neo-cartilage pellets (**Figure 1C**) and culture medium

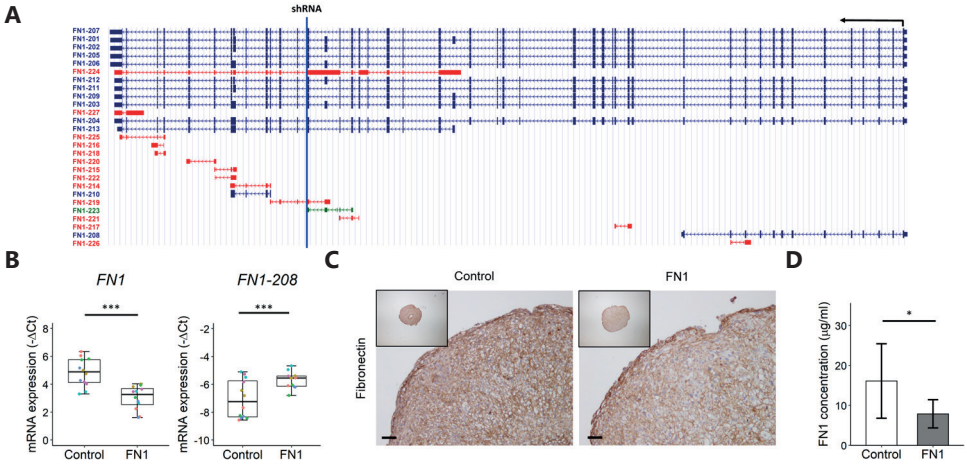



**Figure 1 | Down-regulation of *FN1* gene and protein expression in neo-cartilage pellets after 3 days of chondrogenesis.** (**A**) Schematic representation of *FN1* transcripts, which are transcribed from the antisense strand, represented by the black arrow. The blue line represents the location of the target sequence of the shRNA targeting *FN1* transcripts. Blue transcripts are protein coding, red and green transcripts are non-protein coding Source: https://genome.ucsc.edu/ (**B**) Gene expression levels depicted by box plots of -ΔCt values of *FN1* and *FN1-208* ratio relative to all full-length *FN1* transcripts in neo-cartilage pellets of primary chondrocytes transduced with non-targeting shRNA (control) and *FN1* targeting shRNA (FN1). Individual samples are represented by colored dots; colors of dots represent the different donors (N = 12) (**C**) Representative images of fibronectin staining of control and *FN1* down-regulated pellets, confirming *FN1* down-regulation on protein level. Scale bar, 50 µm. (**D**) Fibronectin concentration in conditioned medium in the control (N=6) and FN1 group (N=5), as determined by ELISA. Data are means ± SD. P values were determined by GEE, with experimental read-out as dependent variable, and donor and group as covariate. * P < 0.05, *** P < 0.005.

(**Figure 1D**). The fibronectin concentration was 50% decreased in the FN1 group (P = 1.0 x $10^{-2}$) compared to the control group, as determined by ELISA.

Subsequently, the effect of *FN1* down-regulation on neo-cartilage deposition was investigated by Alcian blue staining, where we observed a decreased deposition of sGAGs in the *FN1* down-regulated pellets (**Figure 2A**). Quantification of the Alcian blue staining showed that this decreased matrix deposition was 54% (P = 1.3 x $10^{-9}$) (**Figure 2B**). Furthermore, quantification of sGAG content normalized to DNA with the DMMB assay confirmed there was a significant decrease (P = 2.6x$10^{-2}$) in the *FN1* down-regulated group compared with the controls (**Figure 2C**). These data imply that *FN1* down-regulation and concomitant relative up-regulation of *FN1-208* have a negative effect on neo-cartilage deposition.

To investigate the effect of changes in *FN1* transcript ratios on gene expression, RT-qPCR was performed on 17 cartilage-relevant genes (**Supplementary Table 5**). As shown in **Figure 3**, both *ACAN* (fold change 0.5, P = 5.7 x $10^{-3}$) and *COL2A1* (fold change 0.1, P = 7.0 x $10^{-7}$) were strongly down-regulated in the FN1 group compared with controls. Moreover, *ADAMTS-5* was significantly up-regulated (fold change 2.6, P = 9.0 x $10^{-6}$), while *MMP-3* was significantly down-regulated (fold change 0.4, P = 2.9 x $10^{-2}$). These data imply that decreased *FN1* expression results in a more catabolic response of the chondrocytes via *ADAMTS-5*.

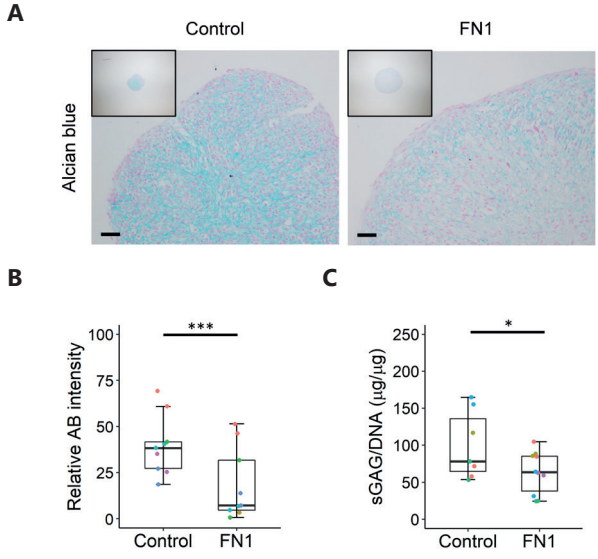



**Figure 2 | Decreased overall *FN1* expression and change of *FN1* transcript ratios results in decreased matrix deposition.** (**A**) Representative images of Alcian blue staining of neo-cartilage pellets of primary chondrocytes transduced with non-targeting shRNA (control) and *FN1* targeting shRNA (FN1). (**B**) Quantification of Alcian blue (AB) pixel intensity staining of control and *FN1* targeting shRNA transduced pellets (N = 9). Colors of dots represent the different donors. (**C**) sGAG content normalized to DNA content in pellets of the control (N = 7) and FN1 group (N = 11) analyzed by dimethylmethylene blue assay. P values were determined by GEE, with experimental read-out as dependent variable, and donor and group as covariate. * P < 0.05, *** P < 0.005.

Subsequently, we aimed to investigate the downstream effects of *FN1* down-regulation on the fibronectin-binding chondrocyte transmembrane integrin receptors. Gene expression levels of *ITGB1* (fold change 2.2, P = 1.7 x $10^{-4}$), and *ITGB5* (fold change 2.9, P = 1.9 x $10^{-4}$) were significantly up-regulated in the FN1 group compared with controls (**Figure 3**).

Finally, we aimed to identify novel *FN1* downstream pathways in OA cartilage. To this end, we calculated correlations between our previously reported differentially expressed genes (N = 2,378) and *FN1* in lesioned and preserved OA cartilage (N = 101) samples (**Supplementary Table 1D**) [4]. As a result, we found 60 genes to be highly correlated (|r| > 0.7) (**Supplementary Table 6**). Pathway enrichment analysis between these highly correlating genes (|r>| 0.7) and *FN1* using STRING resulted in five FDR significantly enriched Gene Ontology terms: cell surface (GO:0009986), plasma membrane (GO:0005886), membrane (GO:0016020), vesicle (GO:0031982), and intrinsic component of membrane (GO:0031224) (**Supplementary Table 7**). These processes are mainly characterized by *ITGA5*, *NT5E*, *BCAM*, *CD55*, and *CD109,* indicating the relation of fibronectin with basic cellular processes such as cell adhesion and cell growth. Of the 60 highly correlated genes, 10 showed correlations > 0.8. When analyzing for interaction among these 10 genes and *FN1*, 3 genes were either directly or indirectly connected to *FN1*, namely *ANKH*, *NT5E*, and



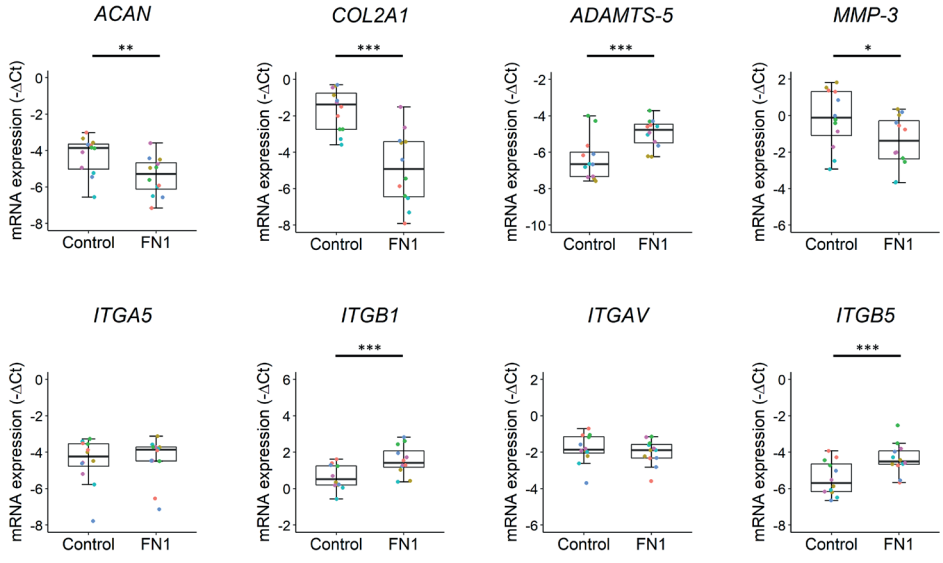



**Figure 3 | Decreased *FN1* expression and change of *FN1* transcript ratios results in catabolic chondrocyte metabolism.** Box plots of -ΔCt values of cartilage matrix-relevant genes *ACAN*, *COL2A1*, *ADAMTS-5*, *MMP-3*, *ITGA5*, *ITGB1*, *ITGAV*, *ITGB5* in neo-cartilage pellets of primary chondrocytes transduced with non-targeting shRNA (control) (N = 12), and *FN1* targeting shRNA (FN1) (N = 12). Individual samples are represented by colored dots; colors of dots represent the different donors. P values were determined by GEE, with experimental read-out as dependent variable and donor and group as covariate. * P < 0.05, ** P < 0.01, *** P < 0.005.

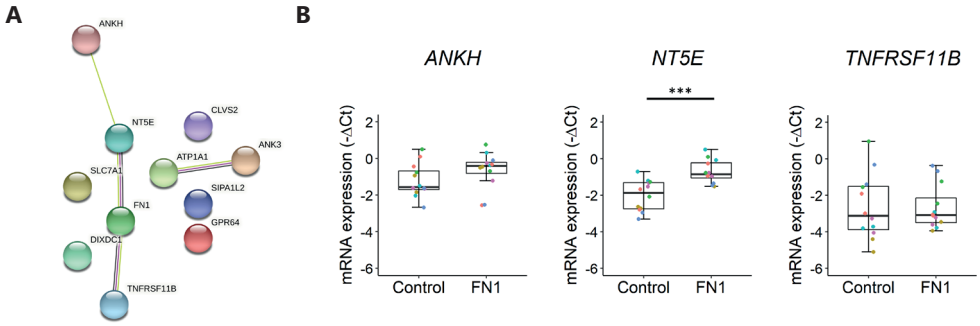



**Figure 4 | Identification of new *FN1* downstream genes** (**A**) Protein-protein interactions between the genes with correlations |r| > 0.8 with *FN1* in preserved and lesioned OA cartilage samples, as determined with STRING. (**B**) Box plots of -ΔCt values of connected genes to *FN1*, *ANKH*, *NT5E*, and *TNFRSF11B* in neo-cartilage pellets of primary chondrocytes transduced with non-targeting shRNA (control) (N = 12), and *FN1* targeting shRNA (FN1) (N = 12). Individual samples are represented by colored dots; colors of dots represent the different donors. P values were determined by GEE, with experimental read-out as dependent variable and donor and group as covariate. *** P < 0.005.

*TNFRSF11B* (**Figure 4A**). As shown in **Figure 4B**, only *NT5E* (fold change 2.5, P = 3.0 x $10^{-6}$) was significantly differentially expressed in the FN1 group compared with controls.

## Discussion

To the best of our knowledge, we are the first to use RNA sequencing to characterize the *FN1* transcriptome in OA cartilage. As a result, we identified 16 *FN1* transcripts FDR significantly up-regulated in lesioned OA cartilage, of which 5 were protein coding and 11 non-protein coding. These results show that considerable changes occur in the *FN1* transcriptome during OA, likely affecting proper function. Moreover, we identified the truncated protein coding transcript *FN1-208* to be significantly up-regulated in lesioned OA cartilage. Upon down-regulation of full-length *FN1* in our human 3D in vitro OA cartilage model, we generated an increased ratio of *FN1-208* relative to the full-length *FN1* transcripts, as such mimicking cartilage in an OA-affected state. This resulted in decreased cartilage deposition compared with the control group, with up-regulation of the β1 and β5 integrin subunits, suggesting a change of integrin heterodimers. Together, our results show that down-regulation of full-length *FN1* is unbeneficial for neo-cartilage deposition, while also highlighting the importance of balance of *FN1* transcripts for healthy cartilage homeostasis.

We identified *FN1-208,* known as MSF, as the most significantly up-regulated protein coding transcript, which has not been previously associated with OA. MSF contains the functional heparin- and gelatin-binding domain of full-length fibronectin and can not bind to transmembrane integrins via the classical RGD binding site. As a result, MSF has distinctive bioactivity compared with full-length fibronectin. MSF has been shown to be a potent motogenic factor and has been associated with cancer pathogenesis as potential driver

of tumor progression by inducing angiogenesis [16, 22]. Furthermore, blocking of MSF suppressed tumor growth through inhibition of tumor-related angiogenesis in an in vitro esophageal cancer model [23]. Angiogenesis contributes to OA pathology, as blood vessels from the subchondral bone invade the articular cartilage, thereby disrupting homeostasis of the chondrocytes. Therefore, we aimed to study whether the identified up-regulation of *FN1-208* is beneficial or unbeneficial to the OA process in cartilage. Since *FN1* is already highly expressed in our established in vitro neo-cartilage model, we could not obtain a significant up-regulation of *FN1-208*. Therefore, we down-regulated full-length *FN1* expression in our model and as a result we obtained an increased *FN1-208* ratio, thus mimicking an OA-related up-regulation. Consequently, in our study design we cannot distinguish between the effect of down-regulating *FN1* and increasing relative amounts of the OA-sensitive transcript *FN1-208*. Nonetheless, the shift in *FN1-208* ratio relative to the total protein coding transcripts resulted in a decreased neo-cartilage deposition, as well as a catabolic state of the chondrocytes. Moreover, we observed up-regulation of *ITGB1* and *ITGB5* gene expression levels. It has been shown that the β1 integrin is up-regulated in osteoarthritic compared to normal cartilage [24]. Therefore, the up-regulation of *ITGB1* and *ITGB5* may represent a higher disease state of the chondrocytes in the *FN1* down-regulated pellets. Together, these data show that decreased availability of the classical integrin-binding site of fibronectin to the cells is detrimental for chondrogenesis, which is likely mediated via β1 and β5 integrin subunits. However, this should be confirmed by quantifying protein expression of the integrin subunits, e.g., by Western Blot. Furthermore, investigating changes in integrin-downstream signaling can shed light on the effects of *ITGB1* and *ITGB5* up-regulation.

The retained intron transcripts *FN1-225* and *FN1-227* were relatively highly expressed and significantly differentially expressed between lesioned and preserved OA cartilage, suggesting they may play a role in OA pathophysiology. Intron retention as alternative splicing mechanism has recently been getting more attention regarding potential regulatory functions as opposed to being merely a consequence of mis-splicing [25]. Intron retention is mostly associated with down-regulation of gene expression via nonsense-mediated decay of the intron-retaining transcript, which has been shown to be a physiological mechanism of gene expression control regulating granulocyte differentiation [26]. However, intron retention has been suggested to potentially regulate non-coding RNAs contained within such introns [27]. Future studies regarding the function of retained intron *FN1* transcripts should address whether they regulate gene expression levels of the protein coding *FN1* transcripts, for example via expression or regulation of non-coding RNAs, such as micro-RNAs or long non-coding RNAs.

Previously, Scanzello *et al.* [17] investigated fibronectin splice variants in joint tissues, including cartilage. $EDB^+$, $EDB^-$, and $EDA^-$ variants were found to be present in cartilage, while $EDA^+$ variants were barely detected, as determined by RT-PCR. In line with these

observations, we identified *FN1-211* and *FN1-209* to be the highest expressed transcripts in OA cartilage, which are both $EDA^-$. The only $EDA^+$ transcript that we identified to be robustly expressed in OA cartilage was *FN1-212* (base mean counts = 37,427.2; quartile 3), which represented only 3.4% of the total transcripts. The EDA domain has been associated with many of the functions ascribed to fibronectin, including cell adhesion, matrix assembly, and dimer formation [13, 28]. However, given its low expression, our results suggest that the EDA domain is not essential for proper functioning of fibronectin in cartilage. On a different note, we observed *FN1-213* to be highly expressed (base mean counts =113,350.0; quartile 4), which is a 5'-truncated transcript, that has not been previously identified in cartilage.

All in all, we showed that RNA sequencing is a powerful technique for identifying involvement of the known *FN1* transcripts in OA cartilage. However, the previously identified cartilage specific $(V+III-15+I-10)^-$, $(V+I-10)^-$ and $(V+III-15)^-$ variants were not present in the Ensembl database and were therefore not detected in our analysis [29, 30]. To circumvent this issue, de novo transcriptome assembly could be performed. Nonetheless, in the current study we prioritized reporting previously unknown *FN1* transcripts present in the Ensembl database involved in OA pathophysiology, as opposed to reporting on previously identified *FN1* transcripts. A drawback of our study design is that we only investigated end-stage OA cartilage. Consequently, we cannot identify *FN1* transcripts that are specific for OA cartilage compared to healthy cartilage and thereby potentially be involved in the early phase of OA pathophysiology. Nonetheless, our paired analysis allows for identification of *FN1* transcripts specific to the pathophysiologic process of OA, independent of confounding factors such as sex and age.

To explore potential *FN1* downstream pathways, correlations were calculated between *FN1* and differentially expressed genes in OA cartilage. Among the highest correlated genes, three genes were interconnected in a protein network. Among these three genes, we identified *NT5E* as a novel *FN1* signaling downstream gene in cartilage. *NT5E* encodes the protein 5'-nucleotidase, also known as CD73, which is a plasma protein that catalyzes the conversion of extracellular nucleotides to membrane-permeable nucleosides and is a marker for mesenchymal stromal cells [31]. In addition to the enzymatic function, CD73 also functions as a receptor molecule that can interact with ECM components [32]. Defects in *NT5E* resulting in CD73 deficiency have been shown to facilitate calcification of joints [33, 34], whereas *NT5E* was significantly up-regulated in between lesioned and preserved OA cartilage [4]. These data imply that its up-regulation as a result of *FN1* down-regulation is a compensatory mechanism as a response to the OA disease state.

In conclusion, we identified multiple novel *FN1* transcripts associated with OA pathophysiology, while showing the potential role of *FN1-208*. We show that down-regulation of full-length *FN1*

was unbeneficial for neo-cartilage deposition and resulted in up-regulation of integrin β1 and β5 expression levels, likely via decreased availability of the classical RGD integrin-binding site of fibronectin.

## Acknowledgements

We thank all the participants of the RAAK study (supported by Leiden University Medical Center). We thank all the members of the MolEpi Osteoarthritis group for valuable discussion and feedback. We also thank Demiën Broekhuis, Robert van der Wal, Anika Rabelink-Hoogenstraaten, Peter van Schie, Shaho Hasan, Maartje Meijer, Daisy Latijnhouwers and Geert Spierenburg for collecting the RAAK material. We thank Martijn Rabelink for kindly providing us with the lentiviral shRNA plasmids and virus from the Mission shRNA library and performing the lentiviral p24 ELISA. Data is generated within the scope of the Medical Delta programs Regenerative Medicine 4D: Generating complex tissues with stem cells and printing technology and Improving Mobility with Technology.

**Funding:**

The study was funded by the Dutch Research council/NWO/ZonMW VICI scheme [nr 91816631/528], Dutch Arthritis Society [grant nr DAF-16-1-405].



## References

1. Litwic, A., et al., *Epidemiology and burden of osteoarthritis*. Br Med Bull, 2013. **105**: p. 185-99.
2. Loeser, R.F., et al., *Osteoarthritis: a disease of the joint as an organ*. Arthritis Rheum, 2012. **64**(6): p. 1697-707.
3. Goldring, S.R. and M.B. Goldring, *Changes in the osteochondral unit during osteoarthritis: structure, function and cartilage-bone crosstalk*. Nat Rev Rheumatol, 2016. **12**(11): p. 632-644.
4. Coutinho de Almeida, R., et al., *RNA sequencing data integration reveals an miRNA interactome of osteoarthritis cartilage*. Ann Rheum Dis, 2019. **78**(2): p. 270-277.
5. Aki, T., et al., *A whole-genome transcriptome analysis of articular chondrocytes in secondary osteoarthritis of the hip*. PLoS One, 2018. **13**(6): p. e0199734.
6. Ramos, Y.F., et al., *Genes involved in the osteoarthritis process identified through genome wide expression analysis in articular cartilage; the RAAK study*. PLoS One, 2014. **9**(7): p. e103056.
7. Singh, P., C. Carraher, and J.E. Schwarzbauer, *Assembly of fibronectin extracellular matrix*. Annu Rev Cell Dev Biol, 2010. **26**: p. 397-419.
8. Pankov, R. and K.M. Yamada, *Fibronectin at a glance*. J Cell Sci, 2002. **115**(Pt 20): p. 3861-3.
9. Almonte-Becerril, M., et al., *Genetic abrogation of the fibronectin-alpha5beta1 integrin interaction in articular cartilage aggravates osteoarthritis in mice*. PLoS One, 2018. **13**(6): p. e0198559.
10. van Hoolwerff, M., et al., *High-impact FN1 mutation decreases chondrogenic potential and affects cartilage deposition via decreased binding to collagen type II*. Sci Adv, 2021. **7**(45): p. eabg8583.

11. Homandberg, G.A., *Potential regulation of cartilage metabolism in osteoarthritis by fibronectin fragments*. Front Biosci, 1999. **4**: p. D713-30.
12. Reed, K.S.M., et al., *Transcriptional response of human articular chondrocytes treated with fibronectin fragments: an in vitro model of the osteoarthritis phenotype*. Osteoarthritis Cartilage, 2020.
13. White, E.S. and A.F. Muro, *Fibronectin splice variants: understanding their multiple roles in health and disease using engineered mouse models*. IUBMB Life, 2011. **63**(7): p. 538-46.
14. Schwarzbauer, J.E. and D.W. DeSimone, *Fibronectins, their fibrillogenesis, and in vivo functions*. Cold Spring Harb Perspect Biol, 2011. **3**(7).
15. Schwarzbauer, J.E., *Alternative splicing of fibronectin: three variants, three functions*. Bioessays, 1991. **13**(10): p. 527-33.
16. Schor, S.L., et al., *Migration-stimulating factor: a genetically truncated onco-fetal fibronectin isoform expressed by carcinoma and tumor-associated stromal cells*. Cancer Res, 2003. **63**(24): p. 8827-36.
17. Scanzello, C.R., et al., *Fibronectin splice variation in human knee cartilage, meniscus and synovial membrane: observations in osteoarthritic knee*. J Orthop Res, 2015. **33**(4): p. 556-62.
18. Bomer, N., et al., *Underlying molecular mechanisms of DIO2 susceptibility in symptomatic osteoarthritis*. Ann Rheum Dis, 2015. **74**(8): p. 1571-9.
19. Farndale, R.W., D.J. Buttle, and A.J. Barrett, *Improved quantitation and discrimination of sulphated glycosaminoglycans by use of dimethylmethylene blue*. Biochim Biophys Acta, 1986. **883**(2): p. 173-7.
20. Szklarczyk, D., et al., *STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets*. Nucleic Acids Res, 2019. **47**(D1): p. D607-D613.
21. Zeger, S.L. and K.Y. Liang, *Longitudinal data analysis for discrete and continuous outcomes*. Biometrics, 1986. **42**(1): p. 121-30.
22. Schor, A.M. and S.L. Schor, *Angiogenesis and tumour progression: migration-stimulating factor as a novel target for clinical intervention*. Eye (Lond), 2010. **24**(3): p. 450-8.
23. Hu, H., et al., *Antibody library-based tumor endothelial cells surface proteomic functional screen reveals migration-stimulating factor as an anti-angiogenic target*. Mol Cell Proteomics, 2009. **8**(4): p. 816-26.
24. Loeser, R.F., C.S. Carlson, and M.P. McGee, *Expression of beta 1 integrins by cultured articular chondrocytes and in osteoarthritic cartilage*. Exp Cell Res, 1995. **217**(2): p. 248-57.
25. Jacob, A.G. and C.W.J. Smith, *Intron retention as a component of regulated gene expression programs*. Hum Genet, 2017. **136**(9): p. 1043-1057.
26. Wong, J.J., et al., *Orchestrated intron retention regulates normal granulocyte differentiation*. Cell, 2013. **154**(3): p. 583-95.
27. Wong, J.J., et al., *Intron retention in mRNA: No longer nonsense: Known and putative roles of intron retention in normal and disease biology*. Bioessays, 2016. **38**(1): p. 41-9.
28. Manabe, R., N. Oh-e, and K. Sekiguchi, *Alternatively spliced EDA segment regulates fibronectin-dependent cell cycle progression and mitogenic signal transduction*. J Biol Chem, 1999. **274**(9): p. 5919-24.
29. Parker, A.E., et al., *Novel cartilage-specific splice variants of fibronectin*. Osteoarthritis Cartilage, 2002. **10**(7): p. 528-34.
30. MacLeod, J.N., et al., *Fibronectin mRNA splice variant in articular cartilage lacks bases encoding the V, III-15, and I-10 protein segments*. J Biol Chem, 1996. **271**(31): p. 18954-60.
31. Dominici, M., et al., *Minimal criteria for defining multipotent mesenchymal stromal cells. The International Society for Cellular Therapy position statement*. Cytotherapy, 2006. **8**(4): p. 315-7.
32. Andrade, C.M., et al., *Ecto-5'-nucleotidase/CD73 knockdown increases cell migration and mRNA level of collagen I in a hepatic stellate cell line*. Cell Tissue Res, 2011. **344**(2): p. 279-86.
33. Ichikawa, N., et al., *Arterial Calcification Due to Deficiency of CD73 (ACDC) As One of Rheumatic Diseases Associated With Periarticular Calcification*. J Clin Rheumatol, 2015. **21**(4): p. 216-20.
34. St Hilaire, C., et al., *NT5E mutations and arterial calcifications*. N Engl J Med, 2011. **364**(5): p. 432-42.

# Supplementary Methods

## RNA sequencing

Reads were aligned to the GrCh38 reference genome with RNA-seq aligner HISAT2 (version 2.1.0) [1]. Thereafter, aligned reads were processed into individual transcripts using Salmon (version 0.10.1) [2], which were then mapped to Ensembl annotated transcripts (version 97) [3]. Transcript abundance was quantified using the tximport R package (version 1.18.0) [4], resulting in normalized counts corrected for transcript length and sequencing depth, which we termed as base mean counts in our work. Subsequently, *FN1* transcripts were filtered. To obtain *FN1* transcripts that were robustly expressed, a cutoff was applied in the lesioned and preserved cartilage samples separately of an average of ≥ 10 counts per 3 samples.

## Differential expression analysis and replication

Differential expression analysis was performed on the expressed *FN1* transcripts in 35 paired samples using the DESeq2 R package (version 1.30.1) [5]. A general linear model assuming a negative binomial distribution was applied, followed by a paired Wald's test comparing lesioned and preserved OA cartilage samples, with the preserved samples as the reference. P values less than 0.05 (after Benjamini-Hochberg correction) were considered significant and are reported as the false discovery rate (FDR). DEXSeq (version 1.36) was used to visualize the exon count data from *FN1* [6]. Moreover, replication of differentially expressed *FN1-208* was performed in an independent cohort of 10 paired cartilage samples (**Supplementary Table 1B**) as described previously [7]. Expression levels of *FN1-208* were measured with QuantStudio 6 Real-Time PCR system (Applied Biosystems) using FastStart SYBR Green Master reaction mix (Roche Applied Science). Relative gene expression levels (-ΔCt) were calculated using *GAPDH* as internal control, primer sequences are shown in **Supplementary Table 2**. A paired t-test was performed on the -ΔCt values, P values less than 0.05 were considered significant.

## Lentiviral production and cell culture

Lentiviral production was performed in HEK293T cells using the Lenti-vpak Lentiviral Packaging kit (Origene Technologies) according to manufacturer's protocol. Virus particles were collected 48 and 72 hours post-transfection and stored at -80 °C. Virus titer was determined by p24 ELISA.

Primary chondrocytes were cultured as previously described [8]. For lentiviral transduction, cells were seeded at a cell density of $3.5x10^5$ cells per 10 cm culture dish. After one day in culture, cells were incubated for 16 hours with the corresponding lentivirus at multiplicity of infection (MOI) of 1 in the presence of 15 µg/ml Polybrene (Sigma-Aldrich), after which the medium was replaced with culture medium. After approximately 24 hours, medium was refreshed with culture medium with 0.5 µg/ml Puromycin (Sigma-Aldrich) to select for transduced cells. Cells were maintained in culture medium supplemented with Puromycin for 1 week,

subsequently the chondrocytes were expanded and *in vitro* chondrogenesis was induced as previously described [8]. Neo-cartilage pellets and medium were collected following three days of chondrogenesis. Surface area of the pellets was measured with CellSens Dimension (Olympus) software (version 3.1.1).

## RNA isolation and relative gene expression analysis

For RNA isolation, two neo-cartilage pellets were pooled and lysed in 200 µl TRIzol reagent (Thermo Fisher Scientific) and homogenized using micro pestles. RNA was extracted with chloroform and purified from the supernatant using the RNeasy Mini kit (Qiagen). RNA concentration was measured using a Nanodrop-1000 photospectrometer (Thermo Scientific). Synthesis of cDNA was performed with 150 ng of total RNA using a First Strand cDNA Synthesis kit according to manufacturer's protocol (Roche Applied Science). The cDNA was diluted five times and pre-amplification with Taqman preamp master mix (Thermo Fisher Scientific) was performed for 16 genes of interest, primer sequences are shown in **Supplementary Table 2**. Expression levels were measured with QuantStudio 6 Real-Time PCR system using FastStart SYBR Green Master reaction mix. Relative gene expression levels (-ΔCt) were calculated using the average of Ct values of *GAPDH* and *SDHA* as internal controls. Fold changes were calculated with the $2^{-\Delta\Delta Ct}$ method. The ratio of specific *FN1* transcripts was calculated as follows: $\Delta Ct_{RATIO} = \Delta Ct_{TRANSCRIPT} - \Delta Ct_{FULL\ LENGTH\ FN1}$.

## Histology and immunohistochemistry

Neo-cartilage pellets were fixed in 4% formaldehyde overnight and stored in 70% EtOH at 4 °C. They were then embedded in paraffin and sectioned at 5 µm. After sectioning, slides were deparaffinized and rehydrated. Sections were stained with 1% Alcian blue 8-GX (Sigma-Aldrich) and Nuclear Fast Red (Sigma-Aldrich) to visualize glycosaminoglycan deposition. To detect fibronectin (ab2413, Abcam, 1:400) immunohistochemistry was performed with 3-diaminobenzidine (DAB)-solution (Sigma-Aldrich) and hematoxylin (Klinipath) as described before [8]. Alcian blue staining was quantified by loading the images in Fiji (ImageJ version 2.1.0) [9] and splitting the color channels. Subsequently, grey values were measured of 3-5 separate squares per pellet and corrected for the grey value of the background.

## References


1. Kim, D., et al., *Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype*. Nat Biotechnol, 2019. **37**(8): p. 907-915.
2. Patro, R., et al., *Salmon provides fast and bias-aware quantification of transcript expression*. Nat Methods, 2017. **14**(4): p. 417-419.
3. Cunningham, F., et al., *Ensembl 2019*. Nucleic acids research, 2019. **47**(D1): p. D745-D751.
4. Soneson, C., M.I. Love, and M.D. Robinson, *Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences*. F1000Res, 2015. **4**: p. 1521.
5. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2*. Genome Biology, 2014. **15**(12).

6. Anders, S., A. Reyes, and W. Huber, *Detecting differential usage of exons from RNA-seq data*. Genome Res, 2012. **22**(10): p. 2008-17.
7. van Hoolwerff, M., et al., *Elucidating Epigenetic Regulation by Identifying Functional cis-Acting Long Noncoding RNAs and Their Targets in Osteoarthritic Articular Cartilage*. Arthritis Rheumatol, 2020. **72**(11): p. 1845-1854.
8. Bomer, N., et al., *Underlying molecular mechanisms of DIO2 susceptibility in symptomatic osteoarthritis*. Ann Rheum Dis, 2015. **74**(8): p. 1571-9.
9. Schneider, C.A., W.S. Rasband, and K.W. Eliceiri, *NIH Image to ImageJ: 25 years of image analysis*. Nat Methods, 2012. **9**(7): p. 671-5.

# Supplementary Materials

## Supplementary Tables

**Supplementary Table 1 | Sample characteristics included in the (A) discovery, (B) replication, (C) transduction of primary chondrocytes with *FN1* targeting shRNA, and (D) correlation analyses.**

**A**

| Discovery | | | | | | |
|---|---|---|---|---|---|---|
| **Tissue type** | **Mean age (range)** | **# Men** | **# Women** | **# Preserved** | **# Lesioned** | **# Total** |
| Knee | 69.5 (48 -79) | 14 | 42 | 28 | 28 | 56 |
| Hip | 64.4 (48 -79) | 2 | 12 | 7 | 7 | 14 |
| All | 68.5 (48 -79) | 16 | 54 | 35 | 35 | 70 |

**B**

| Replication | | | | | | |
|---|---|---|---|---|---|---|
| **Tissue type** | **Mean age (range)** | **# Men** | **# Women** | **# Preserved** | **# Lesioned** | **# Total** |
| Knee | 64.6 (57 - 75) | 3 | 2 | 5 | 5 | 10 |
| Hip | 68.2 (59 - 75) | 2 | 3 | 5 | 5 | 10 |
| All | 66.4 (57 - 75) | 5 | 5 | 10 | 10 | 20 |

**C**

| Transduction shRNA targeting *FN1* | | | | |
|---|---|---|---|---|
| **Tissue type** | **Mean age (range)** | **# Men** | **# Women** | **# Total** |
| Knee | 68.2 (63-77) | 3 | 3 | 6 |

**D**

| Correlations | | | | | | |
|---|---|---|---|---|---|---|
| **Tissue type** | **Mean age (range)** | **# Men** | **# Women** | **# Preserved** | **# Lesioned** | **# Total** |
| Knee | 69.1 (46 - 79) | 56 | 12 | 35 | 33 | 68 |
| Hip | 66.2 (48 - 82) | 27 | 6 | 22 | 11 | 33 |
| All | 68.2 (46 - 82) | 83 | 18 | 57 | 44 | 101 |

**Supplementary Table 2 | Primer sequences to measure mRNA expression levels**

| Gene | Forward primer (5' --> 3') | Reverse primer (5' --> 3') |
|---|---|---|
| *ACAN* | AGAGACTCACACAGTCGAAACAGC | CTATGTTACAGTGCTCGCCAGTG |
| *ADAMTS-5* | GTGGTGAAGGTGGTGGTGCT | CTCATGGTCATCTCCCAGCTG |
| *ANKH* | GTCTGCATGGCTCTGTCACT | AGGCAAAGTCCACTCCGATG |
| *COL1A1* | GTGCTAAAGGTGCCAATGGT | ACCAGGTTCACCGCTGTTAC |
| *COL2A1* | CTACCCCAATCCAGCAAACGT | AGGTGATGTTCTGGGAGCCTT |
| *FN1 full length* | GGATGGGGAGCAGAGTTTGA | GGTGTCATCAACTTGGTCCAC |
| *FN1-208* | AATGCACCACAGCCATCTCA | AGGTTTCTGGGTGGGATACTC |
| *GAPDH* | TGCCATGTAGACCCCTTGAAG | ATGGTACATGACAAGGTGCGG |
| *ITGA5* | GTCGGGGGCTTCAACTTAGAC | AGCACACTGACCCCGTCTG |
| *ITGAV* | TCTCTCGGGACTCCTGCTAC | AAGAAACATCCGGGAAGACGC |
| *ITGB1* | AGCGAAGGCATCCCTGAAAG | AATGTCTACCAACACGCCCT |
| *ITGB5* | CAGCAGCTTCCATGTCCTGA | AGGTTCACGGCAATCTCCTG |
| *MMP-3* | GAGGCATCCACACCCTAGGTT | TCAGAAATGGCTGCATCGATT |
| *MMP-13* | TTGAGCTGGACTCATTGTCG | GGAGCCTCTCAGTCATGGAG |
| *NT5E* | ATTGCACTGGGACATTCGGG | TGGAAGGTGGATTGCCTGTG |
| *RUNX2* | CAATTTCCTCCTTGCCCCTCA | TCGGATCTACGGGAATACGCA |
| *SDHA* | TGGGAACAAGAGGGCATCTG | GCCTACCACCACTGCATCAA |
| *SOX9* | CCCCAACAGATCGCCTACAG | CTGGAGTTCTGGTGGTCGGT |
| *TNFRS11B* | TTGATGGAAAGCTTACCGGGA | TCTGGTCACTGGGTTTGCATG |



**Supplementary Table 3 | False discovery rate significantly differentially expressed *FN1* transcripts in lesioned versus preserved OA cartilage in knee samples.**

| Ensembl ID | Name | Biotype | baseMean | FoldChange | Pvalue | Padj |
|---|---|---|---|---|---|---|
| ENST00000480024.1 | FN1-220 | Retained intron | 277.2 | 2.76 | 2.64E-14 | 5.80E-13 |
| ENST00000494446.1 | FN1-225 | Retained intron | 33536.1 | 2.44 | 3.98E-09 | 4.38E-08 |
| ENST00000473614.1 | FN1-218 | Retained intron | 155.2 | 2.30 | 2.28E-08 | 1.67E-07 |
| ENST00000492816.6 | FN1-224 | Retained intron | 2028.0 | 2.38 | 6.03E-07 | 3.31E-06 |
| ENST00000426059.1 | FN1-208 | Protein coding | 16523.2 | 2.35 | 1.10E-06 | 4.86E-06 |
| ENST00000496542.1 | FN1-226 | Retained intron | 87.6 | 2.88 | 1.49E-06 | 5.45E-06 |
| ENST00000421182.5 | FN1-207 | Protein coding | 30469.0 | 2.57 | 1.72E-04 | 5.39E-04 |
| ENST00000490833.5 | FN1-223 | Processed transcript | 8.8 | 2.30 | 3.91E-04 | 1.08E-03 |
| ENST00000460217.1 | FN1-214 | Retained intron | 447.5 | 1.76 | 8.85E-04 | 2.09E-03 |
| ENST00000471193.1 | FN1-217 | Retained intron | 903.3 | 2.21 | 9.49E-04 | 2.09E-03 |
| ENST00000498719.1 | FN1-227 | Retained intron | 50673.0 | 2.30 | 2.62E-03 | 5.23E-03 |
| ENST00000469569.1 | FN1-216 | Retained intron | 41.5 | 1.70 | 8.08E-03 | 1.48E-02 |
| ENST00000461974.1 | FN1-215 | Retained intron | 2163.2 | 1.80 | 9.75E-03 | 1.65E-02 |
| ENST00000446046.5 | FN1-212 | Protein coding | 37427.2 | 2.11 | 1.90E-02 | 2.99E-02 |
| ENST00000356005.8 | FN1-204 | Protein coding | 112854.7 | 2.14 | 3.59E-02 | 4.99E-02 |
| ENST00000432072.6 | FN1-209 | Protein coding | 234116.5 | 1.96 | 3.63E-02 | 4.99E-02 |

baseMean = mean of normalized counts of all samples normalized for transcript length and sequencing depth, FoldChange = fold change between lesioned and preserved OA cartilage samples, Pvalue = nominal P value, Padj = P value according to false discovery rate.

**Supplementary Table 4 | Significantly differentially expressed *FN1* transcripts in lesioned versus preserved osteoarthritic cartilage in hip samples.**

| Ensembl ID | Name | Biotype | baseMean | FoldChange | Pvalue | Padj |
|---|---|---|---|---|---|---|
| ENST00000480024.1 | FN1-220 | Retained intron | 277.2 | 2.76 | 2.64E-14 | 5.01E-13 |
| ENST00000494446.1 | FN1-225 | Retained intron | 33536.1 | 2.44 | 3.98E-09 | 3.78E-08 |
| ENST00000473614.1 | FN1-218 | Retained intron | 155.2 | 2.30 | 2.28E-08 | 1.45E-07 |
| ENST00000492816.6 | FN1-224 | Retained intron | 2028 | 2.38 | 6.03E-07 | 2.86E-06 |
| ENST00000426059.1 | FN1-208 | Protein coding | 16523.2 | 2.35 | 1.10E-06 | 4.19E-06 |
| ENST00000496542.1 | FN1-226 | Retained intron | 87.6 | 2.88 | 1.49E-06 | 4.71E-06 |
| ENST00000421182.5 | FN1-207 | Protein coding | 30469 | 2.57 | 1.72E-04 | 4.66E-04 |
| ENST00000460217.1 | FN1-214 | Retained intron | 447.5 | 1.76 | 8.85E-04 | 2.00E-03 |
| ENST00000471193.1 | FN1-217 | Retained intron | 903.3 | 2.21 | 9.49E-04 | 2.00E-03 |
| ENST00000498719.1 | FN1-227 | Retained intron | 50673 | 2.30 | 2.62E-03 | 4.97E-03 |
| ENST00000469569.1 | FN1-216 | Retained intron | 41.5 | 1.70 | 8.08E-03 | 1.40E-02 |
| ENST00000461974.1 | FN1-215 | Retained intron | 2163.2 | 1.80 | 9.75E-03 | 1.54E-02 |
| ENST00000446046.5 | FN1-212 | Protein coding | 37427.2 | 2.11 | 1.90E-02 | 2.78E-02 |
| ENST00000356005.8 | FN1-204 | Protein coding | 112854.7 | 2.14 | 3.59E-02 | 4.60E-02 |
| ENST00000432072.6 | FN1-209 | Protein coding | 234116.5 | 1.96 | 3.63E-02 | 4.60E-02 |
| ENST00000357867.8 | FN1-205 | Protein coding | 23405.6 | 0.44 | 3.96E-02 | 4.71E-02 |

baseMean = mean of normalized counts of all samples normalized for transcript length and sequencing depth, FoldChange = fold change between lesioned and preserved OA cartilage samples, Pvalue = nominal P value, Padj = P value according to false discovery rate.

**Supplementary Table 5 | Effect of modulating *FN1* and relative *FN1-208* expression on cartilage relevant genes in pellets from primary chondrocytes transduced with non-targeting (control) and *FN1* targeting shRNA after 3 days of chondrogenesis.** P values were determined by generalized estimation equations, with experimental read-out as dependent variable and donor and group as covariate.

| Gene | Fold change | P value |
|---|---|---|
| *ACAN* | 0.5 | 5.7E-03 |
| *ADAMTS-5* | 2.6 | 9.0E-06 |
| *ANKH* | 1.5 | 1.4E-01 |
| *COL1A1* | 0.7 | 3.2E-01 |
| *COL2A1* | 0.1 | 7.0E-07 |
| *FN1 full length* | 0.3 | 7.6E-07 |
| *FN1-208* | 3.5 | 6.0E-06 |
| *ITGA5* | 1.1 | 8.0E-01 |
| *ITGAV* | 0.9 | 4.8E-01 |
| *ITGB1* | 1.9 | 4.0E-06 |
| *ITGB5* | 2.1 | 1.2E-03 |
| *MMP-3* | 0.4 | 2.9E-02 |
| *MMP-13* | 0.3 | 3.2E-02 |
| *NT5E* | 2.5 | 3.0E-06 |
| *RUNX2* | 0.6 | 6.2E-02 |
| *SOX9* | 0.8 | 2.2E-01 |
| *TNFRS11B* | 1.0 | 9.5E-01 |

**Supplementary Table 6 | Correlations > |0.7| between significant differently expressed genes and *FN1* in the same lesioned and preserved osteoarthritic cartilage samples.**

| Ensembl ID | Gene | FoldChange | Cor | Pval |
|---|---|---|---|---|
| ENSG00000151150 | ANK3 | 2.62 | 0.87 | 2.22E-16 |
| ENSG00000150764 | DIXDC1 | 1.53 | 0.86 | 2.22E-16 |
| ENSG00000116991 | SIPA1L2 | 1.78 | 0.84 | 2.22E-16 |
| ENSG00000164761 | TNFRSF11B | 3.01 | 0.84 | 2.22E-16 |
| ENSG00000139514 | SLC7A1 | 1.44 | 0.83 | 2.22E-16 |
| ENSG00000154122 | ANKH | 1.44 | 0.82 | 2.22E-16 |
| ENSG00000163399 | ATP1A1 | 1.29 | 0.82 | 2.22E-16 |
| ENSG00000135318 | NT5E | 2.00 | 0.81 | 2.22E-16 |
| ENSG00000146352 | CLVS2 | 2.15 | 0.81 | 2.22E-16 |
| ENSG00000173698 | ADGRG2 | 1.59 | 0.80 | 2.22E-16 |
| ENSG00000185634 | SHC4 | 2.10 | 0.78 | 2.22E-16 |
| ENSG00000158966 | CACHD1 | 1.80 | 0.78 | 2.22E-16 |
| ENSG00000173276 | ZBTB21 | 1.36 | 0.78 | 2.22E-16 |
| ENSG00000282121 | AL592430.2 | 1.95 | 0.77 | 2.22E-16 |
| ENSG00000154229 | PRKCA | 1.24 | 0.77 | 2.22E-16 |
| ENSG00000173267 | SNCG | 0.61 | -0.77 | 2.22E-16 |
| ENSG00000137070 | IL11RA | 0.80 | -0.76 | 2.22E-16 |
| ENSG00000168675 | LDLRAD4 | 1.33 | 0.76 | 2.22E-16 |
| ENSG00000188706 | ZDHHC9 | 1.24 | 0.76 | 2.22E-16 |
| ENSG00000204592 | HLA-E | 0.72 | -0.76 | 2.22E-16 |
| ENSG00000167994 | RAB3IL1 | 0.48 | -0.76 | 2.22E-16 |
| ENSG00000211455 | STK38L | 1.65 | 0.76 | 2.22E-16 |
| ENSG00000235750 | KIAA0040 | 1.50 | 0.76 | 2.22E-16 |
| ENSG00000118596 | SLC16A7 | 1.61 | 0.75 | 2.22E-16 |
| ENSG00000204291 | COL15A1 | 1.74 | 0.74 | 2.22E-16 |
| ENSG00000144648 | ACKR2 | 1.80 | 0.74 | 2.22E-16 |
| ENSG00000147642 | SYBU | 1.26 | 0.74 | 2.22E-16 |
| ENSG00000244486 | SCARF2 | 0.62 | -0.74 | 2.22E-16 |
| ENSG00000156535 | CD109 | 1.54 | 0.74 | 2.22E-16 |
| ENSG00000154721 | JAM2 | 0.61 | -0.74 | 2.22E-16 |
| ENSG00000134198 | TSPAN2 | 2.42 | 0.73 | 2.22E-16 |
| ENSG00000182934 | SRPRA | 1.14 | 0.73 | 2.22E-16 |
| ENSG00000161638 | ITGA5 | 1.45 | 0.73 | 2.22E-16 |
| ENSG00000123933 | MXD4 | 0.78 | -0.73 | 2.22E-16 |
| ENSG00000241839 | PLEKHO2 | 0.70 | -0.73 | 2.22E-16 |
| ENSG00000187957 | DNER | 3.37 | 0.73 | 2.22E-16 |
| ENSG00000151693 | ASAP2 | 1.34 | 0.72 | 2.22E-16 |
| ENSG00000135269 | TES | 1.35 | 0.72 | 2.22E-16 |
| ENSG00000196305 | IARS | 1.32 | 0.72 | 2.22E-16 |
| ENSG00000185386 | MAPK11 | 0.74 | -0.72 | 2.22E-16 |
| ENSG00000153823 | PID1 | 1.48 | 0.72 | 2.22E-16 |
| ENSG00000185760 | KCNQ5 | 1.32 | 0.72 | 2.22E-16 |
| ENSG00000134243 | SORT1 | 1.40 | 0.72 | 2.22E-16 |
| ENSG00000178752 | ERFE | 3.44 | 0.72 | 2.22E-16 |
| ENSG00000150995 | ITPR1 | 1.58 | 0.72 | 2.22E-16 |
| ENSG00000171621 | SPSB1 | 1.57 | 0.71 | 2.22E-16 |
| ENSG00000182580 | EPHB3 | 0.52 | -0.71 | 2.22E-16 |
| ENSG00000134802 | SLC43A3 | 0.66 | -0.71 | 2.22E-16 |
| ENSG00000117020 | AKT3 | 1.43 | 0.71 | 2.22E-16 |

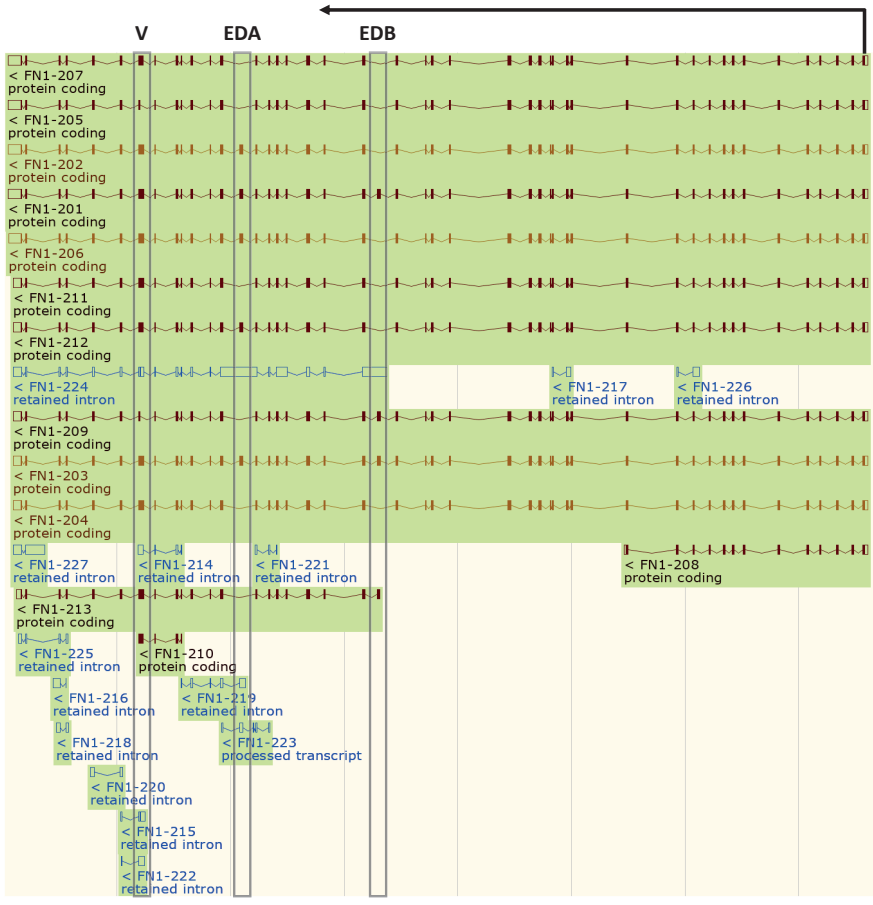
| Ensembl ID | Gene | FoldChange | Cor | Pval |
|---|---|---|---|---|
| ENSG00000154930 | ACSS1 | 0.51 | -0.71 | 2.22E-16 |
| ENSG00000182752 | PAPPA | 3.43 | 0.71 | 2.22E-16 |
| ENSG00000141441 | GAREM1 | 1.30 | 0.71 | 2.22E-16 |
| ENSG00000150907 | FOXO1 | 1.22 | 0.71 | 2.22E-16 |
| ENSG00000196352 | CD55 | 2.96 | 0.71 | 2.22E-16 |
| ENSG00000152503 | TRIM36 | 2.85 | 0.71 | 2.22E-16 |
| ENSG00000140526 | ABHD2 | 1.49 | 0.70 | 2.22E-16 |
| ENSG00000132334 | PTPRE | 1.30 | 0.70 | 4.44E-16 |
| ENSG00000197892 | KIF13B | 1.37 | 0.70 | 4.44E-16 |
| ENSG00000187244 | BCAM | 0.64 | -0.70 | 4.44E-16 |
| ENSG00000165434 | PGM2L1 | 1.88 | 0.70 | 4.44E-16 |

FoldChange = fold change between lesioned and preserved OA cartilage samples, Cor = correlation, Pval = nominal P value

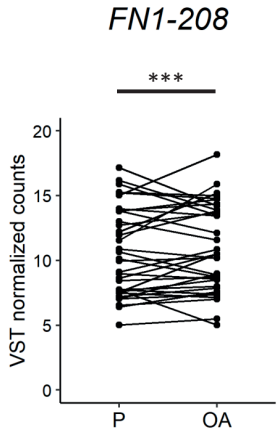
**Supplementary Table 7 | False discovery rate (FDR) significantly enriched Gene Ontology processes between highly correlated genes (|r| > 0.7) and *FN1* analyzed using STRING.**

| Term ID | Term description | Observed gene count | Background gene count | FDR | Matching proteins in network |
|---|---|---|---|---|---|
| GO:0009986 | Cell surface | 12 | 824 | 0.012 | SORT1, NT5E, BCAM, ANK3, CD109, ITGA5, CD55, HLA-E, GPR64, JAM2, ACKR2, IL11RA |
| GO:0005886 | Plasma membrane | 31 | 5314 | 0.043 | PTPRE, SORT1, NT5E, SLC16A7, ABHD2, GAREM, BCAM, ANK3, ASAP2, ANKH, CD109, CACHD1, ITGA5, TNFRSF11B, ITPR1, SHC4, EPHB3, KCNQ5, DNER, FN1, TES, CD55, TSPAN2, HLA-E, GPR64, SLC7A1, JAM2, PRKCA, ACKR2, ATP1A1, IL11RA |
| GO:0016020 | Membrane | 43 | 9072 | 0.043 | PTPRE, SORT1, NT5E, SLC16A7, AKT3, ABHD2, GAREM, BCAM, CLVS2, ANK3, ASAP2, ANKH, CD109, CACHD1, ITGA5, TNFRSF11B, ITPR1, SRPR, SHC4, EPHB3, KCNQ5, DNER, FN1, ZDHHC9, TES, LDLRAD4, CD55, TSPAN2, COL15A1, IARS, HLA-E, GPR64, SLC7A1, STK38L, JAM2, SYBU, PRKCA, ACKR2, SLC43A3, ATP1A1, IL11RA, KIAA0040, SCARF2 |
| GO:0031982 | Vesicle | 25 | 3879 | 0.043 | SORT1, NT5E, ABHD2, BCAM, CLVS2, TRIM36, CD109, ITGA5, ITPR1, PLEKHO2, SRPR, DNER, FN1, LDLRAD4, CD55, SNCG, COL15A1, IARS, HLA-E, GPR64, RAB3IL1, SYBU, PRKCA, ACKR2, ATP1A1 |
| GO:0031224 | Intrinsic component of membrane | 30 | 5345 | 0.046 | PTPRE, SORT1, NT5E, SLC16A7, ABHD2, BCAM, ANKH, CD109, CACHD1, ITGA5, ITPR1, EPHB3, KCNQ5, DNER, ZDHHC9, LDLRAD4, CD55, TSPAN2, COL15A1 ,HLA-E, GPR64, SLC7A1, JAM2, SYBU, ACKR2, SLC43A3, ATP1A1, IL11RA, KIAA0040, SCARF2 |

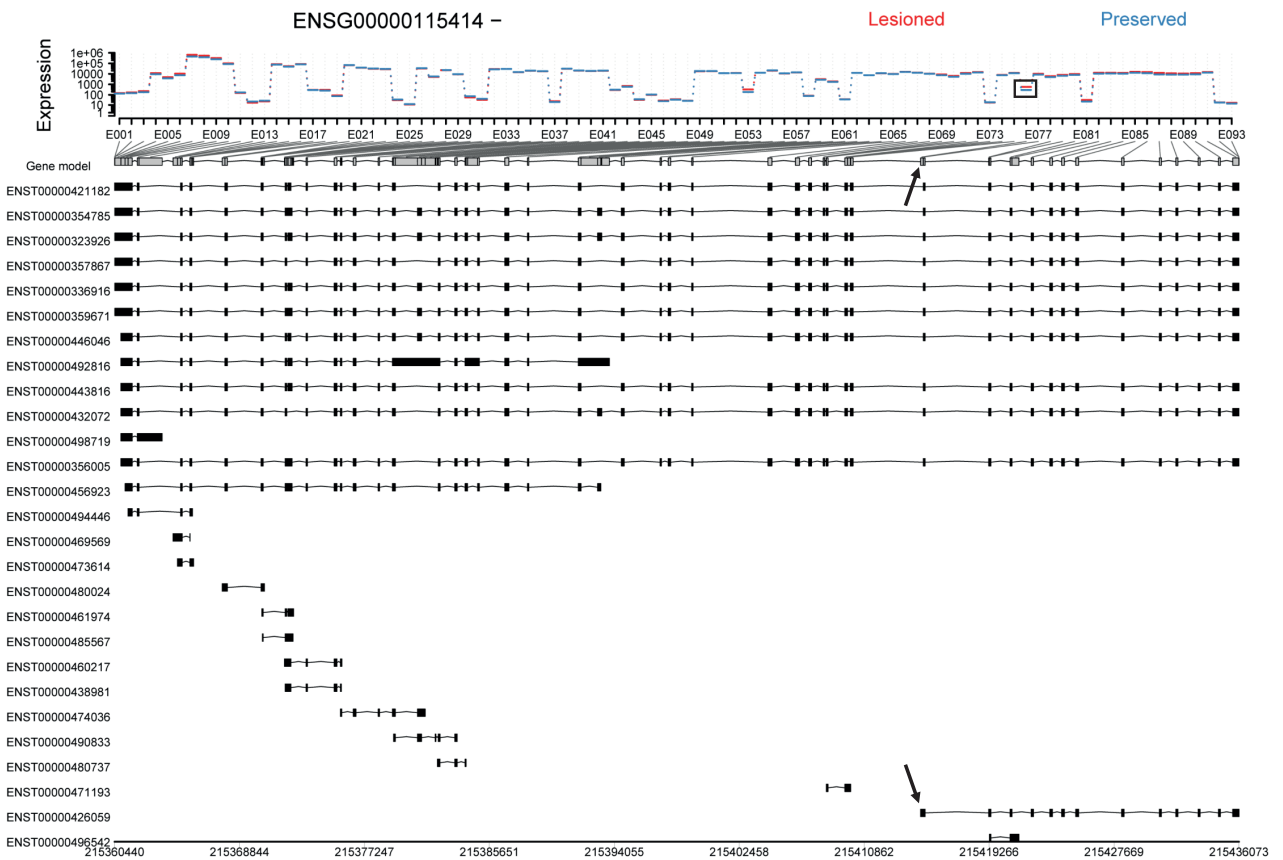
## Supplementary Figures



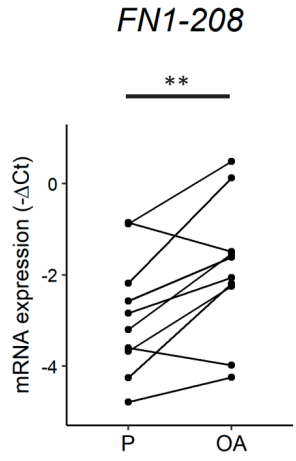

**Supplementary Figure 1 | Overview of *FN1* transcripts as annotated in the Ensembl database, which are transcribed from the antisense strand, represented by the black arrow.** The black squares represent the three regions of alternative splicing, from 5' to 3' the Extra Domain B (EDB), Extra domain A (EDA), and Variable region (V). Source: https://www.ensembl.org/index.html



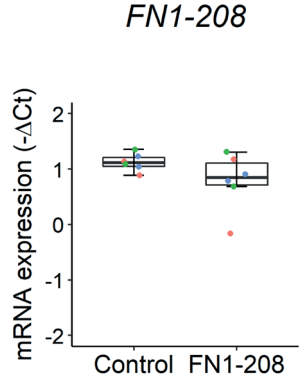

**Supplementary Figure 2 | *FN1-208* was significantly upregulated (FC = 2.3) in lesioned (OA) versus preserved (P) osteoarthritic cartilage samples as determined by RNA-sequencing in 35 paired cartilage samples.** *** P value according to false discovery rate < 0.001.




**Supplementary Figure 3 | Overview of exon count data of all *FN1* transcripts in lesioned versus preserved OA cartilage.** The exons are divided into bins that are identical across all transcripts. The most prominent difference between lesioned and preserved is the bin marked by the black box, which corresponds to the bin marked by the black arrows, which is the specific exon that is unique for *FN1-208* (ENST00000426059 in figure).




**Supplementary Figure 4 | Replication of *FN1-208* upregulation (FC = 2.0) in lesioned (OA) and preserved (P) osteoarthritic cartilage samples by RT-qPCR in 10 independent paired samples.** ** P < 0.01, as determined by paired t-test on -ΔCt values per donor.




**Supplementary Figure 5 | Unsuccessful overexpression of *FN1-208* in neo-cartilage pellets of primary chondrocytes of 3 donors transduced with pLV-CMV-eGFP-FN1-208 (FN1-208) compared to cells transduced with empty vector pLV-CMV-eGFP (control).** Individual samples are represented by colored dots, colors of the dots represent different donors.

# Chapter 5

## Discussion

## Summary

In the field of osteoarthritis (OA), development of effective drugs is considerably hampered by lacking insight into underlying OA pathophysiology and etiology. The aim of this thesis was to combine transcriptomics, genetics and human disease modeling to obtain further insight into molecular processes underlying OA. Performing transcriptome-wide analyses of OA relevant tissue, such as cartilage, has been shown to be a successful method to identify previously unknown genes that mark OA pathophysiology [1-5]. To further expand on this knowledge, in this thesis we aimed to elucidate the role of long noncoding RNAs (lncRNAs) expression changes as aberrant epigenetic mechanism in regulating gene expression in chondrocytes in **chapter 2**. Consequently, we identified previously unknown lncRNAs associated with the OA process in samples obtained from the Research osteoArthritis and Articular Cartilage (RAAK) study. Upon integrating messenger RNA (mRNA) sequencing data, we showed that intergenic and antisense lncRNAs demonstrate high, positive correlations with their respective flanking or sense genes. We functionally validated this *cis*-regulation for the antisense lncRNAs *P3H2-AS1* and its sense gene *P3H2*.

To provide insight in the etiology of OA, causal pathways can be identified by unravelling the substantial genetic component of OA. In **chapter 3**, we identified a high-impact causal mutation in *FN1* in an early-onset OA family, after which we set up an OA disease model to identify underlying pathways. To this end, we introduced the *FN1* mutation in human induced pluripotent stem cells (hiPSCs), followed by chondrogenic differentiation to neo-cartilage producing chondrocytes. We demonstrated that the missense mutation in the gelatin-binding domain of fibronectin resulted in significant decreased binding capacity to collagen type II. Further analyses of formed hiPSC-derived neo-cartilage tissue highlighted that mutated fibronectin affected chondrogenic capacity and enhanced propensity to a procatabolic OA state.

Finally, the common function of *FN1* in cartilage was investigated, since it is also highly up-regulated in lesioned compared to preserved OA cartilage. Moreover, *FN1* can give rise to 27 transcripts, of which 13 are protein coding, which raises the question whether specific *FN1* transcripts play a role in OA pathophysiology. In **chapter 4**, we identified migration-stimulating factor (MSF or *FN1-208*), a truncated isoform of fibronectin, associated with OA pathophysiology and not previously identified in OA cartilage. Down-regulation of full length *FN1* was unbeneficial for neo-cartilage deposition by human primary chondrocytes obtained from the RAAK study in our 3D in vitro chondrogenesis model.

## Role of lncRNAs in osteoarthritic cartilage

OA pathophysiology in cartilage is marked by alterations in gene expression regulation in chondrocytes. Since chondrocytes remain in a maturational arrested state, they rely heavily on epigenetic mechanisms to regulate dynamic changes in gene expression in response to intrinsic and external challenges such as microtraumas and mechanical stress. As a response to these processes, chondrocytes need to become temporarily metabolically active and adjust expression levels of anabolic and catabolic genes, which is controlled by multiple levels of control including DNA methylation, histone modifications and noncoding RNAs [6]. Unraveling aberrant epigenetic mechanisms in chondrocytes thus provides another important level of insight into OA pathophysiology. One of the least characterized levels of epigenetic mechanisms in articular cartilage are lncRNAs. Potentially, lncRNAs could be candidate targets in OA, since their expression can be highly tissue specific [7].

### Identifying long noncoding RNAs associated with OA pathophysiology

Hypothesis-free profiling of lncRNAs in healthy and OA cartilage was first based on microarray data [8-10], but as a consequence of decreasing costs of and significant technical advances in RNA sequencing, studies using this technique gained traction [11-13]. RNA sequencing greatly improved the ability to detect and identify lncRNAs, since they are structurally highly similar to mRNAs but relatively lower expressed. However, annotating lncRNAs remain challenging, since their sequence-function relationship is poorly understood and the number of experimentally characterized lncRNAs is low, namely <1% of identified loci [14]. Therefore, in **chapter 2**, we used a new RNA sequencing in-house pipeline to robustly detect lncRNAs in OA cartilage samples from the RAAK study. Recently, ribosome profiling and bioinformatic studies showed that a large proportion of transcripts has unknown protein coding potential [14]. In order to filter transcripts with unknown protein coding potential, we integrated two machine learning methods, Coding Potential Assessment Tool (CPAT) and the LncFinder R package. Transcripts with protein coding potential predicted by both tools were removed from the dataset. As a result, we identified 5,053 lncRNAs to be robustly expressed in OA cartilage, 191 of which were significantly differentially expressed lncRNAs between lesioned and preserved OA cartilage [15]. Notably, we observed an increase in the percentage of intergenic lncRNAs (lincRNAs), highlighting their general involvement in the OA pathophysiology process. Potential interactions were identified between the differentially expressed lncRNAs and differentially expressed protein coding genes in the same OA cartilage samples, where we observed an enrichment between lincRNAs and their flanking genes and between antisense lncRNAs and their sense genes, implying *cis*-regulation. In vitro functional validation of this *cis*-regulation revealed that the antisense lncRNA *P3H2-AS1* regulates its sense gene *P3H2*.

Of the 191 identified lncRNAs that associated with OA pathophysiology, multiple lncRNAs have been previously identified, such as *MEG3*, *LINC01614*, and *PART1* [12, 16]. However,

multiple lncRNAs previously found associated with OA, including *MALAT1*, *HOTAIR*, and *GAS5*, were not significantly differentially expressed in our study. One explanation could be that our study design comprises a within patient comparison between lesioned and preserved cartilage, as opposed to a cross-sectional design comparing healthy and preserved OA cartilage. The cross-sectional design can give insight into which lncRNAs are involved in the early phase of OA and therefore potentially causal to the process, while our design allows for detection of lncRNAs specific to the OA pathophysiological process, independent of confounding factors such as genetic background, sex, and age. We were able to validate and replicate the direction of effect for five lncRNAs, indicating robustness of our lncRNA mapping strategy. However, upon applying a filter with a cutoff of ≥2 counts per lncRNAs, the number of detected lncRNAs was drastically decreased by ~ 83%. LncRNAs are known to be expressed at very low levels, yet can still be functional. To perform exploratory analyses of lowly expressed lncRNAs, deeper sequencing would have to be performed, with a read-depth of 50-100 million reads per sample. Furthermore, in our study poly-A enrichment was performed for the RNA sequencing library prep, meaning that lncRNAs without a poly-A tail could not be identified in our analysis. To capture transcripts both with and without a poly-A tail, future studies should enrich for poly-A RNAs yet keep the other fraction to obtain non-poly-A lncRNAs, followed by ribosomal RNA depletion, similar to what was done by Yang *et al.* [17].

## Identifying downstream targets of long noncoding RNAs

To be able to potentially use lncRNAs as druggable targets, it is necessary to identify their downstream targets. Currently, no lncRNA-targeting therapeutics have entered clinical development. However, lncRNAs have increasingly been investigated and show promise as RNA interference or CRISPR targets [18]. Amodio *et al.* [19] investigated the function of *MALAT1* in multiple myeloma, where locked nucleic acid-GapmeR (LNA-GapmeR) antisense oligonucleotide (ASO) technology was used to target *MALAT1* expression. Down-regulation of *MALAT1* resulted in antitumor activity in a humanized myeloma mouse model, providing preclinical evidence for the use of this new ASO-targeting of lncRNAs for the treatment of multiple myeloma. In this study the effect of *MALAT1* down-regulation was measured by cell proliferation and viability, however, in this thesis we aimed to identify specific mRNAs downstream of the identified 191 differentially expressed lncRNAs in OA cartilage.

Unlike conserved miRNAs, there is no clear understanding yet of the sequence-function relation of lncRNAs. Functions of lncRNAs can be based on two elements; the base pairing in linear form in direct physical interaction with nucleic acids, proteins or lipids, and the chemical interactions as a consequence of secondary or tertiary structures [18]. Furthermore, lncRNAs can be classified based on whether they regulate the expression of neighboring genes in *cis* or more distant genes in *trans* [20]. *Cis*-acting lncRNAs comprise a considerable portion of known lncRNAs and can be positioned at various distances and orientations relative to

their target genes. Examples are lincRNAs around transcription factor start sites, as well as sense and antisense lncRNAs that overlap with their sense genes [20]. To explore potential regulatory interactions we generated a lncRNA-mRNA coexpression network in cartilage based on correlations. This showed an enrichment of high correlations between lincRNAs and their flanking genes and between antisense lncRNAs and their sense genes, implying *cis*-regulation of these lncRNAs. However, these correlations do not provide evidence for downstream effects of the lncRNAs on the mRNAs.

To functionally validate the observed *cis*-regulation we selected lncRNA *P3H2-AS1* as proof of concept to establish whether it regulates its sense gene. *P3H2-AS1* is an antisense lncRNA and showed the highest correlation with its sense gene *P3H2* (**Figure 1A**). To down-regulate *P3H2-AS1* expression, we used LNA-GapmeR ASO technology, also used by Amodio *et al.* [19]. As a result, *P3H2* expression was also down-regulated, thereby confirming that *P3H2-AS1* positively regulates the expression of its sense gene in *cis*. Antisense lncRNAs can affect biogenesis or mobilization of target mRNA on multiple levels, such as transcription, splicing and translation [21]. *Cis*-acting antisense lncRNAs are known to function at nearly all levels of gene regulation: pre-transcriptional, transcriptional and post-transcriptional [21]. *P3H2-AS1* and *P3H2* have no linear sequence similarities, so it is likely that *P3H2-AS1* regulates gene expression not by binding to *P3H2* mRNA, but functions at the pre-transcription level, e.g. by influencing chromatin state, influencing DNA methylation, or modulating transcription factor activity (**Figure 1B**). Visualization of subcellular localization of lncRNAs by RNA fluorescence in situ hybridization can provide insight into potential function of lncRNA [14, 22]. To investigate more specific lncRNA-protein interactions, RNA immunoprecipitation or crosslinked immunoprecipitation can be performed, which can show whether a lncRNA targets chromatin-modifying enzymes or transcription factors. The more recent development of CRISPR-mediated interference and activation can modulate expression of lncRNAs from their endogenous promoter by blocking or activating transcription, respectively [23]. In this way, lncRNA function can be determined including the production of *cis*-acting RNA transcripts and *cis*-mediated regulation related to lncRNA transcription itself. Liu *et al.* [24] developed a large CRISPR interference platform in multiple cell lines and hiPSCs and identified many lncRNA loci required for robust cellular growth. It would be interesting to perform a comparable study for lncRNAs in chondrocytes. Functionality of lncRNAs can be assigned with more confidence when RNA interference techniques, such as LNA-GapmeRs, are complemented with CRISPR-based experiments [23].

*P3H2* was shown to be significantly up-regulated in lesioned versus preserved OA cartilage samples from the RAAK study [25]. *P3H2* encodes an enzyme that catalyzes post-translational 3-hydroxylation of proline residues and plays a critical role in collagen chain assembly, stability, and crosslinking. Therefore, it seems likely that up-regulation of *P3H2-AS1* with

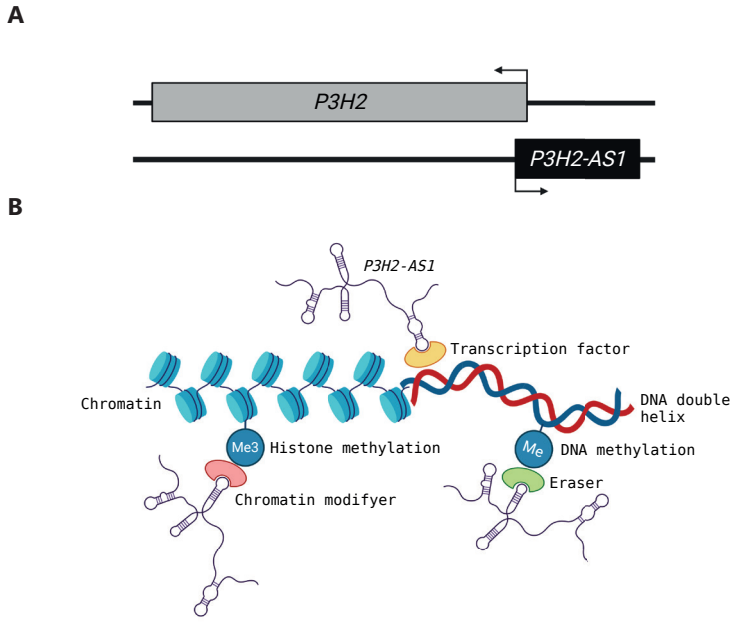



**Figure 1 | The antisense long noncoding RNA *P3H2-AS1* regulates gene expression of its sense gene *P3H2*** (**A**) Relative genomic location of *P3H2* and the antisense lncRNA *P3H2-As1*, the 5' end of *P3H2-AS1* is near the 5' end of *P3H2*, where the arrows indicate direction of transcription. (**B**) Potential mechanisms by which *P3H2-AS1* (purple) pre-transcriptionally regulates *P3H2* gene expression, binding to chromatin modifying enzymes (red), facilitating histone modifications thereby influencing chromatin state, binding to transcription factors (yellow) thereby modulating transcription, or binding to an eraser (green) and removing DNA methylation thereby inducing gene transcription. (Created with Biorender.com)

concurrent up-regulation of *P3H2* is a response to the OA process and beneficial in articular cartilage. This hypothesis could be tested by using CRISPR activation to up-regulate *P3H2-AS1* expression and investigate *P3H2* expression and further downstream effects on neo-cartilage deposition in human primary chondrocytes. Furthermore, additional targets of *P3H2-AS1* can be identified by performing transcriptome-wide analyses after CRISPR activation of *P3H2-AS1*.

Overall, we show that generating coexpression networks between lncRNAs and mRNAs can provide insight in potential regulatory function of lncRNAs. However, future studies regarding lncRNAs in relation to OA should be complemented by functional validation in order to confirm whether a correlation signifies a biologic causal relation between lncRNA and mRNA or is rather consequential. As a result of quickly advancing techniques involving CRISPR, the possibilities to determine the function of lncRNAs are growing steadily, indicating exciting future perspectives for identifying druggable targets for preclinical trials in OA.

## Genetic disease modeling for osteoarthritis

In an effort to elucidate the complex genetic architecture of OA, genome wide association studies have provided evidence for susceptibility loci in common OA pathophysiology [26-28]. It has been shown that developing new drugs with genetic support can double the success rate in clinical development [29]. However, translation to clinically druggable targets is lacking for OA, among others as a result of the small effects of the associated genetic variants. For that matter, identifying rare mutations with large effects in early-onset OA patients can provide insight into genotype-phenotype relations and thereby can elucidate causal OA pathways. However, functional follow-up studies of earlier identified high-impact mutations in OA patients have often not been performed. The quickly developing progress in genomic engineering with CRISPR/Cas9 technology has advanced the field greatly in this aspect, of which we readily took advantage of. Hence, in **chapter 3** we investigated the biological functionality of the high-impact, pathogenic mutation identified in *FN1* in an early-onset OA family [30]. To this end, we introduced the C518F *FN1* mutation in hiPSCs using CRISPR/Cas9 gene editing, thereby creating *FN1* heterozygous and homozygous hiPSC lines. Subsequently, the mutant and isogenic control hiPSCs were used in an established in vitro organoid cartilage model, where we observed a decrease of both chondrogenic potential and neo-cartilage deposition of the *FN1* mutant cells. Moreover, we demonstrated that the underlying pathogenic mechanism of the mutation was caused by a decreased binding of mutant fibronectin to collagen type II.

### Identification of high-impact mutation in early-onset OA family

By applying whole exome sequencing to an affected individual of an early-onset OA family, we obtained over 73,000 candidate variants after quality control. As the phenotype showed a dominant Mendelian inheritance pattern, we hypothesized the causal variant results in an amino acid change, thereby affecting protein structure and functioning. Consequently, we applied a pathogenic prioritization scheme to exclude intergenic, intronic, synonymous, common and tolerated missense variants. Common genic variants were filtered out when they were present in various population-scale variant databases, resulting in over 1,000 variants. Variant prioritization tools Sorts Intolerant From Tolerant (SIFT) and Polymorphism Phenotyping (PolyPhen) were used to remove tolerated missense variants, further reducing the number of variants predicted to have a functional impact on the gene produced to 122 missense variants. SIFT and PolyPhen were shown to have moderate sensitivity and their accuracy is dependent on whether loss-of-function or gain-of-function are being tested, indicating that further evidence to support causality is necessary [31]. Previously, our group showed strong linkage on 2q33.3 with multiple extended early-onset OA families, to which the current family contributed substantially, giving us a further indication of the chromosomal location of the pathogenic variant [32]. Of the 122 variants, three variants were located around the previously mentioned linkage area, namely *ALS2*, *FN1*, and *ABCB6*. Firstly, we investigated relevance of these three genes to OA by exploring gene expression levels in our previously

published RNA sequencing data of lesioned and preserved cartilage and bone samples from the RAAK study [25, 33]. Only *ALS2* and *FN1* were expressed in both cartilage and bone, suggesting these genes are functional in these tissues. Furthermore, *FN1* was significantly up-regulated in lesioned OA cartilage compared to preserved, revealing that this gene is also sensitive to the OA process. Subsequently, de novo genotyping was performed for the *ALS2* and *FN1* variants. Since the investigated family is rather extended, genotyping showed complete linkage of the *FN1* variant in affected individuals, while the *ALS2* variant was not detected, thereby confirming that the C518F mutation in the *FN1* gene is likely causal to the early-onset phenotype in this family. Identifying a causal pathogenic mutation in rare Mendelian disease is not always successful, as the human exome contains thousands of variants [34]. However, in this thesis we exhibit the powerful combination of exome sequencing followed by linkage analysis in an extended family, allowing us to identify the causal mutation to the early-onset phenotype in the family [35]. Consequently, we aimed to set up a relevant in vitro OA disease model to investigate downstream biological pathways.

## In vitro OA disease modeling

In this study, we choose to use hiPSCs in our OA disease model, as opposed to human primary articular chondrocytes. Disadvantages of primary chondrocytes include limited availability and representing end-stage disease state, as they are often obtained from patients who underwent joint replacement surgery due to OA. Moreover, since we sought to introduce a specific mutation, the selection process for the correct clone without off-target effects would result in substantial 2D culturing, which in primary chondrocytes results in significant dedifferentiation and loss of chondrogenic potential [36, 37]. Studies that performed CRISPR/Cas9 genome editing in chondrocytes did so either in a rat chondrosarcoma cell line [38], or performed gene knockout, which has a higher efficiency than precise gene editing [39, 40]. We obtained both hetero- and homozygous *FN1* hiPSC clones which in essence were two separate clones. After chondrogenic differentiation we observed a dose response as a result of the mutation at the molecular level, thereby providing robustness to our obtained results. As hiPSCs can be expanded substantially, we acquired a sustainable cell source, which can be readily used for future experiments.

Differentiation of hiPSCs to chondrocytes has been shown to give variable efficiency, yet progress has been made in establishing reproducible step-wise differentiation protocols [41-43]. In our group we showed that neo-cartilage from hiPSC-derived chondrocytes was almost 70% similar to that of neo-cartilage from human primary articular chondrocytes based on gene expression profiles, indicating suitability of our hiPSC-derived organoid neo-cartilage model [44]. In this thesis we observed that *FN1* mutated organoids contained less cartilage-producing cells relative to the total number of cells compared to wild type organoids, indicating decreased chondrogenic potential. Furthermore, we observed a decrease in the deposition of

neo-cartilage, altogether indicating a less efficient formation of neo-cartilage. However, we could not separate the effect of the *FN1* mutation on the decreased chondrogenic potential and decreased neo-cartilage deposition, confounding our analyses. Dicks *et al.* demonstrated that heterogeneity of the chondroprogenitor cell population is partly due to mesenchymal and neurogenic lineage cells [45]. To circumvent the issue of heterogeneity in the chondroprogenitor population, a *GFP* reporter hiPSC line could be engineered with a specific chondrogenic marker to be able to purify chondroprogenitors, similar to Adkar *et al.* [41]. In this way, the effect of the mutation on the deposition of neo-cartilage can be investigated without the confounding factor of decreased chondrogenic potential. Furthermore, longitudinal analyses of the differentiation would have to be performed to elucidate how the *FN1* mutation affects hiPSC differentiation to chondroprogenitors and thereby chondrogenic potential. Previous findings showed that homogenous inactivation of fibronectin in mice resulted in early embryonic lethality and that fibronectin plays an essential role in mesodermal migration [46]. Additionally, the presence of fibronectin matrix was shown to be essential for mesenchymal stromal condensation and chondrogenic differentiation. Possibly, the mutation negatively affects these processes during the differentiation from hiPSCs to chondrocytes [47].

For decades, animals have served as the most common models of human disease, however, use of animal models is also limited due to genetic background differences, which has led to high rates of translational failure between human and animal models [48, 49]. Moreover, costs, housing and length of experiments are generally more costly with animal experiments and ethical guidelines are to be considered, since usually animals need to be sacrificed for OA studies, while there are no ethical issues regarding hiPSCs. In this thesis, we created isogenic hiPSC clones with the *FN1* mutation lacking off-target effects by precise genetic engineering, after which we applied an established differentiation protocol producing biomimetic human in vitro neo-cartilage. Hence, we consider our conditions near optimal and we are confident that our approach was able to create reliable data highly translating to the human in vivo situation, while contributing to the societal need to reduce animal studies. Taken together, we show the immense potential of combining exome sequencing, hiPSCs, CRISPR/Cas9 and organoid disease modeling in common, complex human genetic diseases such as OA.



## Role of fibronectin in osteoarthritis pathophysiology

### *FN1* mutation in gelatin-binding domain

In **chapter 3** we showed that the C518F mutation in the gelatin-binding domain of fibronectin resulted in a linear reduction in binding of mutant fibronectin to collagen type II. The change from a polar cysteine to a nonpolar phenylalanine was predicted to result in a conformational change of the protein, as determined by RaptorX, whereby the formation of a conserved disulfide bond is abrogated. The gelatin-binding domain of fibronectin consists of six modules,

namely 6FnI, 1-2FnII, and 7-9FnI, which were all shown to contribute to the interaction with gelatin and collagen, either directly or indirectly [50, 51]. As the C518F mutation is located in the 8FnI module, the predicted conformational change as a result of the mutation is likely directly causal to the decreased binding of mutant fibronectin to collagen type II. Collagen and fibronectin fibrillogenesis are thought to be interdependent processes [52], however, we did not observe obvious differences in collagen type II deposition when comparing wild type and mutant neo-cartilage pellets. We therefore hypothesized that the mutation induced structural differences at the fibril level. Unfortunately, we could not observe collagen fibrils by means of transmission electron microscopy in our neo-cartilage model, but it is possible that the conformation of collagen and fibronectin fibrils was affected by the mutation and thereby mechanical properties of the neo-cartilage. Consequently, determining mechanical properties of the wild type and mutant neo-cartilage as a measure of quality could provide insight in the effects of the decreased binding between collagen and fibronectin. Altogether, in **chapter 3** we highlight the importance of the proper binding of fibronectin to collagen type II in articular cartilage (**Figure 2**).Since fibronectin functions as a transducer of biomechanical signals to chondrocytes from the ECM to chondrocytes via integrins, the decreased binding to collagen type II likely results in changed interactions between ECM and chondrocytes. Thus, determining threshold strains of mechanical loading that result in catabolic responses and cartilage degeneration of neo-cartilage produced by wild type and mutant chondrocytes can provide insight into whether the mutation potentially affects mechanotransduction and in that way responsible for the early-onset OA phenotype.

The *FN1* mutation resulted in aberrant chondrocyte gene expression, where anabolic markers were down-regulated and catabolic markers were up-regulated in *FN1* mutant chondrocytes. Moreover, integrin subunits *ITGA3* and *ITGB1* were significantly up-regulated in the homozygous *FN1* mutant chondrocytes. It has been known that integrin expression changes

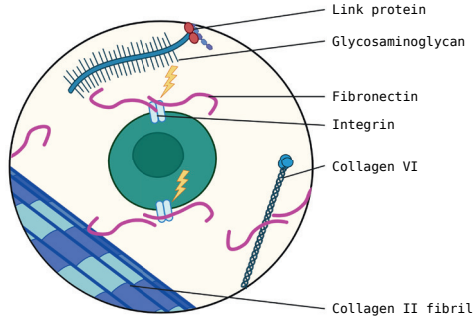



**Figure 2 | Potential mechanism of how the conformational change of C518F mutant fibronectin induces unbeneficial responses in chondrocytes.** Mutant fibronectin has decreased binding to collagen type II and potentially binds to integrin α5β1 as well as α3β1, where it induces unbeneficial gene expression changes, represented by the yellow bolt. (Created with Biorender.com).

during the development of OA, where α3β1 was shown to be up-regulated in OA chondrocytes [53]. Therefore, the up-regulation of *ITGA3* and *ITGB1* may reflect un unbeneficial state of the *FN1* mutant chondrocytes. Co-immunoprecipitation of wild-type and mutant fibronectin with integrins can provide insight into whether the shift in gene expression also resulted in changed interactions of fibronectin with these integrin subunits. Conversely, down-regulation of *ITGA5* occurred solely in the heterozygous *FN1* mutant chondrocytes. We hypothesized that the down-regulation could be a response to wild type-mutant fibronectin dimers binding to integrin α5β1. *ITGA5* was found to be down-regulated in OA cartilage compared to healthy [4], implying that down-regulation is likely unbeneficial for the heterozygous *FN1* mutant chondrocytes. It has been shown that fibronectin-α5β1 adhesion is essential for cartilage remodeling in mice, so quantitative binding assays between (mutant) fibronectin and integrin α5β1 could shed light on whether the mutation affects binding to integrin α5β1 [54]. Furthermore, integrin activation can occur via "outside-in" and "inside-out" signaling, thus up-regulation of integrin α3β1 in homozygous *FN1* mutant chondrocytes could also affect matrix homeostasis by changing chondrocyte adhesion to the ECM via "inside-out" signaling [55, 56].

### Identifying *FN1* transcripts associated with OA pathophysiology

Studies have shown that genes with a larger number of transcripts play biologically more fundamental roles [57]. Fibronectin is a ubiquitous protein in the human body as part of the extracellular matrix, as well as a major component of blood plasma, where it is involved in wound healing [58]. Alternative splicing of *FN1* mRNA can give rise to many transcripts that encode protein molecules with different binding capacities [59]. However, it has not been completely clear what changes occur at the transcript level with OA with respect to *FN1*. As such, in **chapter 4** we aimed to identify *FN1* transcripts annotated in the Ensembl database associated with OA pathophysiology. As a result, we identified sixteen *FN1* transcripts to be significantly up-regulated in lesioned compared to preserved OA cartilage obtained from the RAAK study, of which five were protein coding and eleven non-protein coding. The non-protein coding transcripts are classified as retained introns, which were shorter in length and generally lower expressed in cartilage than protein coding transcripts. Intron retention has recently been getting more attention as alternative splicing mechanism and is mostly associated with down-regulation of gene expression via nonsense-mediated decay of the intron-retaining transcript [60]. In the case of fibronectin this seems unlikely, since the retained intron transcripts are so much smaller than the protein coding transcripts. However, it is suggested that intron retention potentially regulates noncoding RNAs, for example if the retained introns encode miRNAs or contain noncoding RNA-response elements thereby affecting miRNA or lncRNA functioning [61]. Future studies regarding the function of retained intron *FN1* transcripts should address whether they regulate gene expression levels of the protein coding *FN1* transcripts, thereby acting as noncoding RNAs. Regarding the protein

coding *FN1* transcripts, we found $EDA^-$, $EDB^-$ and $EDB^+$ variants to be present in cartilage, while $EDA^+$ variants were less abundant, which is in line with previous findings [62]. The EDA domain has been associated with many functions ascribed to fibronectin, including cell adhesion, matrix assembly, and dimer formation [63]. However, the low abundance suggest that the EDA domain is not essential for proper functioning of fibronectin in cartilage.

Furthermore, we found *FN1-208*, encoding migration-stimulating factor (MSF), to be the most significantly up-regulated protein coding *FN1* transcript, which has not been previously identified in OA cartilage. MSF is a 3' truncated isoform of full length fibronectin of 70 kDa, containing the heparin- and gelatin-binding domain of full length fibronectin. It has been shown to be a potent motogenic factor, meaning it promotes cell motility, and it has been associated with cancer pathogenesis [64]. Consequently, we aimed to functionally investigate MSF in our established human 3D in vitro neo-cartilage model from primary chondrocytes. We could not achieve MSF overexpression in our model, therefore, we aimed to down-regulate full length *FN1*. As such, we were mimicking cartilage in an OA affected state by obtaining an up-regulation of MSF relative to all other *FN1* transcripts. Down-regulation of full length *FN1* transcripts was unbeneficial for neo-cartilage deposition, implying that the observed up-regulation in lesioned versus preserved OA cartilage from the RAAK study is a response to the OA process. Furthermore, *ADAMTS-5*, *ITGB1* and *ITGB5* expressoin levels were increased as a result of *FN1* down-regulation, suggesting a more disease state of the chondrocytes. Both *ADAMTS-5* and *ITGB1* showed similar responses in our *FN1* mutant hiPSC-derived neo-cartilage model, robustly indicating that *ADAMTS-5* and *ITGB1* are part of the fibronectin downstream signaling response. As MSF does not contain the classical arginine-glycine-aspartate (RGD) binding site to bind integrin $\alpha5\beta1$, we hypothesize that decreased availability of this fibronectin domain is unbeneficial for chondrogenesis. However, this remains to be confirmed e.g. by down-regulating full length *FN1* transcripts in parallel to up-regulating MSF and investigating the downstream effects on neo-cartilage deposition. Furthermore, up-regulation of fibronectin in our in vitro neo-cartilage model could confirm whether the observed up-regulation in lesioned versus preserved OA cartilage is a response to the OA process and not causal.

## Fibronectin fragments and migration-stimulating factor

Fibronectin can be cleaved by proteinases into fragments (FN-fs), which have catabolic activities in OA joints [65]. These FN-fs have obtained cryptic binding sites, resulting in altered binding to integrins and disharmonious downstream signaling. There are three main fragments that have been identified in this respect, comprising the 29 kDa N-terminal heparin-binding domain containing fragment, the 45 kDa gelatin-binding domain containing fragment, and the 110-140 kDa cell-binding domain fragment [66]. It has been shown that FN-fs increase aggrecan degradation via up-regulation of MMPs and ADAMTS-5 [67, 68]. Since

MSF is the length of the 29 kDa and 45 kDa fragment, it seems likely that this cell-produced fibronectin isoform has detrimental consequence for cartilage homeostasis. This hypothesis can be tested by adding MSF to chondrocyte pellet cultures and investigating downstream effects on neo-cartilage deposition.

## Future perspectives

In this thesis we showed that identifying rare, high-impact variants and their biological functionality can give insight into underlying pathways of OA in articular cartilage. The usefulness of modifiable human in vitro hiPSC models such as the one established in this thesis can be expanded by using it to test potential therapeutics that act against the molecular pathways that are disrupted in the model (**Figure 3**). The 3D neo-cartilage pellets can relatively easily be scaled up to perform high throughput drug screening.

The complexity of OA pathophysiology is partially because it is a disease of the whole joint. In our current model we focused on cartilage but excluded investigating any effects in bone tissue. For that matter, human ex vivo osteochondral explants can be considered the most accurate 3D model of OA and have been shown to be useful in OA pathophysiology models, potentially for pre-clinical studies [69]. However, one of the drawbacks of explants is the fact that cells cannot be genetically modified. Considering that hiPSCs can be differentiated into any cell type, differentiation of genetically modified cells to both cartilage- and bone-producing cells can overcome this drawback. By seeding multiple hiPSC-derived cell types in microfluidic chips, so-called joint-on-a-chip technology, cross-talk between OA-relevant tissue can be investigated (**Figure 3**).

Apart from rare high-impact mutations, it is also valuable to perform functional follow-up studies of more common genetic variants identified in genome-wide association studies. Finding biological functional consequences, other than expression quantitative loci analyses, and causality of these loci has been shown to be challenging in the field of OA. Partly because of accessibility of disease relevant tissue, as well as the fact that these variants usually have small effect sizes and that multiple, independently associated risk alleles may be responsible for occurrence of the disease [70]. hiPSC technology creates the possibility to use large cohorts of hiPSCs with known genotypes and perform genome-wide analyses of genetic variant-driven cellular phenotypes, both in hiPSCs and hiPSC-differentiated cells. A suitable cohort of participants needs to be considered with relevant phenotypes and known genotypes, after which hiPSC lines can be reprogrammed from participant primary cells. Subsequently, transcriptome, proteome, and metabolome approaches can be applied to correlate genotype to phenotype [71]. This “humanity in a dish” approach could drastically accelerate the elucidation of the molecular basis of common OA (**Figure 3**).

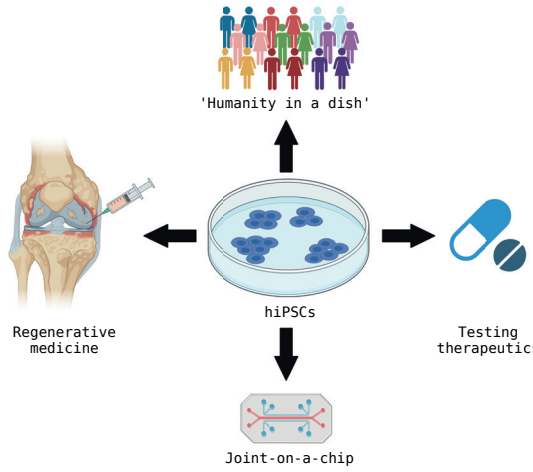



**Figure 3 | Overview of future applications of human induced pluripotent stem cells (hiPSC) in preclinical models (regenerative medicine, testing therapeutics) and disease models ('humanity in a dish', joint-on-a-chip) for osteoarthritis (OA).** For regenerative medicine, differentiated hiPSCs, either genetically engineered or not, can be used to repair damaged tissue by implantation in the osteoarthritic joint. Potential therapeutics that act against specific disrupted molecular pathways can be tested in modifiable human in vitro hiPSC models. Functional follow-up of more common genetic variants associated with OA can be performed in large cohorts of hiPSCs with known genotypes, so-called 'humanity in a dish'. Seeding genetically engineered hiPSC-derived OA-relevant cell types in microfluidic chips can provide valuable insight into the cross-talk between OA-relevant tissues. (Created with Biorender.com).

Next to the applicability of hiPSCs in OA disease modeling, they show promising potential for regenerative medicine, such as stem cell therapy. Regenerative medicine integrates cell biology, materials science and gene therapy, potentially resulting in cell-based implantation methods to repair damaged tissue in OA joints (**Figure 3**) [72]. The current problem is that hiPSC differentiation to chondroprogenitors results in a heterogenous cell population, thereby tempering progress to clinical applications. Gaining insight into chondrogenic lineage commitment of the hiPSCs can provide identification of modifiable factors that determine hiPSC cell fate to chondrogenic lineage. More specifically, insight at the single cell level of hiPSC differentiation gives more information regarding inter-cell variability. Wu *et al.* investigated gene regulatory networks regulating hiPSC differentiation at single-cell level during chondrogenesis, identifying *WNT* and *MITF* as hub genes governing the generation of off-target differentiation [73]. However, a multi-omics approach including epigenetic and proteomic analyses will allow an even more accurate characterization of factors regulating chondroprogenitor cell fate.

Regarding the role of fibronectin in osteoarthritic cartilage, in this thesis we highlighted the importance of proper binding between fibronectin and the ECM in articular cartilage, specifically via collagen type II. Furthermore, decreased deposition of full length fibronectin was

unbeneficial for neo-cartilage deposition. Our work merits further exploration of therapeutic interventions focusing on fibronectin as potential target. Engineering recombinant fibronectin fragments that compensate unbeneficial interactions between the ECM and chondrocytes can be a starting point for tissue engineering [74]. Fibronectin conformational change can influence integrin specificity and we showed that ECM interactions can also be influenced by specific conformational structures of fibronectin, thereby regulating cell behavior. Further functional analyses of the role of MSF in cartilage can provide initial clues for functional recombinant fragments.

The observed extensive changes in the *FN1* transcriptome with OA pathophysiology suggests that there are changes in regulation of these transcripts. The question that arises is how these transcripts are regulated by epigenetic mechanisms. As previously mentioned, the non-protein coding transcripts could act as noncoding RNAs or influence noncoding RNAs. Thus, it would be interesting to use LNA-GapmeR ASO technology to elucidate the function of the non-protein coding transcripts. Furthermore, generating coexpression networks between fibronectin and miRNAs or lncRNAs could provide initial clues for how fibronectin expression is epigenetically regulated.

Looking back on the past years of OA research, it has become clear that fast progress has been made by many exciting technical advancements. This sparks hope for the future of OA research and therapy development. In this thesis we performed multifaceted studies, which can be used as starting points for future OA disease modeling and towards development of new therapeutic strategies.



## References


1. Ramos, Y.F., et al., *Genes involved in the osteoarthritis process identified through genome wide expression analysis in articular cartilage; the RAAK study*. PLoS One, 2014. **9**(7): p. e103056.
2. Dunn, S.L., et al., *Gene expression changes in damaged osteoarthritic cartilage identify a signature of non-chondrogenic and mechanical responses*. Osteoarthritis Cartilage, 2016. **24**(8): p. 1431-40.
3. Fisch, K.M., et al., *Identification of transcription factors responsible for dysregulated networks in human osteoarthritis cartilage by global gene expression analysis*. Osteoarthritis Cartilage, 2018. **26**(11): p. 1531-1538.
4. Aigner, T., et al., *Large-scale gene expression profiling reveals major pathogenetic pathways of cartilage degeneration in osteoarthritis*. Arthritis Rheum, 2006. **54**(11): p. 3533-44.
5. Karlsson, C., et al., *Genome-wide expression profiling reveals new candidate genes associated with osteoarthritis*. Osteoarthritis Cartilage, 2010. **18**(4): p. 581-92.
6. Coutinho De Almeida, R., Y.F.M. Ramos, and I. Meulenbelt, *Involvement of epigenetics in osteoarthritis*. Best Practice & Research Clinical Rheumatology, 2017. **31**(5): p. 634-648.
7. Sun, H., et al., *Emerging roles of long noncoding RNA in chondrogenesis, osteogenesis, and osteoarthritis*. Am J Transl Res, 2019. **11**(1): p. 16-30.
8. Liu, Q., et al., *Long Noncoding RNA Related to Cartilage Injury Promotes Chondrocyte*

*Extracellular Matrix Degradation in Osteoarthritis*. Arthritis & Rheumatology, 2014. **66**(4): p. 969-978.
9. Fu, M., et al., *Expression profile of long noncoding RNAs in cartilage from knee osteoarthritis patients*. Osteoarthritis and Cartilage, 2015. **23**(3): p. 423-432.
10. Xing, D., et al., *Identification of long noncoding RNA associated with osteoarthritis in humans*. Orthop Surg, 2014. **6**(4): p. 288-93.
11. Pearson, M.J., et al., *Long Intergenic Noncoding RNAs Mediate the Human Chondrocyte Inflammatory Response and Are Differentially Expressed in Osteoarthritis Cartilage*. Arthritis Rheumatol, 2016. **68**(4): p. 845-56.
12. Ajekigbe, B., et al., *Identification of long non-coding RNAs expressed in knee and hip osteoarthritic cartilage*. Osteoarthritis Cartilage, 2019. **27**(4): p. 694-702.
13. Xiao, K., et al., *Identification of differentially expressed long noncoding RNAs in human knee osteoarthritis*. J Cell Biochem, 2019. **120**(3): p. 4620-4633.
14. Uszczynska-Ratajczak, B., et al., *Towards a complete map of the human long non-coding RNA transcriptome*. Nature Reviews Genetics, 2018. **19**(9): p. 535-548.
15. van Hoolwerff, M., et al., *Elucidating Epigenetic Regulation by Identifying Functional cis-Acting Long Noncoding RNAs and Their Targets in Osteoarthritic Articular Cartilage*. Arthritis Rheumatol, 2020. **72**(11): p. 1845-1854.
16. Chen, K., et al., *LncRNA MEG3 Inhibits the Degradation of the Extracellular Matrix of Chondrocytes in Osteoarthritis via Targeting miR-93/TGFBR2 Axis*. Cartilage, 2019: p. 1947603519855759.
17. Yang, L., et al., *Genomewide characterization of non-polyadenylated RNAs*. Genome Biol, 2011. **12**(2): p. R16.
18. Winkle, M., et al., *Noncoding RNA therapeutics - challenges and potential solutions*. Nat Rev Drug Discov, 2021. **20**(8): p. 629-651.
19. Amodio, N., et al., *Drugging the lncRNA MALAT1 via LNA gapmeR ASO inhibits gene expression of proteasome subunits and triggers anti-multiple myeloma activity*. Leukemia, 2018. **32**(9): p. 1948-1957.
20. Gil, N. and I. Ulitsky, *Regulation of gene expression by cis-acting long non-coding RNAs*. Nat Rev Genet, 2020. **21**(2): p. 102-117.
21. Villegas, V.E. and P.G. Zaphiropoulos, *Neighboring gene regulation by antisense long non-coding RNAs*. Int J Mol Sci, 2015. **16**(2): p. 3251-66.
22. Kashi, K., et al., *Discovery and functional analysis of lncRNAs: Methodologies to investigate an uncharacterized transcriptome*. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms, 2016. **1859**(1): p. 3-15.
23. Charles Richard, J.L. and P.J.A. Eichhorn, *Platforms for Investigating LncRNA Functions*. SLAS TECHNOLOGY: Translating Life Sciences Innovation, 2018. **23**(6): p. 493-506.
24. Liu, S.J., et al., *CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells*. Science, 2017. **355**(6320).
25. Coutinho de Almeida, R., et al., *RNA sequencing data integration reveals an miRNA interactome of osteoarthritis cartilage*. Ann Rheum Dis, 2019. **78**(2): p. 270-277.
26. Boer, C.G., et al., *Deciphering osteoarthritis genetics across 826,690 individuals from 9 populations*. Cell, 2021. **184**(18): p. 4784-4818 e17.
27. den Hollander, W., et al., *Genome-wide association and functional studies identify a role for matrix Gla protein in osteoarthritis of the hand*. Ann Rheum Dis, 2017. **76**(12): p. 2046-2053.
28. Styrkarsdottir, U., et al., *Meta-analysis of Icelandic and UK data sets identifies missense variants in SMO, IL11, COL11A1 and 13 more new loci associated with osteoarthritis*. Nat Genet, 2018. **50**(12): p. 1681-1687.
29. Nelson, M.R., et al., *The support of human genetic evidence for approved drug indications*. Nat Genet, 2015. **47**(8): p. 856-60.
30. van Hoolwerff, M., et al., *High-impact FN1 mutation decreases chondrogenic potential and affects cartilage deposition via decreased binding to collagen type II*. Sci Adv, 2021. **7**(45): p. eabg8583.
31. Flanagan, S.E., A.M. Patch, and S. Ellard, *Using SIFT and PolyPhen to predict loss-of-function and gain-of-function mutations*. Genet Test Mol Biomarkers, 2010. **14**(4): p. 533-7.
32. Meulenbelt, I., et al., *Strong linkage on 2q33.3 to familial early-onset generalized*

*osteoarthritis and a consideration of two positional candidate genes*. Eur J Hum Genet, 2006. **14**(12): p. 1280-7.
33. Tuerlings, M., et al., *RNA sequencing reveals interacting key determinants of osteoarthritis acting in subchondral bone and articular cartilage*. Arthritis Rheumatol, 2020.
34. Eilbeck, K., A. Quinlan, and M. Yandell, *Settling the score: variant prioritization and Mendelian disease*. Nat Rev Genet, 2017. **18**(10): p. 599-612.
35. Rabbani, B., et al., *Next-generation sequencing: impact of exome sequencing in characterizing Mendelian disorders*. J Hum Genet, 2012. **57**(10): p. 621-32.
36. Lin, Z., et al., *Gene expression profiles of human chondrocytes during passaged monolayer cultivation*. J Orthop Res, 2008. **26**(9): p. 1230-7.
37. Darling, E.M. and K.A. Athanasiou, *Rapid phenotypic changes in passaged articular chondrocyte subpopulations*. J Orthop Res, 2005. **23**(2): p. 425-32.
38. Yang, M., et al., *CRISPR/Cas9 mediated generation of stable chondrocyte cell lines with targeted gene knockouts; analysis of an aggrecan knockout cell line*. Bone, 2014. **69**: p. 118-25.
39. Seidl, C.I., T.A. Fulga, and C.L. Murphy, *CRISPR-Cas9 targeting of MMP13 in human chondrocytes leads to significantly reduced levels of the metalloproteinase and enhanced type II collagen accumulation*. Osteoarthritis Cartilage, 2019. **27**(1): p. 140-147.
40. D'Costa, S., M.J. Rich, and B.O. Diekman, *Engineered Cartilage from Human Chondrocytes with Homozygous Knockout of Cell Cycle Inhibitor p21*. Tissue Eng Part A, 2020. **26**(7-8): p. 441-449.
41. Adkar, S.S., et al., *Step-Wise Chondrogenesis of Human Induced Pluripotent Stem Cells and Purification Via a Reporter Allele Generated by CRISPR-Cas9 Genome Editing*. Stem Cells, 2019. **37**(1): p. 65-76.
42. Nejadnik, H., et al., *Improved approach for chondrogenic differentiation of human induced pluripotent stem cells*. Stem Cell Rev Rep, 2015. **11**(2): p. 242-53.
43. Diekman, B.O., et al., *Cartilage tissue engineering using differentiated and purified induced pluripotent stem cells*. Proc Natl Acad Sci U S A, 2012. **109**(47): p. 19172-7.
44. Rodriguez Ruiz, A., et al., *Cartilage from human-induced pluripotent stem cells: comparison with neo-cartilage from chondrocytes and bone marrow mesenchymal stromal cells*. Cell Tissue Res, 2021.
45. Dicks, A., et al., *Prospective isolation of chondroprogenitors from human iPSCs based on cell surface markers identified using a CRISPR-Cas9-generated reporter*. Stem Cell Res Ther, 2020. **11**(1): p. 66.
46. George, E.L., et al., *Defects in mesoderm, neural tube and vascular development in mouse embryos lacking fibronectin*. Development, 1993. **119**(4): p. 1079-91.
47. Singh, P. and J.E. Schwarzbauer, *Fibronectin and stem cell differentiation - lessons from chondrogenesis*. J Cell Sci, 2012. **125**(Pt 16): p. 3703-12.
48. Liu, H., et al., *The potential of induced pluripotent stem cells as a tool to study skeletal dysplasias and cartilage-related pathologic conditions*. Osteoarthritis Cartilage, 2017. **25**(5): p. 616-624.
49. Hwang, J.J., et al., *Application of Induced Pluripotent Stem Cells for Disease Modeling and 3D Model Construction: Focus on Osteoarthritis*. 2021. **10**(11): p. 3032.
50. Katagiri, Y., S.A. Brew, and K.C. Ingham, *All six modules of the gelatin-binding domain of fibronectin are required for full affinity*. J Biol Chem, 2003. **278**(14): p. 11897-902.
51. Erat, M.C., et al., *Structural analysis of collagen type I interactions with human fibronectin reveals a cooperative binding mode*. J Biol Chem, 2013. **288**(24): p. 17441-50.
52. Kadler, K.E., A. Hill, and E.G. Canty-Laird, *Collagen fibrillogenesis: fibronectin, integrins, and minor collagens as organizers and nucleators*. Curr Opin Cell Biol, 2008. **20**(5): p. 495-501.
53. Loeser, R.F., C.S. Carlson, and M.P. McGee, *Expression of beta 1 integrins by cultured articular chondrocytes and in osteoarthritic cartilage*. Exp Cell Res, 1995. **217**(2): p. 248-57.
54. Almonte-Becerril, M., et al., *Genetic abrogation of the fibronectin-alpha5beta1 integrin interaction in articular cartilage aggravates osteoarthritis in mice*. PLoS One, 2018. **13**(6): p. e0198559.
55. Tian, J., F.J. Zhang, and G.H. Lei, *Role of integrins and their ligands in osteoarthritic cartilage*. Rheumatol Int, 2015. **35**(5): p. 787-98.
56. Hynes, R.O., *Integrins: bidirectional, allosteric signaling machines*. Cell, 2002. **110**(6): p.

673-87.
57. Ryu, J.Y., H.U. Kim, and S.Y. Lee, *Human genes with a greater number of transcript variants tend to show biological features of housekeeping and essential genes*. Mol Biosyst, 2015. **11**(10): p. 2798-807.
58. Patten, J. and K. Wang, *Fibronectin in development and wound healing*. Adv Drug Deliv Rev, 2021. **170**: p. 353-368.
59. White, E.S. and A.F. Muro, *Fibronectin splice variants: understanding their multiple roles in health and disease using engineered mouse models*. IUBMB Life, 2011. **63**(7): p. 538-46.
60. Jacob, A.G. and C.W.J. Smith, *Intron retention as a component of regulated gene expression programs*. Hum Genet, 2017. **136**(9): p. 1043-1057.
61. Wong, J.J., et al., *Intron retention in mRNA: No longer nonsense: Known and putative roles of intron retention in normal and disease biology*. Bioessays, 2016. **38**(1): p. 41-9.
62. Scanzello, C.R., et al., *Fibronectin splice variation in human knee cartilage, meniscus and synovial membrane: observations in osteoarthritic knee*. J Orthop Res, 2015. **33**(4): p. 556-62.
63. Manabe, R., N. Oh-e, and K. Sekiguchi, *Alternatively spliced EDA segment regulates fibronectin-dependent cell cycle progression and mitogenic signal transduction*. J Biol Chem, 1999. **274**(9): p. 5919-24.
64. Schor, S.L., et al., *Migration-stimulating factor: a genetically truncated onco-fetal fibronectin isoform expressed by carcinoma and tumor-associated stromal cells*. Cancer Res, 2003. **63**(24): p. 8827-36.
65. Homandberg, G.A., R. Meyers, and D.L. Xie, *Fibronectin fragments cause chondrolysis of bovine articular cartilage slices in culture*. J Biol Chem, 1992. **267**(6): p. 3597-604.
66. Homandberg, G.A., *Potential regulation of cartilage metabolism in osteoarthritis by fibronectin fragments*. Front Biosci, 1999. **4**: p. D713-30.
67. Ding, L., D. Guo, and G.A. Homandberg, *Fibronectin fragments mediate matrix metalloproteinase upregulation and cartilage damage through proline rich tyrosine kinase 2, c-src, NF-kappaB and protein kinase Cdelta*. Osteoarthritis Cartilage, 2009. **17**(10): p. 1385-92.
68. Stanton, H., L. Ung, and A.J. Fosang, *The 45 kDa collagen-binding fragment of fibronectin induces matrix metalloproteinase-13 synthesis by chondrocytes and aggrecan degradation by aggrecanases*. Biochem J, 2002. **364**(Pt 1): p. 181-90.
69. Houtman, E., et al., *Human Osteochondral Explants: Reliable Biomimetic Models to Investigate Disease Mechanisms and Develop Personalized Treatments for Osteoarthritis*. Rheumatol Ther, 2021. **8**(1): p. 499-515.
70. Freedman, M.L., et al., *Principles for the post-GWAS functional characterization of cancer risk loci*. Nat Genet, 2011. **43**(6): p. 513-8.
71. Warren, C.R. and C.A. Cowan, *Humanity in a Dish: Population Genetics with iPSCs*. Trends Cell Biol, 2018. **28**(1): p. 46-57.
72. Adkar, S.S., et al., *Genome Engineering for Personalized Arthritis Therapeutics*. Trends Mol Med, 2017. **23**(10): p. 917-931.
73. Wu, C.L., et al., *Single cell transcriptomic analysis of human pluripotent stem cell chondrogenesis*. Nat Commun, 2021. **12**(1): p. 362.
74. Bachman, H., et al., *Utilizing Fibronectin Integrin-Binding Specificity to Control Cellular Responses*. Adv Wound Care (New Rochelle), 2015. **4**(8): p. 501-511.

# Chapter 6

## Appendix

Nederlandse samenvatting

List of publications

Curriculum Vitae

Dankwoord

# Nederlandse samenvatting

## Inleiding

Osteoartrose, ook wel artrose genoemd, is een complexe aandoening van de gewrichten, gekenmerkt door veranderingen in kraakbeen en het onderliggende bot. Artrose kan zich ontwikkelen in nagenoeg ieder gewricht, maar komt het meest voor in de knie-, heup- en handgewrichten. Risicofactoren die de kans op het ontwikkelen van artrose vergroten zijn hogere leeftijd, vrouwelijk geslacht, overgewicht, overbelasting en genetische aanleg. Op het moment is artrose de meest voorkomende gewrichtsziekte wereldwijd en alleen al in Nederland hebben ruim 1.2 miljoen mensen artrose in één of meerdere gewrichten. Door de toename van zowel overgewicht als vergrijzing in de samenleving zal de last op de maatschappij naar verwachting exponentieel toenemen de komende jaren. De ziektelast voor een patiënt bestaat uit pijn en stijfheid in het gewricht, wat kan zorgen voor ernstige beperkingen in het dagelijks leven. De zorgkosten van artrose zijn aanzienlijk en waren 1.4% van de totale zorgkosten in Nederland in 2017. Helaas zijn de beschikbare behandelmethoden zijn op het moment beperkt tot pijnstilling en uiteindelijk een kostbare gewrichtsvervangende operatie op hogere leeftijd. De ontwikkeling van effectieve behandelingen voor artrose blijft onsuccesvol doordat er te weinig inzicht is in de onderliggende pathofysiologie van artrose.

Een van de kenmerken van artrose is de afbraak van het kraakbeen in het gewricht. Kraakbeen bedekt de uiteinden van de botten in onze gewrichten en is ontworpen om de krachten op te vangen die tijdens het lopen, rennen en springen op de gewrichten uitgeoefend worden. In het kraakbeen bevindt zich slechts één celtype die het kraakbeen produceert en onderhoudt, de zogeheten chondrocyten. Chondrocyten zijn gespecialiseerde cellen die zich niet meer delen. In gezond kraakbeen speelt de chondrocyte een belangrijke rol in het onderhouden van de balans tussen de productie en afbraak van kraakbeen componenten. Echter, bij artrose is deze balans verstoord en produceren de chondrocyten te veel kraakbeen afbrekende enzymen, waardoor op een gegeven moment het kraakbeen wordt afgebroken en het zijn optimale eigenschappen verliest.

## Doel van dit proefschrift

In dit proefschrift is getracht meer inzicht te krijgen in onderliggende moleculaire processen in artrotisch kraakbeen op verschillende niveaus. Wat het artroseproces in kraakbeen initieert is nog niet bekend, maar wat wel bekend is, is dat de verstoorde balans van chondrocyten veroorzaakt kan worden door veranderingen in genexpressie van de chondrocyten. Aangezien chondrocyten niet meer delen in het kraakbeen, zijn ze sterk afhankelijk van zogeheten epigenetische mechanismes om hun genexpressie te reguleren. In dit proefschrift hebben we een van deze mechanismes in kaart gebracht, om zo te achterhalen wat een van de oorzaken is van de verstoorde balans in genexpressie van chondrocyten.

De verstoorde balans van chondrocyten kan ook op een ander niveau dan genexpressie worden veroorzaakt. Als er afwijkingen in de genetische code zijn, kunnen er foute genproducten, zogenoemde eiwitten, worden gemaakt, die verantwoordelijk zijn voor het ontwikkelen van artrose. Deze genetische afwijkingen kunnen worden doorgegeven aan de volgende generatie en zo is er overerving van de ziekte. Door de genetische afwijking te vinden die zieke familieleden gemeen hebben, kunnen we deze zodanig herkennen als de oorzaak van artrose. Door de moleculaire gevolgen te ontrafelen die ontstaan als gevolg van afwijkingen in de genetische code, kan inzicht worden verkregen over de oorzaak van artrose. Daarom hebben we de genetische code bestudeerd in een familie waar veel familieleden op jonge leeftijd artrose ontwikkelden. Dit resulteerde in de identificatie van een genetische afwijking in het gen *Fibronectin1* (*FN1*) die de oorzaak was voor artrose op jonge leeftijd. Vervolgens hebben we functionele studies uitgevoerd met betrekking tot *FN1* om de genetische bevindingen te vertalen naar een biologisch mechanisme in kraakbeen.

## Rol van lange niet-coderende RNAs in artrotisch kraakbeen

Zoals eerder genoemd, reguleren chondrocyten hun genexpressie met behulp van zogeheten epigenetische mechanismes. Chondrocyten kunnen op deze manier reageren op externe stimuli, zoals mechanische belasting, zodat het kraakbeen herstelt kan worden als dat nodig is. In chondrocyten in artrotisch kraakbeen vertonen deze epigenetische mechanismes afwijkingen ten opzichte van chondrocyten in gezond kraakbeen. Een van deze mechanismes is de klasse lange niet coderende RNAs (lncRNAs). LncRNAs zijn RNA transcripten langer dan 200 nucleotiden, die structureel sterk lijken op messenger RNA (mRNA), maar niet worden vertaald naar een eiwit. LncRNAs kunnen transcriptie en translatie van mRNAs reguleren via meerdere mechanismes, zoals door het beïnvloeden van de beschikbaarheid van het DNA voor transcriptie, activiteit van transcriptiefactoren en mRNA stabiliteit. Als we kunnen ontrafelen welke lncRNAs essentieel zijn in het reguleren van genexpressie in chondrocyten kunnen we daarna gericht zoeken naar manieren om deze regulatie aan te passen, zodat uiteindelijk het artroseproces verminderd. Gebaseerd op de locatie in de genetische code waar de lncRNAs worden afgeschreven, kunnen ze worden ingedeeld in zogeheten biotypes, zoals ‘antisense’ RNAs, ‘sense’ RNAs, intergenische RNAs en pseudogenen. Omdat op dit moment nog niet bekend is of het biotype de functie van een lncRNA bepaald, wilden we dit verder onderzoeken.

Daarom hebben we in **hoofdstuk 2** eerst de verschillende lncRNAs gekarakteriseerd met betrekking tot hun biotypes in kraakbeen van patiënten die een gewrichtsvervangende operatie zijn ondergaan, van de zogeheten Research Artrose en Articulair Kraakbeen (RAAK) studie. Door de expressie van lncRNAs te vergelijken tussen ‘normaal’ en artrotisch kraakbeen van hetzelfde gewricht hebben we een aantal nieuwe lncRNAs geïdentificeerd die associeerden met het ziekteproces van artrose. Vervolgens wilden we weten of de manier waarop deze lncRNAs genexpressie reguleren gelijk is voor de verschillende biotypes. Hierbij legden we

de focus op intergenische en antisense lncRNAs, omdat die het merendeel van de lncRNAs in chondrocyten omvatten. We observeerden dat een groot aantal lncRNAs omringende dan wel overlappende genen reguleren, ongeacht van het biotype. Experimenteel hebben we dit gevalideerd voor één antisense lncRNA. Hieruit concludeerden we dat als de functie van een lncRNA wordt onderzocht, functionele validatie nodig zal zijn, omdat het biotype geen uitsluitsel zal geven over het mechanisme van regulatie van de lncRNA. De mogelijkheden om de functie van lncRNAs te onderzoeken nemen gestaag toe, waardoor ze potentieel het doelwit kunnen zijn van medicijnen die in de kliniek getest kunnen worden.

## Erfelijke factoren betrokken bij artrose

Om de complexe rol van erfelijkheid in artrose te verklaren zijn genoomwijde associatie studies uitgevoerd. Deze studies hebben aanwijzingen gegeven voor risicogenen die de aanleg voor artrose verhogen in de algemene bevolking. Onderzoek heeft aangetoond dat het ontwikkelen van medicijnen met risicogenen als aangrijpingspunt het klinische succes kan verdubbelen. Helaas mist de vertaling naar klinische therapeutische toepassingen voor artrose, onder andere doordat de gevonden risicogenen vaak kleine effecten geven en de onderliggende mechanismes nog regelmatig onbekend zijn. Wat dat betreft kan het identificeren van zeldzame genetische afwijkingen met grote effecten in families met ernstig artrose inzicht geven in genotype-fenotype relaties. Op die manier kunnen moleculaire mechanismes worden gevonden, waarbij mildere afwijkingen in dezelfde genen ook een rol kunnen spelen bij het ontstaan van artrose in de algemene bevolking.

In **hoofdstuk 3** is beschreven hoe we de genetische afwijking hebben gevonden die verantwoordelijk is voor het ontwikkelen van ernstige artrose op jonge leeftijd in een familie. Hiervoor hebben we van één familielid met artrose de code van het DNA bepaald van alle genen, met behulp van zogeheten exome sequencing. De genetische afwijkingen die verantwoordelijk zijn voor het ontstaan van ernstige artrose resulteren vaak in veranderingen van functie van het eiwit dat afgelezen wordt van het gen. Daarom hebben we bij dit familielid geselecteerd op afwijkingen waarvan voorspeld werd dat ze een effect zouden hebben op de functie van het eiwit. Deze analyse resulteerde in de identificatie van een afwijking in het gen *FN1*, die verantwoordelijk was voor artrose op jonge leeftijd in de familie. Dit werd verder bevestigd doordat alle aangedane familieleden inderdaad drager waren van deze afwijking en vice versa. Het identificeren van een pathogene genetische afwijking in zeldzame ziektes is niet altijd succesvol, aangezien het humane genoom duizenden varianten bevat. Echter, in dit proefschrift laten we de krachtige combinatie zien van exome sequencing en strenge selectiecriteria van afwijkingen in het DNA van een familie om tot de identificatie van de causale genetische afwijking te komen.

## De rol van fibronectine in pathofysiologie van artrose

In het onderzoeksveld van artrose zijn er weinig functionele opvolgstudies gedaan van genetische afwijkingen. Recentelijk is er grote vooruitgang geboekt in het veld dankzij de zogeheten CRISPR/Cas9 technologie, waar we in dit proefschrift dan ook van hebben geprofiteerd. CRISPR/Cas9 is de afkorting van 'Clustered Regularly Interspaced Short Palindromic Repeats' en 'CRISPR-associated protein 9' en was ontdekt als het verdedigingsmechanisme van bacteriën tegen bacteriofagen. Hierbij herkent de bacterie stukjes genetisch materiaal van een virus en Cas9 is het eiwit dat dit vervolgens kapot knipt, zodat het virus niet meer functioneel is. Tegenwoordig weten we hoe het systeem zo kunnen ontwerpen dat het Cas9 eiwit precies knipt op een gewenste plek in het DNA en dat het wordt gerepareerd met de gewenste veranderingen. Voor onderzoek in het veld van artrose is het relevant om CRISPR/Cas9 toe te passen op bijvoorbeeld chondrocyten, om te onderzoeken wat het effect is van een genetische afwijking op kraakbeen. Echter, de chondrocyten die we verzamelen met de RAAK studie stoppen op een gegeven moment met zowel delen als nieuw kraakbeen maken in het laboratorium. Daarom hebben we ervoor gekozen 'human induced pluripotent stem cells' (hiPSCs) te gebruiken. hiPSCs zijn pluripotente stamcellen die gemaakt worden van lichaamscellen en zich voor geruime tijd blijven delen. Ook kunnen deze stamcellen zich differentiëren in elk gewenste celtype van het menselijk lichaam en dus ook in chondrocyten.

In **hoofdstuk 3** staat beschreven hoe we het model opgezet hebben om het biologische mechanisme van de eerder genoemde genetische afwijking in het gen *FN1* te onderzoeken. Hiervoor hebben we met CRISPR/Cas9 deze afwijking geïntroduceerd in het DNA van hiPSCs. Vervolgens werden de (gemuteerde en niet-gemuteerde) stamcellen gedifferentieerd naar chondrocyten die kraakbeen gingen produceren. We observeerden dat zowel de potentie om kraakbeencel te worden als de productie van nieuw kraakbeen verminderd was van de *FN1* mutant cellen. Bovendien hebben we laten zien dat het onderliggende mechanisme van de genetische afwijking wordt veroorzaakt door een verlaagde binding tussen mutant fibronectine en collageen type II, wat zorgt voor verstoorde signalen tussen kraakbeen en cel.

Naast de karakterisering van deze zeldzame variant, is het ook van belang om de algemene functie van fibronectine in kraakbeen te onderzoeken. Het was al bekend dat de genexpressie van *FN1* enorm hoog is in kraakbeen en nog verder toeneemt in artrotisch kraakbeen, onder andere van bevindingen uit de RAAK studie. Bovendien kunnen er door alternatieve splicing 27 verschillende transcripten worden geproduceerd van het *FN1* gen. Het is echter niet bekend welke functie deze verschillende transcripten van *FN1* hebben in gezond en artrotisch kraakbeen. Daarom onderzochten we in **hoofdstuk 4** welke specifieke *FN1* transcripten aanwezig zijn in kraakbeen en hoe het patroon verandert met artrose. We vonden onder andere een verkorte isoform genaamd 'migration-stimulating factor' (MSF, of *FN1-208*) die significant verhoogde genexpressie had, die niet eerder was geïdentificeerd in (artrotisch)

kraakbeen. Om te trachten het onderliggende moleculaire mechanisme te onderzoeken van de sterke verhoogde expressie van MSF op het vormen van nieuw kraakbeen hebben we een modelsysteem opgezet met primaire chondrocyten verkregen in de RAAK studie, die kraakbeen in het laboratorium maken en gemoduleerde expressie van MSF hadden. Resultaten van de experimenten lieten zien dat het verlagen van de genexpressie van alle overige *FN1* transcripten ongunstig was voor de productie van nieuw kraakbeen. Dit impliceert dat de eerder geobserveerde verhoogde genexpressie van *FN1* en MSF in de RAAK studie een respons is en niet causaal aan het artroseproces.

## Conclusie en toekomstperspectieven

In dit proefschrift laten we zien dat het identificeren van een zeldzame variant en de biologische gevolgen daarvan inzicht kunnen geven in onderliggende mechanismes die een rol spelen in het ontstaan van artrose. De toepasbaarheid van flexibele humane hiPSC modellen zoals in dit proefschrift omschreven kan vergroot worden voor toekomstige studies. Voorbeelden van nieuwe toepassingen zijn het testen van medicijnen, differentiatie naar meerdere cel types en de interactie daartussen onderzoeken, het functioneel onderzoeken van gangbare genetische varianten en potentieel als stamceltherapie in 'regenerative medicine'. Betreffende de rol van fibronectine in artrotisch kraakbeen, hebben we laten zien dat goede binding tussen fibronectine en de kraakbeen matrix via collageen type II essentieel is voor communicatie tussen chondrocyte en kraakbeen. Onze studies tonen aan dat fibronectine als potentieel doelwit van medicatie moet worden onderzocht.

Terugkijkend op de afgelopen jaren van artrose onderzoek, wordt het duidelijk dat er snelle voortgang is geboekt dankzij het toepassen van nieuwe technieken zoals hiPSCs en CRISPR/Cas9. Dit geeft hoop voor de toekomst van zowel artrose onderzoek als de ontwikkeling van medicijnen. De brede studies beschreven in dit proefschrift kunnen gebruikt worden als startpunt voor toekomstige artrose modelsystemen en het testen van nieuwe medicijnen.

## List of publications

**van Hoolwerff M**, Tuerlings M, Wijnen IJL, Suchiman HED, Cats D, Mei H, Nelissen RGHH, van der Linden - van der Zwaag HMJ, Ramos YFM, Coutinho de Almeida R, Meulenbelt I. Identification and functional characterization of imbalanced osteoarthritis associated fibronectin splice variants. Rheumatology (Oxford). 2022

Rodríguez Ruiz A*, **van Hoolwerff M***, Sprangers S, Suchiman E, Schoenmaker T, Dibbets-Schneider P, Bloem JL, Nelissen RGHH, Freund C, Mummery C, Everts V, de Vries TJ, Ramos YFM, Meulenbelt I. Mutation in the CCAL1 locus accounts for bidirectional process of human subchondral bone turnover and cartilage mineralization. Rheumatology (Oxford). 2022

Timmermans RGM, Bloks NGC, Tuerlings M, **van Hoolwerff M**, Nelissen RGHH, van der Wal RJP, van der Kraan PM, Blom AB, van den Bosch MHJ, Ramos YFM, Meulenbelt I. A human in vitro 3D neo-cartilage model to explore the response of OA risk genes to hyper-physiological mechanical stress. Osteoarthr. Open. 2022

**van Hoolwerff M***, Rodríguez Ruiz A*, Bouma M, Suchiman HED, Koning RI, Jost CR, Mulder AA, Freund C, Guilak F, Ramos YFM, Meulenbelt I. High-impact FN1 mutation decreases chondrogenic potential and affects cartilage deposition via decreased binding to collagen type II. Sci Adv. 2021

Tuerlings M, **van Hoolwerff M**, van Bokkum JM, Suchiman HED, Lakenberg N, Broekhuis D, Nelissen RGHH, Ramos YFM, Mei H, Cats D, Coutinho de Almeida R, Meulenbelt I. Long non-coding RNA expression profiling of subchondral bone reveals AC005165.1 modifying FRZB expression during osteoarthritis. Rheumatology (Oxford). 2021

Tuerlings M, **van Hoolwerff M**, Houtman E, Suchiman HED, Lakenberg N, Mei H, van der Linden HMJ, Nelissen RGHH, Ramos YFM, Coutinho de Almeida R, Meulenbelt I. RNA Sequencing Reveals Interacting Key Determinants of Osteoarthritis Acting in Subchondral Bone and Articular Cartilage: Identification of IL11 and CHADL as Attractive Treatment Targets. Arthritis Rheumatol. 2021

Houtman E, **van Hoolwerff M**, Lakenberg N, Suchiman EHD, van der Linden-van der Zwaag E, Nelissen RGHH, Ramos YFM, Meulenbelt I. Human Osteochondral Explants: Reliable Biomimetic Models to Investigate Disease Mechanisms and Develop Personalized Treatments for Osteoarthritis. Rheumatol Ther. 2021

**van Hoolwerff M**, Metselaar PI, Tuerlings M, Suchiman HED, Lakenberg N, Ramos YFM, Cats D, Nelissen RGHH, Broekhuis D, Mei H, de Almeida RC, Meulenbelt I. Elucidating Epigenetic Regulation by Identifying Functional cis-Acting Long Noncoding RNAs and Their Targets in Osteoarthritic Articular Cartilage. Arthritis Rheumatol. 2020

Coutinho de Almeida R, Ramos YFM, Mahfouz A, den Hollander W, Lakenberg N, Houtman E, **van Hoolwerff M**, Suchiman HED, Rodríguez Ruiz A, Slagboom PE, Mei H, Kiełbasa SM, Nelissen RGHH, Reinders M, Meulenbelt I. RNA sequencing data integration reveals an miRNA interactome of osteoarthritis cartilage. Ann Rheum Dis. 2019

Albulescu IC, **van Hoolwerff M**, Wolters LA, Bottaro E, Nastruzzi C, Yang SC, Tsay SC, Hwu JR, Snijder EJ, van Hemert MJ. Suramin inhibits chikungunya virus replication through multiple mechanisms. Antiviral Res. 2015

* equal contribution

## Curriculum Vitae

Marcella van Hoolwerff was born on September 27th in 1990 in Delft, the Netherlands. She attended secondary school College Het Loo in Voorburg, where she passed her exams in 2008. To pursue her passion for cell- and microbiology, she enrolled in the study Life Science and Technology, which is a collaboration between Technical University Delft in Delft and Leiden University in Leiden and encompasses a broad curriculum. She obtained her bachelor's degree in 2013 and continued het academic career by enrolling in the Life Science and Technology master's program at Leiden University, Leiden.

During her master's, she worked at the department of Medical Microbiology at the Leiden University Medical Center in the group of dr. Martijn van Hemert, where she investigated suramin as antiviral drug. Subsequently, she obtained experience abroad at the Edison Biotechnology Institute at Ohio University in Athens, Ohio, USA, under the supervision of prof. dr. John Kopchick and dr. Ed List. Here she investigated the ageing of the adaptive immune system of growth hormone receptor knockout mice compared to age-matched wildtype controls. She obtained her master's degree in 2016 and started her PhD studies in the group of prof. dr. Ingrid Meulenbelt at the department of Biomedical Data Sciences of the Leiden University Medical Center in the beginning of 2017. Her research involved elucidating the role of fibronectin in the pathophysiology of osteoarthritis, which was funded by the Dutch Arthritis Society. During her PhD she presented her work at national and international conferences, including the Nederlandse Vereniging voor Matrix Biologie (NVMB), international workshop on the epigenetics of Osteoarthritis, annual meeting of the Orthopaedic Research Society (ORS), and annual meeting of the Osteoarthritis Research Society International (OARSI). At the OARSI annual meeting of 2022, Marcella was awarded an Early Career Investigator award.

Currently, Marcella continues her work in the group of prof. dr. Ingrid Meulenbelt as a post-doctoral researcher in a collaborative project with Technical University Twente to unravel the factors that determine key factors in hiPSC differentiation to chondroprogenitors.

## Dankwoord

Na ruim 5 jaar is het einde van mijn promotietraject dan echt daar. In het begin lijkt het zo lang, maar de tijd is toch omgevlogen, met onderweg genoeg uitdagingen en overwinningen. Uiteraard is dit proefschrift met een hoop hulp tot stand gekomen en ik wil graag de mensen bedanken die er aan hebben bijgedragen.

Geachte Professor Meulenbelt, beste Ingrid, bedankt voor de mogelijkheid om een promotie te doen in jouw groep en je begeleiding tijdens dit traject. Ik waardeer het enorm dat jouw deur staat altijd open staat voor vragen en feedback op posters, teksten en presentaties.

Geachte Dr. Ramos, beste Yolande, ik heb mijn basis in het lab van MolEpi te danken aan jouw directe begeleiding, waar ik vervolgens zelf verder op kon doorbouwen. En later in de jaren was jouw feedback op mijn artikelen wat ze tot een hoger niveau tilde.

Geachte leden van mijn promotiecommissie: Professor Heijmans, Professor van der Kraan, en Professor Zaucke. Hartelijk dank dat jullie gelegenheid hebben gevonden dit manuscript te lezen en voor de waardevolle suggesties ter verbetering van dit proefschrift. Geachte Professor Bovée, beste Judith, bedankt voor de meetings met de groepen waarbij we met collega PhD's laagdrempelig ons onderzoek konden bespreken.

Vervolgens wil ik graag mijn collega's bedanken die ook veel werk hebben verricht voor dit proefschrift. Als eerste Eka, dankjewel dat je mijn paranimf wilt zijn! Van samen kweken en mega agarose gels doen tot thee pauzes en borrels, dankzij jou was mijn promotietijd een stuk aangenamer. Alejandro, we started our PhD journey at the same time in the OA group and now we will both defend our thesis this year, what a ride it has been. We shared the struggles of the hiPSC (weekend) work and now we're ready to take on the world, right? Verder gaat mijn dank uit naar alle (oud) collega's van de OA groep en de studenten die ik mocht begeleiden: Ghazaleh, Rodrigo, Nicoline, Evelyn, Margo, Ilja, Mathew, Rick, Niek, Ritchie, Nico, Selina, Paula en Imke.

Verder wil ik alle collega's van MolEpi graag bedanken voor de feedback en vragen tijdens de werkbesprekingen en de fijne sfeer bij de altijd gezellige MolEpi labuitjes. Expliciet wil ik Eline graag bedanken, ik voel me altijd geïnspireerd als we elkaar gesproken hebben op een borrel, Nieuwjaarslunch of bij de voorbereiding van Ingrid's feest voor haar oratie. Inge, hartelijk bedankt voor al je hulp en het altijd beantwoorden van vragen rondom logistieke dingen.

Verder wil ik de mensen bedanken die hebben bijgedragen door onze samenwerkingen: Professor Nelissen, Demiën, Enrike, Davy, Leon, Roman, Carolina, Aat, Christian, en Professor Guilak. Thank you for the great collaborations. Ook wil ik graag de artrose-patiënten

bedanken voor hun medewerking aan de RAAK biobank en daarmee hun bijdrage aan het wetenschappelijke onderzoek.

Tenslotte wil ik graag mijn vrienden en familie bedanken voor hun steun en dankzij wie ik ook groei als persoon. De vriendschappen die ontstonden uit werk, Thies, Pia, en Daniele, af en toe ontmoet je mensen op werk met wie het echt klikt en ik waardeer jullie enorm. Mauricio, met jou kletsen is altijd heerlijk en jouw relaxte kijk op het leven breng je over op mij, “dat is het leven” is nu ook een gevleugelde uitspraak van mij. Mijn vriendinnen met wie mijn avonden en weekenden gezellig en/of sportief werden: Loes, Louisa, Gülben, Safi, en Stèph. De rest van de LST gang, Rik, Laurens, Daan en Ruben en aanhang natuurlijk. Voor de gezellige tijden in Tilburg en Den Haag, maar ook voor de hulp bij klussen in het huis. De DOPC’ers van Plunk, bedankt voor de top borrels, weekendjes, en vakanties. Richard en Larissa, lieve neef en nicht, jullie zijn familie maar vooral eerst vrienden. Selina, thank you so much for your help with the lay-out of this thesis.

Lieve Vincent, lil’ bro. Dankjewel dat je vandaag mijn paranimf wilt zijn. Ik vind het zo fijn dat we zo’n goede band hebben opgebouwd en ik kijk uit naar nog vele jaren en avonturen samen! Lieve mama, dankjewel voor je oneindige liefde en steun. Van mij krijg je altijd een 10 als moeder en ik weet dat je super trots bent. Lieve papa, ik weet dat je er vandaag bij bent, ook al kunnen we je niet zien, en ook heel trots bent. Dit proefschrift draag ik aan jou op, omdat jij als enige in de familie er nog een beetje van zou kunnen begrijpen.