A novel *CCDC91* isoform associated with ossification of the posterior longitudinal ligament of the spine works as a non-coding RNA to regulate osteogenic genes

Authors

Masahiro Nakajima, Masaru Koido, Long Guo, Chikashi Terao, Shiro Ikegawa

Correspondence

chikashi.terao@riken.jp (C.T.),
sikegawa@yahoo.co.jp (S.I.)

Nakajima et al. identified a functional variant in the 5′ UTR of a novel non-coding isoform of *CCDC91* within the known 12p11.22 genome-wide association study locus for ossification of the posterior longitudinal ligament of the spine. The isoform directly interacts with *MIR890*, which binds to *RUNX2* and regulates osteogenesis.

Nakajima et al., 2023, The American Journal of Human Genetics *110*, 638–647
April 6, 2023
https://doi.org/10.1016/j.ajhg.2023.03.004

CellPress

A novel *CCDC91* isoform associated with ossification of the posterior longitudinal ligament of the spine works as a non-coding RNA to regulate osteogenic genes

Masahiro Nakajima,[1] Masaru Koido,[2,3] Long Guo,[1,4] Chikashi Terao,[2,*] and Shiro Ikegawa[1,*]

Summary

Ossification of the posterior longitudinal ligament of the spine (OPLL) is a common intractable disease that causes spinal stenosis and myelopathy. We have previously conducted genome-wide association studies for OPLL and identified 14 significant loci, but their biological implications remain mostly unclear. Here, we examined the 12p11.22 locus and identified a variant in the 5′ UTR of a novel isoform of *CCDC91* that was associated with OPLL. Using machine learning prediction models, we determined that higher expression of the novel *CCDC91* isoform was associated with the G allele of rs35098487. The risk allele of rs35098487 showed higher affinity in the binding of nuclear proteins and transcription activity. Knockdown and overexpression of the *CCDC91* isoform in mesenchymal stem cells and MG-63 cells showed paralleled expression of osteogenic genes, including *RUNX2*, the master transcription factor of osteogenic differentiation. The *CCDC91* isoform directly interacted with *MIR890*, which bound to *RUNX2* and decreased *RUNX2* expression. Our findings suggest that the *CCDC91* isoform acts as a competitive endogenous RNA by sponging *MIR890* to increase *RUNX2* expression.

Introduction

Ossification of the posterior longitudinal ligament of the spine (OPLL) is a pathologic process of bone deposition at the site of the posterior longitudinal ligament.[1] OPLL can cause spinal cord and nerve roots compression,[2] resulting in quadriplegia or other myelopathies.[2] The prevalence of OPLL is reported up to 3.0% in Asian countries.[2] Conversely, a lower prevalence of 0.1%–1.7% is described in comparable European cohorts and a cohort of North American individuals of non-Asian descent.[3] Since the molecular pathogenesis of OPLL has not been understood and no efficient treatment strategies have been proposed, especially pharmacotherapy or preventive interventions, surgical treatment by indirect spinal cord decompression is the only option for symptomatic OPLL-affected individuals.

OPLL is multifactorial, where genetic[1,4] and environmental factors, including diet, obesity, physical strain, age, and diabetes mellitus, are involved in the etiology.[5] To understand the genetic factors for OPLL, we previously conducted a genome-wide association study (GWAS) and identified six loci associated with OPLL.[6] Recently, we conducted a meta-analysis of GWASs and identified 14 significant loci,[7] including eight that had not been previously reported in the GWAS catalog.[8] Among them, we discovered a candidate functional variant upstream of R-spondin 2 (*RSPO2*) on chromosome 8q23.1.[9] In addition, we identified cell division cycle 5-like (*CDC5L*) on chromosome 6p21.2 as a susceptibility gene for OPLL.[10] However, additional causal variants and susceptibility genes remain to be discovered in other loci.

In this study, we identified a functional variant and a susceptibility gene for OPLL at 12p11.22. Using machine learning prediction models, we identified rs35098487, which affects the expression of a cap analysis of gene expression (CAGE) tag (p4@CCDC91, indicating a *CCDC91* promoter with fourth highest read counts after promoter 1 (p1), p2, and p3 in all of the FANTOM5 CAGE profiles[11]) in OPLL-related cells. We discovered a novel isoform of *CCDC91* (coiled-coil domain containing 91) transcripts that started its transcription from p4@CCDC91 and had not been previously reported in Ensembl. rs35098487 was in the 5′ UTR of the *CCDC91* isoform, and the risk allele of rs35098487 showed higher affinity in the binding of nuclear proteins and transcription activity compared with the non-risk allele. Knockdown and overexpression of the *CCDC91* isoform in mesenchymal stem cells (MSCs) showed paralleled expression of osteogenic genes, including *RUNX2*, a master transcription factor regulating osteogenic differentiation. The *CCDC91* isoform directly interacted with *MIR890* that bound to the 3′ UTR of *RUNX2* and inhibited osteogenic differentiation by decreasing RUNX2 transcription and translation. Our study showed that a tissue-/cell-type-specific non-coding RNA regulates ectopic ossification in OPLL by directly controlling a microRNA regulating osteogenic genes.

Material and methods

Rapid amplification of cDNA ends (RACE) and RT-PCR

3′ RACE was performed with SMARTer RACE Kit (Takara Bio) according to the manufacturer's protocol. We used poly $(A)^+$ RNA

[1]Laboratory for Bone and Joint Diseases, Center for Integrative Medical Sciences, RIKEN, Tokyo 108-8639, Japan; [2]Laboratory for Statistical and Translational Genetics, Center for Integrative Medical Sciences, RIKEN, Yokohama 230-0045, Japan; [3]Laboratory of Complex Trait Genomics, Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Tokyo 108-8639, Japan; [4]Department of Laboratory Animal Science, School of Basic Medical Sciences, Xi'an Jiaotong University, Xi'an 710061, China
*Correspondence: chikashi.terao@riken.jp (C.T.), sikegawa@yahoo.co.jp (S.I.)
https://doi.org/10.1016/j.ajhg.2023.03.004.

(1 μg) of MG-63 cells to produce the RACE template. cDNAs of various human tissues were purchased from Takara Bio. Some cDNAs for RT-PCR were synthesized with MultiScribe Reverse Transcriptase and random hexamer primer (Thermo Fisher Scientific). PCR was performed with KOD FX Neo (Toyobo), and the amplicons were separated by agarose electrophoresis for analysis. The primers used in the PCR reactions are listed in Table S1.

Cell culture and osteogenic differentiation

Human osteosarcoma cell lines MG-63 and Saos-2 and normal human fibroblasts were obtained from RIKEN BioResource Research Center and cultured in DMEM containing 10% FBS at 37°C under 5% CO_2. The study was approved by the ethical committees at the RIKEN Yokohama Institute. All individuals gave written informed consent to participate in the study. Human MSCs were purchased from Lonza and cultured in MSC NutriStem XF Medium (Sartorius) at 37°C under 5% CO_2. For osteogenic differentiation, MSCs were cultured with osteogenic induction media containing 100 nM dexamethasone, 50 μg/mL ascorbic acid, and 10 mM β-glycerophosphate sodium (Sigma-Aldrich).[12] The medium was changed every 3 days.

Alkaline phosphatase staining

MSCs were cultured with an osteogenic induction media for 7 days. Alkaline Phosphatase Staining Kit (Cosmo Bio) was used according to the manufacturer's instructions. In brief, the cells were washed with PBS and fixed with 10% neutral buffered formalin for 20 min. Then, the cells were washed with distilled water. The staining reagent was added to the cells and incubated for 15 min at 37°C. After removal of the staining reagent, the cells were washed with distilled water and observed under a phase-contrast microscope (Olympus CKX53; Olympus).

Alizarin red staining

MSCs were cultured with osteogenic induction media for 14 days. Alizarin red S Solution (PG research) was used after fixation with 10% neutral buffered formalin. The images were captured by a phase-contrast microscope.

Small interfering RNA (siRNA) knockdown

siRNA targeting a novel *CCDC91* isoform (Table S1) was obtained from Thermo Fisher Scientific. As a control, we used a Stealth RNAi negative control medium GC duplex (Thermo Fisher Scientific). The siRNAs (10 nM) were transfected with LipofectamineRNAiMAX (Thermo Fisher Scientific) according to the manufacturer's protocol. Knockdown was estimated by quantitative RT-PCR.

Construction of plasmids

The region between −337 and +356 (hg38, chr12: 28,190,642–28,191,334) of exon 1 of the novel *CCDC91* isoform (containing allele G or A of rs35098487) was subcloned into the pGL4.10 [luc2] vector (Promega). Full-length cDNA of the novel *CCDC91* isoform was cloned into pcDNA3.1(−) vector (Thermo Fisher Scientific). Three tandem repeats of the wild-type or mutated sequence of the *MIR890* response element (MRE) in the novel *CCDC91* isoform (Figure 3A) or *RUNX2* (Figure 4B) were subcloned into the pmirGLO Dual-Luciferase miRNA Target Expression Vector (Promega). The constructs were verified by DNA sequencing (Model 3730xl; Thermo Fisher Scientific).

Quantitative RT-PCR

Total RNA from cells was extracted with SV Total RNA Isolation System (Promega), according to the manufacturer's instructions. cDNA was synthesized from total RNA with MultiScribe Reverse Transcriptase (Thermo Fisher Scientific). We performed quantitative real-time PCR was performed by using a StepOnePlus Real-Time PCR System (Thermo Fisher Scientific) with Quantitect SYBR Green PCR Kit (Qiagen) in accordance with the manufacturer's instructions. *GAPDH* expression was analyzed as an internal control. For microRNAs, we used a Mir-X miRNA First-Strand Synthesis Kit (Takara Bio) and PowerTrack SYBR Green Master Mix (Thermo Fisher Scientific). The primers used in PCR reactions are listed in Table S1.

Reporter assay

MG-63 cells were transfected with the reporter plasmid and the pGL4.74 vector (Promega) as an internal control via TransIT-LT1 Transfection Reagent (Mirus Bio). For the microRNA-binding assay, we co-transfected MG-63 cells with luciferase reporter plasmid harboring the wild-type/mutant binding site of the novel *CCDC91* isoform (Figure 3A) along with *MIR890* mimics or negative control (GenePharma) by using Lipofectamine RNAiMAX reagent. After 24 h of transfection, we measured the luciferase activities by using a Dual-Luciferase Reporter Assay System (Promega).

Electrophoretic mobility shift assay (EMSA)

Nuclear extract from cells was prepared with NE-PER Nuclear and Cytoplasmic Extraction Reagents (Thermo Fisher Scientific). The nuclear extract was incubated with 31 bp double-strand digoxigenin-labeled oligonucleotide probes for rs35098487 alleles (sense, 5′-CTTTTACTTTGCCAARTACATTGGGGTAAAA-3′; antisense, 5′-TTTTACCCCAATGTAYTTGGCAAAGTAAAAG-3′) for 30 min at room temperature. DNA-protein complexes were resolved on 6% DNA retardation gels (Thermo Fisher Scientific), and the signal was detected with Digoxigenin Luminescent Detection Kit (Sigma-Aldrich).

Allele-specific quantitative PCR

Allele-specific expression was performed in cDNA from normal fibroblasts that were heterozygous at rs35098487. Each TaqMan probe, which is specific for a different allele, was labeled by a different dye, and fluorescence was detected by real-time PCR, as described previously.[13] To make a linear regression line for the log of fluorescence intensity ratio (VIC/FAM) versus the log of allele ratio, we mixed gDNA from two homozygous individuals as the ratios 8:2, 7:3, 6:4, 5:5, 4:6, 3:7, and 2:8 and we quantified the two alleles in the mixes. We calculated the expression ratio of the G allele to the A allele by comparing those of the cDNAs to those of the genomic DNAs from the same individual.

Immunoblotting

Cell lysates were prepared 72 h after transfection with RIPA buffer (Thermo Fisher Scientific) and Halt Protease Inhibitor Cocktail (Thermo Fisher Scientific). Immunoblot was performed as previously described,[14] and the blot was immunolabeled with RUNX2 (D1L7F) Rabbit mAb (1:1,000; Cell Signaling Technology) or β-Actin pAb-HRP-DirecT (1:4,000; Medical & Biological Laboratories). Immunoreactive bands were exposed by enhanced chemiluminescence (Amersham ECL Prime; Cytiva).

Bioinformatic analyses

For *in silico* prediction of variant effects on gene expression, we used MENTR (mutation effect prediction on non-coding RNA transcription) as described previously.[15] All significant variants from 14 GWAS loci were analyzed by MENTR. We used the pre-determined threshold (absolute mutation effect > 0.05) achieving 80% accuracy to predict expression change.[15] The downstream microRNA targets of the novel *CCDC91* isoform were predicted with RNA22[16] and miRDB.[17] We used miRDB to find the targets of predicted microRNAs.

Statistical analyses

Individual values, mean, and standard deviation (SD) are displayed. We used Mann-Whitney U test to identify significant differences between various treatments. $p < 0.05$ was regarded as statistically significant.

Results

In silico prediction of variant effects on gene expression

We applied machine learning prediction models[15] to prioritize causal variants from 14 GWAS loci. We identified the expression of six CAGE tags associated (absolute mutation effect > 0.05) with 12 variants in OPLL-related cells including MSCs, articular chondrocytes, and osteoblasts (Table S2). For a 12p11.22 locus, rs35098487 was predicted to affect the expression of a CAGE tag (p4@CCDC91) in amnion MSCs and articular chondrocytes. Our previous Bayesian statistical fine-mapping analysis showed that rs35098487 is a member of the calculated 95% credible set at this locus.[7]

Identification of a novel *CCDC91* isoform

A search of the Ensembl Genome Browser revealed 22 *CCDC91* isoforms (splicing variants) in the human genome (Figure S1). When aligning the sequence of p4@CCDC91 with the genomic sequence of *CCDC91*, we found no isoform overlapping p4@CCDC91. We performed RT-PCR and 3′ RACE experiments and found a novel isoform of *CCDC91* transcripts that started its transcription from p4@CCDC91 and was not previously annotated in Ensembl (Figure S1). The novel isoform shared the same translation start codon (ATG) with the major isoform (ENST00000381259.5) that is highly expressed in many tissues (Figure S2). However, the novel isoform had polyadenylation in the middle of the protein-coding sequence, resulting in an aberrant mRNA that lacks a stop codon (non-stop mRNA) (Figure S3B). rs35098487 is in the 5′ UTR of the novel isoform (Figure S3A).

Expression of the novel *CCDC91* isoform

It has been shown that a non-stop mRNA is degraded by a non-stop mRNA decay (NSD) pathway.[18] However, the level of non-stop mRNA has been shown to be reduced only to one-fourth of the wild-type level in an analysis using the reporter gene.[19] We assessed the expression of the novel isoform in various tissues and cell lines by RT-PCR. We amplified the novel isoform with PCR primers designed in exons 1 and 7. The novel isoform was specifically expressed in the brain, kidney, tendon, and ligament but not in the cartilage (Figure S4A). Strong expression was observed in osteoblastic MG-63 cells, whereas no expression could be detected in other osteoblastic Saos-2 cells (Figure S4B). Expression of the novel isoform was not observed in the testis where *CCDC91* was the most expressed (Figure S4A). These results indicate that expression of the novel isoform is tissue and cell specific, unlike the major *CCDC91* isoform.

Functional analysis of rs35098487

To evaluate whether the rs35098487 variant might affect transcription of the novel isoform, we created a reporter construct harboring the region between -337 and $+356$ of exon 1 of the novel isoform (containing rs35098487), placed upstream of the luciferase reporter gene. The construct containing the risk allele (G) showed approximately 1.4-fold higher luciferase activity than that containing the non-risk allele (A) (Figure 1A). We then performed EMSA to examine the difference in the binding of nuclear proteins to the risk (G) and non-risk (A) alleles of rs35098487. We observed higher affinity binding of nuclear proteins to the oligonucleotide containing the risk allele (G) (Figure 1B). To confirm the effect of rs35098487 on transcription, we performed an allele-specific quantitative PCR by using a TaqMan probe on normal fibroblasts, which were heterozygous at rs35098487. The G allele showed 1.4-fold higher expression of the transcript than the A allele (Figure 1C). Importantly, the directionality of the expression change for rs35098487 was correctly predicted *in silico*. From these observations, we concluded that rs35098487 affected the transcription of the novel isoform.

Expression of the novel *CCDC91* isoform during osteogenic differentiation of MSCs

To understand the functional basis of the novel isoform in osteoblasts, we examined the expression of the novel isoform in MSCs subjected to osteogenic-induced differentiation. The novel isoform expression transiently increased at day 3 and then decreased (Figure S5). The increase in the novel isoform expression preceded that of *ALPL*, suggesting an earlier role in osteogenic differentiation. The novel isoform expression also paralleled the expression of *RUNX2*, which is a master transcription factor for osteogenic differentiation (Figure S5).

Role of the novel *CCDC91* isoform in the osteogenic differentiation of MSCs

To explore the role of the novel isoform in osteogenic differentiation, we transfected MSCs with the novel *CCDC91* isoform-specific siRNA and cultured knockdown MSCs under osteogenic conditions. Novel *CCDC91* isoform knockdown significantly decreased *RUNX2* and *ALPL* expression (Figure 2A). Similarly, expressions of *IBSP*, *BGLAP*, and *SP7* were inhibited in knockdown MSCs (Figure 2A), indicating that the novel isoform promotes

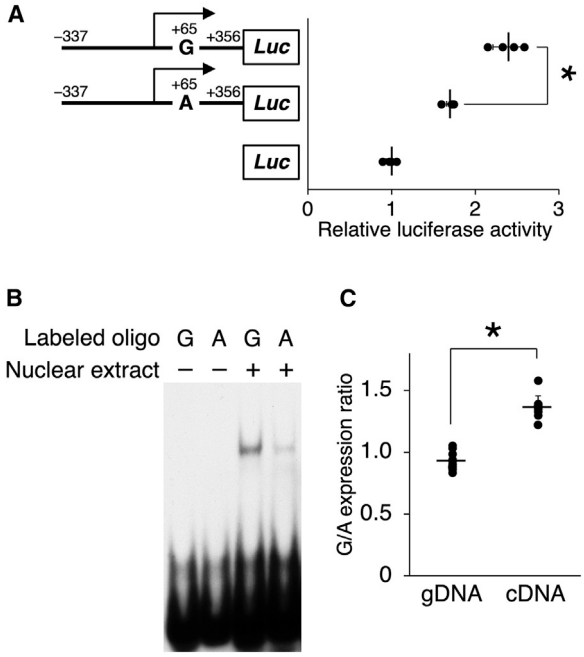

Figure 1. Effects of rs35098487 on the promoter activity, nuclear protein binding, and mRNA expression

(A) Reporter assays in MG-63 cells. There was a significantly increased promoter activity for the risk allele G of rs35098487 compared to the non-risk allele A. Data represent means ± SD from quadruplicate assays. Each dot represents the result from one sample. *, $p < 0.05$.

(B) Electrophoretic mobility shift assays with nuclear protein extracts from MG-63 cells. There was a higher affinity for the probe with the G allele (lane 3) than the probe with the A allele (lane 4) of rs35098487.

(C) The allelic gene expression ratio for the novel *CCDC91* isoform in fibroblasts. The transcript carrying the G allele showed higher expression than that with the A allele. Genomic DNA (gDNA) was used as a control for equal bi-allelic representation. Data are mean ± SD of values from ten samples. Each dot represents the result from one sample. *, $p < 0.05$.

osteogenic differentiation. Similar results were observed in MG-63 cells (Figure S6). Next, we carried out overexpression analysis for the novel isoform in MSCs. The overexpression of the novel isoform increased the osteogenic markers *RUNX2*, *ALPL*, *IBSP*, *BGLAP*, and *SP7* in MSCs (Figure 2B) and MG-63 cells (Figure S7).

To further investigate the regulatory role of the novel isoform in osteogenic differentiation, we next examined the effects of the novel isoform on alkaline phosphatase activity and the capability of mineralization in MSCs. The alkaline phosphatase activity and capability of mineralization were decreased in novel *CCDC91* isoform knockdown cells, whereas the reverse was true in the MSCs transfected with the novel *CCDC91* isoform (Figures 2C and S8). These results suggest that the novel *CCDC91* isoform promoted the early osteogenic differentiation of MSCs.

A novel *CCDC91* isoform binds to *MIR890*

Recently, long non-coding RNAs (lncRNAs) have been reported to work as competing endogenous RNA to sponge microRNAs.[20] By using the online tools RNA22 V2[16] and miRDB,[17] we predicted that the sequence of the novel *CCDC91* isoform contains seven potential microRNA binding sites (Table S3). One of the seven microRNAs, *MIR890*, was predicted to bind to the 3′ UTR of *RUNX2* (Table S3). To confirm the predicted interaction between the novel *CCDC91* isoform and *MIR890*, we constructed the wild-type sequence of isoform harboring the predicted *MIR890* binding sites (WT) or the mutated sequence (mut) into a luciferase reporter vector (Figure 3A) and then transfected these into MG-63 cells. *MIR890* overexpression could reduce the wild-type isoform-driven luciferase activity (Figure 3B), while it was abolished by the mutated sequence (Figure 3C), which indicated that *MIR890* could bind to the isoform directly.

MIR890 inhibits osteogenic differentiation

Quantitative RT-PCR analysis showed that the expression of *MIR890* was reduced at the early stage of osteogenic differentiation in MSCs (Figure S9), whereas the expression of the novel *CCDC91* isoform and *RUNX2* was increased (Figure S5). To investigate the effect of *MIR890* on osteogenic differentiation, we transfected MSCs with *MIR890* and cultured them under osteogenic conditions. Overexpression of *MIR890* significantly decreased *RUNX2*, *ALPL*, *IBSP*, *BGLAP*, and *SP7* expression in MSCs (Figure 4A). The alkaline phosphatase activity and mineralization were also decreased in the MSCs transfected with *MIR890* (Figure S10), indicating that *MIR890* inhibits osteogenic differentiation. Similar results were observed in MG-63 cells (Figure S11).

RUNX2 is the target of *MIR890*

Potential binding sequences between *MIR890* and *RUNX2* were predicted by miRDB (Figure 4B). The luciferase reporter assay showed that co-transfection of *MIR890* and 3′ UTR of *RUNX2* wild type could decrease the luciferase activity (Figure 4C). Hence, the binding of *MIR890* to *RUNX2* was verified. Overexpression of *MIR890* significantly decreased the protein level of RUNX2 (Figure 4D). These data indicated that *MIR890* inhibited osteogenic differentiation via targeting *RUNX2* and that the novel *CCDC91* isoform may negatively regulate *MIR890* by competing with *RUNX2* to bind *MIR890*-binding sites, leading to the up-regulation of *RUNX2* target genes.

Discussion

GWASs have been developed as an excellent tool for identifying genetic variation underlying complex diseases. GWASs have linked tens of thousands of variants to thousands of phenotypes, furthering our understanding of the genetic causes of complex diseases.[21] Elucidating how associated variants modulate disease risk and how they affect cellular phenotypes will provide mechanism-based therapeutic hypotheses and lead to more effective drug discovery.[22] In this study, we identified a functional variant

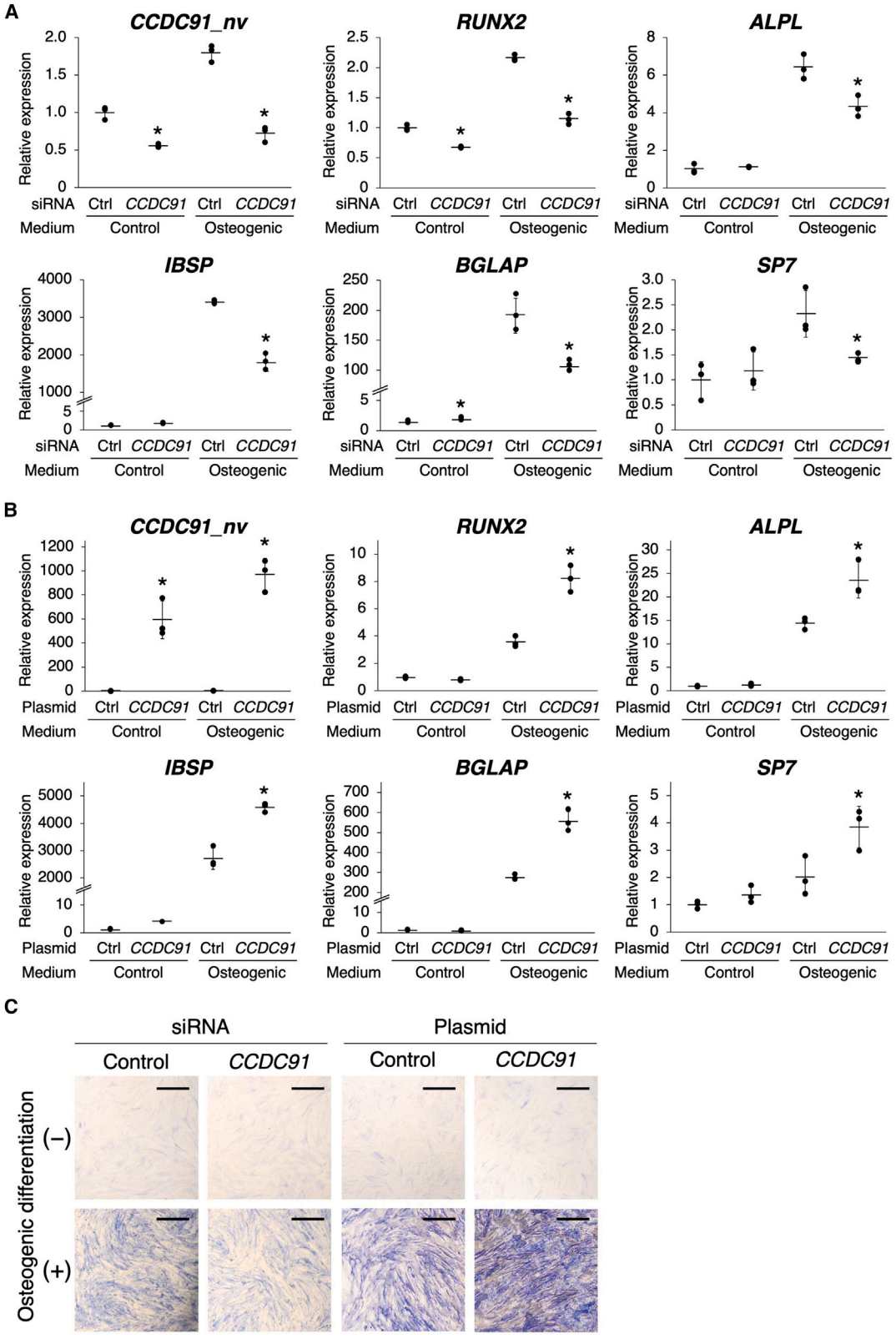

Figure 2. The novel *CCDC91* isoform promotes the expression of osteogenic marker genes

(A) Effects of the novel *CCDC91* isoform knockdown on the expression of osteogenic marker genes.

(B) Effects of overexpression of the novel *CCDC91* isoform on the expression of osteogenic marker genes. Expression of the novel *CCDC91* isoform (*CCDC91_nv*) and osteogenic marker genes in MSCs cultured in osteogenic conditions was measured at 3 days

(legend continued on next page)

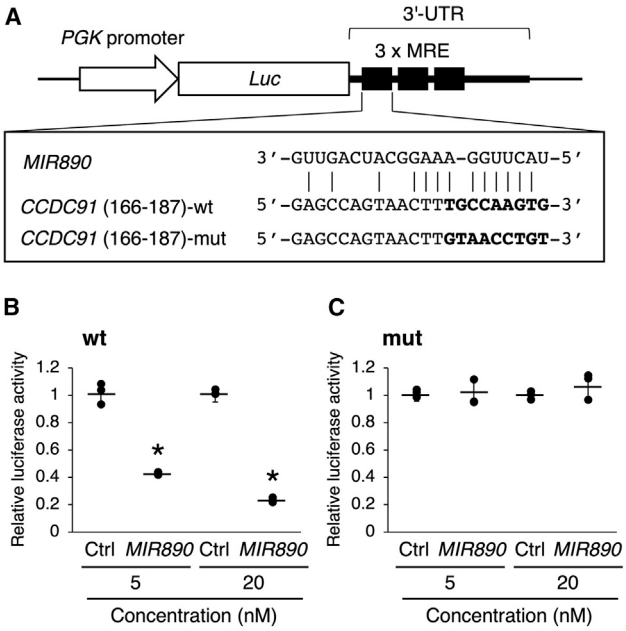

Figure 3. The novel *CCDC91* isoform binds to *MIR890*

(A) Luciferase reporter vectors containing wild-type or mutated sequence of the novel *CCDC91* isoform. Sequences that differ between wild-type and mutated are shown in bold. MRE, *MIR890* response element in the *CCDC91* isoform.

(B and C) Luciferase reporter vectors containing wild-type (B) or mutated (C) sequence of the novel *CCDC91* isoform were co-transfected with *MIR890* or control into MG-63 cells. Data represent means ± SD from quadruplicate assays. Each dot represents the result from one sample. *, $p < 0.05$ versus control microRNA. Ctrl, control.

associated with OPLL, whose risk variant increases the expression of a non-coding *CCDC91* isoform. We then demonstrated that the non-coding *CCDC91* isoform is involved in promoting osteogenic differentiation. Our findings reveal a role for the non-coding *CCDC91* isoform in osteogenesis and contribute to our understanding of the pathogenesis of OPLL.

The previously reported GWASs show that most of the associated variants for complex diseases were located in non-coding regions,[23] suggesting that most of the causal variants in complex diseases, including OPLL, may be regulatory variants that affect gene expression. Therefore, gene expression quantitative trait locus (eQTL) analysis is an effective approach to investigating the genetic basis of complex diseases.[24] Because the eQTL effects would be different among cell types and tissues, cell-type-specific eQTL analysis is necessary to analyze the causal effects of complex diseases. There are several eQTL databases, including the Genotype-Tissue Expression (GTEx) project,[25] but none currently address OPLL-related tissues (cells), such as ligament, bone, and cartilage. In this study, to detect potential variants that regulate gene expression levels in OPLL-related cells, we used a machine learning model trained with cell-type-specific transcription,[15] measured by CAGE,[26,27] that can predict the effect of mutations on gene expression.

We identified putative causal variants in the region at chromosome 12p11.22 containing *CCDC91*, which encodes a protein involved in the *trans*-Golgi network.[28] Most mammalian genes express more than one transcript, often generated as a result of alternative splicing. Furthermore, some genes express lncRNA isoforms in addition to their canonical protein-coding transcripts, making it impossible to strictly define a gene as coding or non-coding.[29] Steroid receptor RNA activator 1 (*SRA1*) is one of the characterized genes that expresses functional lncRNA isoforms in addition to protein-coding transcripts.[30] Alternative splicing and the use of different transcription start sites generate protein-coding transcripts and lncRNAs from this gene locus with variations at the 5′ end.[31] While SRA1 is a transcriptional repressor, the lncRNA is a co-activator for nuclear receptors including the estrogen receptor and peroxisome proliferator-activated receptor γ (PPARγ).[32] *CCDC91* has 22 isoforms (transcripts) in the Ensembl Genome Browser, nine of which are lncRNAs (Figure S1). The novel non-coding *CCDC91* isoform has an osteogenic differentiation-promoting function not known for the canonical CCDC91 protein.

Genetic variants that control alternative splicing, called splicing quantitative trait loci (sQTLs), have been found to be enriched among GWAS loci.[33] rs35098487 was associated with the read counts corresponding to the exons 10–12 junction of *CCDC91* (Figures S1 and S3) in visceral adipose tissue in the GTEx database. However, we did not detect any splice variants starting at p4@CCDC91 other than the novel isoform by RACE and RT-PCR, suggesting that this sQTL is derived from splice variants with other transcription start sites. In the future, an sQTL analysis should focus on the tissues expressing the isoform.

Cells have mRNA surveillance systems to recognize aberrant translation termination or elongation and remove abnormal mRNAs. It is well known that mRNA containing a premature termination codon is eliminated by nonsense-mediated mRNA decay.[34] In contrast, the NSD system degrades non-stop mRNA that lacks a translation termination codon and is produced mainly by polyadenylation within an open reading frame.[35] However, non-stop mRNA levels are reduced to only a quarter of wild-type levels in analyses using reporter genes.[19] In fact, the novel isoform of *CCDC91* had sufficient expression in multiple tissues to be detected by conventional RT-PCR. In addition to mRNA degradation by NSD,

(*CCDC91_nv* and *RUNX2*), 7 days (*ALPL*), or 14 days (*IBSP*, *BGLAP*, and *SP7*) after induction. Data represent means ± SD from triplicate assays. Each dot represents the result from one sample. *, $p < 0.05$ versus control siRNA or control vector. Ctrl, control.

(C) Effects of the novel *CCDC91* isoform knockdown or overexpression on alkaline phosphatase activity in osteogenic differentiation of MSCs. Scale bars, 0.5 mm.

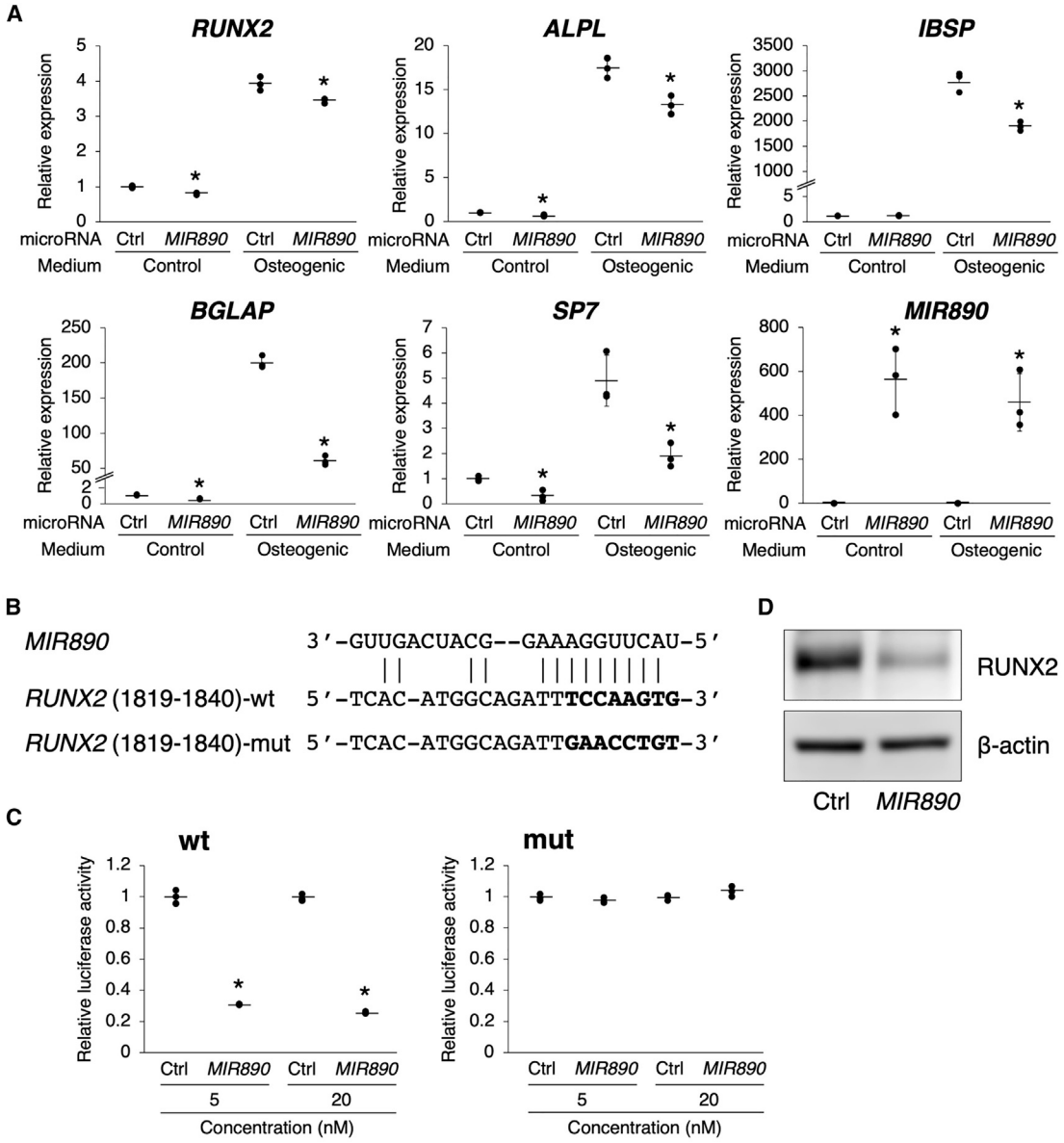

Figure 4. *MIR890* inhibits osteogenic differentiation via targeting *RUNX2*

(A) Effects of overexpression of *MIR890* on the expression of osteogenic marker genes. Expression of osteogenic marker genes and *MIR890* in MSCs cultured in osteogenic conditions was measured 3 days (*RUNX2* and *MIR890*), 7 days (*ALPL*), or 14 days (*IBSP, BGLAP,* and *SP7*) after induction. Data represent means ± SD from triplicate assays. Each dot represents the result from one sample. *, $p < 0.05$ versus control microRNA. Ctrl, control.

(B) Alignment of mature human *MIR890* and 3′-UTR of *RUNX2* sequences. Sequences that differ between wild-type (WT) and mutated (mt) are shown in bold.

(C) Luciferase reporter assay confirmed that co-transfection of *MIR890* and *RUNX2* 3′-UTR WT remarkably decreased the luciferase activity. Data represent means ± SD from quadruplicate assays. Each dot represents the result from one sample. *, $p < 0.05$ versus control microRNA. Ctrl, control.

(D) Effects of overexpression of *MIR890* on protein level of RUNX2.

repression of translation and protein degradation are important for suppressing the level of abnormal proteins.[36,37] We expect that the novel isoform will not produce abnormal proteins because of these repressive mechanisms. Further experiments are necessary to determine whether the novel isoform produces protein.

The expression of the novel isoform is specific to the stage of osteogenic differentiation. The novel isoform was upregulated in the early stage of osteogenic differentiation in MSCs (Figure S5). Both MG-63 and Saos-2 cells are derived from individuals with osteosarcoma, but the isoform is observed only in MG-63 (Figure S4B). MG-63 is an immature osteoblast cell line, while Saos-2 is characterized as a mature osteoblast cell line.[38] This difference in the cell differentiation stage could be related to the difference in the expression of the isoform.

Several microRNAs regulate osteogenic differentiation by modulating osteogenic signaling pathways, such as TGF-β/BMP[39] and PI3K/AKT[40] signaling pathways. The homologous sequence of *MIR890* was screened by alignment. Among the large numbers of sequence alignment data, we focused on *RUNX2*, which was a well-recognized master transcription factor of osteogenic differentiation.[41] *RUNX2* regulates several bone-specific genes, including *ALPL*, *IBSP*, *BGLAP*, and *SPP1*,[42] by binding to the osteoblast-specific element.[43] Previous studies have found that some microRNAs affect osteogenic differentiation by regulating *RUNX2* expression.[44] For example, *Mir133a-1* inhibits the differentiation of mouse-derived C2C12 cells into osteoblasts by targeting *Runx2*.[45]

In this study, we found that *MIR890* suppresses *RUNX2* expression via direct binding to a site on the 3′ UTR of *RUNX2*. Little is known about the function of *MIR890*. *MIR890* inhibits proliferation and invasion in breast cancer cells by targeting *BSG*.[46] Consistent with our findings, *MIR890* expression is significantly repressed during the osteogenic differentiation of MSCs.[47] These findings suggest that *MIR890* could control the osteogenic differentiation of MSCs through modulating *RUNX2*.

Supplemental information

Supplemental information can be found online at https://doi.org/10.1016/j.ajhg.2023.03.004.

Acknowledgments

We thank OPLL-affected individuals and their families who participated in this study and the OPLL network in Japan (Zenkoku sekichujintai-kotsukasho kanja-kazoku renraku-kyougikai). We also thank T. Oguma for her technical assistance. This work was supported by grants from Japan Society for the Promotion of Science KAKENHI (JP20H03811 [M.N.], JP20K15773 [M.K.], JP20H00462 [C.T.], and JP22H03207 [S.I.]), The Japan College of Rheumatology Grant for Promoting Basic Rheumatology (C.T.), and grants from Japan Agency for Medical Research and Development (JP21kk0305013 [C.T.], JP21tm0424220 [C.T.], and JP21ck0106642 [C.T.]).

Author contributions

Conceptualization: M.N. and S.I. Methodology: M.N. and M.K. Software: M.K., L.G., and C.T. Investigation: M.N. and M.K. Visualization: M.N. Funding acquisition: M.N., M.K., C.T., and S.I. Supervision: C.T. and S.I. Writing (original draft): M.N. Writing (review and editing): M.K., C.T., and S.I.

Declaration of interests

The authors declare no competing interests.

Received: January 18, 2023
Accepted: March 7, 2023
Published: March 28, 2023

Web resources

Ensembl Genome Browser, https://www.ensembl.org
GTEx, https://gtexportal.org
miRDB, https://mirdb.org
RNA22 v2, https://cm.jefferson.edu/rna22/Interactive/

References

1. Boody, B.S., Lendner, M., and Vaccaro, A.R. (2019). Ossification of the posterior longitudinal ligament in the cervical spine: a review. Int. Orthop. *43*, 797–805.
2. Matsunaga, S., and Sakou, T. (2012). Ossification of the posterior longitudinal ligament of the cervical spine: etiology and natural history. Spine *37*, E309–E314.
3. Belanger, T.A., Roh, J.S., Hanks, S.E., Kang, J.D., Emery, S.E., and Bohlman, H.H. (2005). Ossification of the posterior longitudinal ligament. Results of anterior cervical decompression and arthrodesis in sixty-one North American patients. J. Bone Joint Surg. Am. *87*, 610–615.
4. Ikegawa, S. (2014). Genetics of ossification of the posterior longitudinal ligament of the spine: a mini review. J. Bone Metab. *21*, 127–132.
5. Saetia, K., Cho, D., Lee, S., Kim, D.H., and Kim, S.D. (2011). Ossification of the posterior longitudinal ligament: a review. Neurosurg. Focus *30*, E1.
6. Nakajima, M., Takahashi, A., Tsuji, T., Karasugi, T., Baba, H., Uchida, K., Kawabata, S., Okawa, A., Shindo, S., Takeuchi, K., et al. (2014). A genome-wide association study identifies susceptibility loci for ossification of the posterior longitudinal ligament of the spine. Nat. Genet. *46*, 1012–1016.
7. Koike, Y., Takahata, M., Nakajima, M., Otomo, N., Suetsugu, H., Liu, X., Endo, T., Imagama, S., Kobayashi, K., Kaito, T., et al. (2022). Genetic insights into ossification of the posterior longitudinal ligament of the spine. Preprint at medRxiv. https://doi.org/10.1101/2022.06.16.22276152.
8. Sollis, E., Mosaku, A., Abid, A., Buniello, A., Cerezo, M., Gil, L., Groza, T., Güneş, O., Hall, P., Hayhurst, J., et al. (2023). The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. Nucleic Acids Res. *51*, D977–D985.
9. Nakajima, M., Kou, I., Ohashi, H., Genetic Study Group of the Investigation Committee on the Ossification of Spinal Ligaments, and Ikegawa, S. (2016). Identification and functional characterization of RSPO2 as a susceptibility gene for ossification of the posterior longitudinal ligament of the spine. Am. J. Hum. Genet. *99*, 202–207.
10. Jokoji, G., Maeda, S., Oishi, K., Ijuin, T., Nakajima, M., Tawaratsumida, H., Kawamura, I., Tominaga, H., Taketomi, E., Ikegawa, S., and Taniguchi, N. (2021). CDC5L promotes early chondrocyte differentiation and proliferation by modulating pre-mRNA splicing of SOX9, COL2A1, and WEE1. J. Biol. Chem. *297*, 100994.
11. Noguchi, S., Arakawa, T., Fukuda, S., Furuno, M., Hasegawa, A., Hori, F., Ishikawa-Kato, S., Kaida, K., Kaiho, A., Kanamori-Katayama, M., et al. (2017). FANTOM5 CAGE profiles of human and mouse samples. Sci. Data *4*, 170112.
12. Park, J.B. (2012). The effects of dexamethasone, ascorbic acid, and beta-glycerophosphate on osteoblastic differentiation by regulating estrogen receptor and osteopontin expression. J. Surg. Res. *173*, 99–104.
13. Ozaki, K., Sato, H., Iida, A., Mizuno, H., Nakamura, T., Miyamoto, Y., Takahashi, A., Tsunoda, T., Ikegawa, S., Kamatani,

N., et al. (2006). A functional SNP in PSMA6 confers risk of myocardial infarction in the Japanese population. Nat. Genet. *38*, 921–925.
14. Nakajima, M., Kizawa, H., Saitoh, M., Kou, I., Miyazono, K., and Ikegawa, S. (2007). Mechanisms for asporin function and regulation in articular cartilage. J. Biol. Chem. *282*, 32185–32192.
15. Koido, M., Hon, C.C., Koyama, S., Kawaji, H., Murakawa, Y., Ishigaki, K., Ito, K., Sese, J., Parrish, N.F., Kamatani, Y., et al. (2022). Prediction of the cell-type-specific transcription of non-coding RNAs from genome sequences via machine learning. Nat. Biomed. Eng., 1–5. https://doi.org/10.1038/s41551-022-00961-8.
16. Miranda, K.C., Huynh, T., Tay, Y., Ang, Y.S., Tam, W.L., Thomson, A.M., Lim, B., and Rigoutsos, I. (2006). A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. Cell *126*, 1203–1217.
17. Chen, Y., and Wang, X. (2020). miRDB: an online database for prediction of functional microRNA targets. Nucleic Acids Res. *48*, D127–D131.
18. Frischmeyer, P.A., van Hoof, A., O'Donnell, K., Guerrerio, A.L., Parker, R., and Dietz, H.C. (2002). An mRNA surveillance mechanism that eliminates transcripts lacking termination codons. Science *295*, 2258–2261.
19. Inada, T., and Aiba, H. (2005). Translation of aberrant mRNAs lacking a termination codon or with a shortened 3′-UTR is repressed after initiation in yeast. EMBO J. *24*, 1584–1595.
20. Thomson, D.W., and Dinger, M.E. (2016). Endogenous microRNA sponges: evidence and controversy. Nat. Rev. Genet. *17*, 272–283.
21. MacArthur, J., Bowler, E., Cerezo, M., Gil, L., Hall, P., Hastings, E., Junkins, H., McMahon, A., Milano, A., Morales, J., et al. (2017). The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). Nucleic Acids Res. *45*, D896–D901.
22. Sonehara, K., and Okada, Y. (2021). Genomics-driven drug discovery based on disease-susceptibility genes. Inflamm. Regen. *41*, 8.
23. Hindorff, L.A., Sethupathy, P., Junkins, H.A., Ramos, E.M., Mehta, J.P., Collins, F.S., and Manolio, T.A. (2009). Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc. Natl. Acad. Sci. USA *106*, 9362–9367.
24. Cano-Gamez, E., and Trynka, G. (2020). From GWAS to Function: Using Functional Genomics to Identify the Mechanisms Underlying Complex Diseases. Front. Genet. *11*, 424.
25. GTEx Consortium (2015). Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science *348*, 648–660.
26. Andersson, R., Gebhard, C., Miguel-Escalada, I., Hoof, I., Bornholdt, J., Boyd, M., Chen, Y., Zhao, X., Schmidl, C., Suzuki, T., et al. (2014). An atlas of active enhancers across human cell types and tissues. Nature *507*, 455–461.
27. FANTOM Consortium and the RIKEN PMI and CLST DGT, Forrest, A.R.R., Kawaji, H., Rehli, M., Baillie, J.K., de Hoon, M.J.L., Haberle, V., Lassmann, T., Kulakovskiy, I.V., Lizio, M., et al. (2014). A promoter-level mammalian expression atlas. Nature *507*, 462–470.
28. Mardones, G.A., Burgos, P.V., Brooks, D.A., Parkinson-Lawrence, E., Mattera, R., and Bonifacino, J.S. (2007). The trans-Golgi network accessory protein p56 promotes long-range movement of GGA/clathrin-containing transport carriers and lysosomal enzyme sorting. Mol. Biol. Cell *18*, 3486–3501.
29. Dhamija, S., and Menon, M.B. (2018). Non-coding transcript variants of protein-coding genes - what are they good for? RNA Biol. *15*, 1025–1031.
30. Kawashima, H., Takano, H., Sugita, S., Takahara, Y., Sugimura, K., and Nakatani, T. (2003). A novel steroid receptor co-activator protein (SRAP) as an alternative form of steroid receptor RNA-activator gene: expression in prostate cancer cells and enhancement of androgen receptor activity. Biochem. J. *369*, 163–171.
31. Hube, F., Guo, J., Chooniedass-Kothari, S., Cooper, C., Hamedani, M.K., Dibrov, A.A., Blanchard, A.A.A., Wang, X., Deng, G., Myal, Y., and Leygue, E. (2006). Alternative splicing of the first intron of the steroid receptor RNA activator (SRA) participates in the generation of coding and noncoding RNA isoforms in breast cancer cell lines. DNA Cell Biol. *25*, 418–428.
32. Leygue, E. (2007). Steroid receptor RNA activator (SRA1): unusual bifaceted gene products with suspected relevance to breast cancer. Nucl. Recept. Signal. *5*, e006.
33. GTEx Consortium (2020). The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science *369*, 1318–1330.
34. Singh, G., and Lykke-Andersen, J. (2003). New insights into the formation of active nonsense-mediated decay complexes. Trends Biochem. Sci. *28*, 464–466.
35. van Hoof, A., Frischmeyer, P.A., Dietz, H.C., and Parker, R. (2002). Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. Science *295*, 2262–2264.
36. Dimitrova, L.N., Kuroha, K., Tatematsu, T., and Inada, T. (2009). Nascent peptide-dependent translation arrest leads to Not4p-mediated protein degradation by the proteasome. J. Biol. Chem. *284*, 10343–10352.
37. Ito-Harashima, S., Kuroha, K., Tatematsu, T., and Inada, T. (2007). Translation of the poly(A) tail plays crucial roles in nonstop mRNA surveillance via translation repression and protein destabilization by proteasome in yeast. Genes Dev. *21*, 519–524.
38. Czekanska, E.M., Stoddart, M.J., Richards, R.G., and Hayes, J.S. (2012). In search of an osteoblast cell model for in vitro research. Eur. Cell. Mater. *24*, 1–17.
39. Garcia, J., and Delany, A.M. (2021). MicroRNAs regulating TGFbeta and BMP signaling in the osteoblast lineage. Bone *143*, 115791.
40. Yang, C., Liu, X., Zhao, K., Zhu, Y., Hu, B., Zhou, Y., Wang, M., Wu, Y., Zhang, C., Xu, J., et al. (2019). miRNA-21 promotes osteogenesis via the PTEN/PI3K/Akt/HIF-1alpha pathway and enhances bone regeneration in critical size defects. Stem Cell Res. Ther. *10*, 65.
41. Chan, W.C.W., Tan, Z., To, M.K.T., and Chan, D. (2021). Regulation and Role of Transcription Factors in Osteogenesis. Int. J. Mol. Sci. *22*, 5445.
42. Bruderer, M., Richards, R.G., Alini, M., and Stoddart, M.J. (2014). Role and regulation of RUNX2 in osteogenesis. Eur. Cell. Mater. *28*, 269–286.
43. Ducy, P., and Karsenty, G. (1995). Two distinct osteoblast-specific cis-acting elements control expression of a mouse osteocalcin gene. Mol. Cell Biol. *15*, 1858–1869.
44. Hensley, A.P., and McAlinden, A. (2021). The role of microRNAs in bone development. Bone *143*, 115760.
45. Li, Z., Hassan, M.Q., Volinia, S., van Wijnen, A.J., Stein, J.L., Croce, C.M., Lian, J.B., and Stein, G.S. (2008). A microRNA

signature for a BMP2-induced osteoblast lineage commitment program. Proc. Natl. Acad. Sci. USA *105*, 13906–13911.

46. Wang, C., Xu, C., Niu, R., Hu, G., Gu, Z., and Zhuang, Z. (2019). MiR-890 inhibits proliferation and invasion and induces apoptosis in triple-negative breast cancer cells by targeting CD147. BMC Cancer *19*, 577.

47. Avendaño-Félix, M., Fuentes-Mera, L., Ramos-Payan, R., Aguilar-Medina, M., Pérez-Silos, V., Moncada-Saucedo, N., Marchat, L.A., González-Barrios, J.A., Ruiz-García, E., Astudillo-de la Vega, H., et al. (2019). A novel osteomirs expression signature for osteoblast differentiation of human amniotic membrane-derived mesenchymal stem cells. BioMed Res. Int. *2019*, 8987268.