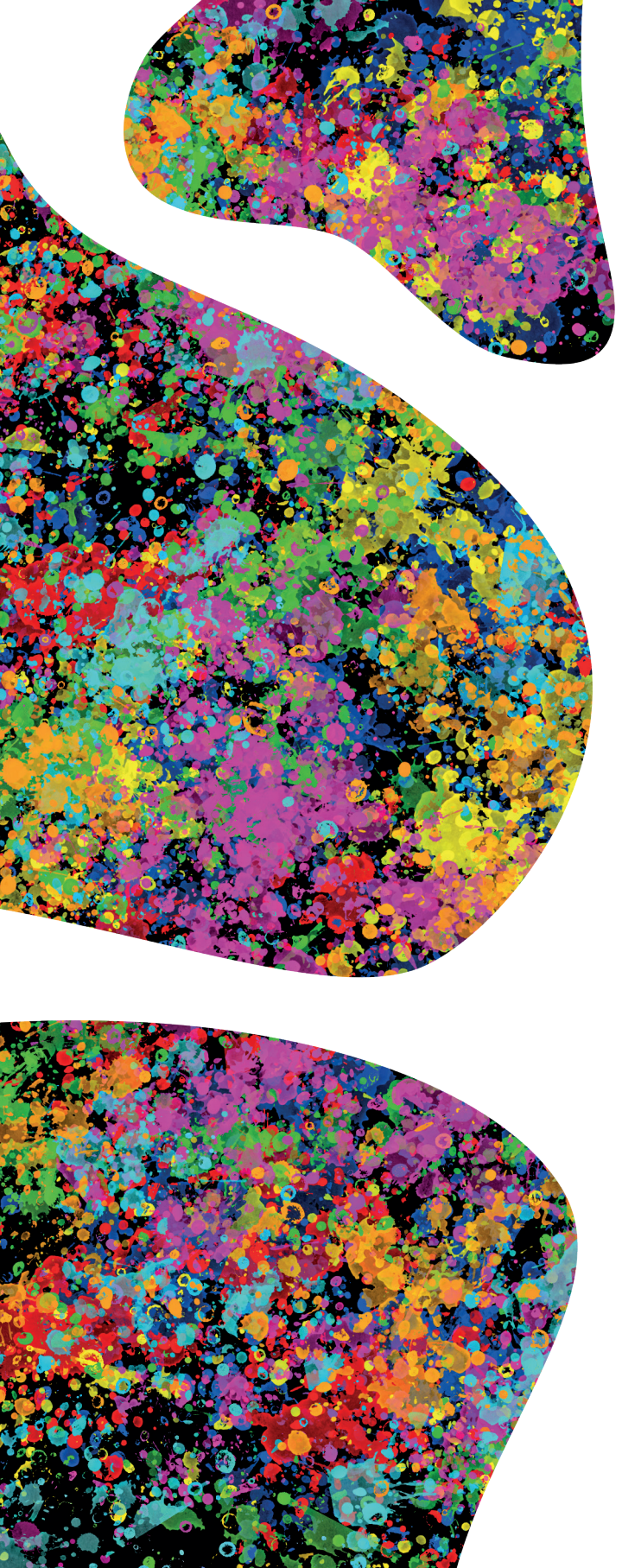
The bone and cartilage interplay in osteoarthritis: key to effective treatment strategy

Tuerlings, M.



CHAPTER 3

Long non-coding RNA expression profiling of subchondral bone reveals AC005165.1 modifying FRZB expression during osteoarthritis

Margo Tuerlings[1], Marcella van Hoolwerff[1], Evelyn Houtman[1], H. Eka D. Suchiman[1], Nico Lakenberg[1], Hailiang Mei[1], Enrike H.M.J. van der Linden[2], Rob G.H.H. Nelissen[2], Yolande F.M. Ramos[1], Rodrigo Coutinho de Almeida[1], Ingrid Meulenbelt[1]

[1] Dept. of Biomedical Data Sciences, Leiden University Medical Center, Leiden, The Netherlands.
[2] Dept. Orthopaedics Leiden University Medical Center, Leiden, The Netherlands.

Rheumatology, volume 61, issue 7, July 2022, pages 3023-3032
DOI: 10.1093/rheumatology/keab826

Abstract

Objective: To gain insight in the expression profile of long non-coding RNAs (lncRNAs) in OA subchondral bone.

Methods: RNA sequencing data of macroscopically preserved and lesioned OA subchondral bone of patients that underwent joint replacement surgery due to OA (N=22 pairs; 5 hips, 17 knees, RAAK-study) was run through an in-house pipeline to detect expression of lncRNAs. Differential expression analysis between preserved and lesioned bone was performed. Spearman correlations were calculated between differentially expressed lncRNAs and differentially expressed mRNAs identified previously in the same samples. Primary osteogenic cells were transfected with Locked nucleic acid (LNA) GapmeRs targeting *AC005165.1* lncRNA, to functionally investigate its potential mRNA targets.

Results: In total, 2816 lncRNAs were well-expressed in subchondral bone and we identified 233 lncRNAs exclusively expressed in knee and 307 lncRNAs exclusively in hip. Differential expression analysis, using all samples (N=22 pairs; 5 hips, 17 knees), resulted in 21 differentially expressed lncRNAs (false discovery rate (FDR)<0.05, Fold change (FC) range:1.19-7.39), including long intergenic non-protein coding RNA (LINC) 1411 (*LINC01411,* FC=7.39, FDR=2.20x10^{-8}), *AC005165.1* (FC=0.44, FDR=2.37x10^{-6}), and embtyp spiracles homeobox 2 opposite strand RNA (*EMX2OS,* FC=0.41, FDR=7.64x10^{-3}). Among the differentially expressed lncRNAs, five were also differentially expressed in articular cartilage, including *AC005165.1,* showing similar direction of effect. Downregulation of *AC005165.1* in primary osteogenic cells resulted in consistent downregulation of highly correlated frizzled related protein (*FRZB*).

Conclusion: The current study identified a novel lncRNA, *AC005165.1,* being dysregulated in OA articular cartilage and subchondral bone. Downregulation of *AC005165.1* caused a decreased expression of OA risk gene *FRZB*, an important member of the wnt pathway, suggesting that *AC005165.1* could be an attractive potential therapeutic target with effects in articular cartilage and subchondral bone.

Introduction

OA is a highly prevalent degenerative joint disease, characterised by articular cartilage degradation and subchondral bone remodelling [1-3]. Since OA is now considered a disease of the whole joint, recently focus has shifted towards characterization of gene expression profiles in OA synovium and subchondral bone [4, 5]. In this respect, we reported on mRNA expression profiling of OA subchondral bone of knee and hip joints [6]. We observed clustering of the samples based on joint site, suggesting distinct subchondral bone OA pathophysiological processes. This indicates that future therapeutic strategies particularly targeting bone should consider such differences between joint sites.

Different epigenetic mechanisms are described in OA, each of them modifying gene expression upon environmental cues such as mechanical stress or disease, without changing the genetic code. Among these, DNA methylation, histone modifications and miRNA expression are most frequently studied in OA articular cartilage [1, 7-11]. In contrast, the role of long non-coding RNAs (lncRNAs) with OA pathophysiology is less explored as they show poor conservation between species [9]. LncRNAs are typically defined as RNAs >200 nucleotides in length, with little or no coding potential, and they are known to be involved in various transcriptional and (post-)translational processes, such as chromatin remodelling, mRNA/protein stabilization, production of short interfering RNAs (siRNAs) and recruitment of scaffolding proteins, or they might act as pseudogenes [12, 13]. Moreover, the expression of lncRNAs can be highly tissue- and disease specific [14, 15]. Due to the fact that OA is a disease of the whole joint, it is of added value to identify disease specific lncRNAs that are expressed in various tissues involved in the OA pathophysiology, since these lncRNAs might serve as a potential druggable target with effects in several disease-relevant tissues.

Upon applying an in-house developed pipeline to reliably detect lncRNAs from RNA sequencing, we recently reported on the characterization of lncRNAs in OA cartilage. Notably, we identified prolyl 3-hydroxylase 2 antisense (*P3H2-AS1*) as being differentially expressed between macroscopically preserved and lesioned OA cartilage and this was shown to regulate prolyl 3-hydroxylase 2 (*P3H2*) in *cis* [16]. Ajekigbe et al. [17] also reported on the expression levels of lncRNAs in OA cartilage, identifying among others *LINC01411* and *AC003090.1* as being differentially expressed between intact and damage OA cartilage from knees. Furthermore, Sun et al. [14] summarized the findings on the identification of lncRNAs involved in osteogenesis, such as maternally expressed 3 (*MEG3*), metastasis associated lung adenocarcinoma transcript 1 (*MALAT1*), and differentiation antagonizing non-protein coding RNA (D*ANCR*). To our knowledge, however, there are no studies yet focussing on the characterization of

lncRNA expression profiles with ongoing OA in subchondral bone.

In the current study, we set out to characterize the lncRNA expression profile in subchondral bone using RNA sequencing data of patients that underwent joint replacement surgery due to OA (RAAK study). First, joint-specific lncRNAs expressed in OA subchondral bone were identified. Differential expression analysis comparing macroscopically preserved and lesioned OA bone (N=22 paired samples) was then performed to identify robust differentially expressed lncRNAs. To investigate the role of the differentially expressed lncRNAs identified herein with OA pathophysiology, we correlated the expression levels of these lncRNAs with the expression levels of our previously identified differentially expressed mRNAs in subchondral bone of the same patients [6]. Finally, we functionally investigated the effect of a specific lncRNA on mRNA expression levels in primary osteogenic cells.

Methods

Sample description

The current study includes N=41 participants of the RAAK study [2], who underwent a joint replacement surgery due to OA (**Supplementary Table 1**). Macroscopically preserved and lesioned subchondral bone were collected from the joints of 37 of the 41 participants, for either RNA-sequencing (N=22) or replication by means of reverse transcriptase-quantitative PCR (RT-qPCR) (N=15) (**Supplementary Table 1A-1B)**. Osteogenic cells were collected from 4 of the 41 participants (**Supplementary Table 1C**). The classification of macroscopically preserved and lesioned OA subchondral bone was based on its preserved and lesioned classified overlying cartilage as described previously [2]. The results reported here were compared to our recently reported results on the expression of lncRNAs in OA articular cartilage [16], in which 98 samples were used (65 knees, 33 hips). Of these OA articular cartilage samples, 10 paired samples did overlap with the OA subchondral bone samples, i.e. of these 10 patients we had preserved and lesioned OA articular cartilage and OA subchondral bone. Written informed consent was obtained from all participants of the RAAK study and ethical approval for the RAAK study was given by the medical ethics committee of the Leiden University Medical Center (P08.239/P19.013).

RNA sequencing

Sequencing was performed on preserved and lesioned OA subchondral on the Illumina HiSeq4000 (San Francisco, California, USA). Detailed information on the RNA isolation, alignment, mapping, and filtering on lncRNAs is available in the **Supplementary methods**. To identify outliers, principal component analysis and hierarchical clustering

on the samples was applied. Three extreme outliers were identified (**Supplementary Figure 1**) and upon performing sensitivity analysis, these outliers were removed from the dataset. Finally, non-paired samples were removed from the dataset resulting in 22 paired samples (N=17 paired knee samples, N=5 paired hip samples) for further analysis, of which 10 paired samples were overlapping with the cartilage samples of our previous study [16].

LncRNA expression

To identify the lncRNAs that are expressed in subchondral bone, we filtered the lncRNAs identified by our in-house pipeline on a minimal average read count of four and a minimal count of two in at least 80% of the samples, indicated as robustly expressed. Cluster analysis was based on Euclidean distance and a heatmap was created using the lncRNAs that were expressed in the total dataset, the knee dataset, and the hip dataset.

Differential expression analysis

Prior to the differential expression analysis, the lncRNAs were filtered on a minimum average read counts of 4 to allow variation. Differential expression analysis was performed between preserved and lesioned OA subchondral bone. The results were validated and replicated by means of RT-qPCR. Additional information is available in the **Supplementary Methods**.

Correlation analysis

Correlation between the expression levels of previously identified differentially expressed mRNAs in subchondral bone [6] and the expression levels of the here identified differentially expressed lncRNAs in subchondral bone was calculated using a Spearman correlation. Additional information is available in the **Supplementary Methods**.

Functional validation of AC005165.1

Primary osteogenic cells were isolated from the OA joints (**Supplementary Table 1C**), resulting in isolation of a mixture of bone cells, which was characterized by measuring osteogenic and chondrogenic markers (**Supplementary Figure 2**). Subsequently, osteogenic cells were transfected with antisense locked nucleic acid (LNA) GapmeRs (Qiagen, Hilden, Germany) targeting *AC005165.1* or GapmeR negative control. RT-qPCR was performed to measure gene expression levels. Additional information is available in **Supplementary methods**.

Data availability

The RNA-sequencing data is deposited at the European Genome-Phenome Archive

(accession number: EGAS00001004476).

A complete overview of the approach applied to identify lncRNAs being expressed in subchondral bone is shown in **Figure 1A**. An overview of the approach applied on identification of differential expressed lncRNAs with OA pathophysiology is shown in **Figure 1B**.

Results

Expression of lncRNAs in OA subchondral bone

Initially, we explored the expression profile of lncRNAs in OA subchondral bone (**Figure 1A**). We applied our in house pipeline [16] on an RNA sequencing dataset of 22 paired samples (5 hips, 17 knees, **Supplementary Table 1A**) of macroscopically lesioned and preserved OA subchondral bone. Henceforth, we filtered on a minimal average read count of four and a minimal count of two in at least 80% of the samples, and we identified 2816 lncRNAs robustly expressed in OA subchondral bone.

Since we observed major differences in mRNA expression levels between knee and hip OA subchondral bone in our previous study [6], we also explored lncRNA expression patterns in knee and hip subchondral bone separately, while including both preserved and lesioned samples. As shown in **Figure 2**, we identified 2057 overlapping lncRNAs commonly expressed in the hip, knee, and total datasets (mean counts between 4.02 and $3.40x10^5$; **Supplementary Table 2**). Moreover, we identified 233 exclusive knee lncRNAs (mean counts between 4.0 and 23; **Supplementary Table 3**) and 307 exclusive hip lncRNAs (mean counts between 4.0- 892; **Supplementary Table 4**).

To investigate differences in expression levels of commonly expressed lncRNAs in knee and hip subchondral bone samples (N=2057 lncRNAs, **Figure 2**), we performed cluster analysis based on these commonly expressed lncRNAs using the Euclidian distance (**Figure 3**). We observed, similar to the mRNA profile of subchondral bone, clustering of lncRNA expression profiles based on joint site. To investigate which lncRNAs are most contributing to this clustering, we performed differential expression analysis between the two clusters, with the hip cluster set as a reference. More specifically, we found 1069 lncRNAs being significantly differentially expressed between the two clusters (**Supplementary Table 5**). The lncRNAs showing the highest fold difference (FD), i.e. lncRNAs highly expressed in knee samples, were *AC068724.4* (FD=158.87), *AL034397.3* (FD=157.82), and *LINC02009* (FD=89.21), while the lncRNAs with the lowest FD, i.e. highly expressed in hip samples, were *AC105046.1* (FD=0.15), *TGFB2-OT1* (FD=0.21), and *LINC02328* (FD=0.21).

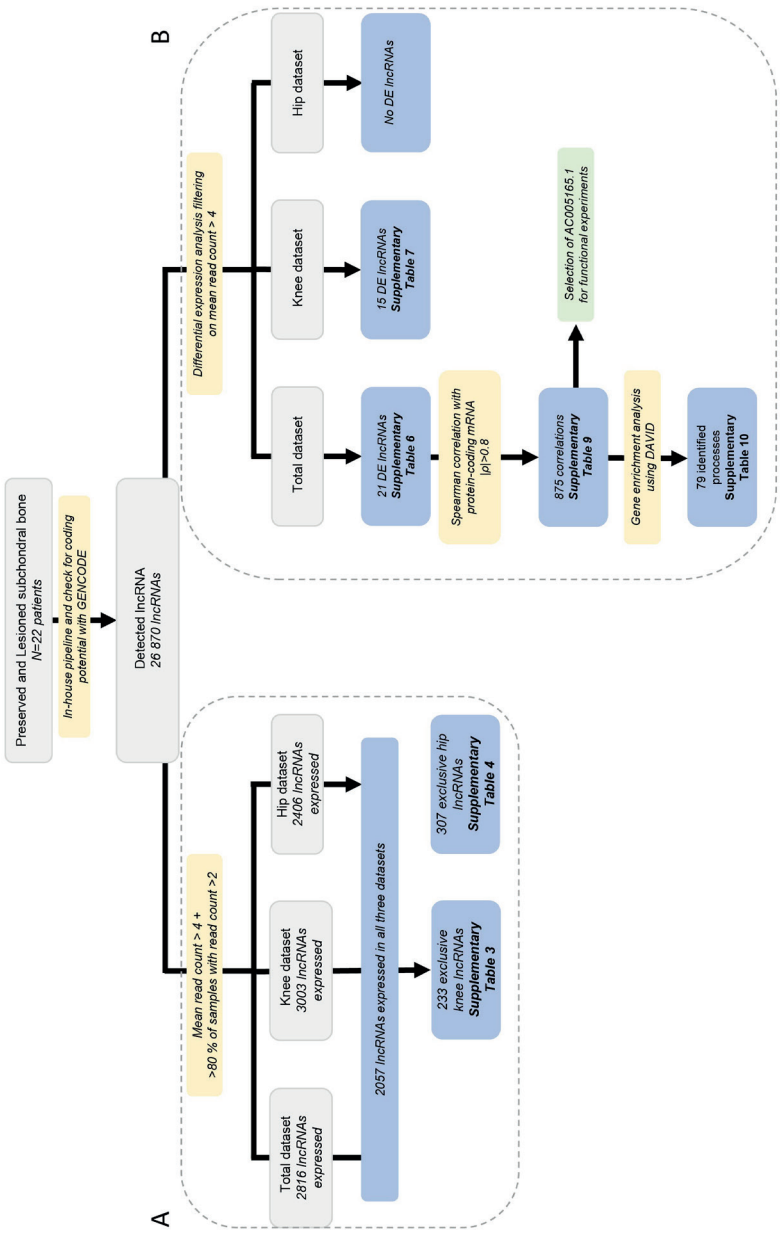


Figure 1 - Schematic overview of applied strategy
(A) Identification of expressed lncRNAs. (B) identification of lncRNAs differentially expressed between macroscopically preserved and lesioned OA subchondral bone. DE: differentially expressed; lncRNA: long non-coding RNA.

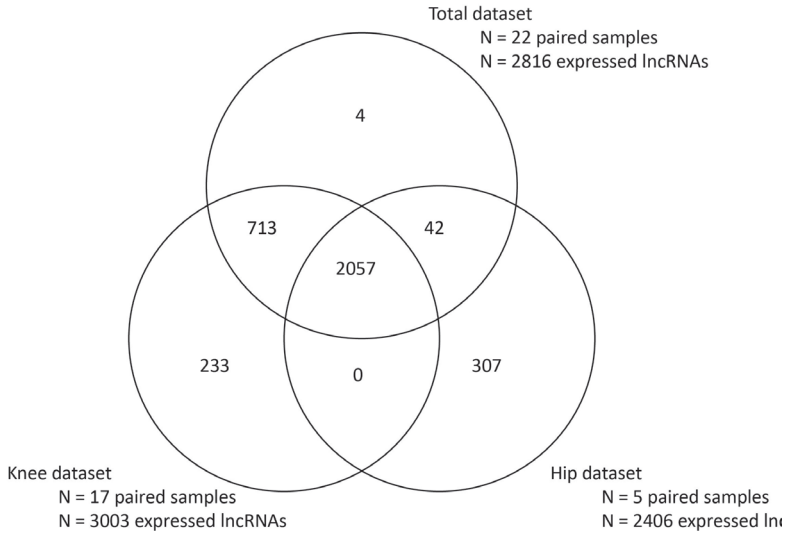


Figure 2 - Venn diagram of lncRNAs being expressed in the total, knee, and hip dataset of preserved and lesioned OA subchondral bone.

Differential expression analysis of lncRNAs in OA subchondral bone

Next, we explored lncRNAs that change expression levels with OA pathophysiology, using a slightly different selection criteria to allow more variation (**Figure 1B**). To identify robust lncRNAs that are associated with the OA pathophysiological process in subchondral bone, we filtered lncRNAs on a minimal average read count of four and we performed differential expression analysis between preserved and lesioned OA subchondral bone samples (knees and hips together). We identified 21 lncRNAs being false discovery rate (FDR) significantly differentially expressed between preserved and lesioned OA subchondral bone (**Figure 4, Supplementary Table 6**). Among these, *LINC01411* (FC=7.39, FDR=2.20x10^{-8}) showed the highest and most significant upregulation, while *AC005165.1* (FC=0.44, FDR=2.37x10^{-6}) showed the most significant downregulation and *EMX2OS* (FC=0.41, FDR=7.64x10^{-3}) the largest downregulation in lesioned compared to preserved OA subchondral bone. Differential expression analysis stratifying for joint site resulted in the identification of 15 lncRNAs being FDR significantly differentially expressed between preserved and lesioned knee samples (N=17 paired samples, **Supplementary Figure 3A**), of which cancer susceptibility 15 (*CASC15,* FC=1.48, FDR=2.67x10^{-2}) and *AL135926.1* (FC=1.70, FDR=9.92x10^{-5}) appeared to be exclusive knee lncRNAs, i.e. not significantly differentially expressed in the total nor the hip dataset (**Supplementary Table 7**). We did not find any significantly differentially expressed lncRNAs between preserved and lesioned hip samples (N=6 paired samples, **Supplementary Figure 3B**). To validate and replicate the results of the differential

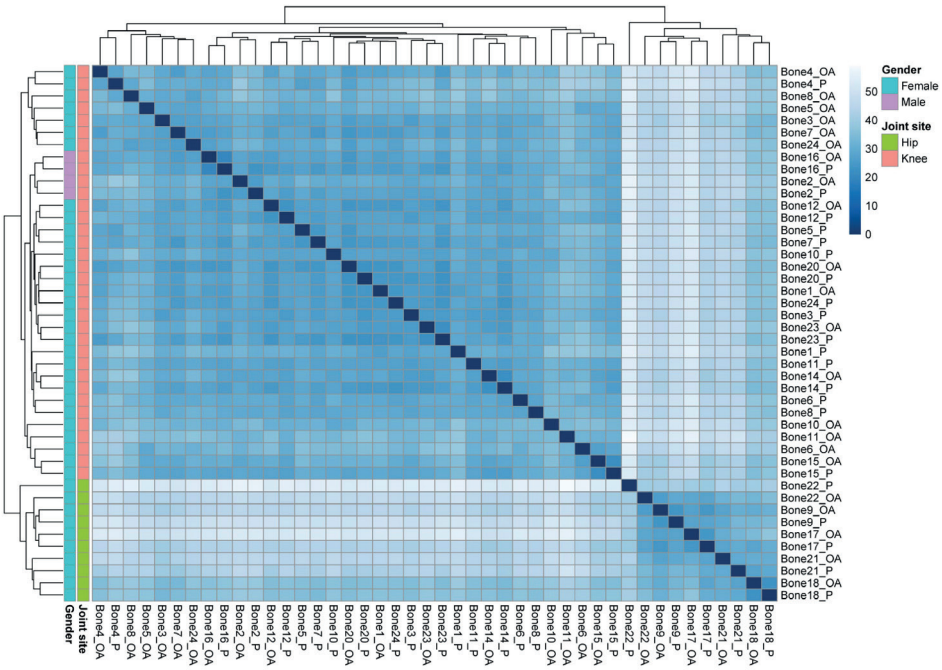


Figure 3 - Heatmap of sample distance
Heatmap is based on lncRNA expression levels of lncRNAs (N = 2057) expressed in all three datasets (i.e. total, hip and knee dataset of preserved and lesioned OA subchondral bone).

expression analysis by means of RT-qPCR, we included 9 paired samples for technical validation, i.e. samples overlapping with the RNA-sequencing dataset, and 15 paired samples for biological validation, i.e. additional preserved and lesioned OA subchondral bone samples (**Supplementary Table 1B**). A selection of seven lncRNAs was measured in these samples: *LINC01411*, growth arrest specific 5 (*GAS5*), *EMX2OS*, *PVT*, *LINC01060*, sciatic injury induced lincRNA upregulator of SOX11 (*SILC1*), and *AC005165.1*. These lncRNAs showed similar directions of effect in the technical validation and the biological replication samples as compared to the direction of effect measured in the RNA-seq data, except for *EMX2OS* (**Supplementary Table 8**).

Correlation of mRNA and lncRNA in OA subchondral bone

To identify possible mRNA targets of the differentially expressed lncRNAs i.e. lncRNAs regulating mRNAs with OA pathophysiology in subchondral bones, we filtered our recently reported differentially expressed mRNAs in subchondral bone [6] for protein-coding mRNAs (N=1417 protein-coding differentially expressed mRNAs) and correlated them with expression levels of the differentially expressed lncRNAs (N=21 lncRNAs) of the same patients (N=22 paired samples). Upon prioritizing on high correlations (-0.8>ρ >0.8) and significance (FDR<0.05), we found 875 significant correlations between 16

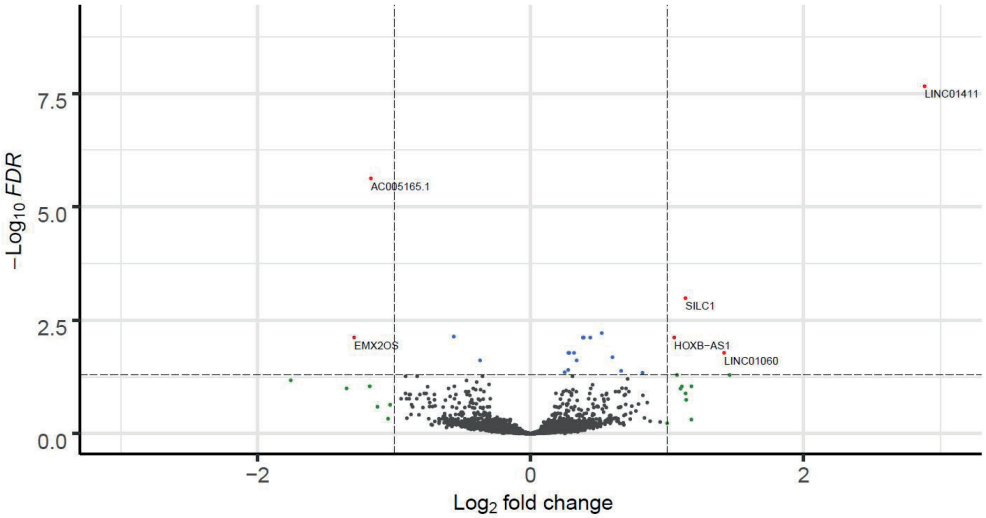


Figure 4 - Volcano plot of differentially expressed lncRNAs in OA subchondral bone.
The dots in the figure represent lncRNAs expressed in bone. Blue dots represent lncRNAs that are significantly differentially expressed, red dots represent lncRNAs that are significantly differentially expressed and have an absolute fold change of ≥2 and green dots represent the lncRNAs with an absolute fold change of ≥2 that are not significantly differentially expressed.

lncRNAs and 378 mRNAs (**Supplementary Table 9**). LncRNA small nucleolar RNA host gene 3 (*SNHG3*) showed the most interactions to mRNAs, with 174 significant correlations. In addition, the highest negative correlation was seen between *SNHG3* and *PTPRM* (ρ= -0.92), encoding Protein Tyrosine Phosphatase Receptor Type M, whereas the highest positive correlation was seen between *AC144548.1* and *ILF2* (ρ=0.92), encoding Interleukin Enhancer-binding Factor 2. Other notable interactions were those between *AC005165.1* and *FRZB* (ρ=0.85), encoding Frizzled Related Protein, and between *SILC1* and *POSTN* (ρ=0.81), encoding Periostin, which are both well-known OA genes.

To explore whether the differentially expressed lncRNAs are involved in certain processes or pathways, we performed gene enrichment analysis on their correlating mRNAs (**Supplementary Table 10**). Genes correlated to 9 out of 16 lncRNAs showed significant enrichment. The genes correlated to *AC006511.5* were enriched for Extracellular exosome (GO:0070062, FDR=3.67x10^{-4}) and Myelin sheath (GO:0043209, FDR=3.67x10^{-4}). Genes correlated to *SILC1* were significantly enriched for the GO-terms proteinaceous extracellular matrix (GO:0005578, FDR= 1.07x10^{-4}) and endoplasmic reticulum lumen (GO:0005788, FDR=4.62x10^{-2}), while for example genes correlated to *AC116533.1*, *AC245033.4* and *GAS5* were all significantly enriched for transcriptional and translational processes such as translational initiation (GO:0006413), poly(A) RNA binding (GO:0044822) and viral transcription (GO:0019083).

Functional investigation of AC005165.1

AC005165.1 was identified as the most significantly downregulated lncRNA (**Supplementary Table 6**) and, among others, it showed high correlation with well-known OA gene *FRZB* (ρ=0.85, **Supplementary Table 9**). Therefore, we selected *AC005165.1* to functionally investigate its possible mRNA targets *in vitro.* As shown in **Figure 5**, upon downregulation of *AC005165.1* (FC=0.55, *P*=0.51) by transfecting primary osteogenic cells (collected from N=4 knees) with an LNA GapmeR targeting *AC005165.1*, we observed consistent downregulation of *FRZB* (FC=0.54), which was in line with the observed positive correlation (ρ=0.85). However, the downregulation of *FRZB* did not reach statistical significance (*P*=0.08). Other mRNAs highly correlating with *AC0051651.1*, such as cysteine rich transmembrane BMP regulator 1 (*CRIM1,* ρ=0.82) and laeverin *LVRN* (ρ=-0.84), showed more donor-dependent variation upon downregulation of *AC005165.1*.

Comparison of lncRNAs between subchondral bone and articular cartilage

Since subchondral bone and the articular cartilage are interacting tissues, we used our previously published results on lncRNAs in OA articular cartilage [16] to compare the identified differentially expressed lncRNAs between preserved and lesioned OA articular cartilage and preserved and lesioned subchondral bone. First, we selected the overlapping samples of which we had RNA-seq data of subchondral bone and articular cartilage (N=10 paired samples, **Supplementary Table 1C**). As shown in **Supplementary Figure 4A**, we found 1763 exclusive subchondral bone lncRNAs, 590 exclusive cartilage lncRNAs, and 1090 lncRNAs that were expressed in both tissues (**Supplementary Table 11**). Upon comparing the here identified differentially expressed lncRNAs in subchondral bone with our previously identified differentially expressed in

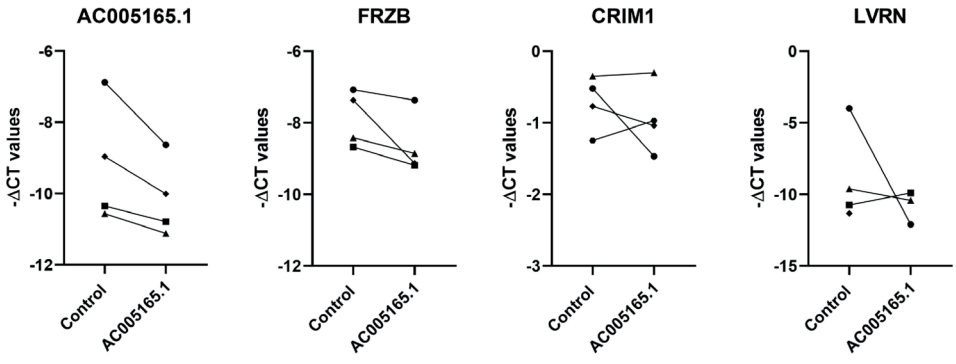


Figure 5 - Expression levels of AC005165.1, FRZB, CRIM1 and LVRN upon either transfecting primary osteogenic cells with LNA GapmeRs targeting AC005165.1 (indicated with AC005165.1) or transfecting primary osteogenic cells with a negative control (cells were collected from N = 4 knee joints).

articular cartilage [16], we found five lncRNAs being differentially expressed in both tissues: *AC005165.1*, *SILC1*, *LINC01411*, *AL590560.2* and *AC079781.5* (**Supplementary Figure 4B, Supplementary Table 12**). These five overlapping lncRNAs showed all similar directions of effect between preserved and lesioned samples in articular cartilage and subchondral bone. .

Discussion

We set out to study lncRNAs in subchondral bone as function of joint site and OA pathophysiology. In doing so, we identified 2057 lncRNAs commonly expressed in subchondral bone of hip and knee joints, 233 exclusive knee lncRNAs and 307 exclusive hip lncRNAs. Moreover, we observed additionally clustering on joint site based on level of lncRNA expression (**Figure 3**) among the commonly expressed lncRNA, signifying the difference between hip and knee OA subchondral bone pathophysiology. Differential expression analysis further identified 21 lncRNAs being differentially expressed between preserved and lesioned OA subchondral bone. Among the 21 differentially expressed lncRNAs we found *AC005165.1,* which was highly correlated to well-known OA gene *FRZB* (ρ=0.86). Upon functional investigation of *AC005165.1 in vitro* by downregulating *AC005165.1* using LNA GapmeRs, we observed a concurrent downregulation of *FRZB*. As lncRNAs tend to be highly tissue specific, lncRNAs, such as *AC005165.1*, could be attractive therapeutic OA targets with tissue specific effects.

Among the 21 differentially expressed lncRNAs, we identified *LINC01411* (FC=6.19, FDR=2.20×10^{-8}) as the most significantly and highest upregulated lncRNA, *AC005165.1* (FC=0.44, FDR=2.37×10^{-6}) as the most significantly downregulated lncRNA, and *EMX2OS* as the most downregulated lncRNA (FC=0.41, FDR=7.64×10^{-3}). The function of *LINC0411* remains unknown, however in a recent study it was found to be differentially expressed between healthy and OA articular cartilage and between healthy and OA synovium, indicating its role in OA across multiple tissues [18]. According to biotype classification of Ensembl v97 [19], *AC005165.1* was classified as a novel transcript and its function is still unknown. *AC005165.1* is genomically located at chromosome 7, with no coding genes lying within a 200-kb window. *EMX2OS* is an antisense RNA to *EMX2*, encoding Empty Spiracles Homeobox 2, which is a transcription factor crucial for the central nervous system. Multiple differentially methylated sites between preserved and lesioned OA articular cartilage have been reported in both *EMX2OS* and its antisense gene *EMX2* [20]. However, we did not find *EMX2* among the differentially expressed genes in our cartilage dataset [1] nor among the differentially expressed genes in bone [6]. Notably, we were not able to either validate or replicate the differential expression of *EMX2OS* by means of RT-qPCR, which might be due to its low expression levels and its consistency across individuals. Other notable differentially expressed lncRNAs were

GAS5 (FC=1.21, FDR=1.66x10^{-2}) and *PVT1* (FC=1.52, FDR=2.07x10^{-2}), as they both have been previously associated with OA pathophysiology [14, 17, 21].

To explore the potential targets and interactions of the 21 differentially expressed lncRNAs identified here, we calculated Spearman correlations between these lncRNAs and the previously identified differentially expressed mRNAs in the same OA subchondral bone samples and gene enrichment analysis was performed (**Supplementary Table 9, Supplementary Table 10**). *AC005165.1* was highly correlated with nine mRNAs, including *FRZB*, *CRIM1*, and *LVRN*. *FRZB* is a known OA gene and absence of FRZB in mice was previously shown to result in increased bone stiffness and increased cartilage degeneration [22]. CRIM1 is involved the TGF-β pathways by its binding to BMP-4 and BMP-7 [23], and LVRN is a metalloprotease which was previously linked to rheumatoid arthritis [24]. Despite the fact that *LINC0411* showed a higher FC than *AC005165.1*, we selected *AC005165.1* for functional investigation to see the functional relation between *AC005165.1* and the correlated mRNAs. Upon downregulation of *AC005165.1* in primary osteogenic cells, we observed consistent downregulation of *FRZB*, while *CRIM1* and *LVRN* expression levels did not change consistently. This suggests that *AC005165.1* directly or indirectly targets *FRZB* gene expression, while *CRIM1* and *LVRN* are functioning upstream of *AC005165.1*.

Similar to our mRNA expression profiling in OA subchondral bone [6], we here identified 233 exclusive knee, and 307 exclusive hip lncRNAs (**Supplementary Table 3, Supplementary Table 4**) indicating that lncRNA are not only tissue specific [14, 15], but also joint site specific. Additionally we showed (**Figure 3**) that such differences are also captured by quantitative differences in expression levels. Consecutively, we showed knee joint specific differentially expressed lncRNAs between preserved and lesioned OA subchondral bone, such as *CASC15* and *AL135926.1* (**Supplementary Table 7**). *CASC15,* which has not previously been associated to OA, is associated to cancer and involved in cell proliferation and migration [25]. *AL135926.1* was classified as novel transcript by Ensembl v97 [19] and its exact function is still unknown. However, *AL135926.1* is genomically located sense to protein-coding gene *DPT*, encoding dermatopontin, which was previously shown to inhibit BMP2 activity in mice [26]. We did not find any FDR significantly differentially expressed lncRNAs when stratifying for hip samples, which is likely due to the low sample size. Together, the here detected tissue and joint site specificity of lncRNA's qualifies them as eligible personalized therapeutic targets.

Although lncRNAs are known for their tissue specificity, we found a relatively large overlap of lncRNAs expressed in both articular cartilage and subchondral bone (N=1090 lncRNAs), which might be due to their common origin. Among the overlapping

differentially expressed between preserved and lesioned OA articular cartilage and subchondral bone, we found *AC005165.1*, making this lncRNA an attractive potential druggable target with effects in both tissues. The relative low number of differentially expressed lncRNAs identified in bone (N=21) compared to those found in cartilage (N=191) might be explained by the fact that cartilage is a single cell type tissue and subchondral bone multicellular and therefore more heterogeneous [27]. Moreover, the analysis on the subchondral bone included a lower sample size (N= 23 paired samples bone, N=32 paired samples cartilage) and stricter threshold for in- or excluding lncRNAs from the analysis.

The RNA-seq dataset that we used in this study was primarily obtained for mRNA expression profiling. Nonetheless, by applying our in-house pipeline we were able to characterize robust lncRNA expression in the same samples. It should be noted, however, that the lncRNA that entered the analyses had relatively high expression levels, while lncRNAs generally tend to be expressed at low levels [28]. To this end, we used two different selection criteria. In our initial, descriptive analyses on the lncRNA being expressed in our (knee and hip) subchondral bone samples we used more stringent selection criteria than in our pairwise differential expression analysis. This because per definition differential pairwise expression analysis is less sensitive for confounding factors. However, in future research the identification of lncRNAs associated with OA pathophysiology might be improved by increasing the sequencing depth.

In conclusion, the current study identified differences between hip and knee OA subchondral bone based on robust lncRNA expression levels. Moreover, *AC005165.1* was identified as an attractive potential therapeutic target, as it was here shown to be differentially expressed between preserved and lesioned OA subchondral bone and previously it was shown to be differentially expressed between preserved and lesioned OA articular cartilage. Furthermore, *AC005165.1* was here shown to regulate well-known OA gene *FRZB in vitro*. Finally, *AC005165.1* was not significantly differentially expressed between the hip and knee clusters, which could make *AC005165.1* a suitable druggable target in OA articular cartilage and OA subchondral bone of both hips and knees. More research is still needed to further elucidate the role and mode of action of *AC005165.1* in the OA pathophysiology. Together, this study shows that lncRNAs could bring new opportunities regarding joint tissue specific therapeutic strategies.

Declarations

Acknowledgements

We thank all the participants of the RAAK study. The LUMC has and is supporting the RAAK study. We thank all the members of our group. We also thank Enrike van der Linden, Robert van der Wal, Anika Rabelink-Hoogenstraaten, Peter van Schie, Shaho

Hasan, Maartje Meijer, Daisy Latijnhouwers and Geert Spierenburg for collecting the RAAK material. Moreover, data was 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. The study was funded by the Dutch Scientific Research council NWO /ZonMW VICI scheme (nr 91816631/528), Dutch Arthritis Society (DAA-10-1-402, DAF-16-1-405).

Funding

The study was funded by the Dutch Scientific Research council NWO /ZonMW VICI scheme (nr 91816631/528), Dutch Arthritis Society (DAA-10-1-402, DAF-16-1-405).

Disclosures

The authors have declared no conflicts of interest.

References

1. 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.
2. 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.
3. 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.
4. Chou, C.H., et al., Direct assessment of articular cartilage and underlying subchondral bone reveals a progressive gene expression change in human osteoarthritic knees. Osteoarthritis Cartilage, 2013. 21(3): p. 450-61.
5. Kuttapitiya, A., et al., Microarray analysis of bone marrow lesions in osteoarthritis demonstrates upregulation of genes implicated in osteochondral turnover, neurogenesis and inflammation. Ann Rheum Dis, 2017. 76(10): p. 1764-1773.
6. Tuerlings, M., et al., RNA sequencing reveals interacting key determinants of osteoarthritis acting in subchondral bone and articular cartilage. Arthritis Rheumatol, 2020.
7. Coutinho de Almeida, R., Y.F.M. Ramos, and I. Meulenbelt, Involvement of epigenetics in osteoarthritis. Best Pract Res Clin Rheumatol, 2017. 31(5): p. 634-648.
8. Reynard, L.N. and M.J. Barter, Osteoarthritis year in review 2019: genetics, genomics and epigenetics. Osteoarthritis Cartilage, 2020. 28(3): p. 275-284.
9. Rice, S.J., et al., Interplay between genetics and epigenetics in osteoarthritis. Nature Reviews Rheumatology, 2020. 16(5): p. 268-281.
10. Yang, J., et al., MiR-140 is co-expressed with Wwp2-C transcript and activated by Sox9 to target Sp1 in maintaining the chondrocyte proliferation. FEBS Lett, 2011. 585(19): p. 2992-7.
11. Endisha, H., et al., The complex landscape of microRNAs in articular cartilage: biology, pathology, and therapeutic targets. JCI Insight, 2018. 3(17).
12. Frankish, A., et al., GENCODE reference annotation for the human and mouse genomes. Nucleic Acids Res, 2019. 47(D1): p. D766-d773.
13. Marchese, F.P., I. Raimondi, and M. Huarte, The multidimensional mechanisms of long noncoding RNA function. Genome Biology, 2017. 18(1): p. 206.
14. 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.
15. Quinn, J.J. and H.Y. Chang, Unique features of long non-coding RNA biogenesis and function. Nature Reviews Genetics, 2016. 17(1): p. 47-62.
16. 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.
17. 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.
18. Li, C. and Z. Zheng, Identification of Novel Targets of Knee Osteoarthritis Shared by Cartilage and Synovial Tissue. Int J Mol Sci, 2020. 21(17).
19. Cunningham, F., et al., Ensembl 2019. Nucleic Acids Research, 2018. 47(D1): p. D745-D751.
20. Zhang, Y., et al., Genome-wide DNA methylation profile implicates potential cartilage regeneration at the late stage of knee osteoarthritis. Osteoarthritis and Cartilage, 2016. 24(5): p. 835-843.
21. Xing, D., et al., Identification of long noncoding RNA associated with osteoarthritis in humans. Orthop Surg, 2014. 6(4): p. 288-93.
22. Lories, R.J., et al., Articular cartilage and biomechanical properties of the long bones in Frzb-knockout mice. Arthritis Rheum, 2007. 56(12): p. 4095-103.
23. Wilkinson, L., et al., CRIM1 regulates the rate of processing and delivery of bone morphogenetic proteins to the cell surface. J Biol Chem, 2003. 278(36): p. 34181-8.
24. Haas, C.S., et al., Identification of genes modulated in rheumatoid arthritis using complementary DNA microarray analysis of lymphoblastoid B cell lines from disease-discordant monozygotic twins. Arthritis Rheum, 2006. 54(7): p. 2047-60.
25. Sheng, L. and R. Wei, Long Non-Coding RNA-CASC15 Promotes Cell Proliferation, Migration, and Invasion by Activating

Wnt/β-Catenin Signaling Pathway in Melanoma. Pathobiology, 2020. 87(1): p. 20-29.
26. Behnam, K., S.S. Murray, and E.J. Brochmann, BMP stimulation of alkaline phosphatase activity in pluripotent mouse C2C12 cells is inhibited by dermatopontin, one of the most abundant low molecular weight proteins in demineralized bone matrix. Connect Tissue Res, 2006. 47(5): p. 271-7.
27. 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.
28. Jarroux, J., A. Morillon, and M. Pinskaya, History, Discovery, and Classification of lncRNAs, in Long Non Coding RNA Biology, M.R.S. Rao, Editor. 2017, Springer Singapore: Singapore. p. 1-46.

Supplementary data

Supplementary methods

RNA sequencing

Preserved and lesioned subchondral bone were collected from the joint and stored in liquid nitrogen. Subsequently, the subchondral bone was pulverized and homogenized in TRIzol reagent (Invitrogen, USA) using a Mixer mill 200 (Retch, Germany). Total RNA was isolated from the subchondral bone using Qiagen RNeasy Mini Kit (Qiagen, Germany). Paired-end 2×100 bp RNA-sequencing (Illumina TruSeq RNA Library Prep Kit, Illumina HiSeq4000) was performed. Strand specific RNA-seq libraries were generated which yielded a mean of 20 million reads per sample. Data from both Illumina platforms were integrated and analysed with the same in-house pipeline. The quality of the raw reads for RNA-sequencing was checked using MultiQC v1.7. [1]. The adaptors were clipped using Cutadapt v1.1 [2] applying default settings (min overlap 3, min length). RNA-seq reads were aligned using Hisat2 v.2.1.0 against GRCh38 using default parameters. The aligned reads were processed into individual transcripts by StringTie v1.3.4 [3] and potential lncRNAs were identified by mapping the transcripts to GENCODE v29 [4] and Ensembl v97 [5]. The annotation of GENCODE v34 was used to filter the transcript on coding potential.

Differential expression analysis

Differential expression analysis was then performed in two ways: between hip and knee subchondral bone samples (preserved and lesioned tissue together); and between paired lesioned and preserved subchondral bone samples. The DESeq2 R package version 1.26.0 [6] was used to apply a general linear model assuming a negative binomial distribution, followed by a paired Wald-test. The Benjamini-Hochberg method was used to correct for multiple testing, as indicated by the false discovery rate (FDR), with a significance cut-off value of 0.05. Hip samples were set as a reference in the differential expression analysis between hip and knee subchondral bone and preserved samples were set as a reference in the analysis between preserved and lesioned OA subchondral bone.

RT-qPCR validation and replication

cDNA synthesis was done using Transcriptor First Strand cDNA Synthesis Kit (Roche, Switzerland), using 400 ng of RNA (Supplementary Table 1B). RT-qPCR was performed to quantitatively determine the lncRNA expression levels. The relative gene expression was evaluated by the -ΔCT values, using GAPDH and SDHA as internal controls. Paired T-test was performed to calculate the statistical difference in -ΔCT values between the lesioned and preserved OA subchondral bone samples.

Spearman correlation and gene ontology enrichment analysis

Prior to correlation, the differentially expressed mRNAs were filtered on protein-coding mRNAs using Ensembl v97 21. The Benjamini-Hochberg method was used to correct for multiple testing, as indicated by the FDR. Correlations with an FDR below 0.05 and an absolute correlation coefficient of 0.8 or higher were considered significantly correlated. Gene ontology enrichment was performed using the online functional annotation tool DAVID, selecting for the gene ontology terms Biological Processes (GOTERM_BP_DIRECT), Cellular Component (GOTERM_CC_DIRECT) and Molecular Function (GOTERM_MF_DIRECT). Genes expressed in OA subchondral bone were used as a background in the gene ontology enrichment analysis. Gene ontology terms with an FDR<0.05 were considered significant.

Functional validation of AC005165.1

Primary osteogenic cells were isolated from the preserved subchondral bone part of OA joints as described previously [7] and expanded in 2D (Supplementary Table 1C). This osteogenic cell isolation results in a mixture of bone cells, i.e. MSCs, osteoblasts, and osteocytes. To characterize this mixture of cells, we measured osteogenic and chondrogenic markers (SPP1, COL1A1, BGLAP, COL2A1, and COMP) using RT-qPCR and we compared these expression levels to the expression levels in preserved subchondral bone (Supplementary Figure-2), showing similar expression profiles. Then, the osteogenic cells were transfected with antisense locked nucleic acid (LNA) GapmeRs (Qiagen, Hilden, Germany) targeting AC005165.1 (GATAAAACCTGTAACT) or GapmeR negative control (AACACGTCTATACGC) at 10 nM final concentration using Lipofectamine RNAiMax Transfection reagent (Invitrogen, USA) according to manufacturer's protocol. After 30 hours, the cells were lysed using TRIzol (Invitrogen, USA) for RNA isolation. cDNA synthesis and RT-qPCR were performed to measure gene expression levels.

Comparison subchondral bone and articular cartilage

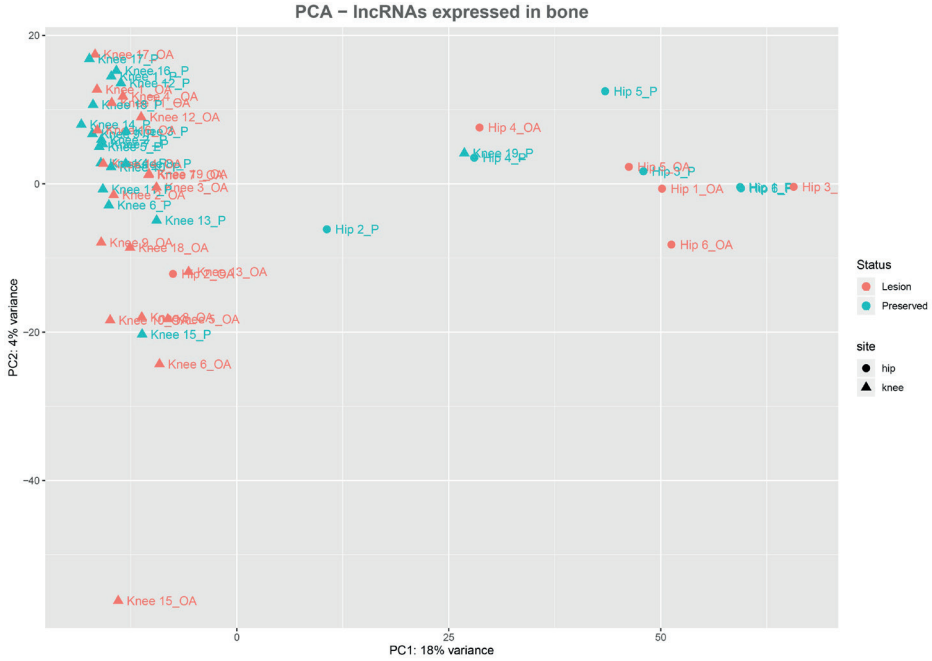
LncRNA expression between subchondral bone and articular cartilage was compared in the overlapping samples (N=10 paired samples: preserved and lesioned OA cartilage and bone, Supplementary Table 1D). Mapping of the RNA sequencing data was done using different versions of Ensembl between subchondral bone (v97) and articular cartilage (v94). To be able to compare the expressed and differentially expressed lncRNAs between the two tissues, we selected the lncRNAs that were mapped with both versions.

References methods

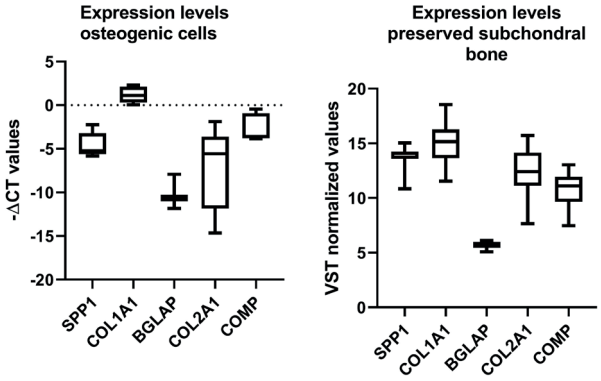
1. Ewels, P., et al., MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics, 2016. 32(19): p. 3047-8.
2. Martin, M., Cutadapt removes adapter sequences from high-throughput sequencing reads. 2011, 2011. 17(1): p. 3 %J EMBnet.journal.
3. Pertea, M., et al., StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nature

Biotechnology, 2015. 33(3): p. 290-295.
4. Frankish, A., et al., GENCODE reference annotation for the human and mouse genomes. Nucleic Acids Res, 2019. 47(D1): p. D766-d773.
5. Cunningham, F., et al., Ensembl 2019. Nucleic Acids Research, 2018. 47(D1): p. D745-D751.
6. 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): p. 550.
7. Stern, A.R., et al., Isolation and culture of primary osteocytes from the long bones of skeletally mature and aged mice. Biotechniques, 2012. 52(6): p. 361-73.

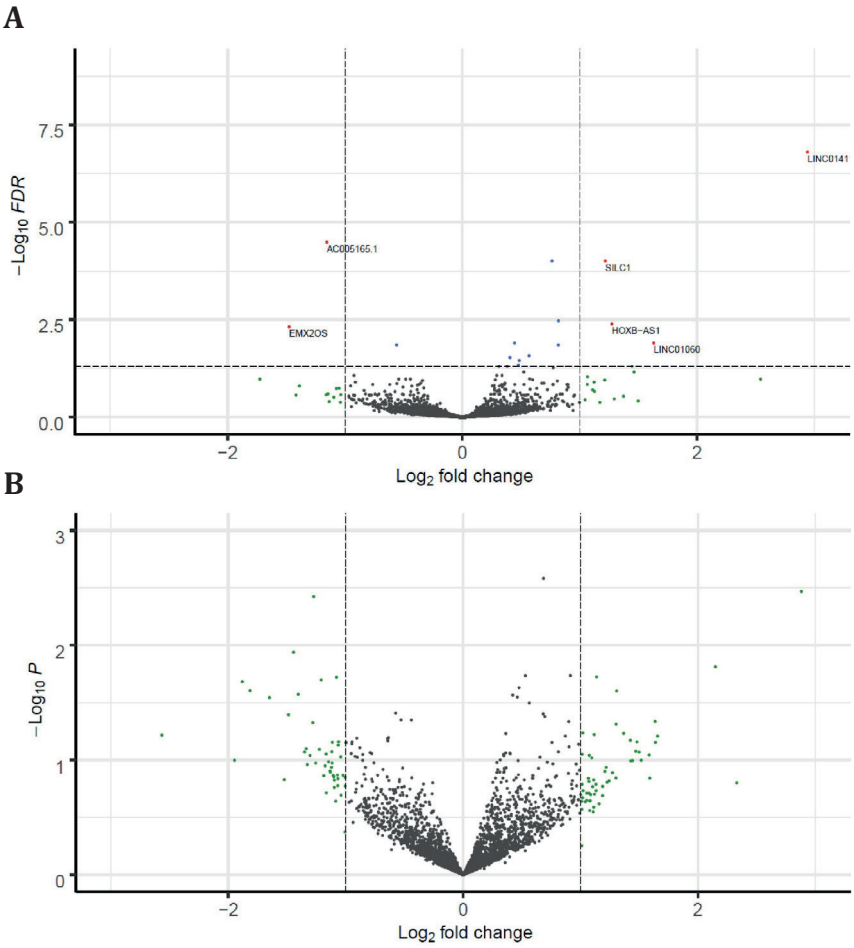
Supplementary figures



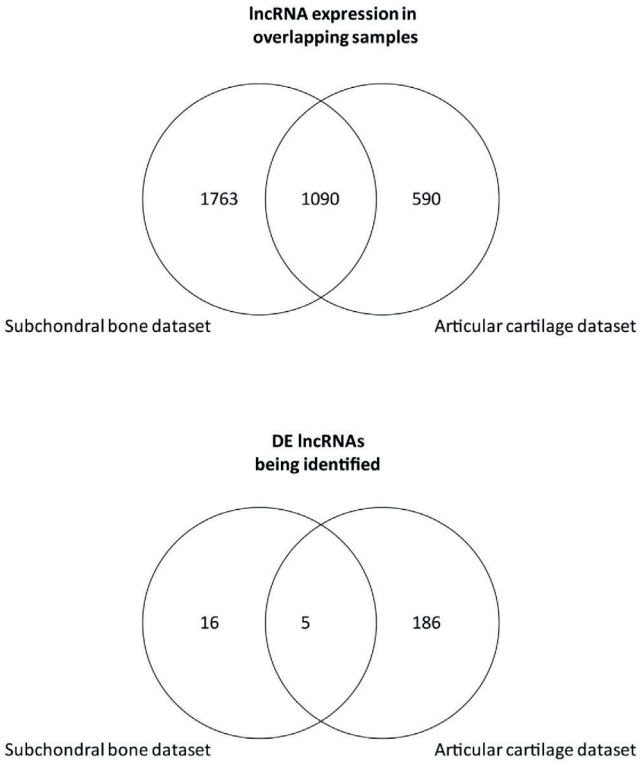
Supplementary Figure 1 – PCA in quality control identifying Knee_15, Knee-19, and Hip_2 as outliers.



Supplementary Figure 2 – Expression levels of osteogenic and chondrogenic markers.
(A) Expression levels in -ΔCT values of the primary osteogenic cells. (B) Expression levels in VST normalized values of preserved subchondral bone.



Supplementary Figure 3 – Volcano plot of differentially expressed lncRNAs in OA subchondral bone of knees (A) and hips (B).
The dots in the figure represent lncRNAs expressed. Blue dots represent lncRNAs that are significantly differentially expressed, red dots represent lncRNAs that are significantly differentially expressed and have an absolute fold change of 2 or higher, and green dots represent the lncRNAs with an absolute fold change of two or higher that are not significantly differentially expressed.



Supplementary Figure 4 – (A) Venn diagram of expressed lncRNAs in articular cartilage and subchondral bone, with 1090 lncRNAs expressed in both tissues. (B) Venn diagram of significantly differentially expressed lncRNAs between preserved and lesioned articular cartilage and subchondral bone, with 5 lncRNAs differentially expressed in both tissues.

Supplementary tables

Supplementary Table 1 – Baseline characteristics of material included in the current study

Supplementary Table 1A – Sample characteristics of lncRNA- and mRNA-seq data subchondral bone (N=44 samples, 22 pairs of preserved and lesioned subchondral bone)

	Hip	Knee	Total
Participants	5	17	22
Mean age	68.40	66.12	66.64
SD age	9.65	8.44	8.67
Female (%)	100.00	88.24	91.00

Supplementary Table 1B – Characteristics of samples used for technical validation (N= 18 samples, 9 pairs of preserved and lesioned subchondral bone) and biological replication (N=30 samples, 15 pairs of preserved and lesioned subchondral bone).

	Technical validation			Biological replication		
	Hip	**Knee**	**Total**	**Hip**	**Knee**	**Total**
Participants	-	9	9	5	10	15
Mean age	-	68.67	68.67	67.80	72.40	70.87
SD age	-	7.42	7.42	6.06	10.07	8.98
Female (%)	-	77.78	77.78	20.00	50.00	60.00

Supplementary Table 1C – Characteristics of samples used to transfect primary cells with LNA GapmeRs (N=4 participants)

	Hip	**Knee**
Participants	-	4
Mean age	-	67.50
SD age	-	8.10
Female (%)	-	25.00

Supplementary Table 1D – Characteristics of samples used in overlapping RNA-seq data of articular cartilage and subchondral bone (N=20 samples, 10 pairs of preserved and lesioned articular cartilage and subchondral bone)

	Hip	**Knee**	**Total**
participants	1	9	10
Mean age	56.00	67.44	66.30
SD age	-	8.81	9.05
Female (%)	100.00	88.89	90.00

Supplementary Table 2 (partially) – LncRNAs expressed in the knee, the hip, and the total datasets of subchondral bone.
The top 50 lncRNAs with highest expression levels in subchondral bone are shown here, the rest of the table can be found in the online supplement: https://doi.org/10.1093/rheumatology/keab826

Ensembl ID	lncRNA	Mean Expression
ENSG00000282885	AL627171.2	340390.89
ENSG00000259001	AL355075.4	48251.77
ENSG00000269900	RMRP	42482.57
ENSG00000251562	MALAT1	38905.93
ENSG00000229807	XIST	7801.50
ENSG00000281181	FP236383.3	6396.73
ENSG00000245532	NEAT1	5659.95
ENSG00000260032	NORAD	4049.39
ENSG00000264772	AC016876.2	3596.00
ENSG00000270066	AL356488.2	3231.50
ENSG00000284803	AC245033.4	2501.48
ENSG00000259976	AC093010.3	1924.68
ENSG00000276232	AC006064.5	1758.39
ENSG00000242125	SNHG3	1665.55
ENSG00000240801	AC132217.1	1634.43
ENSG00000175061	SNHG29	1504.66
ENSG00000253352	TUG1	1412.27
ENSG00000247092	SNHG10	1393.16
ENSG00000247556	OIP5-AS1	1320.75
ENSG00000224078	SNHG14	1233.25
ENSG00000261771	DNAAF4-CCPG1	1201.68
ENSG00000280614	FP236383.2	1188.66
ENSG00000225733	FGD5-AS1	1095.93
ENSG00000257379	AC023509.1	1091.86
ENSG00000274536	AL034397.3	1073.93
ENSG00000279738	AL022311.1	958.75
ENSG00000203930	LINC00632	934.77
ENSG00000263244	AC087190.3	930.82
ENSG00000230590	FTX	861.39
ENSG00000273189	AC010619.2	773.16
ENSG00000249669	CARMN	766.27
ENSG00000230551	AC021078.1	748.86

Ensembl ID	lncRNA	Mean Expression
ENSG00000234456	MAGI2-AS3	660.55
ENSG00000285565	AL671762.1	642.93
ENSG00000267009	AC007780.1	595.41
ENSG00000262202	AC007952.4	576.36
ENSG00000256028	AC026362.1	575.20
ENSG00000163597	SNHG16	570.59
ENSG00000234741	GAS5	554.84
ENSG00000269821	KCNQ1OT1	501.52
ENSG00000237298	TTN-AS1	501.23
ENSG00000263798	AC018521.1	467.14
ENSG00000203875	SNHG5	460.61
ENSG00000272888	LINC01578	431.84
ENSG00000196295	GARS-DT	424.18
ENSG00000267519	AC020916.1	411.95
ENSG00000215386	MIR99AHG	400.34
ENSG00000263753	LINC00667	398.73
ENSG00000285622	AL135926.1	397.48

Supplementary Table 3 (partially) – LncRNAs exclusively expressed in knee OA subchondral bone.
The top 50 lncRNAs with highest expression levels exclusively in knee OA subchondral bone are shown here, the rest of the table can be found in the online supplement: https://doi.org/10.1093/rheumatology/keab826

Ensembl ID	lncRNA	Mean Expression
ENSG00000258820	AF111167.1	22.76
ENSG00000260653	AC237221.1	12.21
ENSG00000261713	SSTR5-AS1	11.82
ENSG00000285051	SLC7A14-AS1	11.12
ENSG00000287415	AC099541.2	10.91
ENSG00000271239	AC007423.1	10.65
ENSG00000287620	AC092053.4	10.65
ENSG00000286113	AC022868.2	10.50
ENSG00000272256	AC044849.1	9.94
ENSG00000213025	COX20P1	9.00
ENSG00000260192	LINC02240	8.76
ENSG00000238042	LINC02257	8.59
ENSG00000258910	LINC01956	8.47
ENSG00000286598	AC100756.4	8.38
ENSG00000267737	AC087645.2	8.29
ENSG00000226581	AC092634.3	8.26
ENSG00000258334	AC125611.4	8.21
ENSG00000261083	LINC02516	8.03
ENSG00000260278	AC098818.2	7.91
ENSG00000253434	LINC02237	7.71
ENSG00000275830	AL390755.1	7.56
ENSG00000265485	LINC01915	7.53
ENSG00000235619	RPL36AP33	7.24
ENSG00000234626	AL021937.3	7.21
ENSG00000256984	AC008013.2	7.12
ENSG00000260364	AC009055.1	7.00
ENSG00000247416	AP000802.1	6.97
ENSG00000275894	AL021578.1	6.97
ENSG00000255008	AP000442.1	6.85
ENSG00000278716	AC133540.1	6.74
ENSG00000267275	AC020911.2	6.47

Ensembl ID	lncRNA	Mean Expression
ENSG00000274213	AC015912.3	6.44
ENSG00000245651	AC083805.1	6.35
ENSG00000254109	RBPMS-AS1	6.26
ENSG00000283945	LINC00032	6.26
ENSG00000239263	RBM43P1	6.15
ENSG00000251314	AC104123.1	6.15
ENSG00000225096	AL445250.1	6.09
ENSG00000242986	RPL21P99	6.09
ENSG00000248896	AC105001.1	6.06
ENSG00000257398	AC126177.3	6.06
ENSG00000287059	AC090004.2	6.06
ENSG00000257277	AC092652.2	6.03
ENSG00000271538	LINC02427	6.03
ENSG00000245384	CXXC4-AS1	5.97
ENSG00000287129	AC097500.1	5.97
ENSG00000255399	TBX5-AS1	5.94
ENSG00000262223	AC110285.1	5.94
ENSG00000236047	AC073410.1	5.91

Supplementary Table 4 (partially) – LncRNAs exclusively expressed in hip OA subchondral bone.
The top 50 lncRNAs with highest expression levels exclusively in hip OA subchondral bone are shown here, the rest of the table can be found in the online supplement: https://doi.org/10.1093/rheumatology/keab826

Ensembl ID	lncRNA	Mean Expression
ENSG00000285783	AC098588.2	892.00
ENSG00000285144	AL359555.3	378.90
ENSG00000275527	AC100835.2	183.80
ENSG00000271736	AL138900.2	138.40
ENSG00000268119	AC010615.2	112.00
ENSG00000264066	AC024267.1	109.30
ENSG00000254006	AC104232.1	96.80
ENSG00000136315	AL355922.1	71.40
ENSG00000268734	AC245128.3	59.10
ENSG00000268833	AC243967.2	58.40
ENSG00000223629	DEFA8P	50.70
ENSG00000260592	AC130456.3	48.30
ENSG00000225345	SNX18P3	46.60
ENSG00000273812	BX640514.2	45.00
ENSG00000255929	AP000943.3	41.10
ENSG00000260188	AC002464.1	40.50
ENSG00000226281	AL031123.1	39.90
ENSG00000259986	AC103876.1	38.20
ENSG00000286342	AC073210.3	36.90
ENSG00000268658	LINC00664	31.40
ENSG00000284138	ATP6V0CP4	30.50
ENSG00000283839	AC096667.1	30.30
ENSG00000251002	AC244502.1	29.70
ENSG00000224177	LINC00570	28.00
ENSG00000238160	AC116366.2	27.60
ENSG00000250155	AC008957.1	27.60
ENSG00000224429	LINC00539	27.10
ENSG00000239219	AC008040.1	26.60
ENSG00000204860	FAM201A	24.60
ENSG00000175746	C15orf54	24.20
ENSG00000230773	AC092650.1	23.80

Ensembl ID	lncRNA	Mean Expression
ENSG00000185275	CD24P4	23.30
ENSG00000233038	AC011899.2	23.20
ENSG00000285486	AC003043.2	23.20
ENSG00000249684	AC106795.2	22.90
ENSG00000255571	MIR9-3HG	22.80
ENSG00000267751	AC009005.1	22.40
ENSG00000269243	AC008894.2	21.80
ENSG00000286419	AC097637.3	21.50
ENSG00000287497	AL031123.4	21.30
ENSG00000286602	AC021660.4	20.80
ENSG00000229140	CCDC26	20.70
ENSG00000261218	AC099524.1	20.60
ENSG00000237803	LINC00211	20.10
ENSG00000255733	IFNG-AS1	19.90
ENSG00000236525	AC007278.2	19.60
ENSG00000261804	AC007342.4	19.60
ENSG00000236304	AP001189.1	19.20
ENSG00000242082	SLC5A4-AS1	18.70

Supplementary Table 5 (partially) – LncRNAs being differentially expressed between cluster 1 (containing knee samples) and cluster 2 (containing hip samples).
With cluster 1 set as a reference. The top 50 most significantly differentially expressed lncRNAs are shown here, the rest of the table can be found in the online supplement: https://doi.org/10.1093/rheumatology/keab826

Ensembl ID	lncRNA	Base Mean	P-value	FDR	Log 2 Fold Change	Fold Change
ENSG00000274536	AL034397.3	1246.27	4.33E-233	8.90E-230	7.30	157.82
ENSG00000283646	LINC02009	83.94	4.47E-148	4.60E-145	6.48	89.21
ENSG00000265206	AC004687.1	32.76	4.48E-95	3.07E-92	4.03	16.30
ENSG00000288046	AL031123.5	21.51	3.54E-87	1.82E-84	4.19	18.29
ENSG00000187951	AC091057.1	26.80	2.81E-86	1.16E-83	3.99	15.90
ENSG00000180539	C9orf139	27.20	1.05E-84	3.60E-82	3.99	15.93
ENSG00000283743	Z84466.1	183.04	2.49E-73	7.32E-71	4.45	21.89
ENSG00000285117	AC068724.4	178.42	6.32E-64	1.62E-61	7.31	158.87
ENSG00000274272	AC069281.2	140.81	5.85E-63	1.34E-60	1.80	3.48
ENSG00000286646	AL121933.1	31.18	1.23E-59	2.52E-57	3.36	10.24
ENSG00000228340	MIR646HG	55.03	5.15E-59	9.64E-57	2.91	7.50
ENSG00000213373	LINC00671	16.10	5.97E-57	1.02E-54	3.74	13.39
ENSG00000276570	AC010327.6	76.51	3.49E-55	5.53E-53	2.15	4.43
ENSG00000213468	FIRRE	20.35	4.95E-54	7.28E-52	2.72	6.61
ENSG00000282907	Z98883.1	130.83	4.51E-51	6.18E-49	2.43	5.38
ENSG00000228794	LINC01128	139.02	1.83E-50	2.35E-48	1.56	2.95
ENSG00000238164	TNFRSF14-AS1	38.04	4.47E-50	5.41E-48	1.92	3.79
ENSG00000257167	TMPO-AS1	21.19	8.44E-48	9.65E-46	2.90	7.47
ENSG00000261997	AC007336.1	16.62	6.97E-47	7.54E-45	2.83	7.09
ENSG00000260401	AP002761.4	21.68	6.12E-46	6.29E-44	3.77	13.63
ENSG00000267121	AC008105.3	27.31	1.20E-45	1.18E-43	2.24	4.74
ENSG00000281344	HELLPAR	20.97	4.67E-45	4.37E-43	2.67	6.38
ENSG00000204282	TNRC6C-AS1	77.20	1.67E-43	1.50E-41	2.39	5.24
ENSG00000248774	AC097534.1	10.78	3.15E-43	2.70E-41	2.88	7.34
ENSG00000257698	GIHCG	51.37	6.35E-43	5.22E-41	1.93	3.81
ENSG00000227039	ITGB2-AS1	62.32	2.09E-42	1.65E-40	2.42	5.36
ENSG00000215908	CROCCP2	282.00	4.18E-42	3.18E-40	1.89	3.71
ENSG00000248323	LUCAT1	34.22	1.58E-41	1.16E-39	3.08	8.45
ENSG00000270956	AC009948.3	12.22	1.18E-40	8.35E-39	2.67	6.38
ENSG00000259343	TMC3-AS1	22.94	9.95E-40	6.82E-38	2.48	5.57

Ensembl ID	lncRNA	Base Mean	P-value	FDR	Log 2 Fold Change	Fold Change
ENSG00000284948	AC107959.4	34.17	4.90E-39	3.25E-37	3.47	11.06
ENSG00000238113	LINC01410	21.17	9.21E-39	5.92E-37	4.00	15.98
ENSG00000247982	LINC00926	23.33	5.29E-38	3.29E-36	2.20	4.58
ENSG00000237298	TTN-AS1	498.13	1.93E-37	1.17E-35	1.71	3.28
ENSG00000261008	LINC01572	19.89	2.58E-36	1.52E-34	2.40	5.27
ENSG00000270277	AC009948.2	39.54	4.95E-36	2.83E-34	2.30	4.92
ENSG00000265148	TSPOAP1-AS1	29.37	6.73E-36	3.74E-34	1.90	3.73
ENSG00000237943	PRKCQ-AS1	22.55	6.23E-35	3.37E-33	2.50	5.64
ENSG00000260641	AC114811.2	12.18	1.17E-34	6.15E-33	3.56	11.83
ENSG00000260528	FAM157C	26.15	1.35E-34	6.93E-33	3.47	11.10
ENSG00000224152	AC009506.1	36.54	1.32E-32	6.60E-31	1.80	3.48
ENSG00000267174	AC011472.2	21.03	3.64E-32	1.78E-30	2.82	7.06
ENSG00000235499	AC073046.1	15.92	5.16E-32	2.47E-30	2.07	4.18
ENSG00000286488	AC103858.3	13.44	5.67E-32	2.65E-30	2.78	6.88
ENSG00000246695	RASSF8-AS1	91.07	1.21E-31	5.52E-30	-1.36	0.39
ENSG00000276649	AL117335.1	17.54	1.50E-31	6.69E-30	2.27	4.84
ENSG00000285952	AC020663.4	65.56	6.43E-31	2.81E-29	2.33	5.03
ENSG00000286288	AL109809.5	9.22	1.31E-30	5.63E-29	2.60	6.04
ENSG00000282164	PEG13	19.88	4.19E-30	1.76E-28	2.47	5.55

Supplementary Table 6 – Differentially expressed lncRNAs in OA subchondral bone

Ensembl ID	lncRNA	Base Mean	P-value	FDR	Log 2 Fold Change	Fold Change
ENSG00000249306	LINC01411	6.19	7.02E-12	2.20E-08	2.89	7.39
ENSG00000249378	LINC01060	32.03	7.29E-05	1.66E-02	1.42	2.67
ENSG00000232044	SILC1	19.66	9.87E-07	1.03E-03	1.13	2.20
ENSG00000230148	HOXB-AS1	10.88	2.26E-05	7.64E-03	1.05	2.07
ENSG00000285906	AC083855.2	13.94	3.05E-04	4.56E-02	0.82	1.77
ENSG00000264672	SEPT4-AS1	15.08	2.50E-04	4.14E-02	0.66	1.58
ENSG00000249859	PVT1	58.59	9.89E-05	2.07E-02	0.60	1.52
ENSG00000242125	SNHG3	1648.02	7.76E-06	6.09E-03	0.52	1.44
ENSG00000284707	AC079781.5	82.01	2.43E-05	7.64E-03	0.44	1.35
ENSG00000284697	AC006511.5	124.06	1.52E-05	7.64E-03	0.39	1.31
ENSG00000276232	AC006064.5	1694.57	2.38E-05	7.64E-03	0.38	1.30
ENSG00000264772	AC016876.2	3402.26	1.30E-04	2.43E-02	0.34	1.26
ENSG00000258210	AC144548.1	290.80	7.40E-05	1.66E-02	0.32	1.25
ENSG00000175061	SNHG29	1471.34	6.98E-05	1.66E-02	0.28	1.22
ENSG00000234741	GAS5	525.86	7.11E-05	1.66E-02	0.28	1.21
ENSG00000284803	AC245033.4	2375.02	2.27E-04	3.95E-02	0.28	1.21
ENSG00000244398	AC116533.1	299.55	2.83E-04	4.44E-02	0.25	1.19
ENSG00000272668	AL590560.2	52.11	1.32E-04	2.43E-02	-0.37	0.77
ENSG00000271880	AGAP11	32.92	1.16E-05	7.25E-03	-0.56	0.68
ENSG00000223561	AC005165.1	17.97	1.51E-09	2.37E-06	-1.17	0.44
ENSG00000229847	EMX2OS	28.30	2.43E-05	7.64E-03	-1.29	0.41

Supplementary Table 7 – Differentially expressed lncRNAs in knee OA subchondral bone

Ensembl ID	lncRNA	Base Mean	P-value	FDR	Log 2 Fold Change	Fold Change
ENSG00000249306	LINC01411	8.07	5.00E-11	1.59E-07	2.94	7.67
ENSG00000249378	LINC01060	39.92	3.18E-05	1.25E-02	1.63	3.10
ENSG00000230148	HOXB-AS1	12.26	7.67E-06	4.07E-03	1.27	2.42
ENSG00000232044	SILC1	25.66	1.25E-07	9.92E-05	1.22	2.33
ENSG00000249859	PVT1	44.52	5.34E-06	3.40E-03	0.82	1.76
ENSG00000264672	SEPT4-AS1	17.32	4.83E-05	1.41E-02	0.82	1.76
ENSG00000285622	AL135926.1	468.53	1.22E-07	9.92E-05	0.76	1.70
ENSG00000272168	CASC15	68.73	1.01E-04	2.67E-02	0.57	1.48
ENSG00000242125	SNHG3	1188.28	1.55E-04	3.53E-02	0.48	1.40
ENSG00000284707	AC079781.5	79.77	2.19E-04	4.64E-02	0.48	1.39
ENSG00000284697	AC006511.5	117.42	3.54E-05	1.25E-02	0.44	1.36
ENSG00000264772	AC016876.2	3501.60	1.21E-04	2.97E-02	0.41	1.32
ENSG00000271880	AGAP11	40.72	4.87E-05	1.41E-02	-0.56	0.68
ENSG00000223561	AC005165.1	22.29	2.06E-08	3.28E-05	-1.15	0.45
ENSG00000229847	EMX2OS	36.04	1.07E-05	4.85E-03	-1.48	0.36

Supplementary Table 8 – Validation and replication of selection of identified lncRNAs

	RNA-seq		techincal validation		biological replication		techinical and biological replicates together	
lncRNA	**FC**	**Padj**	**FC**	**Pvalue**	**FC**	**Pvalue**	**FC**	**Pvalue**
LINC01411	7.39	2.20E-08	33.12	1.27E-02	9.60	2.69E-03	17.84	5.80E-04
GAS5	1.21	1.66E-02	3.12	8.64E-01	1.11	6.92E-01	1.61	8.44E-01
EMX2OS	0.41	7.64E-03	0.99	1.06E-01	1.23	2.88E-01	1.14	2.60E-02
PVT1	1.52	2.07E-02	2.76	2.41E-02	1.83	8.51E-02	2.14	2.03E-02
LINC01060	2.67	1.66E-02	2.15	3.13E-01	3.01	9.74E-03	2.89	3.56E-03
SILC1	2.20	9.87E-07	2.41	5.00E-06	1.79	1.88E-10	2.05	1.12E-15
AC005165.1	0.44	2.37E-06	-	-	0.49	4.83E-03	-	-

Supplementary Table 9 (partially) – Significant correlations between differentially expressed protein-coding mRNAs and differentially expressed lncRNAs in bone of the same patients (N=22 paired samples).
The top 50 highest correlations between lncRNAs and mRNAs are shown here, the rest of the table can be found in the online supplement: https://doi.org/10.1093/rheumatology/keab826

lncRNA	lncRNA Ensembl ID	mRNA	mRNA Ensembl ID	rho	P-value	FDR
AC006511.5	ENSG00000284697	SMARCC2	ENSG00000139613	-0.89	4.44E-16	3.78E-13
AC016876.2	ENSG00000264772	GLB1	ENSG00000170266	0.90	2.22E-16	2.26E-16
AC016876.2	ENSG00000264772	ATP6V0B	ENSG00000117410	0.91	2.22E-16	2.26E-16
AC083855.2	ENSG00000285906	ACKR3	ENSG00000144476	-0.89	1.33E-15	8.26E-13
AC083855.2	ENSG00000285906	FST	ENSG00000134363	-0.89	1.33E-15	8.26E-13
AC083855.2	ENSG00000285906	ENTPD7	ENSG00000198018	0.89	4.44E-16	3.78E-13
AC116533.1	ENSG00000244398	RPLP0	ENSG00000089157	0.89	4.44E-16	3.78E-13
AC116533.1	ENSG00000244398	RPS8	ENSG00000142937	0.89	4.44E-16	3.78E-13
AC144548.1	ENSG00000258210	CHD6	ENSG00000124177	-0.91	2.22E-16	2.26E-16
AC144548.1	ENSG00000258210	ABRACL	ENSG00000146386	0.89	8.88E-16	5.87E-13
AC144548.1	ENSG00000258210	RAB32	ENSG00000118508	0.89	6.66E-16	5.22E-13
AC144548.1	ENSG00000258210	CKS2	ENSG00000123975	0.89	4.44E-16	3.78E-13
AC144548.1	ENSG00000258210	C14orf119	ENSG00000179933	0.90	2.22E-16	2.75E-13
AC144548.1	ENSG00000258210	TOR1A	ENSG00000136827	0.90	2.22E-16	2.75E-13
AC144548.1	ENSG00000258210	PCNA	ENSG00000132646	0.90	2.22E-16	2.75E-13
AC144548.1	ENSG00000258210	ARPC3	ENSG00000111229	0.90	2.22E-16	2.26E-16
AC144548.1	ENSG00000258210	CAP1	ENSG00000131236	0.90	2.22E-16	2.26E-16
AC144548.1	ENSG00000258210	CLTC	ENSG00000141367	0.91	2.22E-16	2.26E-16

lncRNA	lncRNA Ensembl ID	mRNA	mRNA Ensembl ID	rho	P-value	FDR
AC144548.1	ENSG00000258210	ILF2	ENSG00000143621	0.92	2.22E-16	2.26E-16
AC245033.4	ENSG00000284803	RPL7A	ENSG00000148303	0.89	6.66E-16	5.22E-13
AC245033.4	ENSG00000284803	RPS29	ENSG00000213741	0.89	4.44E-16	3.78E-13
AC245033.4	ENSG00000284803	RPLP0	ENSG00000089157	0.91	2.22E-16	2.26E-16
AC245033.4	ENSG00000284803	RPS7	ENSG00000171863	0.91	2.22E-16	2.26E-16
AC245033.4	ENSG00000284803	RPS12	ENSG00000112306	0.91	2.22E-16	2.26E-16
AC245033.4	ENSG00000284803	RPS17	ENSG00000182774	0.92	2.22E-16	2.26E-16
AC245033.4	ENSG00000284803	RPS8	ENSG00000142937	0.92	2.22E-16	2.26E-16
LINC01060	ENSG00000249378	GNG4	ENSG00000168243	0.89	8.88E-16	5.87E-13
LINC01060	ENSG00000249378	PRSS35	ENSG00000146250	0.89	4.44E-16	3.78E-13
PVT1	ENSG00000249859	PPP1R16B	ENSG00000101445	-0.90	2.22E-16	2.26E-16
PVT1	ENSG00000249859	ABCC4	ENSG00000125257	0.89	2.22E-16	2.75E-13
SNHG29	ENSG00000175061	PTPRM	ENSG00000173482	-0.89	4.44E-16	3.78E-13
SNHG29	ENSG00000175061	CHD6	ENSG00000124177	-0.89	4.44E-16	3.78E-13
SNHG29	ENSG00000175061	UBE2J1	ENSG00000198833	0.89	8.88E-16	5.87E-13
SNHG29	ENSG00000175061	MGAT2	ENSG00000168282	0.89	8.88E-16	5.87E-13
SNHG29	ENSG00000175061	PCNA	ENSG00000132646	0.89	8.88E-16	5.87E-13
SNHG29	ENSG00000175061	PTDSS1	ENSG00000156471	0.90	2.26E-16	2.26E-16
SNHG3	ENSG00000242125	PTPRM	ENSG00000173482	-0.92	2.22E-16	2.26E-16
SNHG3	ENSG00000242125	AHNAK	ENSG00000124942	-0.90	2.22E-16	2.26E-16
SNHG3	ENSG00000242125	SYT12	ENSG00000173227	-0.90	2.22E-16	2.75E-13

lncRNA	lncRNA Ensembl ID	mRNA	mRNA Ensembl ID	rho	P-value	FDR
SNHG3	ENSG00000242125	SPATA6	ENSG00000132122	-0.90	2.22E-16	2.75E-13
SNHG3	ENSG00000242125	PLEKHM3	ENSG00000178385	-0.89	4.44E-16	3.78E-13
SNHG3	ENSG00000242125	LDHD	ENSG00000166816	-0.89	1.11E-15	7.18E-13
SNHG3	ENSG00000242125	UTRN	ENSG00000152818	-0.88	1.78E-15	1.04E-12
SNHG3	ENSG00000242125	UGGT1	ENSG00000136731	0.89	8.88E-16	5.87E-13
SNHG3	ENSG00000242125	BID	ENSG00000015475	0.89	8.88E-16	5.87E-13
SNHG3	ENSG00000242125	ABCC4	ENSG00000125257	0.89	6.66E-16	5.22E-13
SNHG3	ENSG00000242125	RGS19	ENSG00000171700	0.89	4.44E-16	3.78E-13
SNHG3	ENSG00000242125	CKS2	ENSG00000123975	0.90	2.22E-16	2.26E-16
SNHG3	ENSG00000242125	JPT1	ENSG00000189159	0.91	2.22E-16	2.26E-16

Supplementary Table 10 (partially)– Gene enrichment analysis of correlating genes.
The most significantly enriched GO-term for each lncRNA are shown here, the rest of the table can be found in the online supplement: https://doi.org/10.1093/rheumatology/keab826

lncRNA	Term	Count	%	FDR	Genes
AC006511.5	GO:0070062~extracellular exosome	28	38.89	3.67E-04	SLC25A5, PSMA4, NSF, ATP6V1D, GLA, ARPC3, TARS, GNB4, ATP6V1A, CAP1, RAN, PCNA, TXN, STXBP1, PSMA5, CCT6A, ATP6V1B2, GSTO1, ARPC2, CCT3, CCT2, IGF2, GLB1, INSR, PPIA, IARS, CLIC1, PSMB3
AC116533.1	GO:0006614~SRP-dependent cotranslational protein targeting to membrane	8	26.67	1.90E-07	RPLP0, RPS12, RPL27, RPS8, RPL7A, RPL27A, RPS7, RPS17
AC144548.1	GO:0043209~myelin sheath	8	7.27	2.02E-02	SLC25A5, GDI2, NSF, HSPA8, ATP5PB, CLTC, CCT2, UQCRFS1
AC245033.4	GO:0006614~SRP-dependent cotranslational protein targeting to membrane	18	45.00	8.01E-25	RPL6, RPLP0, RPS13, RPS12, RPL5, RPL23, RPL27, RPS6, RPS24, RPS8, RPL7A, RPL27A, RPS7, RPL4, RPS17, RPL12, RPL10A, RPL39
GAS5	GO:0006614~SRP-dependent cotranslational protein targeting to membrane	12	85.71	1.00E-20	RPL6, RPS13, RPS12, RPL5, RPL23, RPS8, RPL7A, RPS7, RPL4, RPS17, RPL35A, RPL10A
LINC01060	GO:0016021~integral component of membrane	11	78.57	4.97E-03	SLC8A3, SLC9A2, RARRES1, IGSF3, RHBDL2, SGMS2, CDH2, ANO5, LRRC15, GXYLT2, SLC36A2
SILC1	GO:0005578~proteinaceous extracellular matrix	7	25.00	1.07E-04	CCN4, SPARC, OMD, POSTN, CTHRC1, MAMDC2, COL5A2
SNHG29	GO:0018279~protein N-linked glycosylation via asparagine	7	7.29	8.87E-05	UGGT1, ST6GAL2, STT3B, RPN1, MGAT2, UBE2J1, DDOST
SNHG3	GO:0051082~unfolded protein binding	8	4.60	4.40E-02	DNAJB11, CRYAB, HSPA8, UGGT1, CCT6A, CCT3, CCT2, CALR

Supplementary Table 11 (partially) – LncRNAs exclusively expressed in bone, lncRNAs exclusively expressed in cartilage, and lncRNAs expressed in both tissues (N=10 paired samples, preserved and lesioned cartilage and bone).
The top 10 highest expressed lncRNAs exclusively in subchondral bone, the top 10 highest expressed lncRNAs exclusively in articular cartilage and the top 10 highest expressed lncRNAs in both tissues are shown here, the rest of the table can be found in the online supplement: https://doi.org/10.1093/rheumatology/keab826

Tissue	lncRNA	lncRNA Ensembl ID	Mean Expression in bone	Mean expresion in cartilage
Bone	AL627171.2	ENSG00000282885	387446.35	-
Bone	FP236383.3	ENSG00000281181	3240.85	-
Bone	SNHG10	ENSG00000247092	1483.60	-
Bone	FP236383.2	ENSG00000280614	605.90	-
Bone	AL034397.3	ENSG00000274536	507.35	-
Bone	TRHDE-AS1	ENSG00000236333	376.95	-
Bone	HLA-DRB6	ENSG00000229391	276.30	-
Bone	LINC02328	ENSG00000258733	234.10	-
Bone	AC244205.1	ENSG00000240040	229.85	-
Bone	AC242426.2	ENSG00000237188	228.35	-
Cartilage	PART1	ENSG00000152931	-	274.50
Cartilage	SSTR5-AS1	ENSG00000261713	-	137.10
Cartilage	AC087521.3	ENSG00000254409	-	119.90
Cartilage	AL009174.1	ENSG00000227008	-	118.10
Cartilage	MT1P3	ENSG00000229230	-	83.20
Cartilage	AC107075.1	ENSG00000277998	-	82.70
Cartilage	RPL22P2	ENSG00000241081	-	63.70
Cartilage	AL139220.2	ENSG00000230615	-	62.50
Cartilage	LINC01411	ENSG00000249306	-	61.45
Cartilage	AC245060.4	ENSG00000272779	-	47.75
Overlap	AL355075.4	ENSG00000259001	51930.45	32608.05
Overlap	RMRP	ENSG00000269900	46300.60	14252.65
Overlap	MALAT1	ENSG00000251562	45942.75	16241.35
Overlap	XIST	ENSG00000229807	8808.60	4236.35
Overlap	NEAT1	ENSG00000245532	7120.70	2880.80
Overlap	NORAD	ENSG00000260032	4529.90	3223.20
Overlap	AL356488.2	ENSG00000270066	3691.15	1526.30
Overlap	AC016876.2	ENSG00000264772	3800.45	1328.90
Overlap	AC245033.4	ENSG00000284803	2655.90	1237.80
Overlap	AC093010.3	ENSG00000259976	2357.90	1228.40

Supplementary Table 12 – Significantly differentially expressed lncRNAs overlapping between articular cartilage and subchondral bone

lncRNA	Base mean in bone	FC in bone	FDR in bone	Base mean in cartilage	FC in cartilage	FDR in cartilage
AC005165.1	17.97	0.44	2.37E-06	47.47	0.47	1.33E-03
AC079781.5	82.01	1.35	7.64E-03	86.37	1.30	3.08E-02
AL590560.1	52.11	0.77	2.43E-02	69.60	0.65	8.74E-03
LINC01411	6.19	7.39	2.20E-08	50.68	4.48	2.58E-06
SILC1	19.66	2.20	1.03E-03	88.54	2.17	6.39E-07