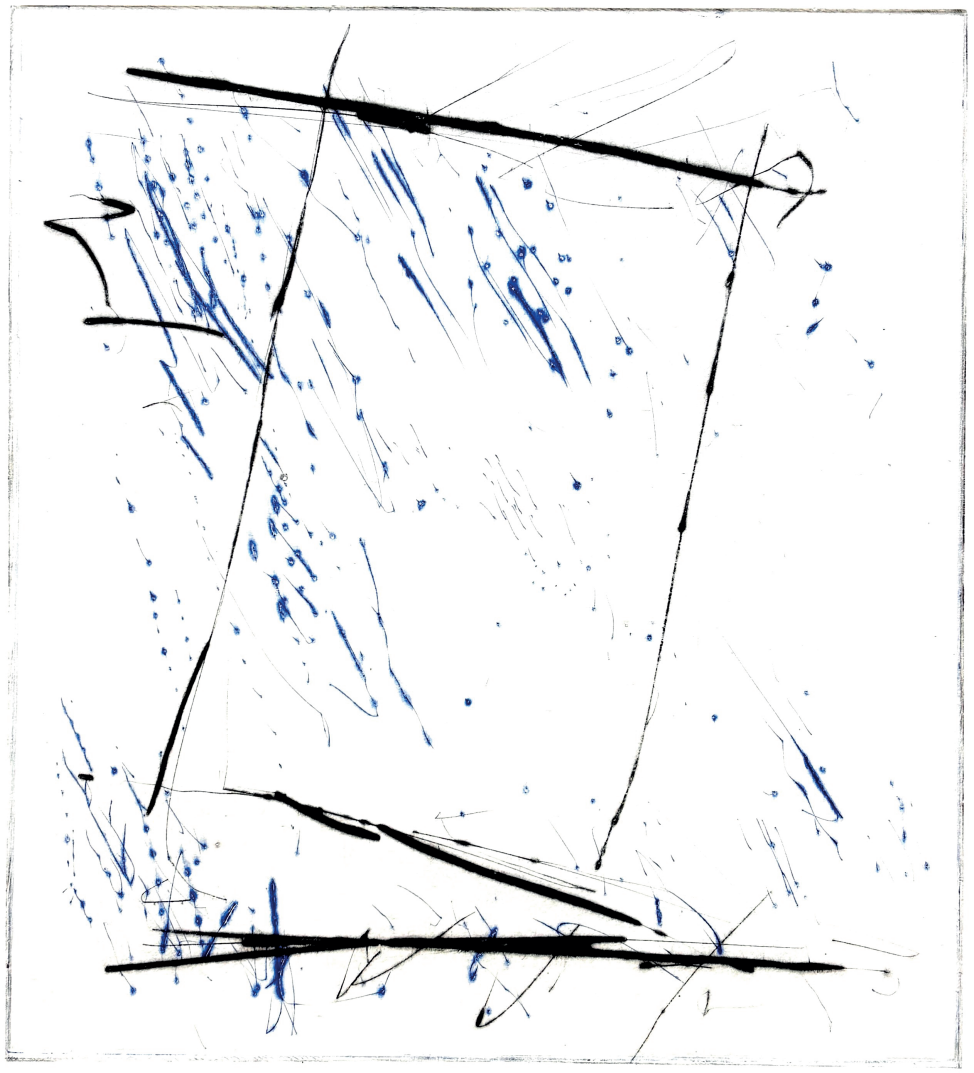
Modeling the genetic and mechanical interplay in osteoarthritis: from in vitro systems to mechanistic insights

Bloks, N.G.C.



PETER M.G COX *1986*

CHAPTER 5

A damaging *COL6A3* variant alters the *MIR31HG*-regulated response of chondrocytes in neo-cartilage organoids to hyper-physiologic mechanical loading

Niek GC Bloks[1], Zainab Harissa[2,3], Giorgia Mazzini[1], Shaunak S Adkar[2,3], Amanda R Dicks[2,3], Ghazaleh Hajmousa[1], Nancy Steward[2,3] Roman I. Koning[1], Aat Mulder[1], Berend B.R. de Koning[1], Margreet Kloppenburg[1], Rodrigo Coutinho de Almeida[1], Yolande FM Ramos[1], Farshid Guilak[2,3,†], Ingrid Meulenbelt[1,†]

† Shared last authors

[1]Dept. Biomedical Data Sciences, Section Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands
[2]Washington University, Saint Louis, MO, USA
[3]Shriners Hospitals for Children, Saint Louis, MO, USA

ADVANCED SCIENCE *11.36 (2024): 2400720.DOI: 10.1002/advs.202400720*

Abstract

The pericellular matrix (PCM), with its hallmark proteins collagen type VI (COLVI) and fibronectin (FN), surrounds chondrocytes and is critical in transducing the biomechanical cues. To identify genetic variants that change protein function, exome sequencing is performed in a patient with symptomatic OA at multiple joint sites. A predicted damaging variant in *COL6A3* is identified and introduced by CRISPR-Cas9 genome engineering in two established human induced pluripotent stem cell-derived *in-vitro* neo-cartilage organoid models. The downstream effects of the *COL6A3* variant on the chondrocyte phenotypic state is studied by a multi-omics (mRNA and lncRNA) approach in interaction with hyper-physiological mechanical loading conditions. The damaging variant in *COL6A3* resulted in significantly lower binding between the PCM proteins COLVI and FN and provoked an osteoarthritic chondrocyte state. By subsequently exposing the neo-cartilage organoids to hyper-physiological mechanical stress, we demonstrate that the *COL6A3* variant in chondrocytes abolished the characteristic inflammatory signaling response after mechanical loading with *PTGS2*, *PECAM1*, and *ADAMTS5*, as central genes. Finally, by integrating epigenetic regulation, we identified the lncRNA *MIR31HG* as key regulator of the characteristic inflammatory signaling response to mechanical loading.

Introduction

Osteoarthritis (OA) is a complex, multifactorial disease revolving around the interplay between genetic and environmental risk factors. In particular, biomechanical cues play a critical role in not only joint health but also drive the onset and progression of disease (*1*). Indeed, physiologic levels of mechanical loading by virtue of physical exercise can slow down OA disease progression (*2, 3*). In contrast, hyper-physiological loading, as seen with post-traumatic injuries such as articular fracture, meniscal tear, or rupture of the anterior cruciate ligament, concomitant with altered joint kinematics, is a major risk factor for the onset and progression of OA (*4*). Both physiological and hyper-physiological loading results in alterations in the structural composition of the cartilage extracellular matrix (*1*). These findings suggest that the balance of chondrocyte anabolic and catabolic processes that maintain cartilage homeostasis is tightly regulated by biomechanical cues.

Chondrocytes reside in a pericellular matrix (PCM), which modulates the transduction of mechanical cues from the extracellular matrix (ECM) towards the chondrocyte. The PCM for that matter contains specific molecular components such as collagen 6 (COLVI) that are known to regulate the biomechanical environment of the chondrocyte e.g. via calcium signaling in response to mechanical stress (*5-7*). This is further evidenced by knock-out of a COLVI sub-unit alpha 1 in a murine model, which causes an osteoarthritic phenotype through dysregulation of mechano-transduction (*5, 8*). Nonetheless, the lasting effects of dysregulated mechano-transduction after injurious mechanical loading conditions on the cellular phenotype of the chondrocyte is, however, less well studied and remains unclear.

Here we report the identification of a *COL6A3* missense variant through exome sequencing of patients with generalized OA (Genetics osteoArthritis Research and Progression (GARP) study) (*9*). *COL6A3,* coding for one of the monomeric sub-units of COLVI (*10*) that resides in and interacts with other PCM proteins (*11, 12*), is likely involved in the regulation of cartilage structural composition and mechanical properties. We hypothesized that the identified *COL6A3* variant would perturb mechano-transduction in response to hyper-physiologic loading, thereby affecting the chondrocyte cellular phenotype. Recently, OA disease modeling platforms utilizing human induced pluripotent stem cells (hiPSCs) have been used to facilitate the interrogation of these pathogenic variants (*13*). To gain an understanding of the role of COLVI during chondrogenesis and chondrocyte function in exposure to environmental stressors, we studied this variant in an *in vitro* model of OA using hiPSC-derived cartilage. Hereto, we employed genetically *COL6A3*-edited hiPSCs to an established *in vitro* cartilage organoid model, and these neo-cartilage organoids were exposed to hyper-physiological mechanical

loading conditions. To study the effects of this variant on the chondrocyte phenotype in response to hyper-physiologic mechanical loading conditions, we characterized downstream molecular pathways based on RNA transcriptome wide profiles (mRNA and lncRNA) to identify changes in the chondrocyte phenotypic state and associated regulatory epigenetic mechanisms, for druggable target discovery.

Materials and Methods

Exome sequencing

Exome sequencing of a patient with generalized OA at multiple joint sites was performed by Illumina HiSeq 2000 technology (Beijing Genome Institute) using the protocol described in the supplementary methods.

Patient and public involvement

The patient participation OA Leiden group was consented during regular meetings on the setting and outcome measures of the research question. They were central to dissemination of the research.

Ethics approval

Any necessary ethics approval for the GARP study was secured by the committee medical ethics (CME) of the Leiden University Medical Centre reference number P76/98

hiPSC line and cell culture

An hiPSC line as described earlier was used as the unedited isogenic control (*14*). In short, RVR-iPSC line was retrovirally reprogrammed from BJ fibroblasts. Cells were characterized and their pluripotency (to three germ layers) was confirmed previously (*14, 15*). Further culture conditions can be found in the supplementary methods.

Genome editing of hiPSCs

We employed CRISPR-Cas9 single-stranded oligonucleotide-mediated homology-directed repair in hiPSCs to attain biallelic modification of rs144223596 (c.4510C>T) in an isogenic background, Additional information on the genomic target, guide sequences, and transfection protocol can be found in the supplementary methods.

hiPSC differentiation to induced chondrocytes

Generation of induced chondroprogenitor cells (hiCPCs) was based on a protocol previously described (*14*) which was shown to produce similar neo-cartilage to that produced by human primary articular chondrocytes (*16*). In short: First, hiPSCs were progressed through the anterior primitive streak on day 0, paraxial mesoderm on day 1, early somite on day 2, sclerotome on days 3 to 5, and chondroprogenitor on days 6

to 14. Subsequently, these were washed with MD medium, dissociated with Gentle Cell dissociation medium (Stem Cell), and centrifuged for 5 min at 1200 rpm. Cell aggregates were subsequently maintained in chondrogenic differentiation (CD) medium. A more detailed description of the hiPSC differentiation process and creation of neo-cartilage organoids in described in the supplementary methods.

Mechanical loading

The spherical-shaped neo-cartilage constructs were mechanically loaded using a MACH-1 mechanical testing device (Biomomentum), at a rate of 5hz with 20% sinusoidal peak-to-peak strain for 10 minutes, as described earlier (*17*)

sGAG measurement

Sulfated glycosaminoglycan (sGAG) concentrations in the neo-cartilage organoids (µg sGAG/µg DNA) were measured using the Farndale Dimethyl Methylene Blue (DMMB, Sigma) method (*18*). Additional information can be found in the supplementary methods.

Histology and immunohistochemistry

Neo-cartilage samples were fixed in 4% formaldehyde and embedded in paraffin. Sections were stained with Alcian Blue (Sigma-Aldrich) and Nuclear Fast Red (Sigma-Aldrich). Deposition of collagen II and collagen VI in the neo-cartilage constructs was visualized immunohistochemically according to the protocol described in the supplementary methods.

RT-qPCR

Per sample, two replicate mRNA samples were measured in triplicates in a MicroAmp™ Optical 384-Well Reaction Plate (ThermoFisher Scientific), using the QuantStudio™ Flex Real-Time PCR system (Applied Biosystems™). Additional information can be found in the supplementary methods.

RNAseq

RNA from neo-cartilage constructs was extracted 12 hours post mechanical loading and analyzed using the Illumina NOVAseq 6000. Additional information on RNA isolation, mapping, alignment, and data processing and analysis can be found in the supplementary methods.

Solid-phase binding assay

Conditioned medium of wild-type and *COL6A3* variant organoids was collected and concentrated in preparation for the binding assay. To this end, 450 µl of medium was

collected in 100 K molecular weight cutoff Pierce Protein Concentrators (Thermo Scientific) and centrifuged for 10 min at 12,000g. Subsequently, COL6 concentration was determined using the Human *COL6A3* ELISA Kit (Assay Genie) according to the manufacturer's protocol which can be found in the supplementary methods.

Transmission electron microscopy

For the neo-cartilage organoids, consisting of cells and matrix, a previously established protocol was used to perform TEM. Additional information on sample processing, image acquisition, processing and analysis can be found in the supplementary methods.

Validation of regulatory effects MIR31HG on protein-coding gene expression

Primary chondrocytes were isolated from 12 independent donors and passaged twice, as previously described (*19*). Chondrocytes were transfected in duplo with LNA GapmeR (Qiagen) targeting *MIR31HG* (AGTGCAGCAAAATTAG) or GapmeR negative control (AACACGTCTATACGC) at 10 nM final concentration using Lipofectamine RNAiMax Transfection Reagent according to instructions of the manufacturer (Invitrogen). Cells were lysed 24 hours post transfection and RT-qPCR was performed as described above for *GAPDH, SDHA*, *MIR31HG, PTGS2, IRAK2, PTGER4, IL1R1*. Relative gene expression levels were calculated with the 2-ΔΔCt method, using *GAPDH* and *SDHA* as internal control. A t-test was performed on the −ΔCt values, and P values less than 0.05 were considered significant.

Statistical analysis

For differential expression analysis we have used negative binomial generalized linear model using the R DESeq2 package. Please find the supplementary materials for the description in full of our RNAseq data processing and differential gene expression analysis. For all other data analyses, we have used a generalized linear model including the factors hyper-physiological loading and the *COL6A3* variant using R statistical software version 4.1.1. We have used a two-sided test with a P-value of 0.05 as a significance cut-off. The alignment of the data with the presumptions of the generalized linear model has been tested both visually using boxplots/QQ-plots, as well as formally using the Kolmogorov-Smirnov test. The reported beta value represents the standardized coefficient, indicating the change in the dependent variable's standard deviations for each standard deviation change in the predictor variable. Statistics in figure legends are reported as beta ± standard error

Results

Identification of a damaging variant in *COL6A3*

Whole exome-sequencing was applied to a Caucasian OA patient of Dutch ancestry at the age of 61 affected predominantly with symptomatic OA at multiple sites (Genetics osteoArthritis Research and Progression (GARP) study) (*9*). The exome sequencing resulted in the detection of 81,416 genetic variants, after which a prioritization scheme was followed to identify dominant pathogenic variants as previously described (*13, 20*). In short, first, we selected novel variants in in-house whole genome sequencing projects (N=222) and the BBMRI-Genome of the Netherlands project (GoNL, N=473) (*20*). And second, damaging missense variants were selected based on the prediction of sorting intolerant from tolerant (SIFT) (*21, 22*). This prioritization generated a dataset of 38 novel coding variants that were predicted to have a functional impact on the protein of each identified coding variant **(table S1)**. And third, these 38 novel coding variants were then further prioritized based on their expression patterns in disease-relevant tissue by means of *in silico* analysis of a previously published RNA-sequencing dataset containing 58 paired preserved and lesioned OA cartilage and subchondral bone samples of patients undergoing joint replacement surgery (*23, 24*). We have prioritized genes that were both highly expressed in joint tissues since these genes likely determine characteristics of the tissue and/or the cellular phenotype of chondrocytes and were differentially expressed between preserved and lesioned OA tissues (cartilage) since these genes mark the OA pathophysiological phenotype of chondrocytes. This resulted in the prioritization of two heterozygous variants affecting protein function of genes that are highly expressed in cartilage and differentially expressed between lesioned and preserved cartilage and bone; *MTHFR* (c.1667C>T, p.Pro597Leu, P597L); cartilage: FC=0.84, FDR=0.04, bone: not detected) and *COL6A3* (c.4510C>T, p.Arg1504Trp, R1504W); cartilage: FC=1.75, FDR=$2.25x10^{-5}$, bone: FC=1.69, FDR=0.042) (*24*). *MTHFR* encodes for methylenetetrahydrofolate reductase, which is involved in folate metabolism (*25*). *COL6A3* encodes for collagen VI subunit A3, which together with collagen VI subunit A1 and A2 forms a triple helical COLVI, a primary component of the cartilage PCM, and interacts with other proteins such as fibronectin and hyaluronan (*11, 12, 26*). We then prioritized the *COL6A3* variant because of its relevance to OA and involvement in mechano-transduction.(*5*) The identified *COL6A3* variant is located in the 3rd N-terminal VWA domains that are known to protrude away from the triple helical structure. Moreover, upon exploring the effect of the R1504W variant on protein function by additional *in silico* tools using sequence homology and evolutionary conservation; Polyphen 2 (*27*), Provean , and PANTHER-PSEP (*28*) it was shown that predicted damaging effect on protein function was with high confidence **(table S2)**. Furthermore, the *in silico* tools i-mutant V2.0 and MUpro , using support vector machines, predicted

that the R1504W variant decreases the stability of the protein **(table S2)**. To gain an understanding of the effects of aberrant COLVI function on chondrocyte phenotype in interaction with mechanical loading, we selected this predicted damaging variant for further study.

Introducing the variant in hiPSCs

To investigate the effects of the aberrant COLVI on matrix deposition and molecular pathways, a gene-edited human induced pluripotent stem cell (hiPSC)-derived neo-cartilage organoid model was generated. We employed CRISPR-Cas9 single-stranded oligonucleotide-mediated homology-directed repair in hiPSCs to attain. heterozygous modification of rs144223596 (c.4510C>T) in an isogenic background, which was confirmed by Sanger sequencing **(Fig. S2)**. To further prevent bias in our results, we have screened for potential off-target effects using CRISPOR resulting in four sites that could be affected, be it with low likely-hood (*29*) **(table S3).** We then confirmed

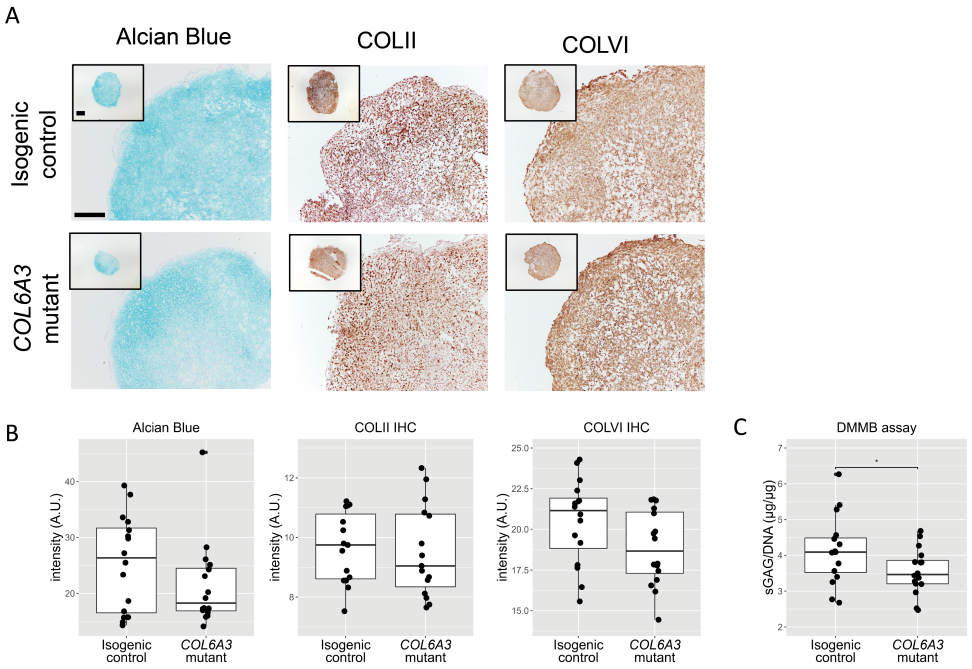


Figure 1 | Effect of the variant on neo-cartilage matrix deposition. (A) Representative images of Alcian blue staining marking sulfated glycosaminoglycans (sGAGs) and immunohistological staining of collagen II (COL II) and collagen VI (COLVI). Scale bar 200µm. **(B)** Quantification of Alcian blue, COLII, and COLVI in isogenic control and *COL6A3*-variant neo-cartilage organoids showed no significant effect of the variant (n=16). **(C)** Quantification of sGAG deposition in neo-cartilage organoids. sGAG deposition in COLVI-variant neo-cartilage organoids is reduced in comparison to isogenic controls (beta=0.55 ± 0.28, t=1.96, P=4.90x10^{-2}, n=16). Statistics: **B, C -** P values were attained using a generalized linear model with **(B)** intensity (for immunostaining) and **(C)** (sGAGs/DNA) as dependent variable and genotype as independent variable. The box plots represent 25th, 50th, and 75th percentiles, and whiskers extend to 1.5 times the interquartile range. Individual samples are depicted by black dots in each graph. * P<0.05.
model, C,D,E – Fisher exact test.

by Sanger sequencing that these sites were not affected by off-target CRISPR-Cas9 single-stranded oligonucleotide-mediated homology-directed repair **(Fig. S2H)**. Next, the *COL6A3*-edited and the unedited isogenic control hiPSCs were differentiated into chondrocytes using a previously established chondrogenic differentiation protocol (*14*). In short, hiPSCs followed a step-wise differentiation protocol via mesodermal lineage differentiation towards chondroprogenitor cells (*14*). These cells were then dissociated, and chondrogenesis was initiated in our organoid pellet model.

Characterization of the matrix of isogenic control and *COL6A3* variant hiPSC-derived cartilage organoids

Successful differentiation towards chondrocytes and the production of neo-cartilage was confirmed by protein staining of collagen II (COLII), collagen VI (COLVI), and sulfated glycosaminoglycans (sGAGs) **(Fig. 1A),** and comparable to isotype controls as shown previously (*17*). COLII and COLVI deposition was not affected by the variant as measured by staining intensity **(Fig. 1B)**. However, quantification of sGAG deposition, using the dimethyl methylene blue (DMMB) assay normalized to DNA content, showed a significant reduction in the COLVI-variant compared to the isogenic control neo-cartilage organoids **(Fig. 1C)**. Next, the isogenic and COLVI-variant neo-cartilage

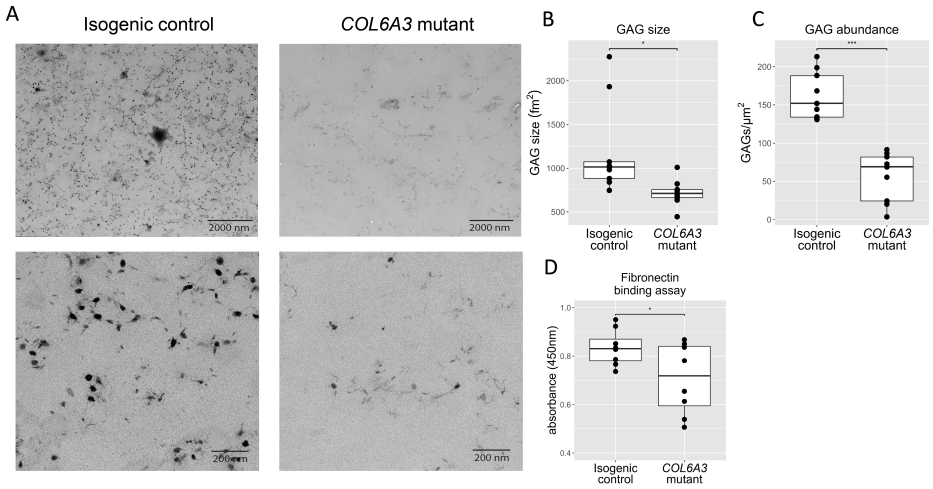


Figure 2 | The COLVI variant reduces sGAG aggregate size and reduces binding to fibronectin. (A) TEM image of isogenic control (left) and a *COL6A3*-variant (right) neo-cartilage organoid. **(B)** Using a deep-learning algorithm, the sGAG aggregate size was measured, showed a reduction in the *COL6A3*-variant neo-cartilage organoids (beta=-479.7±184.4, t=-2.60, P=1.93X10^{-2}). **(C)** GAG amount was decreased in the *COL6A3*-variant neo-cartilage organoids (beta=-106.14 ± 14.95, t=-7.10, P=2.52X10^{-6}). **(D)** Solid-phase binding assay with full-length fibronectin-coated wells and wild-type and R1504W variant COLVI. R1504W showed a reduced binding to fibronectin (beta=-0.13 ± 0.06, t=-2.17, P=4.37 x10^{-2}, N=8). This assay did not show binding of COLVI to hyaluronan. Statistics: **B, C, D** - P values were attained using a generalized linear model, with sGAG size, sGAGs/mm^2, or absorbance at 450nm as a dependent variable and genotype as an independent variable. The box plots represent 25th, 50th, and 75th percentiles, and whiskers extend to 1.5 times the interquartile range. Individual samples are depicted by black dots in each graph. * P<0.05, *** P<0.001.

organoids were characterized by targeted gene expression analysis using RT-qPCR of markers relevant for cartilage homeostasis which were previously shown to be mechano-responsive **(table S3)**(17). The *COL6A3* variant reduced expression of the catabolic marker *ADAMTS5*, while also reducing expression of the anabolic markers *COL2A1* as well as *ACAN*.

Transmission electron microscopy of the cartilage neo-matrix

To further investigate the effect of the variant on the structural properties of the PCM and ECM, we performed transmission electron microscopy. As shown in **Figure 2A,** there was a reduced abundance of sGAG-like structures in the *COL6A3*-variant neo-cartilage organoids versus the isogenic control. Consistent with the DMMB assay, upon performing quantitative analysis of these sGAG-like structures using a deep learning algorithm (*30*), we confirmed that variant COLVI cartilage had significantly decreased aggregate size **(Fig.2B)** and sGAG abundance **(Fig. 2C)**.

Binding of isogenic control and *COL6A3* variant COLVI to pericellular matrix proteins

The identified R1504W variant is located in the N-terminal domain of *COL6A3* coding for a von Willebrand Factor A domain, which is involved in binding two important constituents in cartilage PCM and ECM, fibronectin and hyaluronan (*12, 13, 26*). Arginine at this position in the *COL6A3* protein is evolutionary highly conserved, indicating that an amino acid change at this position is likely to affect protein function **(Fig. S1)**. Thus, we hypothesized that the change from a polar arginine to a non-polar tryptophan affects the interaction between COLVI and PCM/ECM proteins. To this end, we performed a fibronectin and hyaluronan solid-phase binding assay with wild-type and variant COLVI, which was extracted from the chondrogenic differentiation medium of neo-cartilage organoids. Results, as shown in **Figure 2D**, demonstrated that binding to fibronectin was reduced for variant COLVI. It must be noted that this binding could be indirect, via other proteins in the medium. However, we could not detect any binding of wild-type nor variant COLVI to hyaluronan in our assay.

Effect of missense *COL6A3* variant on the transcriptomic landscape

Next, being one of the most sensitive and informative measures of cellular responses to genetic perturbations, we performed mRNA sequencing, to determine the downstream effects of these changes in the PCM and ECM secondary to the damaging *COL6A3* variant. Multifactorial analysis, using surrogate variable analysis (SVA) correction **(Fig. S3)** (*31*) revealed 3700 significant DEGs between the *COL6A3* variant and isogenic controls (FDR<0.05) **(Fig. 3A; table S4)**. Among these genes, 55% were upregulated and 45% were downregulated. Notable highly significant DEGs were related to development and cartilage metabolism, with the downregulation of structural PCM and ECM proteins,

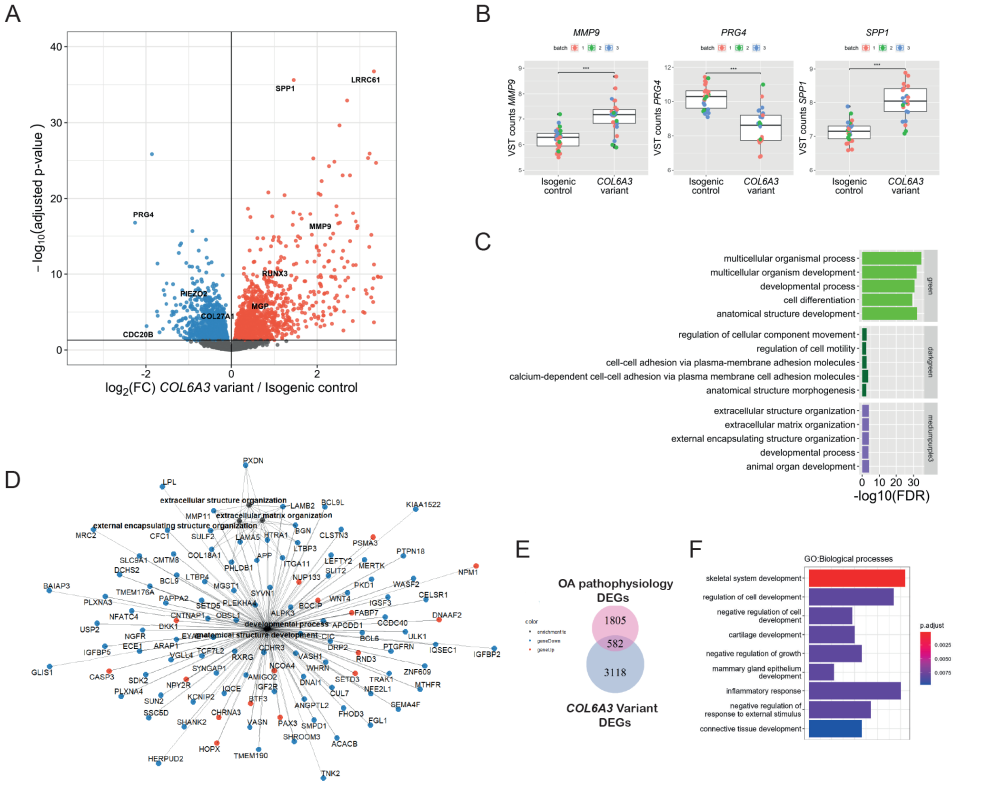


Figure 3 | Transcriptomic profile in presence of the *COL6A3* variant (A) Volcano plot showing the differentially expressed genes (DEGs) in response to the *COL6A3* variant (n=25-26). Red dots denote DEGs with an FDR<0.05 that are upregulated, and blue dots represent DEGs that are downregulated, as determined by DESeq2 analysis. **(B)** Notable examples of DEGs between isogenic control and *COL6A3*-variant neo-cartilage pellets. The box plots represent 25th, 50th, and 75th percentiles, and whiskers extend to 1.5 times the interquartile range. Individual samples are depicted by black dots in each graph. ***FDR<0.001 **(C)** Overrepresentation enrichment analysis (KEGG, REACTOME, GO biological processes) of the top 3 weighted gene co-expression network analysis (WGCNA) co-expression modules where the first principal component is significantly associated with the *COL6A3* variant. **(D)** Pathway – gene network of the enrichment analysis of the mediumpurple3 cluster. Lines depict the relationship between the genes and the pathways determined by enrichment analysis. Blue dots depict downregulated DEGs with the *COL6A3* variant, while red dots depict upregulated DEGs with the *COL6A3* variant. **(E)** Overlap between the *COL6A3* variant DEGs and differential expression between lesioned and preserved cartilage from a previously published dataset (RAAK study, Ramos et al., 2014). **(F)** Over-representation enrichment analysis of overlapping DEGs between the variant and lesioned versus preserved cartilage. Count depicts the number of genes that are categorized to each pathway. Statistics: **A, B –** negative binomial generalized linear model, **C, D, F –** Fisher exact test.

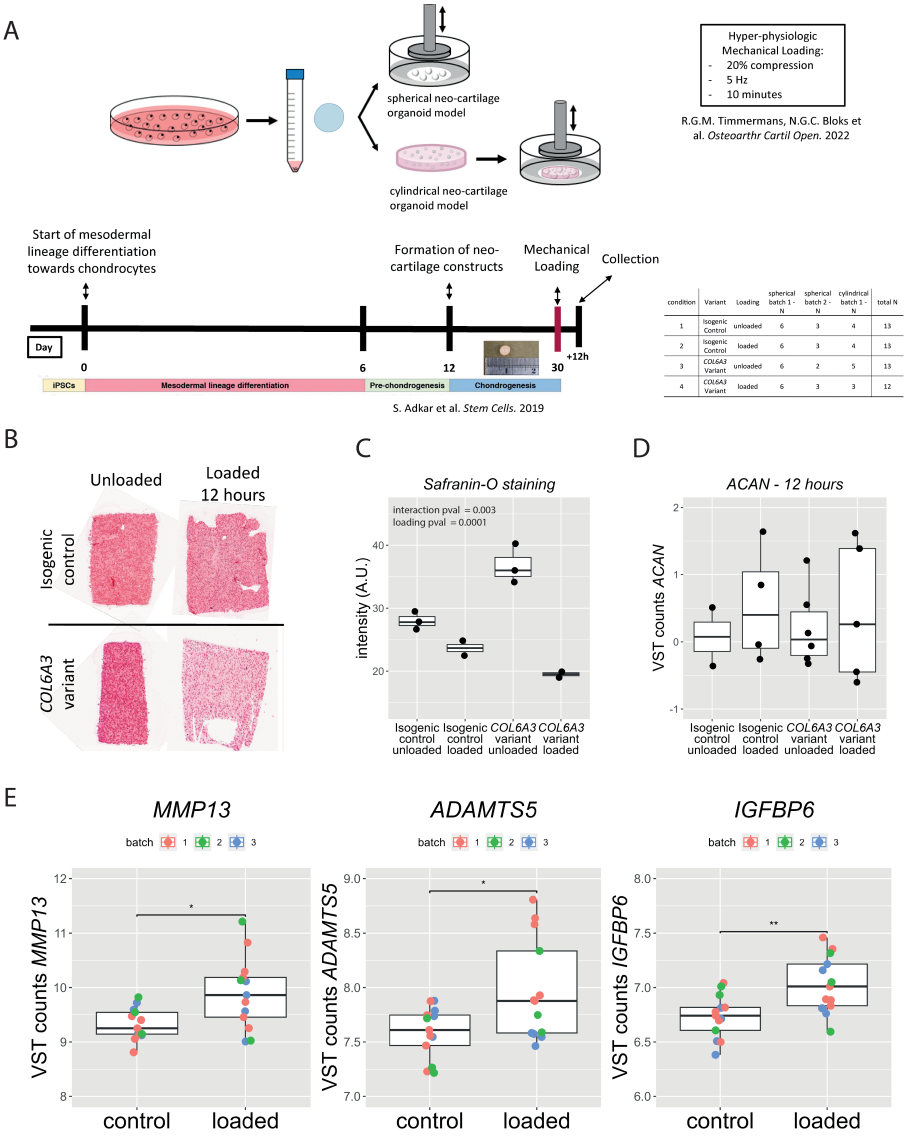
such as *COL27A1* and *PRG4*, increased catabolic activity, such as *MMP9*, and ECM mineralization, such as *SPP1* **(Fig. 3A-B)**.

To determine the biological processes associated with the *COL6A3* variant, we performed a weighted gene co-expression network analysis (WGCNA) on the RNA-sequencing data set. This resulted in the detection of 20 distinct co-expression networks **(Fig. S4)**. Multifactorial regression analysis revealed 10 co-expression networks with the first principal component significantly associated with the *COL6A3* genotype **(Fig. S4)**. The

top 3 most significant co-expression networks associated with the *COL6A3*-variant genotype were enriched for developmental processes, cell-cell adhesion/anatomical structure morphogenesis, and ECM organization **(Fig. 3C; table S5)**. Enrichment of ECM organization again underlines the detrimental effects of the *COL6A3* variant on the development of cartilage, affecting genes such as *LAMA5*, *LAMB2,* and *COL18A1* that code for structural PCM components **(Fig. 3D)**. While enrichment of cell-cell adhesion pathways suggests that the variant also affects mechano-transduction via altered cell-cell adhesion by downregulation of cadherin-associated genes**,** such as *CDH18*, *CDH6*, *PTGER4*, *MAPK14,* and *ADAMTS5* **(Fig. S5)**. Of note is that *ADAMTS5* is downregulated in the *COL6A3* variant while quantification of sGAGs, the target peptides of ADAMTS5 cleavage, are reduced at both the gene level (*ACAN*, **Table 1**) and protein level **(Fig. 1C; Fig. 2A-C).**

To study the relevance of the *COL6A3* variant to OA pathophysiology, we determined the overlap of differentially expressed genes with the *COL6A3* variant and those previously reported between lesioned and preserved cartilage from OA patients who underwent a joint replacement surgery. The latter marks the heterogeneous overall pathophysiologic state of OA chondrocytes. This revealed a subset of 582 overlapping DEGs **(table S6)**. We then performed pathway analysis on this subset of genes **(Fig. 3F)**. The most significant enriched pathway was skeletal system development, including transcription factors (*RUNX3*, *TNFRSF11B*), ECM components (such as *COL27A1*, *COL11A2*), and anabolic factors (such as *IGF1*, *BMP3*). Well-known pathways related to inflammation "inflammatory response" and "acute inflammatory response" were also enriched containing inflammatory cytokines (*IL31A*) and general inflammatory response factors (*PTGES*, *HLA-e*). Also, the "ossification" pathway was enriched, containing OA risk genes (*SPP1*, *TNC,* and *MGP*). Genes related to cartilage metabolism and mechano-sensing were confirmed by quantitative reverse transcription polymerase chain reaction (RT-qPCR) **(table S7).** Together, this data suggests that the *COL6A3* variant results in downstream expression changes, in part describing the osteoarthritis pathophysiology.

The stranded RNA sequencing allowed us to confirm the genotyping as performed with Sanger sequencing **(Fig S1).** This showed a dispersion of the variant allele frequencies (VAFs) in the included CRISPR-Cas9 edited samples with an average proportion of 0.70 towards the damaging *COL6A3* variant allele **(table S8)**. To explore possible dose-response effects of the dispersed *COL6A3* variant frequency, the VAF was considered as a covariate and analyzed for differential expression. This revealed a robust dose-response effect (FDR<0.05) of the damaging *COL6A3* variant on downstream gene expression in 2728 genes out of the 3700 DEGs (74%), including genes such as *MMP9*, *PRG4*, and *SPP1* **(table S9; Fig. S7).**



condition	Variant	Loading	spherical batch 1 - N	spherical batch 2 - N	cylindrical batch 1 - N	total N
1	Isogenic Control	unloaded	6	3	4	13
2	Isogenic Control	loaded	6	3	4	13
3	COL6A3 Variant	unloaded	6	2	5	13
4	COL6A3 Variant	loaded	6	3	3	12

Figure 4 | Chondrogenic models to study the effects of aberrant COLVI function in interaction with hyper-physiologic mechanical loading conditions. (A) hiPSCs in which the R1504W variant was introduced using CRISPR-Cas9 genome editing. These cells were differentiated using an established differentiation protocol to produce neo-cartilage organoids. Two different organoid models were employed and jointly analyzed; 1. A spherical pellet model harnessing the original matrix produced by the hiPSCs. 2. A cylindrical organoid model in which the hiPSC-derived chondrocytes were embedded in an agarose construct, ideally suited for testing the effects of mechanical loading conditions. These constructs were both exposed to hyper-physiological loading conditions, after which the organoids were harvested for downstream analysis. (B) Safranin 'O' staining at 12 hours after hyper physiological mechanical loading. **(C)** Safranin 'O' staining intensity at 12 hours after hyper physiological mechanical loading. (n=2-3) **(D)** RT-qPCR data of ACAN and *COL2A1* expression at 12 hours and 4 days after hyper physiological mechanical loading. (n=2-6) **(E)** Upregulation of post-traumatic markers in isogenic control samples measured by RNA-sequencing (n=13). Statistics: C, D – generalized linear model. E, negative binomial generalized linear model. The box plots represent 25th, 50th, and 75th percentiles, and whiskers extend to 1.5 times the interquartile range. *FDR<0.05

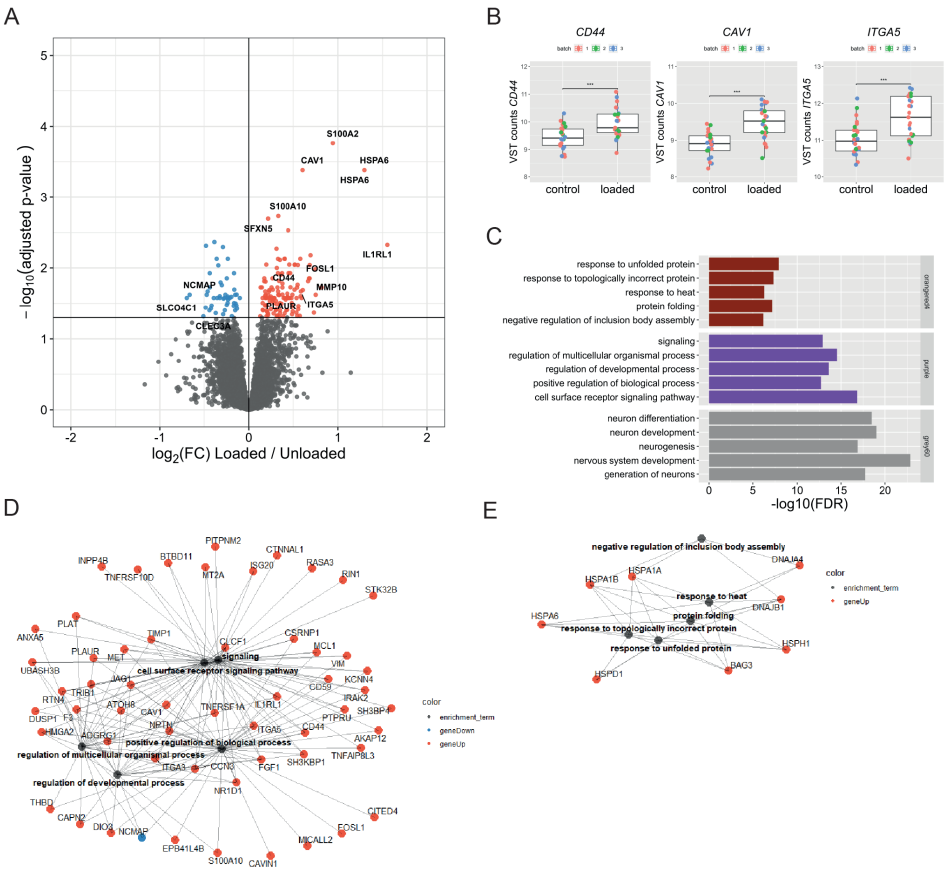


Figure 5 | Transcriptomic profile in response to hyper-physiologic mechanical loading conditions. (A) Volcano plot of differentially expressed genes (DEGs) in response to hyper-physiologic loading conditions. (n=25-26) Red dots denote differentially DEGs with an FDR<0.05 that are upregulated, and blue dots represent DEGs that are downregulated as determined by DESeq2 analysis. **(B)** Notable examples of mechano-sensor genes upregulated in response to hyper-physiological mechanical loading conditions. The box plots represent the 25th, 50th, and 75th percentiles, and the whiskers extend to 1.5 times the interquartile range. Individual samples are depicted by black dots in each graph. ***FDR<0.001 **(C)** Overrepresentation enrichment analysis (KEGG, REACTOME, GO biological processes) of the top 3 weighted gene co-expression network analysis (WGCNA) co-expression modules of which the first principal component is significantly associated with hyper-physiologic mechanical loading conditions. **(D-E)** Gene-pathway network of the enrichment analysis of the purple **(D)** and the orangered4 **€** co-expression network where lines depict the relationship between the genes and the pathways determined by enrichment analysis. Blue dots depict downregulated DEGs in response to hyper-physiologic mechanical loading conditions, red dots depict upregulated DEGs in response to hyper-physiologic mechanical loading conditions. Statistics:A,B - negative binomial generalized linear model, C,D,E – Fisher exact test.

Effect of mechanical loading on the transcriptomic landscape

Next, we characterized the effect of hyper-physiologic loading conditions as previously defined (*32*) on these neo-cartilage organoids (**Fig. 4A**). Hereto, two different organoid models were applied and jointly analyzed; cylindrical constructs in which hiPSC-derived pellets were digested to obtain single-cell chondrocytes that were then

encapsulated in an agarose gel, and spherical constructs with neo-cartilage deposited by hiPSC-derived chondrocytes. Both these organoid models were exposed to hyper-physiologic mechanical loading conditions (20% sinusoidal peak-to-peak strain at 5hz for 10 minutes) cartilage organoids were harvested at 12 hours post loading. As shown in **Figure 4B** and as measured by of Safranin-O staining, hyper-physiological loading resulted in a moderately reduction of proteoglycan content whereas in the *COL6A3* mutated organoids large reduction was observed. These observations were however not confirmed by qPCR data (**Fig. 4B; table S7**). Induction of a post-traumatic response by hyper-physiologic loading conditions was confirmed by upregulation of *MMP13, ADAMTS5*, and *IGBP6* gene-expression **(Fig. 4E)**. This was in line with previous results in an *ex vivo* model measuring genome-wide expression after injurious mechanical loading(*17, 33, 34*). To obtain in depth insight into phenotypic changes of the chondrocyte in response to hyper-physiological loading conditions, transcriptome wide activity was subsequently measured by RNA sequencing. Joint multifactorial analyses of the two models revealed 177 DEGs (FDR<0.05) between unloaded and mechanically loaded organoids, of which 74% were upregulated and 26% were downregulated **(Fig.5A, table S10)**. Notable genes were, amongst others; *CD44*, *CAV1*, *ITGA5* **(Fig. 5B**), which are all genes encoding for mechano-sensors involved in OA pathophysiology. To determine the biological processes affected by hyper-physiologic mechanical loading WGCNA co-expression networks were associated with mechanical loading, revealing 5 significantly associated co-expression networks **(Fig. S4; table S6)**. The top 3 most significantly associated modules were enriched for developmental and signaling processes, stress responses**,** and neuronal pathways **(Fig. 5C)**. While enrichment of signaling responses containing genes such as *IL1RL1* and *FOSL1* show an adaptive response to hyper-physiologic mechanical loading conditions **(Fig. 5D)**, enrichment of stress responses with genes such as *HSPA1B* and *DNAJA4* underlines the damaging effects of the hyper-physiologic mechanical loading conditions **(Fig 5E).**

Effect of the *COL6A3* missense variant on the response to hyper-physiological mechanical loading conditions

Next, the effects of the damaging *COL6A3* variant on the response to hyper-physiological mechanical loading were investigated. To this end, a multifactorial analysis was performed resulting in a set of 135 genes with a significant interaction effect (P<0.01) indicating that the variant affected the response to hyper-physiological mechanical loading **(table S11)**. Of these 135 genes, 70 proteins show a significant protein-protein interaction (PPI) (FDR<0.05) as determined by STRING-DB **(Fig. 6A)**. Notable is that highly connected genes in this PPI network such as; *PTGS2*, *IL1R1, IRAK2, PECAM1,* and *ADAMTS5* are all related to catabolic and inflammatory signaling. *PTGS2* is known to be regulated by TRPV4 signaling in response to mechanical stress and is involved

in inducing an inflammatory response to stimuli (*35*). *IL1R1* codes for the receptor of interleukin-1, one of the key inflammatory markers in OA, and thus is involved in inflammatory signaling, while *IRAK2* encodes for the interleukin-1 receptor-associated kinase 2, which is involved in interleukin-1 (IL-1) induced upregulation of NF-κβ

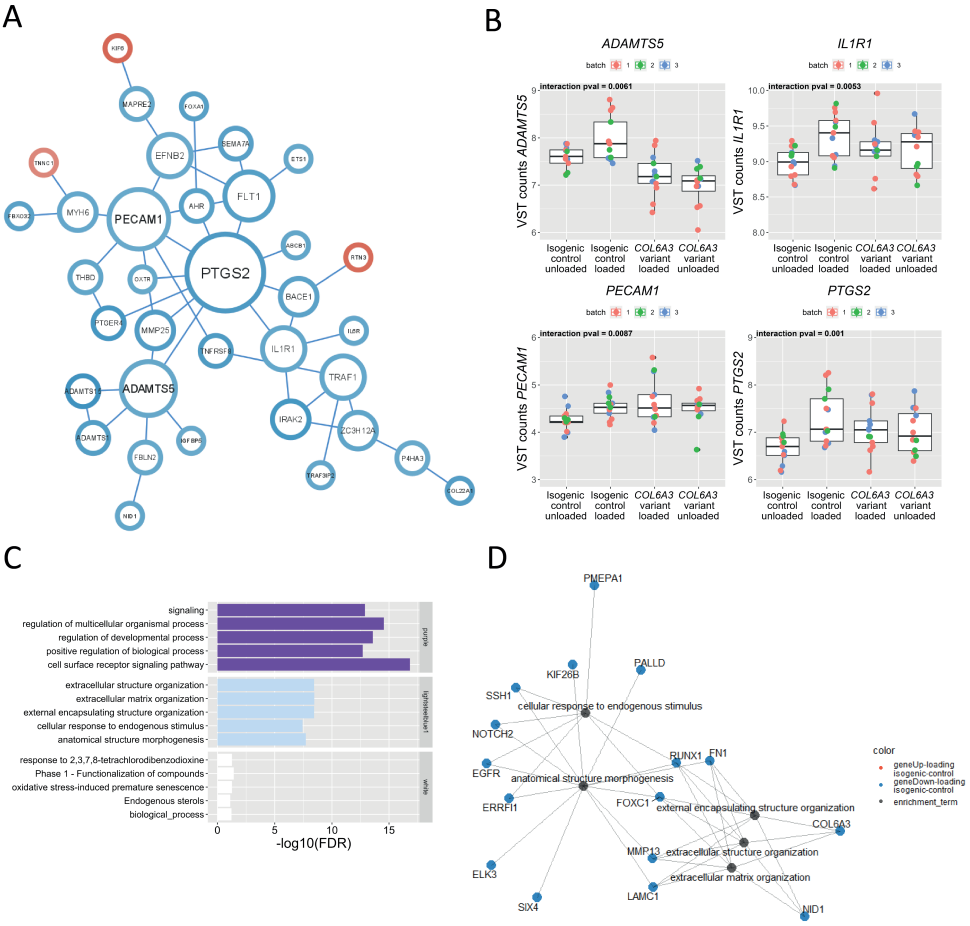


Figure 6 | The *COL6A3* variant affects the biological response to hyper-physiologic loading conditions. **(A)** Protein-protein interaction (PPI) network based on the STRING-DB of genes that show an interaction (P<0.01) between the *COL6A3* variant and hyper-physiological mechanical loading conditions, as determined by DESeq2 analysis. Red circles denote an increased response to hyper-physiological mechanical loading due to the *COL6A3* variant, while blue circles denote a reduced response to mechanical stress due to the *COL6A3* variant (only showing connected nodes). Relative node size depicts number of connections for each gene within the network (n=25-26) **(B)** Examples of central DEGs in the PPI that are related to an inflammatory response. The box plots represent 25th, 50th, and 75th percentiles, and whiskers extend to 1.5 times the interquartile range. Individual samples are depicted by black dots in each graph. **(C)** Gene enrichment analysis of WGCNA hubs significantly related to the interaction between the *COL6A3* variant and the response to hyper-physiological mechanical loading conditions. **(D)** Pathway – gene network of the enrichment analysis of the purple cluster. Lines depict the relationship between the genes and the pathways determined by enrichment analysis. Blue and red dots depict, respectively, downregulated and upregulated DEGs in response to hyper-physiological loading conditions in the isogenic control neo-cartilage organoids. Statistics: **A** – log-likelihood score for PPI, A,B - negative binomial generalized linear model for differential expression analysis, **C,D** – Fisher exact test.

(*36*). Upon performing stratified analysis, it was shown that the interaction effect was particularly caused by an upregulation of gene expression observed in isogenic controls in response to hyper-physiological loading conditions, that appeared absent in the *COL6A3* variant neo-cartilage organoids **(Fig. 6B; table S12).**
To study the biological processes in which the response to hyper-physiological loading differed between the isogenic control and the *COL6A3* variants, associate this interaction effect with the principal component of each detected co-expression network. This revealed three distinct co-expression networks associated with the interaction effect, which were enriched for processes related to signaling, ECM organization, and general biological processes **(Fig. 6C)**. Enrichment of pathways related to signaling shows an aberrant response in genes such as *INHBA* which acts downstream of TRPV4 calcium channels, *CAV1*, *SEMA7A*, and *MMP3,* which are all genes that respond to mechanical loading in isogenic controls, with an absent loading effect in the *COL6A3* variant **(table S5)**. Again, the enrichment of processes related to ECM formation and organization, containing genes such as; *MMP13, RUNX1, FN1, LAMC1, EGFR,* and *PMEPA1* **(Fig. 6D)**, underlines the effect of an impaired repair response to hyper-physiological mechanical loading conditions in the *COL6A3* variant. Also, the enrichment of general biological processes including *IL1R1* highlights the aberrant response to mechanical loading in the *COL6A3* variant neo-cartilage organoids.

Finally, we studied longer term effect of the response to hyper-physiological mechanical loading conditions in the *COL6A3* variant compared to the isogenic controls. To this end, gene expression analysis and safranin-o staining were performed 12 hours and 4 days after mechanical loading in both models (**Fig. 4)**. As shown in **Figure S8A** the transient inflammatory response appeared in remission in isogenic controls and was still absent in the *COL6A3* mutated chondrocytes at day 4 after hyper-physiological mechanical loading.as. Furthermore, as shown in **Figure. S8B-C** a consistent pattern in the proteoglycan content, as measured by safranin-o staining is shown in both models. More specifically, in the isogenic controls, the proteoglycan content 12 hours after loading dropped and was restored at days 4 after loading to levels observed in the unloaded isogenic controls. In the *COL6A3* variant, however, , a strong reduction of proteoglycan content was observed at 12 hours after hyper physiological loading that was not fully restored at day 4. These observations were confirmed by *ACAN* gene expression in the cylindrical model, in contrast to the spherical organoid model. (**Fig S8D).** Regarding the ongoing anabolic responses, as measured by *COL2A1* expression, it was shown in the isogenic control that the initial reparative anabolic response 12 hrs after hyper physiological loading has turned to unloaded 'steady state' levels 4 days after hyper physiological mechanical loading. In the *COL6A3* mutated chondrocytes, however, it was shown in both models that 4 days after loading *COL2A1* expression is

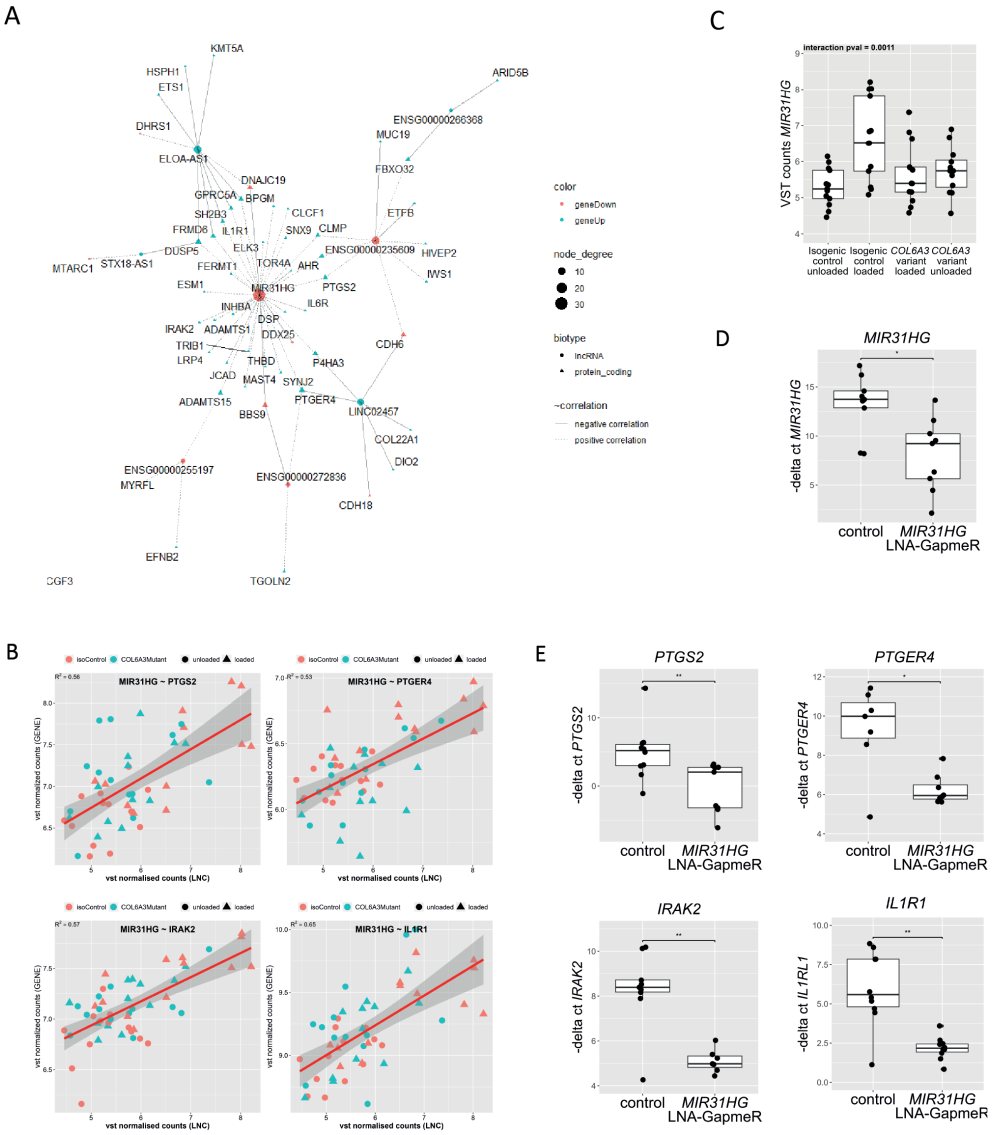


Figure 7 | Druggable target discovery of *MIR31HG* as a regulator of the inflammatory response to hyper-physiological mechanical loading (A) *MIR31HG* as a central hub in a lncRNA – protein coding correlation network (FDR<0.05, r^2 >0.5) of genes showing a significant interaction between hyper-physiological mechanical loading and the *COL6A3* variant. **(B)** Individual gene expression plot of VST normalized gene expression plot of *MIR31HG* showing interaction effect between mechanical loading and the *COL6A3* variant. (n=26) **(C)** Correlation plots of VST normalized counts of protein coding genes and MIR31HG. (n=13 per condition) **(D)** Consistent downregulation of *MIR31HG* by LNA-GapmeR (FC=0.02, P=6.4X10^{-3}, n=10) **(E)** *MIR31HG* inhibition using a LNA-GapmeR consistently downregulates *PTGS2* (FC=0.03, P=9.14X10^{-3}), *IRAK2* (FC=0.11, P=4.44X10^{-4}), *PTGER4* (FC=0.12, P=9.32X10^{-3}), *IL1R1* (FC=0.07, P=5.71X10^{-3}) (n=10). Statistics: A, B – spearman's rank correlation, C, negative binomial generalized linear model, D, E – generalized linear model. The box plots represent the 25th, 50th, and 75th percentiles, and the whiskers extend to 1.5 times the interquartile range. Individual samples are depicted by black dots in each graph. *P<0.05, **P<0.01, ***P<0.001.

still increased, highlighting an activated chondrocyte state (**Fig. S8D**). Together, this data suggests that after hyper physiological mechanical cues the *COL6A3* neo-cartilage organoids are subject to long-term differential effects as compared to the isogenic controls.

Epigenetic regulation by *MIR31HG* as a driver of the interacting effect between the *COL6A3* missense variant and the hyper-physiological mechanical loading induced stress response

Finally, we aimed to find key regulators behind the interacting effect of the *COL6A3* variant and the hyper-physiological mechanical loading-induced stress response. We chose to integrate long-non-coding (lnc)RNAs as epigenetic regulatory elements, as these previously have been shown to regulate gene expression in OA pathology (*37*). To this end, expression analysis of lncRNAs as annotated by the GENCODE V41 (*38*) was performed. Multifactorial analysis resulted in a significant interaction effect between the *COL6A3* variant and the response to hyper-physiological mechanical loading in a set of 17 lncRNAs (P<0.01) **(table S13).** Notable examples are; *LINC00574* as well as *MIR31HG,* which were previously reported as being higher expressed in lesioned cartilage compared to preserved cartilage from OA patients who underwent joint replacement surgery (*37*). To screen for potential regulatory effects of lncRNAs on protein-coding gene expression, we integrated the lncRNA data with the protein-coding RNA expression data. This resulted in a correlation network of 12 unique lncRNAs with 55 unique protein-coding genes for a total of 72 significant correlations (r^2>0.5, FDR<0.05) **(Fig. 7A)**. Of particular interest was *MIR31HG* as it appears as a central interacting node in the correlation network, accounting for 32 out of the 72 significant correlations, interacting with stress response and inflammation related genes such as *PTGS2, PTGER4, IRAK2,* and *IL1R1***(Fig. 7B)**. Genomically, *MIR31HG* is located near *IFNA8*, *IFNA1* and *IFNE*. and potentially directly interacts with its inflammatory signaling. Additionally, *MIR31HG* strongly responds to mechanical loading in isogenic controls, whereas this response is absent in *COL6A3* variants. Hence, given *MIR31HG* is a central interaction node in the correlation network **(Fig. 7A-B)** and its lack of response to mechanical loading in *COL6A3* variants **(Fig 7C)** we hypothesized that it regulates these inflammatory genes. Thus, finally, we validated this regulatory role of *MIR31HG* in the expression of these genes by transfection of *MIR31HG* targeting LNA-GapmeR in primary chondrocytes. As shown in figure 7D this resulted in a significant knock-down of *MIR31HG* expression in comparison to a nontargeting LNA GapmeR. Subsequently, we have measured the expression of stress response and inflammation-related genes in response to *MIR31HG* targeting LNA-GapmeRs, which showed significant downregulation of *PTGS2, PTGER4, IRAK2,* and *IL1R1* **(Fig. 7E)**. Together these results provide experimental validation that *MIR31HG* is a key regulator of the initial stress response to hyper-physiologic

mechanical loading and that this regulatory action of *MIR31HG* in the damaging *COL6A3* variant neo-cartilage organoids is absent.

Discussion

In the current study, we identified a missense variant in *COL6A3* (c.4510C>T, R1504W) in a subject of the GARP study affected with symptomatic OA in two or more joint sites (*9*). By introducing this damaging variant in hiPSCs using CRISPR-Cas9 genome engineering and employing these cells in two different established 3D *in vitro* neo-cartilage organoid models, we showed that the variant decreases cartilage matrix integrity, as reflected by a reduction in abundance and size of sGAGs. Moreover, by subsequently isolating mutated COLVI protein, we showed that it had reduced binding to FN. Being both important and particular proteins of the chondrocyte PCM, the reduced binding is likely affecting PCM function. Analysis of the transcriptome-wide gene expression changes with the *COL6A3* variant, showed overlap to those observed with OA-pathophysiology. Together these data indicated that the *COL6A3* variant is likely affecting the propensity of chondrocytes to enter an osteoarthritic state, secondary to its effect on PCM function. By subsequently exposing the neo-cartilage organoids to hyper-physiological mechanical stress, we demonstrated that *COL6A3* variant in chondrocytes abolished the characteristic upregulation of inflammatory signaling after mechanical loading (*17, 39, 40*) with *PTGS2*, *PECAM1*, and *ADAMTS5*, as most central genes. We also showed that an aberrant chondrocyte phenotype in the *COL6A3* variant persisted at least 4 days. Finally, by integrating epigenetic regulation of protein coding gene expression we identified lncRNA *MIR31HG* as a key regulator of the characteristic inflammatory signaling observed in response to mechanical loading, a response that was abolished in the *COL6A3* variant. Taken together, our findings suggest that the identified variant in *COL6A3* resulted in impaired binding between COLVI and the PCM protein FN. Secondary to these alterations in PCM function the initial stress response to hyper-physiologic mechanical loading conditions was abolished that in turn affected the propensity of the chondrocyte to enter an OA disease state.

By integration of lncRNA's driven epigenetic regulation of protein-coding genes, we were able to identify *MIR31HG* as a key regulator in the aberrant response to mechano-transduction in the *COL6A3* variants. *MIR31HG* expression has previously been reported in OA pathophysiology (*37*). Mechanistically, in other cell types, *MIR31HG* has been shown to activate the expression of similar inflammatory genes such as *IL1R1*, *PTGS2*, and *TNFRSF11B (41)*, but also genes that have previously been linked to damaging loading in cartilage such as *IGFBP5* and *IGFBP7*, via promoting phosphorylation of *YBX*1, thereby initiating translation of *IL1A.* This further underlines the relevance of *MIR31HG* for osteoarthritis. We would like to postulate that the loss of the initial response of

MIR31HG to mechanical loading, and hence, the loss of the transient inflammatory response as marked by, amongst others, *PTGS2* and *IL1R1,* inhibit subsequent repair of cartilage. By studying a damaging *COL6A3* variant in two hiPSC-derived neo-cartilage organoid models, we have been able to study in detail the downstream chondrocyte mechano-biologic effects to hyper-physiological mechanical stress. Nonetheless it should be noted that upon checking the heterozygous introduction of the damaging *COL6A3* variant by genotyping the RNA sequencing data, an allelic imbalance of the *COL6A3* variant relative to the reference allele was observed. This predominance of *COL6A3* variant expression could indicate increased stability of the *COL6A3* variant mRNA allele or alternatively could be caused due to an unrecognized mixed population of homo- and heterozygous hiPSC clones. Irrespectively, the allelic imbalanced expression of the *COL6A3* variant confirmed a dose-response effect of the presence *COL6A3* variant in 74% of identified DEGs including notable genes such as *SPP1*, *MMP9*, and *PGR4*. Henceforth, we believe that such a dose response effect adds to the validity of downstream effects of the *COL6A3* variant.

Our previous work on *Col6a1−/−* mice reported an increased progression of OA pathophysiology, secondary to impaired PCM function. The latter is reflected by reduced PCM stiffness, increased cell swelling, and altered calcium response to osmotic stress (*5, 8*). Following the altered mechano-transduction associated with aberrant COLVI function, we here focused particularly on changes in the cellular phenotype of chondrocytes governed by the damaging *COL6A3* R1504W variant in interaction with hyper-physiological mechanical cues via transcriptome-wide (co)-expression and biological pathway analyses.

By combining and jointly analyzing hiPSC derived chondrocytes in two models namely embedded in cylindrical disc-shaped agarose organoids (*39*) and spherical neo-cartilage pellets (*13, 17*) we aimed to obtain most consistent and robust results while taking advantage of the specific aspects of either model. The advantage of the first model is that it ensured equal distribution of mechanical stress throughout the disc-shaped sample. The advantage of the second model is that it allows harnessing the original matrix deposited by the hiPSC-derived chondrocytes thereby permitting the characterization of changes in the ECM. Despite the fact that we addressed reproducibility by combining and jointly analyzing data of two cartilage organoid models that were performed in two different laboratories with multiple differentiations, a weakness of our study is that the *COL6A3* R1504W mutation was introduced in only one hiPSC line. However, *in silico* off-target analysis showed only five potential, but highly unlikely, off target sites. Moreover, Sanger sequencing confirmed that these were not affected by the CRISPR/Cas9 gene editing.

By isolating mutated COLVI protein from the neo-cartilage organoids we could show that it had reduced binding to fibronectin. Based on this finding it is tempting to hypothesize that this reduced binding is underlying the observed lack of the inflammatory response to hyper-physiological loading. In contrast to previous findings (*11*), we could not detect any binding of COLVI with hyaluronan, which might be explained by the source of the COLVI proteins, as they were extracted from the culture medium. Alternatively, other intermediary proteins are necessary for binding COLVI to hyaluronan which might not have been present in the medium.

We have used supervised machine learning with a recently developed convolutional neural network to quantify the differences in both sGAG size and numbers between isogenic controls and *COL6A3* variants (*30, 42*). This procedure relied on initial manual annotation of sGAG structures for supervised classification because the staining protocols used for TEM (a protocol including osmium tetroxide and potassium ferrocyanide, uranyl acetate, and lead citrate) are rather nonspecific and stain almost any cellular structure. Nonetheless, given that our annotated sGAG structures are very similar to previously annotated sGAGs in rat and mouse tissue (*43, 44*), our TEM results are in line with the results of the DMMB assay, and the Alcian blue staining, we are confident that the applied supervised machine learning approach has reliably annotated sGAG structures. Nonetheless, more specific staining that enhances contrast e.g. carbon double bonds (Osmium) or charged structures (Ruthenium salts) could have reduced our initial manual annotation effort and increased sensitivity. Together, our data implies that the reduced sGAG found with the *COL6A3* variant is likely to be explained by a chondrocyte phenotype marked by increased expression of MMPs and reduced expression of *ACAN.* Next to a lacking initial upregulation of inflammatory signaling after mechanical loading, the aberrant phenotype of *COL6A3* chondrocytes persisted also 4 days after the hyper-physiological mechanical loading regime. We can, however, not exclude the involvement of other collagen type VI-mediated changes in the PCM function that could result in sGAGs reduction. Notable in this respect is the recently reported COLVI-mediated role of Decorin in aggrecan retention to the PCM.(*45*)
A limitation of the current study is that we have initially relied on the predicted damaging effects of the identified R1504W variant on COLVI protein function by *in silico* tools (SIFT, Polyphen, Provean, i-mutant V2.0, MUpro, and PANTHER-PSEP) for prioritization on the damaging effect on protein function. Because the damaging effect was predicted with high confidence, we studied the effect of the variant on the chondrocyte phenotypic state using human *in vitro* cartilage organoid models. Due to ethical constraints, we did not address the clinical genotype-phenotype relationship of carriers as this requires an in-depth study of the (extended) pedigree, penetrance, age of onset, and focused phenotyping.

On a different note, applying human *in vitro* cartilage organoid models has refrained us from studying the effect of the R1504W *COL6A3* variant more physiological loading regimes, that could exert additional effects on the genotype – phenotype relationship. The latter because aberrant responses to physiological cues are likely subtle, that would require repetitive and long-term loading. This was outside the scope of our study, as we aimed to investigate the short- and long-term effects of acute, and injurious hyper-physiological loading as outlined by previously(*46*) and as marked by induction of pro-inflammatory mediators. Such acute injurious loading is generally accepted as a major cause for onset of post-traumatic OA. Additionally, long-term cultures of our cartilage organoids could evoke confounding factors such as dedifferentiation of the chondrocyte phenotype, or cell death that would refrain robust data generation. Additionally, an *in vivo* mouse model would likely be eligible to study the effect of such physiological loading regimes and that is currently under investigation.

In characterizing the chondrocyte phenotype associated with the R1504W *COL6A3* variant, we have determined the overlapping transcriptome-wide profile with that of OA chondrocytes. We must note that, while significant in implication, this overlap was only limited. This is likely explained by the fact that the transcriptome-wide differences between lesioned and preserved OA cartilage particularly describe the overall OA-associated chondrocyte phenotype in the general OA population which is likely different from the chondrocyte phenotype due to a specific variant in *COL6A3*. Nonetheless, we would like to argue that the significance of the DEG overlap between the *COL6A3* variant, and OA lies in the underlying pathways and biological functions of the genes involved such as ECM components (*COL27A1*, *COL11A2)*, and ossification (*SPP1*, *MGP*) as they describe critical aspects of the OA chondrocyte phenotype.

Together, using genetic engineering of hiPSC-derived neo-cartilage organoid models while implementing hyper-physiological mechanical loading conditions, we established a tailored model to study the biological function of proteins in the transduction of mechanical cues from ECM to chondrocytes, while complying with the societal wish to reduce animal models. By using this model, we showed that the COLVI protein variant had reduced binding to fibronectin, and we advocated that the observed lack of the initial inflammatory response to hyper-physiological loading is likely secondary to this aberrant function of COLVI in the PCM. Additionally, we demonstrated that the initial inflammatory response to hyper-physiological loading is particularly epigenetically regulated by the lncRNA *MIR31HG*.

Acknowledgments

Funding

Leiden University Medical Centre, Pfizer Groton, Connecticut, USA and the Dutch Arthritis Society have supported the GARP study. Furthermore, the research leading to these results has received funding from the Biobanking and BioMolecular resources Resvvearch Infrastructure the Netherlands (BBMRI-NL) (complementation project CP2013-84), the Dutch Scientific Research council NWO/ZonMW VICI scheme (nr. 91816631/528) and the Dutch Arthritis Society. We are indebted to drs. N. Riyazi, J. Bijsterbosch, H.M. Kroon and I. Watt for collection of data. Furthermore, this work was supported by the Shriners Hospitals for Children and the NIH (grants AG46927, AG15768, AR080902, AR072999, AR073752, and AR074992). Additionally, the data is generated within the scope of the Medical Delta programs Regenerative Medicine 4D: Generating complex tissues with stem cells and printing technology and Improving Mobility with Technology. This publication is part of the LoaD project (NWA.1389.20.009) of the NWA-research program Research by Consortia (ORC) that is financed by the Netherlands Organisation of Scientific Research (NWO).

Author Contributions

N.G.C.B., Z.H., A.R.D., Y.F.M.H, F.G., I.M. developed the concept of this study
N.G.C.B., Z.H., G.M., S.S.A., A.R.D., G.H., N.S. M.K., R.I.K., A.A.M., B.B.R.K. and F.G., acquired materials and data
N.G.C.B., R.C.D.A., B.B.R.K., Y.F.M.H, F.G. and I.M. analyzed the data
all authors contributed to the writing of the manuscript.

Competing interests

None

Data and materials availability

All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

References

1. F. Guilak, Biomechanical factors in osteoarthritis. *Best Pract Res Clin Rheumatol* **25**, 815-823 (2011).
2. C. Juhl, R. Christensen, E. M. Roos, W. Zhang, H. Lund, Impact of exercise type and dose on pain and disability in knee osteoarthritis: a systematic review and meta-regression analysis of randomized controlled trials. *Arthritis Rheumatol* **66**, 622-636 (2014).
3. O. A. Uthman, D. A. van der Windt, J. L. Jordan, K. S. Dziedzic, E. L. Healey, G. M. Peat, N. E. Foster, Exercise for lower limb osteoarthritis: systematic review incorporating trial sequential analysis and network meta-analysis. *BMJ* **347**, f5555 (2013).
4. S. A. Olson, B. D. Furman, V. B. Kraus, J. L. Huebner, F. Guilak, Therapeutic opportunities to prevent post-traumatic arthritis: Lessons from the natural history of arthritis after articular fracture. *J Orthop Res* **33**, 1266-1277 (2015).
5. N. A. Zelenski, H. A. Leddy, J. Sanchez-Adams, J. Zhang, P. Bonaldo, W. Liedtke, F. Guilak, Type VI Collagen Regulates Pericellular Matrix Properties, Chondrocyte Swelling, and Mechanotransduction in Mouse Articular Cartilage. *Arthritis Rheumatol* **67**, 1286-1294 (2015).
6. F. Guilak, L. G. Alexopoulos, M. L. Upton, I. Youn, J. B. Choi, L. Cao, L. A. Setton, M. A. Haider, The pericellular matrix as a transducer of biomechanical and biochemical signals in articular cartilage. *Ann N Y Acad Sci* **1068**, 498-512 (2006).
7. F. Guilak, R. J. Nims, A. Dicks, C. L. Wu, I. Meulenbelt, Osteoarthritis as a disease of the cartilage pericellular matrix. *Matrix Biol* **71-72**, 40-50 (2018).
8. L. G. Alexopoulos, I. Youn, P. Bonaldo, F. Guilak, Developmental and osteoarthritic changes in Col6a1-knockout mice: biomechanics of type VI collagen in the cartilage pericellular matrix. *Arthritis Rheum* **60**, 771-779 (2009).
9. N. Riyazi, F. R. Rosendaal, E. Slagboom, H. M. Kroon, F. C. Breedveld, M. Kloppenburg, Risk factors in familial osteoarthritis: the GARP sibling study. *Osteoarthritis Cartilage* **16**, 654-659 (2008).
10. N. Beecher, A. M. Roseman, T. A. Jowitt, R. Berry, H. Troilo, R. A. Kammerer, C. A. Shuttleworth, C. M. Kielty, C. Baldock, Collagen VI, conformation of A-domain arrays and microfibril architecture. *J Biol Chem* **286**, 40266-40275 (2011).
11. C. Wiberg, A. R. Klatt, R. Wagener, M. Paulsson, J. F. Bateman, D. Heinegard, M. Morgelin, Complexes of matrilin-1 and biglycan or decorin connect collagen VI microfibrils to both collagen II and aggrecan. *J Biol Chem* **278**, 37698-37704 (2003).
12. U. Specks, U. Mayer, R. Nischt, T. Spissinger, K. Mann, R. Timpl, J. Engel, M. L. Chu, Structure of recombinant N-terminal globule of type VI collagen alpha 3 chain and its binding to heparin and hyaluronan. *EMBO J* **11**, 4281-4290 (1992).
13. M. van Hoolwerff, A. Rodriguez Ruiz, M. Bouma, H. E. D. Suchiman, R. I. Koning, C. R. Jost, A. A. Mulder, C. Freund, F. Guilak, Y. F. M. Ramos, I. Meulenbelt, High-impact FN1 mutation decreases chondrogenic potential and affects cartilage deposition via decreased binding to collagen type II. *Sci Adv* **7**, eabg8583 (2021).
14. S. S. Adkar, C. L. Wu, V. P. Willard, A. Dicks, A. Ettyreddy, N. Steward, N. Bhutani, C. A. Gersbach, F. Guilak, Step-Wise Chondrogenesis of Human Induced Pluripotent Stem Cells and Purification Via a Reporter Allele Generated by CRISPR-Cas9 Genome Editing. *Stem Cells* **37**, 65-76 (2019).
15. J. Lee, S. E. Taylor, P. Smeriglio, J. Lai, W. J. Maloney, F. Yang, N. Bhutani, Early induction of a prechondrogenic population allows efficient generation of stable chondrocytes from human induced pluripotent stem cells. *FASEB J* **29**, 3399-3410 (2015).
16. A. Rodriguez Ruiz, A. Dicks, M. Tuerlings, K. Schepers, M. van Pel, R. Nelissen, C. Freund, C. L. Mummery, V. Orlova, F. Guilak, I. Meulenbelt, Y. F. M. Ramos, Cartilage from human-induced pluripotent stem cells: comparison with neo-cartilage from chondrocytes and bone marrow mesenchymal stromal cells. *Cell Tissue Res* **386**, 309-320 (2021).
17. R. G. M. Timmermans, N. G. C. Bloks, M. Tuerlings, M. van Hoolwerff, R. Nelissen, R. J. P. van der Wal, P. M. van der Kraan, A. B. Blom, M. H. J. van den Bosch, Y. F. M. Ramos, I. Meulenbelt, A human in vitro 3D neo-cartilage model to explore the response of OA risk genes to hyper-physiological mechanical stress. *Osteoarthr Cartil Open* **4**, 100231 (2022).
18. R. W. Farndale, C. A. Sayers, A. J. Barrett, A Direct Spectrophotometric Microassay for Sulfated Glycosaminoglycans in Cartilage Cultures. *Connective Tissue Research* **9**, 247-248 (1982).
19. M. Tuerlings, M. van Hoolwerff, J. M. van Bokkum, H. E. D. Suchiman, N. Lakenberg, D. Broekhuis, R. Nelissen, Y. F. M. Ramos, H. Mei, D. Cats, R. Coutinho de Almeida, I. Meulenbelt, Long non-coding RNA expression profiling of subchondral bone reveals AC005165.1 modifying FRZB expression during osteoarthritis. *Rheumatology (Oxford)* **61**, 3023-3032 (2022).
20. A. Kiezun, S. L. Pulit, L. C. Francioli, F. van Dijk, M. Swertz, D. I. Boomsma, C. M. van Duijn, P. E. Slagboom, G. J. van Ommen, C. Wijmenga, C. Genome of the Netherlands, P. I. de Bakker, S. R. Sunyaev, Deleterious alleles in the human genome are on average younger than neutral alleles of the same frequency. *PLoS Genet* **9**, e1003301 (2013).
21. R. Vaser, S. Adusumalli, S. N. Leng, M. Sikic, P. C. Ng, SIFT missense predictions for genomes. *Nat Protoc* **11**, 1-9 (2016).
22. N. L. Sim, P. Kumar, J. Hu, S. Henikoff, G. Schneider, P. C. Ng, SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res* **40**, W452-457 (2012).
23. Y. F. Ramos, W. den Hollander, J. V. Bovee, N. Bomer, R. van der Breggen, N. Lakenberg, J. C. Keurentjes, J. J. Goeman, P. E. Slagboom, R. G. Nelissen, S. D. Bos, I. Meulenbelt, Genes involved in the osteoarthritis process identified through genome wide expression analysis in articular cartilage; the RAAK study. *PLoS One* **9**, e103056 (2014).
24. R. Coutinho de Almeida, Y. F. M. Ramos, A. Mahfouz, W. den Hollander, N. Lakenberg, E. Houtman, M. van Hoolwerff, H. E. D. Suchiman, A. Rodriguez Ruiz, P. E. Slagboom, H. Mei, S. M. Kielbasa, R. Nelissen, M. Reinders, I. Meulenbelt, RNA sequencing data integration reveals an miRNA interactome of osteoarthritis cartilage. *Ann Rheum Dis* **78**, 270-277 (2019).
25. G. N. Welch, J. Loscalzo, Homocysteine and atherothrombosis. *N Engl J Med* **338**, 1042-1050 (1998).
26. D. A. Keesler, T. L. Slobodianuk, C. E. Kochelek, C. W. Skaer, S. L. Haberichter, V. H. Flood, Fibronectin binding to von Willebrand factor occurs via the A1 domain. *Res Pract Thromb Haemost* **5**, e12534 (2021).
27. I. Adzhubei, D. M. Jordan, S. R. Sunyaev, Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet* **Chapter 7**, Unit7 20 (2013).
28. H. Tang, P. D. Thomas, PANTHER-PSEP: predicting disease-causing genetic variants using position-specific evolutionary preservation. *Bioinformatics* **32**, 2230-2232 (2016).
29. J. P. Concordet, M. Haeussler, CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Res* **46**, W242-W245 (2018).
30. C. Xiao, X. Chen, W. Li, L. Li, L. Wang, Q. Xie, H. Han, Automatic Mitochondria Segmentation for EM Data Using a 3D Supervised Convolutional Network. *Front Neuroanat* **12**, 92 (2018).
31. J. T. Leek, svaseq: removing batch effects and other unwanted noise from sequencing data. *Nucleic Acids Res* **42**, e161 (2014).
32. R. G. M. Timmermans, N. G. C. Bloks, M. Tuerlings, M. van Hoolwerff, R. G. H. H. Nelissen, R. J. P. van der Wal, P. M. van der

Kraan, A. B. Blom, M. H. J. van den Bosch, Y. F. M. Ramos, I. Meulenbelt, A human in vitro 3D neo-cartilage model to explore the response of OA risk genes to hyper-physiological mechanical stress. *Osteoarthritis and Cartilage Open* **4**, 100231 (2022).
33. P. Occhetta, A. Mainardi, E. Votta, Q. Vallmajo-Martin, M. Ehrbar, I. Martin, A. Barbero, M. Rasponi, Hyperphysiological compression of articular cartilage induces an osteoarthritic phenotype in a cartilage-on-a-chip model. *Nat Biomed Eng* **3**, 545-557 (2019).
34. E. Houtman, M. Tuerlings, J. Riechelman, E. Suchiman, R. J. P. van der Wal, R. Nelissen, H. Mei, Y. F. M. Ramos, R. Coutinho de Almeida, I. Meulenbelt, Elucidating mechano-pathology of osteoarthritis: transcriptome-wide differences in mechanically stressed aged human cartilage explants. *Arthritis Res Ther* **23**, 215 (2021).
35. L. S. Simon, Role and regulation of cyclooxygenase-2 during inflammation. *The American journal of medicine* **106**, 37S-42S (1999).
36. F. Guo, Y. Li, S. J. I. Wu, Antisense IRAK-2 Oligonucleotide Blocks Il-1-Stimulated NF-κB Activation and ICAM-1 Expression in Cultured Endothelial Cells. *J Inflammation* **23**, 535-543 (1999).
37. M. van Hoolwerff, P. I. Metselaar, M. Tuerlings, H. E. D. Suchiman, N. Lakenberg, Y. F. M. Ramos, D. Cats, R. Nelissen, D. Broekhuis, H. Mei, R. C. de Almeida, I. Meulenbelt, Elucidating Epigenetic Regulation by Identifying Functional cis-Acting Long Noncoding RNAs and Their Targets in Osteoarthritic Articular Cartilage. *Arthritis Rheumatol* **72**, 1845-1854 (2020).
38. A. Frankish, M. Diekhans, A. M. Ferreira, R. Johnson, I. Jungreis, J. Loveland, J. M. Mudge, C. Sisu, J. Wright, J. Armstrong, I. Barnes, A. Berry, A. Bignell, S. Carbonell Sala, J. Chrast, F. Cunningham, T. Di Domenico, S. Donaldson, I. T. Fiddes, C. Garcia Giron, J. M. Gonzalez, T. Grego, M. Hardy, T. Hourlier, T. Hunt, O. G. Izuogu, J. Lagarde, F. J. Martin, L. Martinez, S. Mohanan, P. Muir, F. C. P. Navarro, A. Parker, B. Pei, F. Pozo, M. Ruffier, B. M. Schmitt, E. Stapleton, M. M. Suner, I. Sycheva, B. Uszczynska-Ratajczak, J. Xu, A. Yates, D. Zerbino, Y. Zhang, B. Aken, J. S. Choudhary, M. Gerstein, R. Guigo, T. J. P. Hubbard, M. Kellis, B. Paten, A. Reymond, M. L. Tress, P. Flicek, GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res* **47**, D766-D773 (2019).
39. R. J. Nims, L. Pferdehirt, N. B. Ho, A. Savadipour, J. Lorentz, S. Sohi, J. Kassab, A. K. Ross, C. J. O'Conor, W. B. Liedtke, B. Zhang, A. L. McNulty, F. Guilak, A synthetic mechanogenetic gene circuit for autonomous drug delivery in engineered tissues. *Sci Adv* **7**, eabd9858 (2021).
40. T. L. Vincent, Mechanoflammation in osteoarthritis pathogenesis. *Semin Arthritis Rheum* **49**, S36-S38 (2019).
41. M. Montes, M. Lubas, F. S. Arendrup, B. Mentz, N. Rohatgi, S. Tumas, L. M. Harder, A. J. Skanderup, J. S. Andersen, A. H. Lund, The long non-coding RNA MIR31HG regulates the senescence associated secretory phenotype. *Nat Commun* 12, 2459 (2021).
42. O. Dzyubachyk, R. I. Koning, A. A. Mulder, M. C. Avramut, F. G. Faas, A. J. Koster, in *Medical Imaging with Deep Learning*. (PMLR, 2021), pp. 148-157.
43. B. Coltoff-Schiller, S. Goldfischer, Glycosaminoglycans in the rat aorta. Ultrastructural localization with toluidine blue O and osmium--ferrocyanide procedure. *Am J Pathol* 105, 232-240 (1981).
44. K. M. Kwan, M. K. Pang, S. Zhou, S. K. Cowan, R. Y. Kong, T. Pfordte, B. R. Olsen, D. O. Sillence, P. P. Tam, K. S. Cheah, Abnormal compartmentalization of cartilage matrix components in mice lacking collagen X: implications for function. *J Cell Biol* 136, 459-471 (1997).
45. D. R. Chery, B. Han, Y. Zhou, C. Wang, S. M. Adams, P. Chandrasekaran, B. Kwok, S. J. Heo, M. Enomoto-Iwamoto, X. L. Lu, D. Kong, R. V. Iozzo, D. E. Birk, R. L. Mauck, L. Han, Decorin regulates cartilage pericellular matrix micromechanobiology. *Matrix Biol* 96, 1-17 (2021).
46. J. Sanchez-Adams, H. A. Leddy, A. L. McNulty, C. J. O'Conor, F. Guilak, The mechanobiology of articular cartilage: bearing the burden of osteoarthritis. *Curr Rheumatol Rep* 16, 451 (2014).

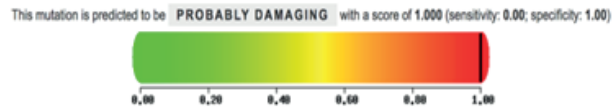
Supplementary Materials

Supplementary figures

A

substitution	preservation time	Message	Pdel
R1504W	361	possibly damaging	0.5

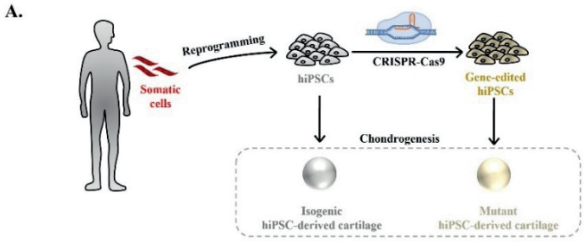
B

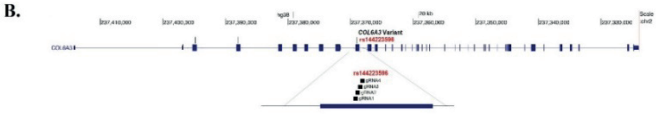


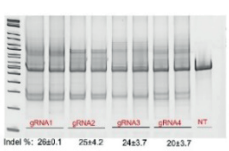
C

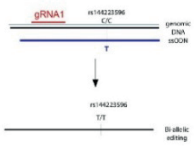
AMINO ACID CHANGE	GENE NAME	GENE ID	TRANSCRIPT ID	PROTEIN ID	SIFT PREDICTION
R1504W	COL6A3	ENSG00000163359	ENST00000295550	ENSP00000295550	DELETERIOUS

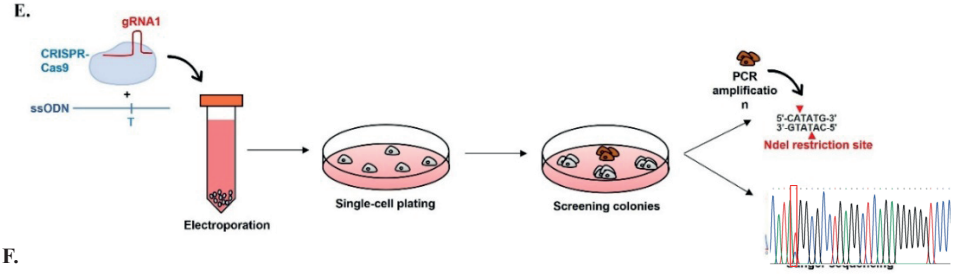
Figure S1 | In-situ predictions of damaging effects COL6A3 R1504W variant **(A)** Panther predicting a possibly damaging effect of the variant using evolutionary conservation **(B)** Polyphen 2 score of 1, which predicts a reduced stability and function of the protein following the impact of amino acid substitutions using structural protein modeling and evolutionary conservation **(C)** SIFT predicting a deleterious effect of the variant based on sequence homology to predict whether an amino acid substitution will have an adverse effect on protein function.





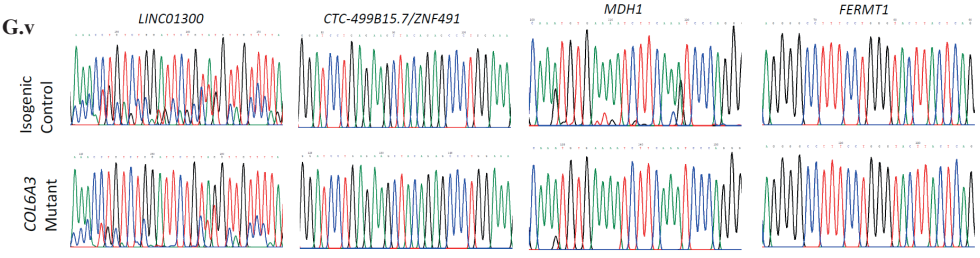






F.

OfftargetSeq	mismatchPos	mismatchCount	cfd Offtarget Score	Chr	start	end	strand	LocusDesc
TTTGAGAACCTTCAGATCCAAGG	.*...*.....*.....*	4	0.11	chr8	141342993	141343015	-	exon:LINC01300
CCTGAGAAGCTACAGAGCCCTGG	*....*..*.....*...	4	0.04	chr19	11797755	11797777	+	exon:CTC-499B15.7/ZNF491
TCTGAAAACCCACTGGACCCTGG	.....*..*.**...	4	0.04	chr20	6115964	6115986	-	exon:FERMT1
TCTGCAAAGCTTCAGATCCTAGG	....*...*..*.....*	4	0.03	chr19	9295432	9295454	-	exon:ZNF699/CTC-325H20.4
TGTGAAAATCTTCAAATCCCAGG	.*.....*..*..*.....	4	0.19	chr2	63597514	63597536	+	exon:MDH1



H.

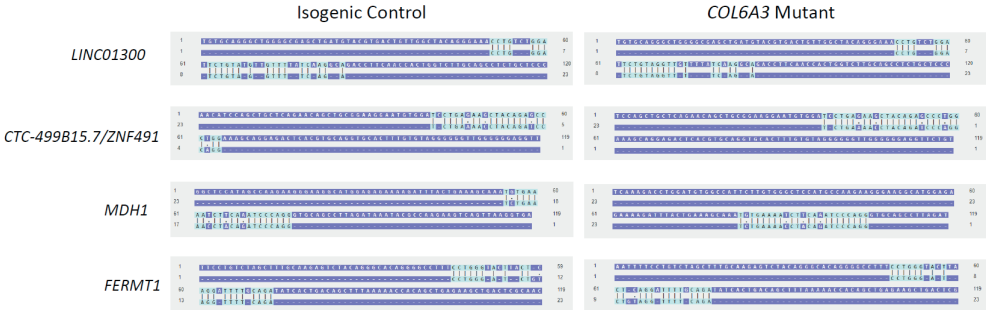


Figure S2 | Overview of experimental set-up CRISPR-Cas9 gene editing of rs144223596 C>T. (A) Schematic of hiPSCs editing OA variants. hiPSCs are gene-edited using CRIPSR/Cas9 to introduce OA disease variants found using exome sequencing studies. Gene-edited hiPSCs alongside their isogenic controls are then differentiated into chondrogenic lineage to obtain hiPSC-derived cartilage, which can be used in downstream analysis to determine the mechanisms linking the variant to OA. **(B)**. gRNAs were designed targeting genomic DNA flanking the COl6A3 variant rs144223596 (*http://genome.ucsc.edu*) *(34)* and **(C)**. Screened in HEK293T cells to evaluate cutting efficiency using Surveyor assay. **(D)**. Schematic of biallelic editing using optimized gRNA targeting the risk variant and ssODN harboring T risk allele. **(E)** Schematic of single-cell expansion into colonies following electroporation with gRNA complexed to Cas9 and ssODN. Colonies were screened by PCR amplification and digestion with an NdeI and evaluated using sanger sequencing **(F)** Potential off-target loci as determined by CRISPOR *in silico* screening. **(G)** Chromatogram of Sanger sequencing confirming no off-target gene editing **(H)** Sanger sequencing results of potential off-target loci confirming no off-target gene-editing. The second row in the figure displays the guide RNA and shows the mismatches with the target sequence.

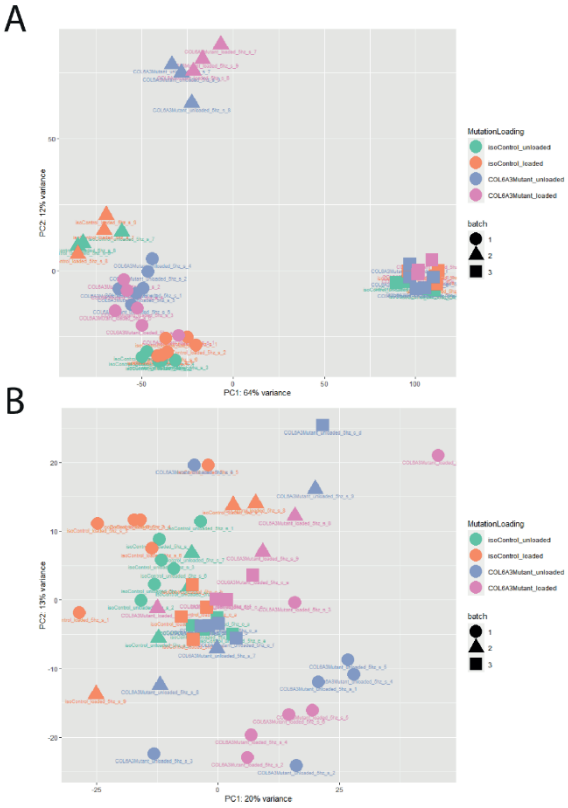


Figure S3 | Principal component analysis (PCA) plot showing the RNA-sequencing sample-to-sample distribution of the three independent differentiations (total N=53) **(A)** PCA plot of uncorrected counts data. **(B)** PCA plot after correction using surrogate variable analysis.

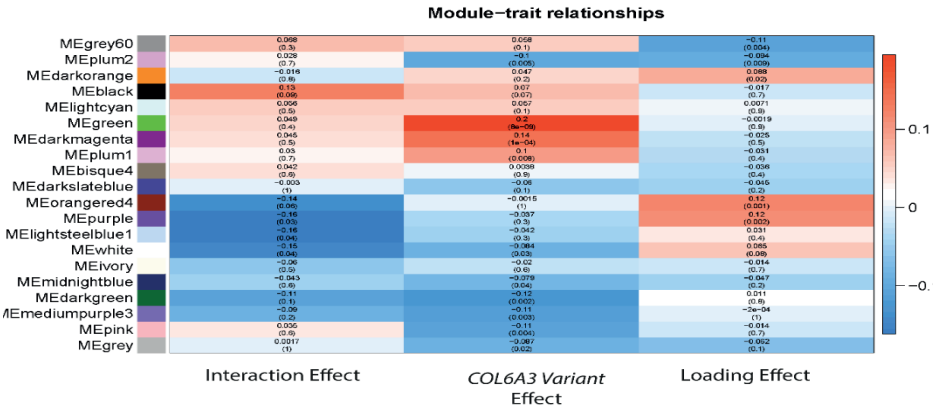


Figure S4 | Association of co-expression modules as determined by WGCNA with the *COL6A3* variant, hyperphysiologic mechanical loading conditions, and the interaction between the *COL6A3* variant and hyperphysiologic mechanical loading conditions. Data is noted as beta (p-value). Statistics: generalized linear model, total N=53.

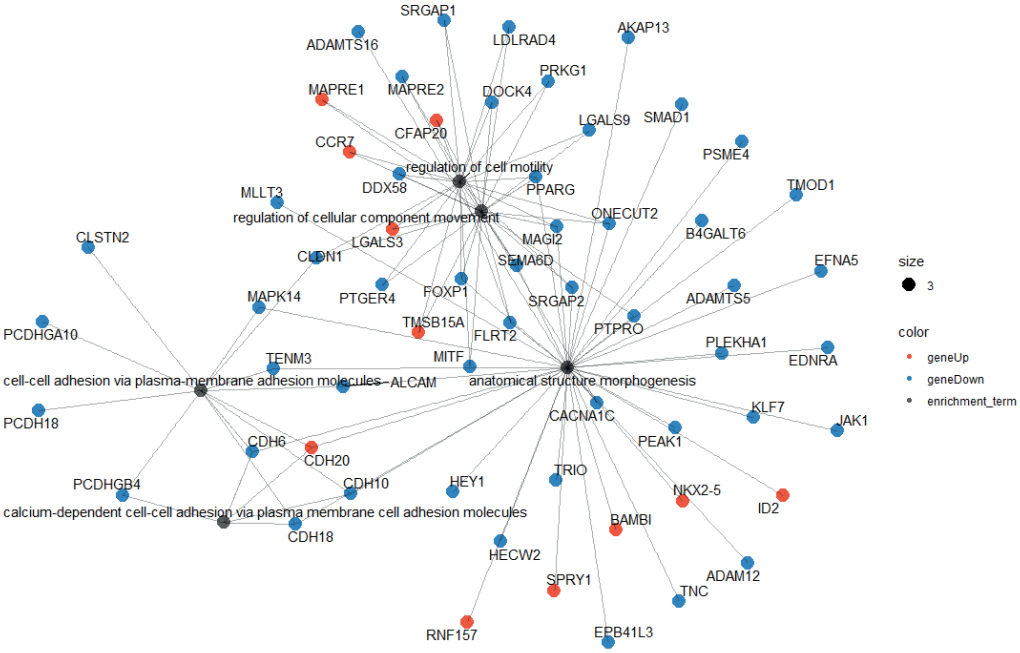


Figure S5 | Enrichment network of the darkgreen module associated with the *COL6A3* variant. (n=25-26). Statistics: Fisher Exact test for enrichment analysis.

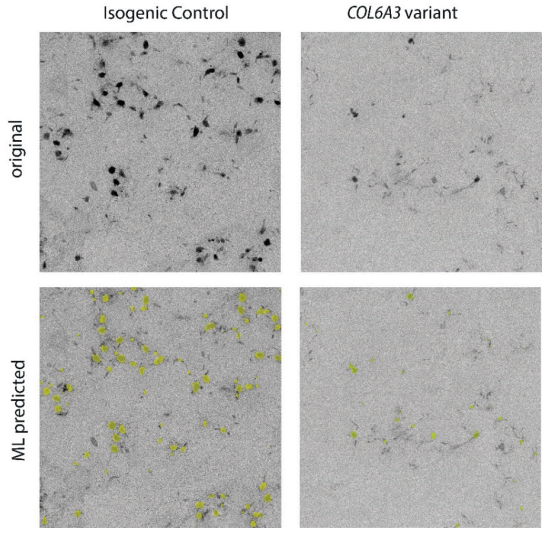


Figure S6 | Prediction of GAGs in TEM data. Yellow denotes the predicted annotation using a machine learning algorithm. (Spherical model, n=6)

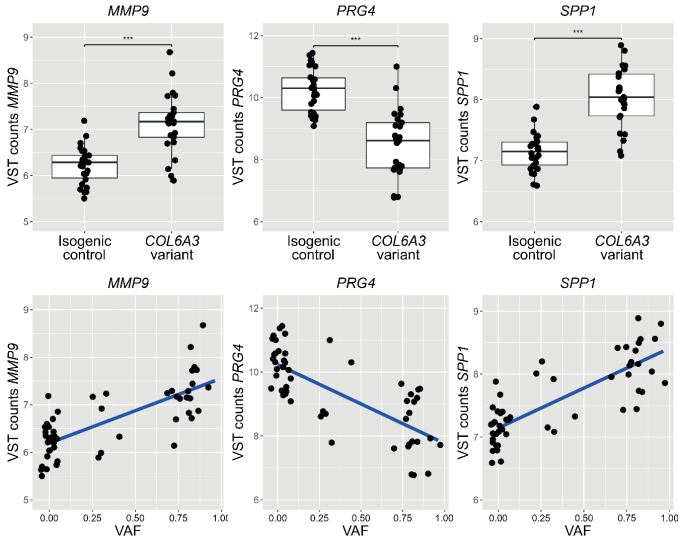
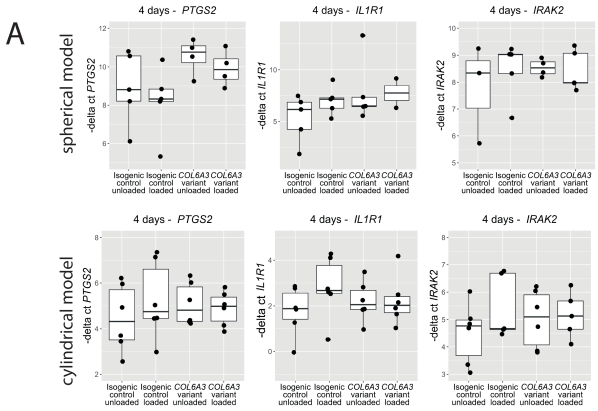
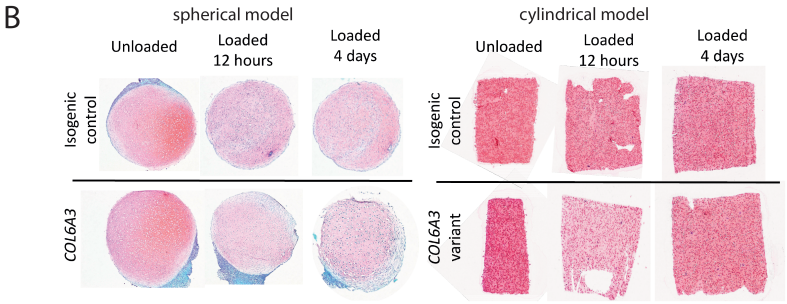
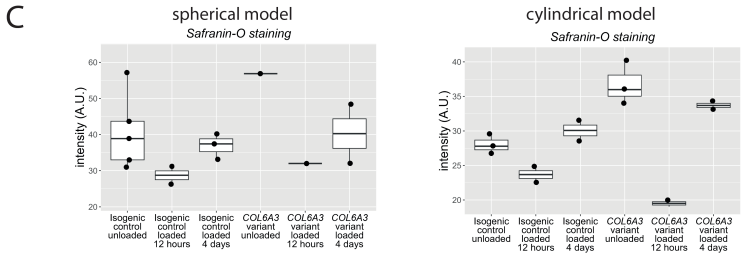


Figure S7| Example of dose response effect of the hetero- and homozygous variant *COL6A3* samples, as analyzed with DEseq2. n = 25-26, The box plots represent 25th, 50th, and 75th percentiles, and whiskers extend to 1.5 times the interquartile range. ***FDR<0.001


A
4 days - PTGS2
4 days - IL1R1
4 days - IRAK2
spherical model
cylindrical model
-delta ct PTGS2
-delta ct IL1R1
-delta ct IRAK2
Isogenic control unloaded
Isogenic control loaded
COL6A3 variant unloaded
COL6A3 variant loaded


B
spherical model
cylindrical model
Unloaded
Loaded 12 hours
Loaded 4 days
Isogenic control
COL6A3 variant


C
spherical model
cylindrical model
Safranin-O staining
intensity (A.U.)
Isogenic control unloaded
Isogenic control loaded 12 hours
Isogenic control loaded 4 days
COL6A3 variant unloaded
COL6A3 variant loaded 12 hours
COL6A3 variant loaded 4 days

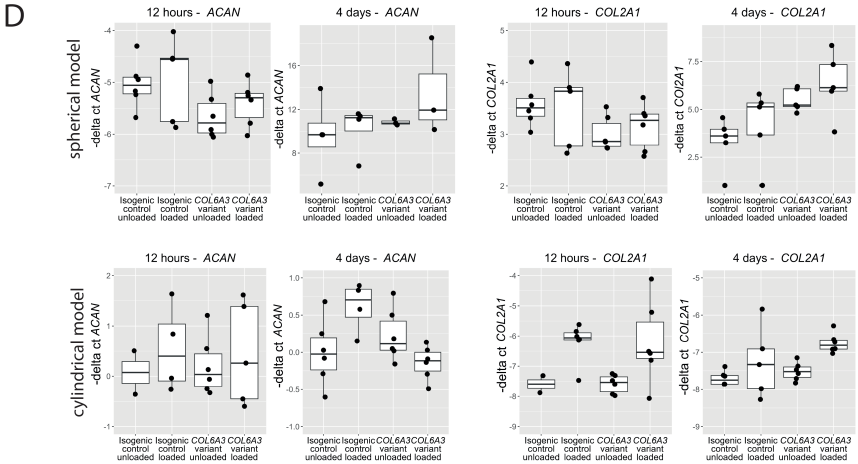

D
12 hours - ACAN
4 days - ACAN
12 hours - COL2A1
4 days - COL2A1
spherical model
cylindrical model
-delta ct ACAN
-delta ct COL2A1
Isogenic control unloaded
Isogenic control loaded
COL6A3 variant unloaded
COL6A3 variant loaded

Figure S8 | (A) RT-qPCR data of inflammatory response markers to hyper-physiological stress at day 4 after loading. (n=6) **(B)** Safranin 'O' staining at 12 hours (reproduced from Fig. 4) and 4 days after hyper physiological mechanical loading. (n=2-3) **(C)** Safranin 'O" staining intensity at 12 hours (reproduced from Fig. 4) and 4 days after hyper physiological mechanical loading. (n=2-3) **(D)** RT-qPCR data of ACAN and *COL2A1* expression at 12 hours and 4 days after hyper physiological mechanical loading. (n=6). In the cylindrical model, stratified analysis of *COL2A1* at day for showed a significant effect of loading in the *COL6A3* variant (P=0.0004, beta = 0.76, t=5.20) in contrast to the isogenic controls at day 4 (P=0.29, beta = 0.44, t = -1.10). The box plots represent 25th, 50th, and 75th percentiles, and whiskers extend to 1.5 times the interquartile range. Statistics are described in supplementary table S7.

Supplementary Table S1 (partially) - Identified novel coding variants that were predicted to have a functional impact on the protein

Ensembl ID	Gene Name	chr	Base change	Base position	Codon change	substitution	Mutation Type	Sequenced	Prediction
ENSG00000135046	ANXA1	9	T/A	75783988	GTT-GaT	V301D	heterozygous	support=40/26	DAMAGING
ENSG00000100852	ARHGAP5	14	T/C	32561296	GTA-GcA	V474A	heterozygous	support=23/5	DAMAGING
ENSG00000213214	ARHGEF35	7	G/C	143885448	GCC-GgC	A10G	heterozygous	support=6/33	DAMAGING
ENSG00000248919	ATP5J2-P	7	A/G	99022752	ATA-AcA	I517T	heterozygous		DAMAGING
ENSG00000168763	CNNM3	2	C/G	97490892	ATC-ATg	I441M	heterozygous	support=8/4	DAMAGING
ENSG00000163359	COL6A3	2	G/A	238277596	CGG-tGG	R1504W	heterozygous	support=23/15	DAMAGING
ENSG00000196730	DAPK1	9	A/C	90322017	AAC-AcC	N1344T	heterozygous	support=8/5	DAMAGING
ENSG00000165732	DDX21	10	T/C	70742332	TCT-cCT	S706P	heterozygous	support=28/4	DAMAGING
ENSG00000086061	DNAJA1	9	G/C	33026847	GAT-cAT	D57H	heterozygous	support=34/33	DAMAGING
ENSG00000116641	DOCK7	1	C/T	62941520	CGA-CaA	R1929Q	heterozygous	support=37/34	DAMAGING
ENSG00000154269	ENPP3	6	T/C	131979558	CTG-CcG	L187P	heterozygous	support=92/85	DAMAGING
ENSG00000217128	FNIP1	5	T/C	130815231	AGT-gGT	S736G			DAMAGING
ENSG00000109536	FRG1	4	C/A	190878563	GCT-GaT	A148D	heterozygous	support=67/22	DAMAGING
ENSG00000054983	GALC	14	A/G	88401092	GTG-GcG	V681A	heterozygous	support=18/21	DAMAGING
ENSG00000096968	JAK2	9	T/A	5064976	TGT-aGT	C384S	heterozygous	support=48/4	DAMAGING
ENSG00000123444	KBTBD4	11	C/T	47595086	CGG-CaG	R318Q	heterozygous	support=6/5	DAMAGING
ENSG00000162929	KIAA1841	2	G/T	61300623	TGT-TtT	C123F	heterozygous	support=62/46	DAMAGING
ENSG00000105976	MET	7	G/T	116436087	TGT-TtT	C1379F	heterozygous	support=25/21	DAMAGING
ENSG00000177000	MTHFR	1	G/A	11851349	CCG-CtG	P597L	heterozygous	support=11/5	DAMAGING
ENSG00000183486	MX2	21	C/T	42770828	TCG-TtG	S385L	heterozygous	support=31/24	DAMAGING
ENSG00000144426	NBEAL1	2	C/G	204000647	TCT-TgT	S1325C	heterozygous	support=29/18	DAMAGING
ENSG00000163531	NFASC	1	G/A	204951061	GTG-aTG	V806M	heterozygous	support=15/16	DAMAGING
ENSG00000164037	NHEDC1	4	C/T	103870560	TGT-TaT	C79Y	heterozygous	support=94/24	DAMAGING

Ensembl ID	Gene Name	chr	Base change	Base position	Codon change	substitution	Mutation Type	Sequenced	Prediction
ENSG00000012504	NR1H4	12	A/T	100926348	AAA-AAt	K196N	heterozygous	support=81/62	DAMAGING
ENSG00000115758	ODC1	2	C/T	10584626	GGG-aGG	G84R	heterozygous	support=33/29	DAMAGING
ENSG00000183310	OR2T34	1	C/T	248737664	TGC-TaC	C132Y	heterozygous	support=92/21	DAMAGING
ENSG00000168300	PCMTD1	8	C/G	52732981	AGA-AcA	R335T	heterozygous	support=190/8	DAMAGING
ENSG00000152527	PLEKHH2	2	C/T	43927635	TCC-TtC	S513F	heterozygous	support=30/15	DAMAGING
ENSG00000137492	PRKRIR	11	G/A	76072076	GCA-GtA	A81V	heterozygous	support=141/21	DAMAGING
ENSG00000164169	PRMT10	4	G/C	148575268	CTG-gTG	L594V	heterozygous	support=29/28	DAMAGING
ENSG00000174373	RALGAPA1	14	C/A	36103954	GGG-tGG	G1482W	heterozygous	support=21/17	DAMAGING
ENSG00000204764	RANBP17	5	G/A	170380604	CGT-CaT	R491H	heterozygous	support=76/54	DAMAGING
ENSG00000213516	RBMXL1	1	T/G	89448812	GAT-GcT	D233A	heterozygous	support=148/17	DAMAGING
ENSG00000100191	SLC5A4	22	T/C	32626996	AAC-AgC	N363S	heterozygous	support=15/14	DAMAGING
ENSG00000136891	TEX10	9	G/A	103090222	CGT-tGT	R550C	heterozygous	support=40/32	DAMAGING
ENSG00000188056	TREML4	6	T/C	41196551	TGT-cGT	C55R	heterozygous	support=23/17	DAMAGING
ENSG00000197888	UGT2B17	4	T/A	69433763	GAT-GtT	D147V	heterozygous	support=43	DAMAGING
ENSG00000251192	ZNF674	X	A/G	46359328	TCA-cCA	S560P	heterozygous	support=50/52	DAMAGING

Supplementary Table 2 (partially) - *in silico* analysis of the R1504W COL6A3 variant

Prediction Model	Type	Predicted Effect	Score	Reference Range benign - damaging (damaging threshold)
Sorting Intolerant From Tolerant (SIFT)	Sequence homology (protein function)	deleterious	0.00	1-0 (<0.05)
Polymorphism phenotyping v2 (PolyPhen-2)	Sequence homology (protein function)	probably damaging	1.00	0-1 (>0.908)
Protein Variation Effect Analyzer (Provean)	Sequence homology (protein function)	deleterious	NA	Binary
i-mutant V2.0	Support vector machine (protein stability)	decreased stability	NA	Binary
Mupro	Support vector machine (protein stability)	decreased stability	decreased stability	Binary
position-specific evolutionary preservation (PANTHER-PSEP)	Evolutionary Conservation (protein function)	Possibly damaging	361M	possibly damaging (450-200 Million), Damaging (>450M)

Supplementary Table 3 - *Effect of the COL6A3 variant on markers of cartilage metabolism measured with rt-qPCR*

	Effect of COL6A3 genotype		
Gene	**beta**	**t**	***p*-value**
Catabolic			
MMP13	0.02	0.07	0.94
MMP3	-1.19	1.70	0.09
ADAMTS5	-1.01	-7.21	**<0.001**
Anabolic			
COL2A1	-0.47	-2.35	**0.02**
ACAN	-0.54	-2.57	**0.01**
Hypertrophic			
COL10A1	-0.07	0.44	0.65

Supplementary Table 4 (partially) - *FDR significant DEGs related to the COL6A3 variant*

Feature	geneName	entrez	baseMean	log2 FC	lfcSE	FDR	FC
ENSG00000177432	NAP1L5	266812	87.1	-5.084	0.20	3.07E-137	0.03
ENSG00000253731	PCDHGA6	56109	45.0	2.303	0.12	3.56E-73	4.94
ENSG00000188707	ZBED6CL	113763	18.8	3.528	0.23	3.48E-49	11.54
ENSG00000127399	LRRC61	65999	13.4	3.325	0.25	1.80E-37	10.02
ENSG00000118785	SPP1	6696	292.6	1.445	0.11	2.45E-36	2.72
ENSG00000145708	CRHBP	1393	23.5	2.693	0.21	1.22E-33	6.47
ENSG00000181965	NEUROG1	4762	14.8	3.580	0.29	8.05E-31	11.96
ENSG00000196092	PAX5	5079	36.7	2.517	0.21	2.31E-30	5.72
ENSG00000165970	SLC6A5	9152	41.5	3.225	0.28	1.19E-26	9.35
ENSG00000115665	SLC5A7	60482	26.6	-1.861	0.16	1.42E-26	0.28
ENSG00000115221	ITGB6	3694	103.7	4.159	0.37	2.45E-26	17.86
ENSG00000165118	C9orf64	84267	11.4	3.189	0.28	4.57E-26	9.12
ENSG00000168779	SHOX2	6474	162.8	1.910	0.17	5.10E-26	3.76
ENSG00000164651	SP8	221833	185.0	2.398	0.22	1.40E-25	5.27
ENSG00000180828	BHLHE22	27319	24.3	3.374	0.31	1.99E-25	10.37
ENSG00000273706	LHX1	3975	87.2	2.273	0.21	5.41E-25	4.83
ENSG00000205420	KRT6A	3853	259.5	5.467	0.50	1.86E-24	44.23
ENSG00000124610	H1-1	3024	349.0	2.759	0.26	8.80E-24	6.77
ENSG00000163064	EN1	2019	15.0	2.589	0.24	8.80E-24	6.02
ENSG00000171786	NHLH1	4807	46.2	2.103	0.20	4.95E-23	4.30
ENSG00000101680	LAMA1	284217	341.1	0.858	0.08	1.56E-21	1.81
ENSG00000168875	SOX14	8403	32.4	2.030	0.20	2.14E-21	4.08
ENSG00000171956	FOXB1	27023	42.5	2.078	0.21	3.16E-21	4.22
ENSG00000143858	SYT2	127833	70.2	1.233	0.12	3.56E-21	2.35
ENSG00000115194	SLC30A3	7781	18.0	2.445	0.25	1.02E-19	5.45
ENSG00000128487	SPECC1	92521	1349.5	0.385	0.04	2.18E-19	1.31
ENSG00000186081	KRT5	3852	123.7	6.128	0.64	6.84E-19	69.92
ENSG00000126803	HSPA2	3306	82.4	1.436	0.15	6.87E-19	2.71
ENSG00000075891	PAX2	5076	77.5	2.831	0.30	9.12E-19	7.11
ENSG00000177508	IRX3	79191	36.7	1.840	0.19	1.14E-18	3.58
ENSG00000188848	BEND4	389206	206.9	1.589	0.17	1.24E-18	3.01
ENSG00000181143	MUC16	94025	100.2	4.633	0.49	1.30E-18	24.80
ENSG00000089116	LHX5	64211	112.8	2.450	0.26	1.66E-18	5.46
ENSG00000171291	ZNF439	90594	47.0	0.948	0.10	2.54E-18	1.93
ENSG00000198825	INPP5F	22876	2603.2	0.442	0.05	3.08E-18	1.36
ENSG00000187135	VSTM2B	342865	31.4	1.629	0.17	5.04E-18	3.09
ENSG00000198205	ZXDA	7789	38.3	2.901	0.31	1.17E-17	7.47
ENSG00000135480	KRT7	3855	52.2	2.527	0.28	1.70E-17	5.76
ENSG00000116690	PRG4	10216	2424.1	-2.252	0.25	1.74E-17	0.21
ENSG00000254245	PCDHGA3	56112	14.4	2.487	0.27	2.16E-17	5.61
ENSG00000177551	NHLH2	4808	51.7	1.422	0.16	3.28E-17	2.68
ENSG00000123307	NEUROD4	58158	25.5	2.946	0.33	4.70E-17	7.71

Supplementary Table 5 (partially) – *Gene Overrepresentation Enrichment analysis of WGCNA modules*

class	rank	dataSetID	inGroups	DataSet Name	nCommon Genes	Bonferroni FDR
bisque4	1	GO:0003676	GO|GO.MF|GO	nucleic acid binding	391	3.95E-09
bisque4	2	BioSystems.REACTOME.R-HSA-74160	REACTOME	Gene Expression	215	1.15E-08
bisque4	3	GO:0046907	GO|GO.BP|GO	intracellular transport	198	9.79E-07
black	1	BioSystems.KEGG.hsa03010	KEGG	Ribosome	83	4.95E-81
black	2	BioSystems.REACTOME.R-HSA-156842	REACTOME	Eukaryotic Translation Elongation	71	5.74E-79
black	3	BioSystems.REACTOME.R-HSA-156902	REACTOME	Peptide chain elongation	70	5.74E-79
darkgreen	1	GO:0016339	GO|GO.BP|GO	calcium-dependent cell-cell adhesion via plasma membrane cell adhesion molecules	8	0.0004019
darkgreen	2	GO:0051270	GO|GO.BP|GO	regulation of cellular component movement	40	0.0027893
darkgreen	3	GO:0098742	GO|GO.BP|GO	cell-cell adhesion via plasma-membrane adhesion molecules	17	0.0030319
darkgreen	4	GO:0009653	GO|GO.BP|GO	anatomical structure morphogenesis	79	0.0048628
darkmagenta	1	BioSystems.REACTOME.R-HSA-1640170	REACTOME	Cell Cycle	147	9.08E-91
darkmagenta	2	BioSystems.REACTOME.R-HSA-69278	REACTOME	Cell Cycle, Mitotic	135	1.31E-88
darkmagenta	3	GO:0051276	GO|GO.BP|GO	chromosome organization	163	2.99E-65
darkmagenta	4	GO:0071103	GO|GO.BP|GO	DNA conformation change	91	2.08E-61
darkorange	1	GO:0003723	GO|GO.MF|GO	RNA binding	85	7.03E-18
darkorange	2	GO:0022613	GO|GO.BP|GO	ribonucleoprotein complex biogenesis	38	5.13E-13
darkorange	3	GO:0034660	GO|GO.BP|GO	ncRNA metabolic process	39	9.34E-13
green	1	GO:0032501	GO|GO.BP|GO	multicellular organismal process	527	3.76E-35
green	2	GO:0048856	GO|GO.BP|GO	anatomical structure development	457	1.30E-32
green	3	GO:0007275	GO|GO.BP|GO	multicellular organism development	431	2.17E-32
grey60	1	GO:0007399	GO|GO.BP|GO	nervous system development	89	1.70E-23
grey60	2	GO:0048666	GO|GO.BP|GO	neuron development	58	1.26E-19

class	rank	dataSetID	inGroups	DataSet Name	nCommon Genes	Bonferroni FDR
grey60	3	GO:0030182	GO|GO.BP|GO	neuron differentiation	63	4.28E-19
ivory	1	GO:0001666	GO|GO.BP|GO	response to hypoxia	16	6.93E-09
ivory	2	GO:0036293	GO|GO.BP|GO	response to decreased oxygen levels	16	1.13E-08
ivory	3	GO:0070482	GO|GO.BP|GO	response to oxygen levels	16	2.95E-08
lightcyan	1	GO:0044281	GO|GO.BP|GO	small molecule metabolic process	438	7.53E-11
lightcyan	2	GO:0003824	GO|GO.MF|GO	catalytic activity	1150	2.34E-10
lightcyan	3	BioSystems.KEGG.hsa01100	KEGG	Metabolic pathways	310	4.07E-09
lightsteelblue1	1	GO:0005201	GO|GO.MF|GO	extracellular matrix structural constituent	13	8.90E-10
lightsteelblue1	2	GO:0030198	GO|GO.BP|GO	extracellular matrix organization	17	4.07E-09
lightsteelblue1	3	GO:0043062	GO|GO.BP|GO	extracellular structure organization	17	4.24E-09
mediumpurple3	1	GO:0048513	GO|GO.BP|GO	animal organ development	162	0.0001243
mediumpurple3	2	GO:0030198	GO|GO.BP|GO	extracellular matrix organization	35	0.0001577
mediumpurple3	3	GO:0043062	GO|GO.BP|GO	extracellular structure organization	35	0.0001673
white	1	GO:0004897	GO|GO.MF|GO	ciliary neurotrophic factor receptor activity	3	0.0054274
white	2	GO:0004896	GO|GO.MF|GO	cytokine receptor activity	6	0.0063461
white	3	GO:0003674	GO|GO.MF|GO	molecular_function	142	0.0078444
purple	1	GO:0007166	GO|GO.BP|GO	cell surface receptor signaling pathway	223	1.85E-17
purple	2	GO:0051239	GO|GO.BP|GO	regulation of multicellular organismal process	200	3.52E-15
purple	3	GO:0050793	GO|GO.BP|GO	regulation of developmental process	191	3.29E-14
plum1	1	BioSystems.REACTOME.R-HSA-1640170	REACTOME	Cell Cycle	54	7.59E-29
plum1	2	GO:0007049	GO|GO.BP|GO	cell cycle	82	7.15E-27
plum1	3	GO:0022402	GO|GO.BP|GO	cell cycle process	68	1.86E-23
pink	1	GO:2000427	GO|GO.BP|GO	positive regulation of apoptotic cell clearance	3	0.0367816
pink	2	GO:2000425	GO|GO.BP|GO	regulation of apoptotic cell clearance	3	0.0571026
pink	3	GO:0016098	GO|GO.BP|GO	monoterpenoid metabolic process	2	0.0692057
midnightblue	1	GO:0005201	GO|GO.MF|GO	extracellular matrix structural constituent	35	1.02E-30
midnightblue	2	GO:0030198	GO|GO.BP|GO	extracellular matrix organization	43	1.14E-25
midnightblue	3	GO:0043062	GO|GO.BP|GO	extracellular structure organization	43	1.27E-25

Supplementary Table 6 (partially) – *Overlap between COL6A3 variant associated DEGs and OA pathophysiology DEGs*

Emsembl_ID	GeneNames	FC RAAK	FDR RAAK	FC COL6A3 Variant	FDR COL6A3 Variant
ENSG00000134259	NGF	4.91	8.8E-18	0.73	3.33E-02
ENSG00000183421	RIPK4	5.22	2.4E-16	0.71	4.91E-03
ENSG00000101074	R3HDML	5.67	7.9E-16	0.49	6.40E-06
ENSG00000148344	PTGES	3.06	2.3E-15	1.45	7.33E-05
ENSG00000164761	TNFRSF11B	3.01	4.9E-15	0.65	8.00E-04
ENSG00000166033	HTRA1	2.39	1.8E-14	0.82	8.79E-05
ENSG00000187323	DCC	0.13	4.0E-14	1.49	6.02E-03
ENSG00000187957	DNER	3.37	5.3E-14	1.14	2.96E-02
ENSG00000106089	STX1A	2.29	7.7E-14	0.76	4.07E-03
ENSG00000135636	DYSF	2.05	1.8E-12	1.21	4.24E-02
ENSG00000165238	WNK2	0.21	7.3E-12	0.81	5.04E-03
ENSG00000092445	TYRO3	1.92	1.2E-11	1.12	1.44E-02
ENSG00000054938	CHRDL2	0.13	1.8E-11	0.40	1.41E-07
ENSG00000175093	SPSB4	2.01	2.5E-11	1.12	2.47E-02
ENSG00000003989	SLC7A2	2.10	2.8E-11	0.77	3.98E-02
ENSG00000040731	CDH10	4.02	4.4E-11	0.79	1.52E-03
ENSG00000134198	TSPAN2	2.42	4.7E-11	0.74	3.50E-03
ENSG00000176887	SOX11	2.99	5.2E-11	1.16	3.79E-03
ENSG00000128487	SPECC1	1.54	6.5E-11	1.31	2.18E-19
ENSG00000123610	TNFAIP6	3.58	9.0E-11	0.69	5.64E-03
ENSG00000009830	POMT2	1.41	1.3E-10	0.92	3.20E-02
ENSG00000178573	MAF	0.58	1.4E-10	0.77	8.22E-04
ENSG00000046889	PREX2	2.53	1.7E-10	1.52	1.89E-15
ENSG00000171488	LRRC8C	2.36	2.7E-10	0.77	3.80E-02
ENSG00000169239	CA5B	1.51	4.2E-10	0.87	4.98E-02
ENSG00000120658	ENOX1	1.88	4.9E-10	0.75	3.08E-06
ENSG00000143816	WNT9A	2.42	4.9E-10	0.79	1.53E-02
ENSG00000077942	FBLN1	0.44	6.7E-10	1.15	3.43E-02
ENSG00000269113	TRABD2B	0.49	7.5E-10	0.64	5.38E-05
ENSG00000158966	CACHD1	1.80	1.0E-09	1.22	9.30E-11
ENSG00000154319	FAM167A	2.67	1.1E-09	0.84	1.56E-02
ENSG00000198075	SULT1C4	0.41	1.7E-09	0.73	7.50E-05
ENSG00000120659	TNFSF11	2.40	1.7E-09	1.43	4.46E-04
ENSG00000077063	CTTNBP2	0.50	1.8E-09	1.20	7.36E-03
ENSG00000167191	GPRC5B	0.49	2.1E-09	1.21	3.22E-05
ENSG00000154864	PIEZO2	0.44	2.3E-09	0.70	4.82E-07
ENSG00000181656	GPR88	0.53	2.6E-09	0.58	1.36E-02
ENSG00000139910	NOVA1	2.15	2.7E-09	1.17	1.20E-02
ENSG00000066468	FGFR2	0.58	2.8E-09	0.91	4.03E-02
ENSG00000144749	LRIG1	0.60	3.3E-09	1.18	1.01E-05
ENSG00000149131	SERPING1	0.63	3.4E-09	0.89	3.24E-02
ENSG00000204592	HLA-E	0.72	4.4E-09	0.88	7.40E-03
ENSG00000165434	PGM2L1	1.88	5.9E-09	0.80	2.22E-02
ENSG00000118785	SPP1	3.14	7.3E-09	2.72	2.45E-36
ENSG00000183098	GPC6	0.63	8.3E-09	0.86	2.54E-02
ENSG00000113739	STC2	0.49	1.0E-08	0.68	2.32E-04
ENSG00000180616	SSTR2	0.53	1.1E-08	0.72	6.38E-05
ENSG00000167037	SGSM1	0.47	1.1E-08	0.83	2.71E-02
ENSG00000146374	RSPO3	0.28	1.3E-08	0.50	2.91E-04

Supplementary Table 7– part 1:Technical validation and long-term replication of RNAseq data using RT-qPCR

RT-qPCR

	Effect of Mechanical Loading		Effect of *COL6A3* genotype	
Gene	**beta**	***p*-value**	**beta**	***p*-value**
Catabolic				
MMP13	0.41	0.94	0.02	0.94
MMP3	1.71	0.08	-1.19	0.09
ADAMTS5	0.14	**<0.001**	-1.01	**<0.001**
Anabolic				
COL2A1	0.03	0.87	-0.47	**0.02**
ACAN	0.17	0.72	-0.54	**0.01**
Hypertrophic				
COL10A1	-0.2	0.18	-0.07	0.65
Mechano-sensors				
PIEZO1	0.33	**0.03**	-0.21	0.16
PIEZO2	0.09	0.61	-0.67	**<0.001**
TRPV4	0.02	0.91	-0.94	**<0.001**
Other				
COL6A3	0.35	0.16	-1.01	**<0.001**

RNA-seq

	decile unloaded	mechanical loading		*COL6A3* genotype	
Gene		**FC**	***p*-value**	**FC**	***p*-value**
Catabolic					
MMP13	*8th*	1.28	**0.017**	0.9	0.332
MMP3	*2nd*	1.66	**0.005**	0.71	0.083
ADAMTS5	*5th*	1.07	0.445	0.48	**2.21E-14**
Anabolic					
COL2A1	*9th*	0.88	0.19	0.78	**0.022**
ACAN	*9th*	0.8	0.051	0.76	**0.023**
Hypertrophic					
COL10A1	*1th*	0.92	0.474	1.1	0.429
Mechano-sensors					
PIEZO1	*9th*	1.1	0.083	0.89	0.063
PIEZO2	*7th*	0.95	0.372	0.7	**1.10E-08**
TRPV4	*8th*	0.93	0.906	0.75	**0.003**
Other					
COL6A3	*9th*	1.13	0.169	0.92	0.374

Supplementary Table for Supplementary Figure S8– Gene expression and Saf-O 12 hours and 4 days after hyper-physiologic mechanical loading.

RT-qPCR Fig. 8A - cylinder model - statistics

	Effect of *COL6A3* genotype		Effect of Mechanical Loading		Interaction effect	
Gene	**beta**	***p*-value**	**beta**	***p*-value**	**beta**	***p*-value**
IL1R1	0.41	0.94	0.02	0.94	0.02	0.94
PTGS2	0.35	0.64	0.2	0.78	-0.968	0.362
IRAK2	0.41	0.94	0.02	0.94	0.02	0.94

SAFO Fig. 8B - cylinder model - statistics

	Effect of *COL6A3* genotype		Effect of Mechanical Loading		Interaction effect	
SAFO	**beta**	***p*-value**	**beta**	***p*-value**	**beta**	***p*-value**
Intensity (A.U.)						
12 hours	4.17	0.09	17.2	**0.0001**	-12.89	**0.003**
4 days	-3.61	0.15	3.05	**0.001**	-0.968	**0.004**

RT-qPCR Fig. 8D - cylinder - statistics

	Effect of *COL6A3* genotype		Effect of Mechanical Loading		Interaction effect	
Gene	**beta**	***p*-value**	**beta**	***p*-value**	**beta**	***p*-value**
12 hours						
COL2A1	0.01	0.97	-1.38	0.01	0.02	0.94
ACAN	0.35	0.64	0.2	0.78	-0.968	0.362
4 days						
COL2A1	-0.51	0.09	-0.76	**0.01**	0.32	0.43
ACAN	0.76	0.004	0.37	0.08	-0.98	**0.004**

RT-qPCR Fig. 8A - spherical model - statistics

	Effect of *COL6A3* genotype		Effect of Mechanical Loading		Interaction effect	
Gene	**beta**	***p*-value**	**beta**	***p*-value**	**beta**	***p*-value**
IL1R1	-0.74	0.71	0.08	0.96	-1.77	0.49
PTGS2	-1.71	0.12	0.63	0.57	0.06	0.97
IRAK2	0.04	0.95	0.12	0.86	-0.81	0.44

SAFO Fig. 8B - spherical model - statistics

	Effect of *COL6A3* genotype		Effect of Mechanical Loading		Interaction effect	
SAFO	**beta**	***p*-value**	**beta**	***p*-value**	**beta**	***p*-value**
Intensity (A.U.)						
12 hours	-3.26	0.78	24.89	0.12	-12.89	0.45
4 days	-3.35	0.7	16.64	0.18	-12.79	0.36

RT-qPCR Fig. 8D - spherical model - statistics

	Effect of *COL6A3* genotype		Effect of Mechanical Loading		Interaction effect	
Gene	**beta**	***p*-value**	**beta**	***p*-value**	**beta**	***p*-value**
12 hours						
COL2A1	-0.35	0.27	0.14	0.64	-0.23	0.6
ACAN	-0.47	0.17	0.24	0.45	-0.15	0.74
4 days						
COL2A1	0.35	0.26	-0.14	0.64	0.23	0.6
ACAN	0.46	0.173	-0.24	0.45	0.15	0.74

Supplementary Table 8– Variant calling on RNAseq samples. G = reference allele, A = variant allele

sample number	Variant	Loading	model	G count	A count	VAF
1	Isogenic Control	unloaded	spherical	122	0	0.00
2	Isogenic Control	unloaded	spherical	16	0	0.00
3	Isogenic Control	unloaded	spherical	40	0	0.00
4	Isogenic Control	unloaded	spherical	135	0	0.00
5	Isogenic Control	unloaded	spherical	73	0	0.00
6	Isogenic Control	unloaded	spherical	35	0	0.00
7	Isogenic Control	loaded	spherical	145	0	0.00
8	Isogenic Control	loaded	spherical	158	0	0.00
9	Isogenic Control	loaded	spherical	61	0	0.00
10	Isogenic Control	loaded	spherical	115	0	0.00
11	Isogenic Control	loaded	spherical	68	0	0.00
12	Isogenic Control	loaded	spherical	79	0	0.00
13	Isogenic Control	unloaded	spherical	125	0	0.00
14	Isogenic Control	unloaded	spherical	101	0	0.00
15	Isogenic Control	unloaded	spherical	62	0	0.00
16	Isogenic Control	loaded	spherical	118	0	0.00
17	Isogenic Control	loaded	spherical	49	0	0.00
18	Isogenic Control	loaded	spherical	41	0	0.00
19	Isogenic Control	unloaded	cylindrical	174	3	0.02
20	Isogenic Control	unloaded	cylindrical	174	3	0.02
21	Isogenic Control	unloaded	cylindrical	171	2	0.01
22	Isogenic Control	unloaded	cylindrical	169	3	0.02
23	Isogenic Control	unloaded	cylindrical	167	5	0.03
24	Isogenic Control	unloaded	cylindrical	169	3	0.02
25	Isogenic Control	loaded	cylindrical	186	1	0.01
26	Isogenic Control	loaded	cylindrical	184	1	0.01
27	Isogenic Control	loaded	cylindrical	171	5	0.03
28	Isogenic Control	loaded	cylindrical	146	6	0.04
29	Isogenic Control	loaded	cylindrical	184	1	0.01
30	Isogenic Control	loaded	cylindrical	156	6	0.04
31	*COL6A3* Variant	unloaded	spherical	14	89	0.86
32	*COL6A3* Variant	unloaded	spherical	1	26	0.96
33	*COL6A3* Variant	unloaded	spherical	12	79	0.87
34	*COL6A3* Variant	unloaded	spherical	8	35	0.81
35	*COL6A3* Variant	unloaded	spherical	7	26	0.79
36	*COL6A3* Variant	unloaded	spherical	4	3	0.43
37	*COL6A3* Variant	loaded	spherical	7	65	0.90
38	*COL6A3* Variant	loaded	spherical	1	4	0.80
39	*COL6A3* Variant	loaded	spherical	13	113	0.90
40	*COL6A3* Variant	loaded	spherical	18	68	0.79
41	*COL6A3* Variant	loaded	spherical	3	7	0.70
42	*COL6A3* Variant	loaded	spherical	4	22	0.85
43	*COL6A3* Variant	unloaded	spherical	61	21	0.26
44	*COL6A3* Variant	unloaded	spherical	29	27	0.48
45	*COL6A3* Variant	unloaded	spherical	122	70	0.36
46	*COL6A3* Variant	loaded	spherical	19	8	0.30
47	*COL6A3* Variant	loaded	spherical	122	52	0.30
48	*COL6A3* Variant	loaded	spherical	166	59	0.26
49	*COL6A3* Variant	unloaded	cylindrical	39	117	0.75
50	*COL6A3* Variant	unloaded	cylindrical	38	129	0.77
51	*COL6A3* Variant	unloaded	cylindrical	33	142	0.81

sample number	Variant	Loading	model	G count	A count	VAF
52	*COL6A3* Variant	unloaded	cylindrical	50	131	0.72
53	*COL6A3* Variant	unloaded	cylindrical	34	142	0.81
54	*COL6A3* Variant	unloaded	cylindrical	34	144	0.81
55	*COL6A3* Variant	loaded	cylindrical	41	141	0.77
56	*COL6A3* Variant	loaded	cylindrical	41	141	0.77
57	*COL6A3* Variant	loaded	cylindrical	33	144	0.81
58	*COL6A3* Variant	loaded	cylindrical	33	144	0.81
59	*COL6A3* Variant	loaded	cylindrical	39	132	0.77
60	*COL6A3* Variant	loaded	cylindrical	46	145	0.76

Supplementary Table 9. *Dose response effect of variant allele frequency COL6A3 on gene expression*

feature	geneName	baseMean	FC	log2 FoldChange	lfcSE	padj
ENSG00000177432	SCYL3	206.6	0.88	-0.18	0.075	1.46E-77
ENSG00000253731	C1orf112	209.1	1.03	0.05	0.092	1.30E-37
ENSG00000188707	NIPAL3	374.7	0.82	-0.29	0.074	9.53E-31
ENSG00000127399	LAP3	570.1	1.04	0.05	0.053	1.53E-25
ENSG00000145708	HECW1	260.1	1.13	0.17	0.087	1.53E-25
ENSG00000168779	TMEM176A	32.7	0.69	-0.53	0.190	7.33E-25
ENSG00000058335	KLHL13	227.3	1.32	0.40	0.101	1.69E-24
ENSG00000163064	CYP26B1	511.7	1.94	0.95	0.130	3.96E-22
ENSG00000180828	ALS2	784.7	0.87	-0.20	0.053	5.46E-22
ENSG00000118785	RBM5	1312.7	0.82	-0.29	0.075	1.07E-20
ENSG00000273706	SLC7A2	472.6	0.69	-0.54	0.193	1.06E-19
ENSG00000115221	SARM1	799.6	0.88	-0.18	0.057	1.61E-18
ENSG00000177508	POLDIP2	1304.3	1.09	0.12	0.058	1.61E-18
ENSG00000075891	KDM1A	2293.3	1.10	0.14	0.041	2.41E-18
ENSG00000165118	CAMKK1	431.3	0.88	-0.19	0.061	2.84E-18
ENSG00000196092	HSPB6	396.7	1.54	0.62	0.172	3.85E-18
ENSG00000115194	PDK4	97.7	0.79	-0.35	0.164	1.16E-17
ENSG00000205420	ARX	434.2	0.58	-0.79	0.246	1.88E-17
ENSG00000198205	ST7	427.2	0.78	-0.35	0.077	1.29E-16
ENSG00000165970	SLC25A5	1655.4	1.14	0.19	0.071	1.99E-16
ENSG00000124610	ACSM3	11.9	1.62	0.70	0.265	3.00E-16
ENSG00000181965	REXO5	83.1	1.10	0.13	0.103	4.06E-16
ENSG00000164651	FAM214B	498.2	0.93	-0.10	0.059	1.45E-15
ENSG00000187559	CROT	251.4	0.88	-0.19	0.100	3.44E-15
ENSG00000171786	RHBDD2	1262.2	0.98	-0.03	0.060	4.08E-15
ENSG00000171956	PDK2	232.3	0.82	-0.29	0.071	4.08E-15
ENSG00000108001	OSBPL7	169.3	0.84	-0.25	0.077	4.68E-15
ENSG00000188848	TMEM98	731.4	1.12	0.16	0.059	7.50E-15
ENSG00000101680	CACNG3	10.9	2.19	1.13	0.325	1.03E-14
ENSG00000171291	TAC1	64.3	1.61	0.69	0.189	1.63E-14
ENSG00000164778	CX3CL1	317.1	0.92	-0.13	0.093	3.21E-14
ENSG00000123307	DLX6	79.4	0.53	-0.90	0.253	5.82E-14
ENSG00000185960	ETV1	475.1	1.23	0.29	0.089	5.82E-14
ENSG00000115665	TTC22	21.5	1.88	0.91	0.225	5.91E-14
ENSG00000143858	USH1C	181.4	1.50	0.59	0.182	5.91E-14
ENSG00000168875	DBF4	232.9	1.04	0.06	0.079	1.00E-13
ENSG00000100987	IFRD1	194.1	0.74	-0.43	0.141	1.58E-13
ENSG00000254245	ELAC2	995.5	1.03	0.04	0.058	1.58E-13
ENSG00000204531	ARSD	192.0	0.82	-0.28	0.067	1.65E-13
ENSG00000053438	PROM1	277.2	1.12	0.16	0.085	1.72E-13
ENSG00000146833	CCDC124	436.0	1.19	0.25	0.084	1.97E-13
ENSG00000124785	CEACAM21	11.5	0.42	-1.24	0.280	2.07E-13
ENSG00000105996	PAFAH1B1	2816.6	0.92	-0.13	0.060	2.77E-13
ENSG00000181143	NOS2	46.5	0.54	-0.89	0.205	3.32E-13
ENSG00000165973	GAS7	1238.1	0.83	-0.26	0.102	4.06E-13
ENSG00000186081	MATK	31.2	1.51	0.59	0.273	4.69E-13
ENSG00000125869	PAX6	309.5	6.45	2.69	0.300	8.98E-13
ENSG00000134138	BAIAP3	316.2	0.81	-0.31	0.104	1.02E-12
ENSG00000135480	TSR3	409.4	1.17	0.22	0.078	1.30E-12

Supplementary Table 10 (partially). *Differentially expressed genes in response to hyper-physiologic mechanical loading conditions*

	Joint analysis		Spherical model		Cylindrical model	
Gene	**FC**	**FDR**	**FC**	**Pvalue**	**FC**	**Pvalue**
FKBP4	1.253	1.19E-02	1.48	2.01E-08	1.02	7.61E-01
MCUB	1.274	4.35E-02	1.35	4.55E-03	1.28	6.65E-02
ITGA3	1.297	1.88E-02	1.35	1.03E-03	1.41	4.66E-03
PLAUR	1.498	2.50E-02	1.78	6.88E-04	1.71	4.75E-04
CNTN1	0.798	2.50E-02	0.88	9.42E-02	1.03	8.76E-01
NRXN3	0.803	1.60E-02	0.80	2.96E-03	1.02	8.68E-01
VIM	1.169	2.14E-02	1.30	1.38E-05	1.12	6.19E-02
CD44	1.458	1.88E-02	1.54	3.09E-03	1.38	1.63E-01
TLL1	1.310	2.64E-02	1.31	3.75E-03	1.17	3.95E-01
SLC4A8	0.849	2.52E-02	0.85	7.86E-03	0.95	6.09E-01
SYNE2	0.852	2.79E-02	0.78	4.44E-04	0.97	6.00E-01
GYG2	0.818	5.03E-03	0.82	2.62E-03	0.76	2.66E-03
PTPRU	1.147	4.81E-02	1.27	6.08E-04	1.08	2.36E-01
SNAP91	0.847	2.64E-02	0.90	1.12E-01	1.04	8.29E-01
TNFRSF1A	1.138	2.64E-02	1.17	5.37E-03	1.08	1.65E-01
ATP1B3	1.096	1.87E-02	1.09	1.21E-02	1.13	6.03E-02
WSCD2	0.830	3.96E-02	0.93	4.83E-01	0.80	8.48E-04
PAK3	0.862	2.64E-02	0.93	1.95E-01	0.80	3.83E-02
COL5A3	1.385	1.87E-02	1.58	2.73E-04	1.62	1.70E-05
BZW1	1.138	3.34E-02	1.11	3.89E-02	1.25	3.08E-03
OXCT1	0.912	3.74E-02	0.89	5.72E-03	0.96	6.16E-01
CD59	1.264	1.52E-02	1.36	6.61E-04	1.29	9.28E-02
RTRAF	0.909	2.37E-02	0.87	4.76E-03	0.90	7.13E-02
PHACTR3	0.811	1.75E-02	0.88	5.47E-02	0.72	1.14E-01
P3H2	1.265	4.87E-02	1.49	4.15E-05	1.40	6.36E-08
PITPNM2	1.297	1.65E-02	1.29	3.27E-04	1.49	1.26E-03
PITPNM3	1.254	2.14E-02	1.26	2.38E-03	1.44	2.53E-03
EPB41L4B	1.207	2.73E-02	1.09	1.46E-01	1.45	8.82E-03
YPEL1	0.822	2.70E-02	0.84	3.22E-02	1.11	5.91E-01
HMOX1	1.495	7.46E-03	1.44	3.42E-02	1.34	1.20E-01
PYGB	1.239	5.31E-03	1.25	4.05E-04	1.31	1.27E-04
JAG1	1.430	9.02E-03	1.47	1.11E-04	1.45	1.11E-01
TIMP1	1.160	4.64E-02	1.28	5.56E-04	1.18	5.98E-03
BEX4	0.891	1.19E-02	0.86	3.36E-05	0.94	5.15E-01
FGF14	0.789	9.20E-03	0.87	2.58E-02	0.74	6.74E-03
NDRG4	0.900	2.38E-02	0.94	4.96E-02	0.93	4.16E-01
CRISPLD2	0.717	4.83E-03	0.76	2.58E-02	0.75	2.92E-04
CAPN15	1.178	2.94E-02	1.26	2.94E-04	1.15	3.90E-02
OCA2	0.739	1.19E-02	0.92	3.88E-01	0.60	6.90E-10
PLAT	1.294	1.19E-02	1.16	6.18E-02	1.47	3.70E-03
STMN2	0.721	2.38E-02	0.81	9.40E-03	1.02	8.97E-01
KCNN4	1.618	6.64E-03	1.54	6.62E-03	2.10	1.18E-04
TFPI2	1.524	2.04E-02	1.72	2.41E-05	1.58	4.98E-02
CAV1	1.522	4.14E-04	1.52	4.15E-03	1.77	8.74E-03
MET	1.474	3.46E-02	1.49	2.34E-02	1.93	3.00E-02
NPTX2	1.185	3.82E-02	1.25	2.67E-03	1.26	4.21E-02
AGFG2	1.139	2.38E-02	1.17	6.78E-03	1.00	9.60E-01

Supplementary Table 11 (partially). *Genes showing an interaction effect between the COL6A3 variant and hyper-physiologic mechanical loading conditions.*

Gene	baseMean Unloaded	baseMean Loaded	BaseMean Iso	baseMean Mut	stat	pvalue
FAM214B	492.71	498.69	533.12	458.05	-3.102	1.92E-03
CRLF1	820.82	843.37	774.15	889.18	-3.040	2.37E-03
MMP25	46.05	50.08	14.48	81.49	-3.216	1.30E-03
HIVEP2	846.14	923.70	999.35	767.51	-3.155	1.60E-03
GPRC5A	131.63	162.79	109.00	184.23	-2.587	9.67E-03
FSTL4	91.91	99.20	114.00	76.83	-2.798	5.15E-03
TRAF1	172.02	169.84	174.54	167.40	-2.741	6.12E-03
TRAF3IP2	271.23	318.15	281.86	305.72	-2.766	5.67E-03
MAST4	472.81	547.82	558.51	459.24	-3.269	1.08E-03
PTGS2	109.19	142.46	112.25	138.12	-3.282	1.03E-03
SYNJ2	228.50	291.45	251.67	265.86	-2.718	6.56E-03
VDAC3	662.33	637.65	646.75	654.18	2.624	8.68E-03
SLC4A4	182.88	161.61	188.75	156.56	-2.699	6.96E-03
ULK2	683.01	642.13	701.75	624.97	-2.579	9.91E-03
ABCB1	14.86	19.98	21.32	13.33	-2.812	4.92E-03
DSP	1352.30	1620.91	904.37	2058.50	-2.733	6.27E-03
FKBP3	461.41	435.49	412.07	485.83	2.700	6.93E-03
SIX4	522.72	538.28	482.19	578.22	-3.100	1.94E-03
FERMT1	131.63	175.82	98.42	207.33	-2.716	6.61E-03
FLT1	170.37	199.20	90.82	277.64	-3.124	1.78E-03
ETFB	182.22	183.52	189.01	176.69	2.835	4.58E-03
AHR	105.32	116.87	108.78	112.96	-3.038	2.38E-03
DDX25	78.44	78.58	80.90	76.10	3.114	1.85E-03
ELK3	1078.04	1193.13	1088.96	1177.79	-2.651	8.02E-03
SH2B3	410.63	453.89	397.57	465.29	-2.795	5.19E-03
NEDD9	1503.86	1510.82	1779.41	1235.00	-2.812	4.93E-03
CDH6	1157.01	1238.25	1662.41	729.73	-2.949	3.19E-03
TNNC1	6.50	5.21	5.37	6.39	2.609	9.07E-03
IGFBP5	7044.53	8637.19	10304.75	5315.72	-2.634	8.44E-03
IL1R1	801.64	876.46	776.07	899.16	-2.788	5.30E-03
NID1	308.55	296.40	253.47	351.95	-2.815	4.87E-03
APOA1	36.53	37.14	46.47	27.17	-2.601	9.30E-03
CCDC92	600.23	603.74	697.48	506.35	-2.586	9.70E-03
ABITRAM	169.80	150.78	155.38	165.94	2.749	5.98E-03
HSPH1	2678.16	3719.03	3615.13	2742.03	-2.688	7.18E-03
TNFRSF8	68.53	67.94	24.90	111.59	-3.254	1.14E-03
MTERF4	307.79	327.10	333.42	300.73	2.612	9.00E-03
BBS9	557.68	507.92	462.83	604.68	2.851	4.35E-03
INHBA	132.32	155.78	126.79	160.41	-2.756	5.85E-03
CLYBL	20.15	18.08	19.56	18.75	3.678	2.35E-04
EFNB2	1142.50	1278.27	1447.29	968.25	-2.629	8.57E-03
TXNDC17	213.96	214.61	227.02	201.52	2.742	6.12E-03
FOXA1	607.35	724.89	896.20	431.52	-3.095	1.97E-03
CDC16	683.29	706.19	648.72	739.88	2.602	9.26E-03
SNX9	715.16	756.72	744.70	725.58	-2.664	7.71E-03
PER2	229.43	237.37	255.30	211.20	-2.996	2.74E-03
RTN3	2187.19	2167.63	2198.50	2157.07	3.439	5.84E-04
ACTR3B	181.43	190.25	191.39	179.95	2.639	8.31E-03
IRAK2	129.50	148.92	137.59	140.08	-3.401	6.70E-04

Supplementary Table 12 (partially). *Stratified analysis of hyper-physiological mechanical loading conditions for genes that show an interaction effect between mechanical loading and the COL6A3 variant.*

Gene	P value interaction	Variant Loaded padj	Variant Loaded pvalue	Variant Loaded FC	Isogenic Control Loaded padj	Isogenic Control Loaded pvalue	Isogenic Control Loaded FC
KIRREL3	2.23E-04	3.02E-01	2.68E-02	0.790	7.46E-02	3.29E-02	1.188
CLYBL	2.35E-04	7.26E-01	4.11E-01	1.120	5.46E-04	5.03E-06	0.597
PCCA	3.02E-04	4.48E-01	8.58E-02	1.078	3.67E-02	1.09E-02	0.901
ADAMTS15	3.08E-04	7.19E-01	3.60E-01	0.879	6.29E-05	1.14E-07	1.452
KIF6	3.58E-04	3.20E-01	3.49E-02	1.473	4.77E-02	1.62E-02	0.705
RTN3	5.84E-04	5.90E-01	2.10E-01	1.042	4.33E-02	1.41E-02	0.923
IRAK2	6.70E-04	7.83E-01	4.98E-01	1.058	2.32E-03	9.28E-05	1.419
PTGER4	7.09E-04	7.05E-01	3.45E-01	0.918	1.38E-03	2.97E-05	1.400
MTARC1	7.34E-04	4.87E-01	1.07E-01	1.139	5.18E-02	1.88E-02	0.817
COMMD6	8.02E-04	5.86E-01	1.98E-01	1.091	4.26E-02	1.35E-02	0.875
PLCH2	8.07E-04	3.02E-01	2.09E-02	0.823	4.53E-02	1.50E-02	1.158
GPR132	1.00E-03	2.59E-01	6.52E-03	0.593	3.13E-01	2.29E-01	1.274
PTGS2	1.03E-03	8.24E-01	5.67E-01	0.921	3.67E-03	1.97E-04	1.642
MAST4	1.08E-03	9.19E-01	7.42E-01	0.974	1.38E-03	2.76E-05	1.279
STPG2	1.10E-03	4.48E-01	9.00E-02	1.362	9.89E-02	5.03E-02	0.730
TNFRSF8	1.14E-03	4.94E-03	2.31E-05	0.513	2.12E-01	1.39E-01	1.302
KATNAL2	1.27E-03	5.30E-01	1.47E-01	0.872	2.24E-02	4.68E-03	1.254
MMP25	1.30E-03	2.81E-01	1.32E-02	0.625	4.22E-02	1.32E-02	1.488
FAM118B	1.30E-03	3.02E-01	2.01E-02	1.108	2.65E-01	1.87E-01	0.948
CLMP	1.35E-03	9.01E-01	7.12E-01	1.040	5.25E-03	3.71E-04	1.413
ADAMTS1	1.43E-03	6.39E-01	2.65E-01	0.900	3.01E-02	8.10E-03	1.211
COL22A1	1.51E-03	5.86E-01	1.98E-01	0.880	6.49E-02	2.69E-02	1.195
HIVEP2	1.60E-03	7.19E-01	3.65E-01	0.932	1.35E-02	1.95E-03	1.167
INSYN2A	1.76E-03	3.02E-01	2.33E-02	0.804	4.32E-01	3.34E-01	1.080
FOXO3B	1.77E-03	3.20E-01	3.49E-02	0.891	2.72E-01	1.93E-01	1.058
FLT1	1.78E-03	7.26E-01	3.94E-01	0.911	2.84E-02	7.28E-03	1.284
DDX25	1.85E-03	6.87E-01	3.06E-01	1.110	1.20E-02	1.54E-03	0.800
FAM214B	1.92E-03	7.26E-01	4.00E-01	0.965	1.32E-02	1.81E-03	1.146
SIX4	1.94E-03	2.81E-01	1.33E-02	0.857	9.60E-01	9.49E-01	0.997
FOXA1	1.97E-03	3.02E-01	2.77E-02	0.694	4.37E-02	1.44E-02	1.478
SEMA7A	2.23E-03	7.37E-01	4.24E-01	0.932	4.85E-02	1.67E-02	1.175
CRLF1	2.37E-03	5.94E-01	2.19E-01	0.879	8.45E-03	8.43E-04	1.414
AHR	2.38E-03	7.19E-01	3.64E-01	0.914	2.46E-02	5.64E-03	1.204
FBXO32	2.41E-03	5.84E-01	1.83E-01	0.856	6.95E-02	2.98E-02	1.184
ATP5MD	2.43E-03	4.02E-01	6.39E-02	1.144	1.53E-01	8.93E-02	0.882
DUSP5	2.70E-03	9.59E-01	8.66E-01	1.020	6.30E-03	5.08E-04	1.737
PER2	2.74E-03	4.48E-01	9.17E-02	0.917	1.25E-01	6.76E-02	1.108
SERTM1	2.88E-03	3.02E-01	1.97E-02	0.745	6.34E-01	5.51E-01	1.066
BACE1	2.90E-03	4.41E-01	7.69E-02	0.930	2.06E-01	1.35E-01	1.054
FOXF2	2.92E-03	9.89E-01	9.83E-01	1.002	1.01E-02	1.11E-03	1.277
UGT8	2.95E-03	4.02E-01	6.28E-02	0.799	6.95E-02	2.99E-02	1.37
DHRS1	2.99E-03	4.62E-01	9.65E-02	1.115	1.84E-01	1.15E-01	0.901

Supplementary Table 13. *Long noncoding RNAs showing an interaction effect between the COL6A3 variant and hyper-physiologic mechanical loading.*

feature	baseMean	log2FoldChange	lfcSE	stat	pvalue
ELOA-AS1	157.270	0.362	0.135	2.673	0.008
ENSG00000228509	37.767	-1.113	0.374	-2.977	0.003
ENSG00000235609	115.813	-0.450	0.163	-2.759	0.006
ENSG00000255197	6.748	-1.021	0.387	-2.636	0.008
ENSG00000257545	24.531	0.688	0.251	2.742	0.006
ENSG00000266368	6.783	0.974	0.359	2.714	0.007
ENSG00000272836	129.246	-0.351	0.132	-2.657	0.008
ENSG00000278472	5.697	1.324	0.466	2.841	0.005
ENSG00000287281	11.540	0.783	0.287	2.726	0.006
LINC00574	36.725	-0.557	0.193	-2.889	0.004
LINC02457	34.829	0.813	0.302	2.696	0.007
LINC02482	69.224	0.443	0.146	3.025	0.002
MIR31HG	134.360	-1.434	0.441	-3.255	0.001
MIS18A-AS1	8.727	1.021	0.343	2.977	0.003
NEXN-AS1	31.990	-0.611	0.213	-2.871	0.004
SNHG26	83.815	-0.832	0.271	-3.071	0.002
STX18-AS1	323.987	0.278	0.102	2.731	0.006

Supplementary methods

Experimental design

The objective of the current study was to study the effects of an OA associated variants, thereby elucidating the effects of aberrant collagen VI functions. Exome sequencing was applied to identify a pathogenic variant. To study the underlying effect of this variant, the variant was introduced in hiPSCs using CRISPR-Cas9 genome engineering, of which a neo-cartilage model was created, followed by mechanical loading and functional analysis.

Exome sequencing

Exome sequencing of a patient with generalized OA at multiple joint sites was performed by Illumina HiSeq 2000 technology (Beijing Genome Institute). The sequences were generated as 100–base pair paired-end reads, after enrichment of 44-Mb exonic sequences by NimbleGen EZ (Roche NimbleGen). Raw imaging files were processed by Illumina base-calling software v1.7 with default parameters. SOAPaligner/SOAP2.21 was used to align reads to the GRCh37 reference genome at the UCSC Genome Browser website (http://genome.ucsc.edu/) (1).

hiPSC line and cell culture

An hiPSC line as described earlier was used as the unedited isogenic control (2, 3). HEK293FT cells were plated at a density of 6 × 106 cells per T225 flask and incubated overnight. The cells were transfected with 10 μg VSV-G (envelope protein), 15 μg pUMVC (packaging plasmid), and 10 μg of the gene of interest [SRY (sex-determining region) box (Sox)2, octamer-binding transcription factor (Oct)4, oncogene of the avian myelocytomatosis virus (c-Myc), or Kruppellike factor (Klf)4] with Lipofectamine (Life Technologies, Carlsbad, CA, USA). The supernatant was collected 48 h after transfection and filtered through a 0.45 μm filter. After it was spun at 17,100 rpm for 2 h 20 min, the viral pellet was resuspended to make 100× stock solutions. To generate hiPSCs, retrovirally transduced human fibroblasts were seeded at 5 × 104 cells per well of a 6-well dish 1 d before transduction. The medium was replaced with virus-containing supernatant supplemented with 8 μg/ml polybrene and incubated for 24 h. The transduced fibroblasts were then cultured in iPSC medium [DMEM/F12 (Life Technologies-Gibco, Grand Island, NY, USA), 20% Knockout Serum Replacement (Life Technologies), nonessential amino acids, penicillin, streptomycin, β-mercaptoethanol, and 10 ng/ml fibroblast growth factor (FGF)-2] on mitomycin-treated MEFs. These iPSCs were then characterized, confirming their pluripotency. The hiPSCs were maintained under standard conditions (37 °C, 5% CO_2) on Matrigel (Corning) coated plates and refreshed daily with TeSR-E8 medium (STEMCELL Technologies) upon

reaching approximately 70% confluence.

Genome editing of hiPSCs

gRNA targeting *COL6A3* (5'-TCTGAAAACCTACAGATCCC -3') were complexed with Cas9 protein at 25 C for 10-15' and added to a RVR-hiPSC cell suspension of 500K cells in 400 µL mTeSR to generate the edited hiPSCs. 500 pmol of ssODN (5'-CTTGCCACAAATTCGA-GAGCCTTGCCAGTG TTCAGTGGGGACCCCCCTCTGAGCCTCAGGCGCCATATGGCGTCCAG-CACCGGGGCGTGGGATCTGTAGGTTTTCAGATAGAATTCTGGGAAGACAT-3') harboring the risk variant was added to cell suspension. Cells were electroporated with BioRad Gene Pulser system with the following conditions: [250 V, 750 µF, ¥ W, 0.4 cm]. Cells were plated at low density and colonies were picked to obtain clones derived from single cells. As the risk variant created an NDEI site, clones were screened by PCR amplification of the risk allele and digestion with NDEI (New England Biolabs) in CutSmart Buffer for 1 hour at 37C. Successful editing of the targeted variant was confirmed by Sanger sequencing **(Fig S1)**.

hiPSC differentiation to induced chondrocytes

Two different chondrogenic constructs were used for downstream analysis; these chondrogenic pellets were directly used for further experiments, or they were dissociated using collagenase II, encapsulated in 2% w/v agarose at 30 million cells/ml, and cultured for 14 days with CD creating cylindrical shaped constructs. When hiPSCs reached 60% confluence, the culture medium was switched to mesodermal differentiation (MD) medium, composed of IMDM GlutaMAX (IMDM; Thermo Fisher Scientific) and Ham's F12 Nutrient Mix (F12; Sigma-Aldrich) with 1% chemically defined lipid concentrate (Gibco), 1% insulin/human transferrin/selenous (ITS+; Corning), 0.5% penicillin-streptomycin (P/S; Gibco), and 450 µM 1-thioglycerol (Sigma-Aldrich). Before induction of anterior primitive streak (day 0), hiPSCs were washed with wash medium (IMDM/F12 and 0.5% P/S) and then fed with MD medium supplemented with activin A (30 ng/ml; Stemgent), 4 µM CHIR99021 (CHIR; Stemgent), and human fibroblast growth factor (20 ng/ml; FGF-2; R&D Systems) for 24 hours. Subsequently, the cells were washed again with wash medium, and paraxial mesoderm was induced on day 1, by MD medium supplemented with 2 µM SB-505124 (Tocris), 3 µM CHIR, FGF-2 (20 ng/ml), and 4 µM dorsomorphin (Tocris) for 24 hours. Before induction of early somite (day 2), cells were washed with wash medium, and then cells were fed with MD medium supplemented with 2 µM SB-505124, 4 µM dorsomorphin, 1 µM C59 (Cellagen Technology), and 500 nM PD173074 (Tocris) for 24 hours. Subsequently, cells were washed with wash medium, and for induction of sclerotome, cells (days 3 to 5) were fed daily with MD medium supplemented with 2 µM purmorphamine (Stemgent) and 1 µM C59. To induce chondroprogenitor cells (days 6 to 14), cells were washed

briefly with wash medium and fed daily with MD medium supplemented with human bone morphogenetic protein 4 (BMP-4; 20 ng/ml; Miltenyi Biotec). Three independent differentiations were performed.
Monolayer cultured hiCPC aggregates present at day 14 of the differentiation were washed with MD medium, dissociated with Gentle Cell dissociation medium (Stem Cell), and centrifuged for 5 min at 1200 rpm. Cell aggregates were subsequently maintained in chondrogenic differentiation (CD) medium containing Dulbecco's modified Eagle's medium/F12 (Gibco), supplemented with 1% ITS+, 55 μM 2-mercaptoethanol (Gibco), 1% non-essential amino acids (Gibco), 0.5% P/S, L-ascorbate-2-phosphate (50 μg/ml; Sigma-Aldrich), L-proline (40 μg/ml; Sigma-Aldrich), ML329 (1μM; CSNpharm), C59 (1μM; Tocris), and transforming growth factor–β3 (10 ng/ml; PeproTech) under hypoxic conditions for 30 days while refreshing medium every 3 to 4 days.

sGAG measurement

Sulphated glycosaminoglycan (sGAG) concentrations in the neo-cartilage organoids (μg sGAG/μg DNA) was measured using the Farndale Dimethyl Methylene Blue (DMMB, Sigma) method (4). Chondroitin sulphate (Sigma) was used as a reference standard. Absorbance was measured at 535 and 595 using a microplate reader (Synergy HT, Biotek). Neo-cartilage sGAG concentrations were corrected for DNA content measured with the Qubit® 2.0 Fluorometer (Invitrogen™) using the dsDNA HS Assay Kit (Invitrogen™).

Histology and immunohistochemistry

Neo-cartilage samples were fixed in 4% formaldehyde and embedded in paraffin. Sections were stained with Alcian Blue (Sigma-Aldrich) and Nuclear Fast Red (Sigma-Aldrich). Deposition of collagen VI and collagen II in the neo-cartilage constructs was visualized immunohistochemically using a polyclonal antibody for COL6A1 (abcam ab6588), a primary sub-unit of COLVI, and a polyclonal antibody for COL2A1 (abcam ab34712), a primary sub-unit of COLII., antigen retrieval was done by treating deparaffinized sections with proteinase K (5 μg/ml, Qiagen) and hyaluronidase (5 mg/ml, Sigma). Sections were incubated overnight with a primary antibody raised against human collagen VI α1 (1:100, abcam), followed by incubation with a HRP conjugated secondary antibody (ImmunoLogic). Peroxidase binding for collagen VI was visualized using diaminobenzidine, and sections were counterstained with haematoxylin. For long-term loading effects, pellets were fixed in 10% formalin overnight and stored in 70% ethanol at 4°C. They were then embedded in paraffin wax and sectioned at 8 μm thickness. Sections were stained with Safranin-O/hematoxylin standard protocol to reveal proteoglycan matrix. These images were quantified using ImageJ. Hereto, we splitted the color channels and selected the red channel for further processing. Next,

using a rolling ball algorithm the background noise was removed and the sample was segmented. Within this segmentation of the sample, we calculated the average staining intensity.

RT-qPCR

Per sample, two replicate neo-cartilage pellets were collected in TRIzol (Invitrogen™) and RNA was isolated using the RNeasy Mini Kit (Qiagen) according to manufacturer's protocol. DNA contamination was removed by treating the RNA with RNase-Free DNase RNA quality (A260/280: 1.7-2.0) was assessed using the Nanodrop. RNA concentrations were measured with the Qubit® 2.0 Fluorometer (Invitrogen™) using the RNA HS Assay Kit (Invitrogen™)., respectively, with an A260/280 between 1.7-2.0. RNA was reverse transcribed into cDNA using the Transcriptor First Strand cDNA Synthesis Kit (Roche). cDNA was amplified using FastStart SYBR Green Master (Roche) and mRNA expression was measured in triplicates in a MicroAmp™ Optical 384-Well Reaction Plate (ThermoFisher Scientific), using the QuantStudio™ Flex Real-Time PCR system (Applied Biosystems™), with the following cycling conditions: 10 min 95 °C; 10 sec 95 °C, 30 sec 60 °C, 20 sec 72 °C (45 cycles); 1 min 65 °C and 15 sec 95 °C . Primer efficiency was tested using a cDNA dilution series, and primers were considered efficient with an efficiency between 90% and 110%. -ΔCt expression levels were calculated using two housekeeping genes *GAPDH* and *SDHA*, with the following formula: $\Delta Ct = Ct\ (gene\ of\ interest) - Ct\ (average\ housekeeping\ genes)$. Both housekeeping genes were stably expressed in this model. Fold changes were calculated using the $2^{-\Delta\Delta Ct}$ method with $\Delta\Delta Ct = \Delta Ct\ (MS) - \Delta Ct\ (Control)$.

RNAseq

RNA from neo-cartilage constructs was extracted 12 hours post mechanical loading. RNA from spherical neo-cartilage constructs was extracted and processed using a pestle homogenizer in TRIzol reagent (Invitrogen). RNA was extracted using chloroform, followed by precipitation using ethanol, and purified with the RNeasy Mini Kit (Qiagen). Genomic DNA was removed by DNase digestion (Qiagen). Paired-end 2 × 150 base pair RNA sequencing (Illumina NEBNext Ultra II Directional RNA Library Prep, Illumina NOVAseq 6000) was performed. Strand-specific RNA-sequencing libraries were generated which yielded on average 25 million reads per sample. Data from the Illumina platform was analyzed with an in-house pipeline as previously described (5). The adapters were clipped using Cutadapt v1.1. RNA-seq reads were then aligned using GSNAP against GRCh38 (6). Read abundances per sample were estimated using HTSeq count v0.11.1 (7) with Ensembl gene annotation version 94. Only uniquely mapping reads were used for estimating expression. The quality of the raw reads and

initial processing for RNA sequencing was checked using MulitQC v1.9 (8). Samples containing > 50% genes with zero values and average read count < 4 were removed from further analysis. The datasets from both neo-cartilage models were combined using surrogate variable analysis using the R package SVA version 3.42.0 (Fig. S3) (9). Using SVA each individual data point was assigned a correction factor per identified surrogate variable, thus regressing out the technical variation and effects of the different organoid models. This allowed for joint analysis and visualization using boxplots. Outliers were identified (n=2) using hierarchical clustering and principal component analysis (PCA) which were removed from further analysis. In total, 52 samples were included of which; isogenic controls (free-swelling isogenic controls=13, mechanically stimulated isogenic controls=13) as well as COL6A3-mutant (free-swelling COL6A3 mutants=14 and mechanically loaded COL6A3 mutants, n=12) organoids (Fig S3). Differential expression analysis was performed using the R package DESeq version 1.34 (10). All samples were combined in a multifactorial analysis. To determine effects of the variant and hyper-physiological mechanical loading conditions only the main effects were included into the model. To determine the interaction effect an interaction term to this model was added. P-values were corrected for their false discovery rate using the Benjami and Hochman method (11). WGCNA analysis was performed using the R package WGCNA version 1.71 (12, 13). A generalized linear model was used to determine the association between the COL6A3 variant, hyper-physiological mechanical loading conditions, the interaction effect and the identified WGCNA co-expression networks. Over representation enrichment analyses of these co-expression networks using the KEGG, Reactome and gene ontologie biological processes databases was performed using the anRichment R package version 1.22. Protein-protein network analysis was performed using the online tool STRING version 11.0 (14).

Solid-phase binding assay

Conditioned medium of wild-type and *COL6A3* mutant organoids was collected and concentrated in preparation for the binding assay. To this end, 450 µl of medium was collected in 100 K molecular weight cutoff Pierce Protein Concentrators (Thermo Scientific) and centrifuged for 10 min at 12,000g. Subsequently, COL6 concentration was determined using the Human COL6A3 ELISA Kit (Assay Genie) according to the manufacturer's protocol.

Clear multiwell plates (R&D Systems) were coated overnight with 100 µl of purified fibronectin (10 µg/ml; Merck) in phosphate-buffered saline (PBS) at 4°C, followed by four wash steps with wash buffer (0.05% Tween 20 in PBS). Nonspecific binding was blocked for 1 hour with 3% (w/v) bovine serum albumin (BSA) in PBS. After washing with wash buffer, the plates were incubated with 100 µl of concentrated medium

samples at COL6A3 concentration of 1.3 ng/ml in assay buffer (0.05% Tween 20 and 0.5% BSA in PBS) for 2 hours. Plates were then washed four times with wash buffer and incubated with rabbit anti-COL6A3 biotin-conjugated antibody (Assay Genie) at 0.2 µg/ml in assay buffer for 1 hour. Plates were washed, after which the plates were incubated with streptavidin–horseradish peroxidase (Thermo Scientific) at 0.1 µg/ml in assay buffer for 1 hour. After washing, color development was performed with 100 µl of tetramethylbenzidine substrate (Thermo Fisher Scientific) for 10 min, reaction was stopped with 100 µl of 1 M HCl, and absorbance was measured at 450 nm. Assays were performed in triplicate.

Transmission electron microscopy

To the neo-cartilage organoids, consisting of cells and matrix, double concentrated fixative was added to the culture medium resulting in a final concentration of 1,5% glutaraldehyde solution in 0,1M cacodylate buffer. The spheres were kept in fixative for at least an hour at room temperature. After rinsing 3 times with 0,1M cacodylate buffer the spheres were postfixed in 1% osmium tetroxide / 1,5% potassium ferricyanide / 0,1M cacodylate buffer at 4°C for an hour. After 3 times rinsing with 0,1M cacodylate buffer the spheres were divided into 4 quarters. The quarters were dehydrated in a series of ethanol (70%, 80%, 90% and 100%), followed by an infiltration series of acetone / EPON LX112 (Ladd Research Industries) mixtures (2:1, 1:1 and 1:2), each step 30 min, and finally in pure EPON LX 112 for 1 hour. The quarters were positioned in a mold with the wide side towards the cutting surface, filled up with EPON and put in an 70 °C oven to polymerize for 48 hours. Ultrathin sections (90 nm) were made with a Leica EM UC6 ultramicrotome and collected on 50 mesh grids. Sections were stained with 7% uranyl acetate in MilliQ for 10 minutes and lead citrate (15). Sections were then imaged in a Tecnai T12 twin (FEI / Thermo Fisher Scientific) with a Gatan 4kx4k OneView camera (Gatan) at 21.000x magnification (at a pixel size of 1.03 nm) for machine learning analysis using automated tiled imaging (MyTEM) and images were stitched together into virtual slides (MyStitch) (16).

Image Analysis

For image analysis, in total 18 virtual slides were used from three isogenic control and three *COL6A3* mutant neo-cartilage organoids (3 positions each). Moreover, we have used supervised machine learning to quantify the differences in both GAG size and numbers between isogenic controls and COL6A3 mutants. Supervised machine learning was performed after manual annotation of GAG structures using a custom-written interface (manuscript in preparation), and training and prediction on 9 control and 9 mutation image stitches (of 10x10 images) using TensorFlow and Keras using a 2D adaptation of the 3D neural network. (17) The structures were predicted in all 18

slides (Fig. S6). Both the surface of the GAGs and their number was extracted from the predictions.

References

1. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. Genome Res. 2002;12(6):996-1006.
2. Adkar SS, Wu CL, Willard VP, Dicks A, Ettyreddy A, Steward N, et al. Step-Wise Chondrogenesis of Human Induced Pluripotent Stem Cells and Purification Via a Reporter Allele Generated by CRISPR-Cas9 Genome Editing. Stem Cells. 2019;37(1):65-76.
3. Lee J, Taylor SE, Smeriglio P, Lai J, Maloney WJ, Yang F, et al. Early induction of a prechondrogenic population allows efficient generation of stable chondrocytes from human induced pluripotent stem cells. FASEB J. 2015;29(8):3399-410.
4. Farndale RW, Sayers CA, Barrett AJ. A Direct Spectrophotometric Microassay for Sulfated Glycosaminoglycans in Cartilage Cultures. Connective Tissue Research. 1982;9(4):247-8.
5. Coutinho de Almeida R, Ramos YFM, Mahfouz A, den Hollander W, Lakenberg N, Houtman E, et al. RNA sequencing data integration reveals an miRNA interactome of osteoarthritis cartilage. Ann Rheum Dis. 2019;78(2):270-7.
6. Wu TD, Watanabe CK. GMAP: a genomic mapping and alignment program for mRNA and EST sequences. Bioinformatics. 2005;21(9):1859-75.
7. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31(2):166-9.
8. Ewels P, Magnusson M, Lundin S, Kaller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics. 2016;32(19):3047-8.
9. Leek JT. svaseq: removing batch effects and other unwanted noise from sequencing data. Nucleic Acids Res. 2014;42(21):e161.
10. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.
11. Hochberg YBaY. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B (Methodological). 1995;Vol. 57:pp. 289-300.
12. Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. Stat Appl Genet Mol Biol. 2005;4(1):Article17.
13. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9(1):559.
14. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic acids research. 2019;47(D1):D607-D13.
15. Reynolds, S. E. The use of lead citrate at high pH as an electron-opaque stain in electron microscopy. The Journal of cell biology. 1963;17.
16. Faas FG, Avramut MC, van den Berg BM, Mommaas AM, Koster AJ, Ravelli RB. Virtual nanoscopy: generation of ultra-large high resolution electron microscopy maps. J Cell Biol. 2012;198(3):457-69.
17. Xiao C, Chen X, Li W, Li L, Wang L, Xie Q, et al. Automatic Mitochondria Segmentation for EM Data Using a 3D Supervised Convolutional Network. Front Neuroanat. 2018;12:92.