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

Bloks, N.G.C.

PETER M.G COX *1988*

CHAPTER 3

The role of DNA methylation in chondrogenesis of human iPSCs as a stable marker of cartilage quality

Ghazaleh Hajmousa[1]†, Rodrigo Coutinho de Almeida[1]†, **Niek G.C. Bloks**[1]†, Alejandro Rodríguez Ruiz[1], Marga Bouma[2], Roderick Slieker[3], Thomas B Kuipers[3], Rob GHH Nelissen[4], Keita Ito[5], Christian Freund[2], Yolande FM Ramos[1]§, Ingrid Meulenbelt[1]§*

† These authors contributed equally

[1]Dept. Biomedical Data Sciences, Section Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands
[2]Department of Anatomy and Embryology and Human iPSC Hotel, Leiden, The Netherlands
[3]Department of Cell and Chemical Biology, Leiden University Medical Center, Leiden, The Netherlands.
[4]Department of Orthopedics, Leiden University Medical Center, Leiden, The Netherlands.
[5]Orthopaedic Biomechanics, Dept. of Biomedical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands

CLINICAL EPIGENETICS *2024, 16(1), 141 - DOI: 10.10.1186/s13148-024-01759-y*

Abstract

Background

Lack of insight into factors that determine purity and quality of human iPSC (hiPSC)-derived neo-cartilage precludes applications of this powerful technology towards regenerative solutions in the clinical setting. Here we set out to generate methylome-wide landscapes of hiPSC derived neo-cartilages from different tissues-of-origin and integrated transcriptome-wide data to identify dissimilarities in set-points of methylation with associated transcription and the respective pathways in which these genes act.

Methods

We applied *in vitro* chondrogenesis using hiPSCs generated from two different tissue sources; skin fibroblasts and articular cartilage. Upon differentiation towards chondrocytes these are referred to as hFiCs and hCiC, respectively. Genome-wide DNA methylation and RNA sequencing datasets were generated of the hiPSC-derived neo-cartilages, and the epigenetically-regulated transcriptome was compared to that of neo-cartilage deposited by human primary articular cartilage (hPAC).

Results

Methylome-wide landscapes of neo-cartilages of hiPSCs reprogrammed from two different somatic tissues were 85% similar to that of hPACs. By integration of transcriptome-wide data, differences in transcriptionally active CpGs between hCiC relative to hPAC were prioritized. Among the CpG-gene pairs lower expressed in hCiCs relative to hPACs, we identified genes such as *MGP, GDF5,* and *CHAD* enriched in closely related pathways and involved in cartilage development that likely mark phenotypic differences in chondrocyte states. Vice versa, among the CpG-gene pairs higher expressed, we identified genes such as *KIF1A* or *NKX2-2* enriched in neurogenic pathways and likely reflecting off target differentiation.

Conclusions

We did not find significant variation between the neo-cartilages derived from hiPSCs of different tissue sources, suggesting that application of a robust differentiation protocol such as we applied here is more important as compared to the epigenetic memory of the cells of origin. Results of our study could be further exploited to improve quality, purity, and maturity of hiPSC derived neo-cartilage matrix, ultimately to realize introduction of sustainable, hiPSC derived neo-cartilage implantation into clinical practice.

Introduction

Osteoarthritis (OA) is an age-related, heterogeneous joint disease, with degeneration of articular cartilage as an important hallmark of the pathophysiological process (120). Clinical manifestations of OA are synovitis, chronic pain, stiffness, and joint deformities with associated dysfunction of the joint and loss of mobility that severely hampers the daily activities of patients (121). Despite the critical impact on patients, there are currently no effective treatments that either stop or reverse OA except for costly joint replacement surgery at end-stage disease. Drawback of such replacements is the relatively limited lifespan of the implants, frequently leading to postponed surgery of patients particularly between 50-70 years of age with concurrent prolonged chronic pain and absence from work.

A major hurdle for the development of therapy in OA is the deficient inherent repair capacity of chondrocytes, the single-cell type present in articular cartilage (122). On the other hand, articular cartilage is considered immunotolerant and therefore eligible for regenerative treatment strategies such as implantation of neo-cartilage (123). To date, neo-cartilage implantation has mainly relied on autologous cell sources such as chondrocytes or mesenchymal stem cells (MSCs). Indeed, we showed previously *in vitro* that differentiated human primary articular chondrocytes (hPAC) readily deposit neo-cartilage that exhibits a DNA methylation landscape with an almost identical (99% similarity) profile to autologous cartilage (110). Upon expansion, chondrocytes are, however, prone to dedifferentiate and lose their characteristic articular phenotype, resulting in deposition of a non-specific, mechanically inferior extracellular matrix which is less suitable for application in larger defects (124). Similarly, autologous MSCs as chondrocyte precursor source is subjected to limited proliferation capacity resulting in decreased chondrogenic potential. Even more important, it has been generally acknowledged that due to its developmental fate, MSC-derived neo-cartilage is prone to deposit hypertrophic, mineralized cartilage (125).

To accomplish strong, durable, neo-cartilage implants at large scale, human induced pluripotent stem cells (hiPSCs), emerge as an exciting novel cell source. hiPSCs have infinite expansion capacity enabling the generation of large quantities of differentiated chondrocytes for cartilage regeneration (126), while their use is minimally invasive (127). To accommodate hiPSCs as sustainable source for articular cartilage regeneration, differentiation protocols resulting in human induced chondrocytes that readily deposit cartilaginous tissue have been developed (128, 129). Nonetheless, lack of insight, hence solutions, into aberrant cell fate decisions during hiPSC chondrogenesis hence quality and purity of deposited neo-cartilages, precludes application of this powerful technology towards regenerative solutions in the clinical setting. Such cell fate decisions

are particularly facilitated by deposited methylation at transcriptionally active CpG sites that evoke stable set-points of transcription (130). Moreover, it was shown that the majority of such lineage-specific methylation patterns remain stable throughout life (131). Owing to the notion that such an epigenetic memory of the somatic cell types exists (132-135), hiPSCs derived from chondrocytes may differentiate more readily and with higher similarity to articular cartilage than hiPSCs derived from cells of other tissues.

In the current study, we set out to test whether the epigenetic memory of the somatic cell type used for generation of hiPSCs is retained as a unique DNA methylation signature. Hereto, we applied established robust stepwise chondrogenesis protocol via mesodermal lineage (129) using hiPSCs from two different tissue sources, namely skin (fibroblasts) and articular cartilage (chondrocytes). *In vitro* hPAC derived neo-cartilage was used as the golden standard (110). To assess similarities, the methylome wide landscape of neo-cartilages derived from these different cell sources, were generated and compared. To subsequently explore and biologically interpret differences in cell identity between hPACs and hiPSC derived chondrocytes in the neo-cartilages, we set out to specifically study discordant set points of DNA methylation that likely act on expression of positional genes. Since this layer of molecular information is key in critical lineage decisions, hence cell fates (130) data generated in this study could pave the way for large-scale hiPSCs derived neo-cartilage formation of superior quality.

Methods

Tissue culture and chondrogenesis

Cell culture of hiPSCs and primary articular chondrocytes

To analyze potential differences in chondrogenic differentiation potential among hiPSCs generated from different cell sources (skin fibroblasts and articular cartilage; **Figure 1A**), five independent hiPSCs lines/clones were used in the current study. All hiPSC cell-lines were generated by the LUMC hiPSC Hotel while applying transformation with polycistronic lentivirus to induce expression of the Yamanaka factors as outlined in detail by Dambrot et al.(136). The two skin fibroblast derived hiPSCs have been officially registered at the Human pluripotent stem cell registry (hPSC^{REG}) and are from healthy donors without known genetic diseases. Lines are from a female donor (LUMC0030iCTRL12 or hPSCreg line LUMC004-B) and from a male donor (LUMC0004iCTRL10 or hPSCreg line LUMC029-B). The three chondrocyte derived hiPSCs were generated from macroscopically preserved chondrocytes of a female OA donor who underwent a joint replacement surgery, from the RAAK study (137) and are

referred to as LUMC0131iCTRL02, LUMC0131iCTRL04, and LUMC0131iCTRL05. The generation of the hiPSCs line was approved by the Leiden University Medical Centre ethical committee under P13.080. Cells were characterized according to pluripotent potential and spontaneous differentiation capacity by the hiPSC core Hotel (136) and were karyotyped after 15 passages in culture (**Figure S1**). hiPSCs were maintained under standard conditions (37°C, 5% CO_2) in TeSR-plus medium (STEMCELL Technologies) on VitronectinXF-coated plates (STEMCELL Technologies). The medium was refreshed three times a week, and cells were passaged in aggregates using Gentle Cell Dissociation Reagent (STEMCELL Technologies) upon reaching approximately 80% confluency. hiPSCs were used for our study between passages 21 and 27. Human primary articular cartilage was collected from the hip of N=10 different OA patients that underwent joint replacement surgery as part of the RAAK study (137). Of note, classification of OA as macroscopically preserved or lesioned for collection, expansion, and differentiation of the primary articular chondrocytes has been previously described (110). Ethical permission for the described studies was obtained from the appropriate medical ethical committee under protocol numbers P08.239 and P19.013. Written informed consent was obtained from all participants.

hiPSC differentiation to human chondroprogenitor cells

Generation of chondroprogenitor cells was based on a protocol previously described (129). 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). Five independent differentiations were performed per clone, and two neo-cartilages were pooled for the RNA and DNA isolations. **Figure 1B** shows a schematic overview of the timelines and times of collection for the analyses.

Chondrogenic differentiation of chondroprogenitor cells

Monolayer cultured chondroprogenitor cell aggregates present at day 14 of the differentiation were washed with MD medium, dissociated using Gentle Cell Dissociation Reagent (STEMCELL Technologies), and centrifuged for 4 min at 1200 rpm. Cell aggregates (250,000 cells per pellet) were subsequently maintained in Dulbecco's modified Eagle's medium/F12 (Gibco), supplemented with 1% ITS+, 55 μM 2-mercaptoethanol (Gibco), 1% non-essential amino acids (Gibco), 0.5% P/S, L-ascorbate-2-phosphate (50 μg/ml; Sigma-Aldrich), L-proline (40 μg/ml; Sigma-Aldrich), and transforming growth factor–β1 (10 ng/ml; PeproTech) for 7/21/35 days while refreshing medium every 3 to 4 days.

Histology and immunohistochemistry

3D neo-cartilage were fixed in 4% formaldehyde overnight and stored in 70% ethanol at 4°C. The pellets were embedded in paraffin and sectioned at 5 μm. After sectioning, slides were deparaffinized and rehydrated prior to histology or immunohistochemistry. Overall cellular and tissue structure was visualized with hematoxylin–eosin (HE) staining. Glycosaminoglycan depositions were visualized by staining with 1% Alcian Blue 8-GX (Sigma-Aldrich) and Nuclear Fast red staining (Sigma-Aldrich). Alcian blue staining was quantified by loading the images in Fiji and splitting the color channels. Subsequently, gray values were measured of three to five separate squares per pellet and corrected for the gray value of the background. To detect collagen type ll (1:100; ab34712, Abcam) immunohistochemistry was performed with 3-diaminobenzidine (DAB) solution (Sigma-Aldrich) and hematoxylin (Klinipath) as described before (110). Pixel intensity quantification was performed by Fiji, and surface area of the pellets were measured with the CellSens Dimension software (Olympus, Leiderdorp, The Netherlands).

RNA analyses

RNA isolation and RT-qPCR

For RNA isolations neo-cartilages at day 35 of differentiation were used. Hereto, two pellets were pooled, and isolation was performed as described previously (110). Total mRNA (150ng) was processed with a first strand cDNA kit according to the manufacturer's protocol (Roche Applied Science). cDNA was further diluted five times, and preamplification with TaqMan preamp master mix (Thermo Fisher Scientific Inc.) was performed for genes of interest. Gene expression was measured with RT-qPCR,

and average of the two biological replicates was determined as relative levels (−ΔCt values) using expression levels of glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) and Acidic ribosomal phosphoprotein P0 (*ARP*) as housekeeping genes. $P < 0.05$ were considered statistically significant. Quality control of the results was performed as described before (18).

RNA sequencing

RNA was extracted from chondrocytes isolated from neo-cartilages at day 14 (hPACs) and day 35 (hCiCs and hFiCs) of differentiation. Hereto, RNA was extracted with chloroform and purified using the RNeasy Mini Kit (QIAGEN). RNA sequencing (polyA enriched) of 22 total RNA samples (hCiC; n=9, hFiC; n= 3, and hPAC; *n*= 10 derived neo-cartilage) was performed using Illumina NovaSeq6000 sequencing, according to the standard operating procedures based on the Illumina protocol for Paired-End Sequencing (Per sample ~6 Gb, 20 million Paired-End reads RNA-sequencing mapping and quality control (QC) was performed using the in house pipeline BioWDL (https://biowdl.github.io/RNA-seq/v4.0.0/index.html) developed by the SASC team at Leiden University Medical Center. This pipeline was used to process FastQ files using Picard. v2.23.2 and Samtools.v1.10, adapter clipping using cutadapt.v2.10. Moreover, QC was performed using FastQC.v0.11.9 and MultiQCv. 1.9. Furthermore, mapping was performed with STAR.v2.7.5a software (138) and expression quantification and transcript assembly using HTSeq-Count.v.0.12.4 (139). The reads were aligned to the reference transcriptome Ensemble GRChr38 release 102. RNA-seq data was normalized using the DESeq2_v.1.30.0 R package (140). The data was further transformed using the variance-stabilizing transforming (VST) method.

Differential expression analysis

Differential expression analysis was performed using the DESeq2 package v. 1.30.0 using R version 4.0.2. A general linear model (GLM) assuming a negative binomial distribution was applied followed by a Wald-test between hCiC and hPAC samples. Benjamini-Hochberg multiple testing corrected P-values with significance cut-off of 0.05 are reported as False Discovery Rate (FDR).

DNA analyses

DNA isolation

Snap frozen neo-cartilage was powdered using a Mixer Mill 200 (Retsch, Germany) with continuous liquid nitrogen cooling. DNA was isolated using the Wizard Genomic DNA Purification kit (Promega) according to the manufacturer's protocol.

Methylation analysis

DNA methylation was assessed in chondrocytes isolated from neo-cartilages at day 14 (hPACs) and day 35 (hCiCs and hFiCs) of differentiation. Hereto, the Illumina Infinium Methylation EPIC (850K) BeadChip were used according to standard operating procedures based on the Illumina Infinium II Protocol. To analyze methylation array data (MethylationEPIC 850k array) the MethylAid R script (141) with the default settings, was used for data quality assessment. All samples showed a detected CpG above 95%. The minfi.v_1.36.0 R package (142) was used to pre-process the data. We removed any probe that failed in one or more samples ($p < 0.01$). Probe level intensities were quantile normalized across samples prior to calculation of the ß-values.
MethylToSNP v0.99.0 R package was used to filter SNPs. This method looked for patterns in Methylation array data and identified methylation probes with SNP-like patterns. This method allows to remove outliers, which adds robustness to the analysis and is enabled by default. A confidence score was calculated to show how close the observed pattern of methylation beta values was to a canonical case of a SNP in a homozygously methylated CpG locus. Additionally, MethylToSNP can overlap the SNPs identified in methylation data with known SNPs from dbSNP. The probes that have shown to be cross-reactive (demonstrated to map to multiple places in the genome) were filtered out (143). The probes that were overlapping with rare SNPs (probes in TFBS that showed extreme methylation pattern) were filtered out (144). To minimize the unwanted variation within and between samples, we used the Functional Normalization method from the minfi.1.36.0 R package (145).

Differential methylation analysis

We run differential mean analysis using t-moderated statistics. Using the MEAL.1.20.3 R package pipeline, which, relies on the lmFit from limma R package (design model = ~ phenotype). CpGs after Bonferroni correction $P < 6.243109e-08$ (0.05/800883) was considered significant.

Similarities

To identify similarities between the cells we used the Jaccard method with the jaccard_0.1.0 R package. Visualization was done using the corrplot_0.84 R package.

Integration methylation and RNA-seq

To identify transcriptive active CpGs, we integrated Significant CpGs with expression of genes using the MEAL_v1.20.3 R package. The ß-values of the methylation data were used. We adapted the CorrelationMethExprs function, which estimates the correlation between methylation and expression. For each CpG, a range was defined by the position of the CpG plus the flank parameter (250kb upstream and 250kb downstream). Only

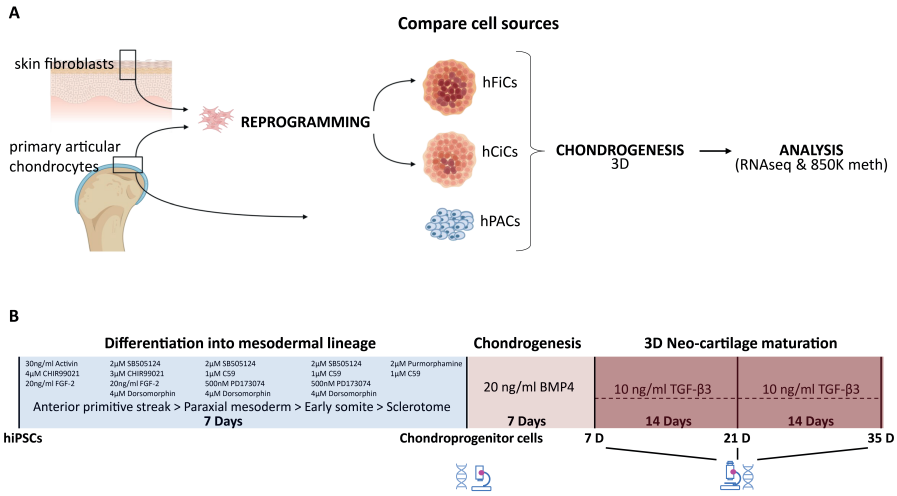


Figure 1 | Generation of 3D neo-cartilage derived from hiPSCs sources relative to hPAC. (**A**) Schematic representation of the study. Samples consisting of neo-cartilage from human chondrocyte-derived iPSCs (hCiC, N=14), neo-cartilage from human fibroblast-derived iPSCs (hFiC, N=3), and neo-cartilage from human primary articular chondrocytes (hPAC, N=10). (**B**) Timelines and times of collection for the analyses of stepwise differentiation from hiPSCs towards paraxial mesoderm lineage and chondrocyte-like cells.

those expressed genes that were entirely in this range were selected. After multiple testing correction by False Discovery Rate (FDR) and considering r2 > 0.5, CpGs correlated with genes were shown.

Gene ontology enrichment analysis

Gene ontology (GO) enrichment analysis was performed using clusterProfiler v4.0.5 R package to identify GO terms. Bonferroni multiple testing-corrected P-values with a significance cut-off of 0.05 are reported as FDR.

Statistical analysis

Statistical analyses were performed in R version 4.0.2 (2020-06-22) and using SPSS version 25 (IBM). Results were considered significant for FDR ≤ 0.05.

Results

Generation and characterization of 3D neo-cartilage derived from different hiPSC sources relative to hPACs

After confirming adequacy of hiPSCs morphology, pluripotency, spontaneous differentiation capacity, and a normal karyotype (136)(**Figure S1A-D**), human skin fibroblast-derived iPSCs (two cell-lines; from two healthy donors) and human cartilage-

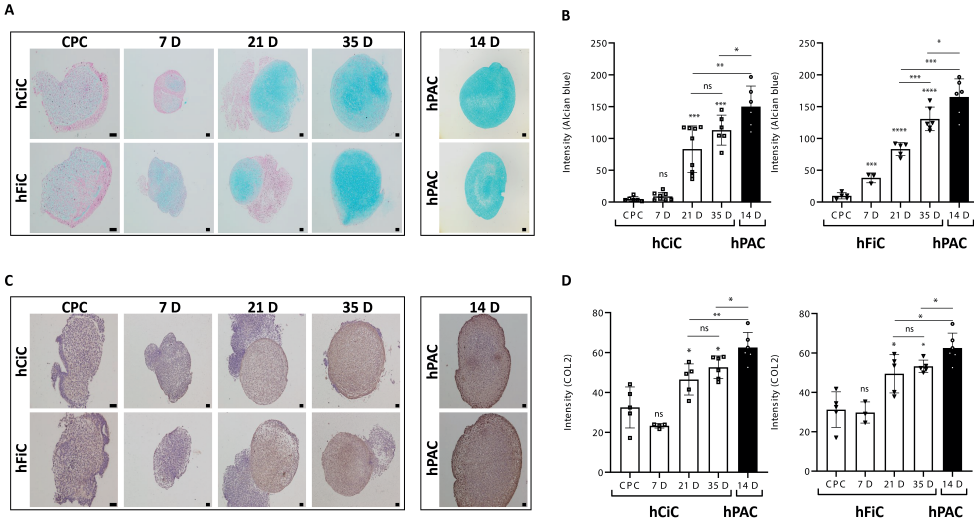
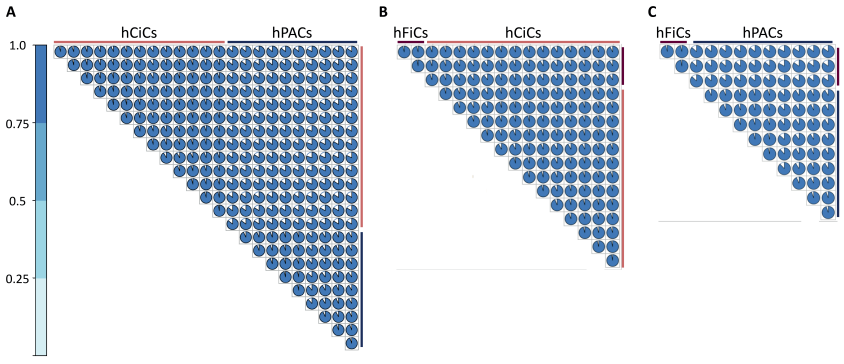


Figure 2 | Characterization of 3D neo-cartilage derived from hiPSCs sources relative to hPAC. (A) Representative **histological staining images of the neo-cartilages by Alcian Blue staining showing glycosaminoglycan deposition. (B**) Quantification of Alcian Blue histology of hCiC and hFiC, respectively, in different time-points of differentiation compared to hPAC. (**C**) Representative immunohistochemistry images of neo-cartilage stained with Collagen-2 antibody. (**D**) Quantification of Collagen-2 immunohistochemistry of hCiC and hFiC, respectively, in different time-points of differentiation process compared to hPAC. Scale bars at day 7: 50 µm, other scale bars: 100 µm. All data are presented as mean ± STD. Two-tailed Student's t-tests were used to compare 1) each group relative to chondroprogenitor cells (stars on top of the bars) and 2) each group relative to hPAC (indicated by lines). *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001 (N=6-9 pellets each).

derived iPSCs (three independent clones hence hiPSC lines; from macroscopically preserved cartilage chondrocytes of a patient who underwent joint replacement surgery due to OA) were differentiated towards chondrocytes (**Figure 1**). Cellular and tissue structure during formation of the hiPSCs derived neo-cartilages in comparison to that of hPACs was visualized by hematoxylin and eosin (H&E) staining at day 0, 7, 21, and 35 of differentiation from chondroprogenitor cells. At day 35, HE staining of both human cartilage-derived iPSCs differentiated to chondrocyte (hCi-Chondrocyte: hCiC) and human fibroblast-derived iPSCs differentiated to Chondrocyte (hFi-Chondrocyte: hFiC) showed similar matrix structure as compared to hPAC controls (**Figure S2A**). Nevertheless, until day 21 of chondrogenic differentiation of both hCiC and hFiC also formation of off-target tissues was detected in developing neo-cartilage pellets as reflected by the darker purple edge of the neo-cartilages. Staining with cytokeratin (CK) showed that some of the off-target cells were keratinocytes. These were largely decreased at day 21 and no longer apparent at day 35 of neo-cartilage generation (**Figure S2B-D**).

To validate glycosaminoglycan deposition in hCiC and hFiC relative to hPACs, Alcian blue staining was performed and quantified (**Figure 2A**). Neo-cartilage from both sources of hiPSCs showed increasing amounts of glycosaminoglycan deposition, with 23-fold increase in hCiC, and 13-fold increase in hFiC from chondroprogenitor stage to day



D

Comparison	Average similarity*	SD
hFIC - hFIC (N=3)	0.96	0.006
hCIC - hCIC (N= 13)	0.94	0.014
hPAC-hPAC (N=10)	0.92	0.024

*Jaccard similarity score

E

Comparison	Average similarity*	SD
hFIC (N=3) - hPAC (N=10)	0.86	0.012
hCIC (N=13) - hPAC (N=10)	0.84	0.013
hCIC (N=13) - hFIC (N=3)	0.92	0.014

*Jaccard similarity score

Figure 3 | DNA methylation landscape of 3D neo-cartilage. (**A**) Plot showing the Jaccard similarity within hCiC (N=14), within hPAC (N=10), and between hCiC vs hPAC . (**B**) Plot showing the Jaccard similarity within hCiC (N=14), within hFiC (N=3), and between two sources of hiPSCs-derived neo-cartilage hCiC vs hFiC . (**C**) Plot showing the Jaccard similarity within hFiC (N=3), within hPAC (N=10), and between hFiC vs hPAC. (**D**) Summary of the average Jaccard similarity index within the hCiC, hFiC and hPAC cell sources upon independent consecutive rounds of differentiation. (**E**) Summary of the average Jaccard similarity index between the hCiC, hFiC, and hPAC cell sources.

35 of differentiation (**Figure 2B-C**). Nonetheless, the final level of glycosaminoglycan deposition in hPACs derived neo-cartilage was 30% higher compared to that in hCiC (P<0.05) and 20% higher compared to that hFIC (P<0.05; **Figure 2B**). Similarly, deposition of cartilaginous extracellular matrix determined by immunohistochemistry of collagen 2 protein (COL2; **Figure 2D**) showed a significant increase from chondroprogenitor to day 35 of neo-cartilage across hiPSCs sources. However, final expression levels of COL2 in neo-cartilage in hPAC was 10% higher relative to both hCiC (P=0.05) and hFiC (P=0.05).

Next, neo-cartilage derived from hFiC and hCiC in comparison to hPAC were characterized by targeted RT-qPCR gene expression of well-known chondrogenic (*COL2A1, SOX9 and ACAN*), fibrotic (*COL1A1*), and hypertrophic (*COL10A1*) cartilage markers at day 35 of differentiation. In line with the histology, this showed that hPACs as compared to hiPSC-derived chondrocytes express significant higher levels of chondrogenic markers *COL2A1, SOX9 and ACAN*, but also of hypertrophic marker *COL10A1*, and fibrotic marker *COL1A1* (**Figure S3**). However, the COL2:COL1 ratios appeared highly comparable among hPACs, hCiCs and hFiCs (**Figure S3).** Together, it can be concluded that neo-cartilage from both hiPSC sources at day 35 of differentiation have high similarity to the hPAC control pellets, and was therefore used for further study.

Methylome wide landscape of 3D neo-cartilage derived from hiPSCs sources relative to hPAC

Founded by the hypothesis that the epigenetic memory of the hiPSCs cartilage tissue-of-origin could contribute to chondrogenic potential of hiPSCs during differentiation, genome-wide DNA methylation datasets were generated from hiPSC derived neo-cartilage at day 35 of differentiation and from hPACs at day 14 of differentiation. Following quality control (QC), we obtained robust methylation data of 800,883 CpGs in total. As shown in **Figure 3A-C**, we observe high overall Jaccard similarity scores within and between the DNA methylome wide landscape from neo-cartilages of different cell sources. More specifically, as outlined in **Figure 3D** Jaccard similarities of 96% and 94% were observed in the DNA methylome wide landscape from neo-cartilages of independent differentiations of, respectively, hFiC (N=3) or hCiCs (N=13), reflecting high consistency in the hiPSC chondrogenic differentiation protocol. Also, a Jaccard similarity of 92% was observed in the DNA methylome wide landscape of neo-cartilages generated from hPACs of N=10 independent donors, indicating high consistency and similarity in the *in vitro* neo-cartilages deposited by hPACs isolated from heterogeneous preserved cartilages of OA patients. As summarized in **Figure 3E**, the Jaccard similarity between the methylome of the neo-cartilage of the hCiC and hFiC cell sources was also high with 92%, indicating consistency in the hiPSC chondrogenic differentiation protocol between hiPSC clones even when generated from different cell origins. Finally, in **Figure 3E**, we outlined the average Jaccard similarities of 86% and 84% between the methylome wide landscape of the neo-cartilages deposited by hPACs and, respectively, hFiC and hCiC. Together these data indicate that the applied hiPSC chondrogenic differentiation protocol is highly consistent within (~95%) and between (~92%) hiPSC lines, and that it resulted in very similar neo-cartilage as deposited by primary hPACs (~85%). Our data suggest that the epigenetic memory of hCiC (hiPSCs from cartilage as tissue-of-origin), did not have an additional positive effect on the chondrogenic differentiations. Despite the overall high similarities, principal component (PC) analysis on the methylome wide landscape showed three specific clusters according to cell sources. Herein, hCiCs and hFiCs separated from hPACs by particularly PC1 explaining 96% of the variance and hCiCs separated from hFiCs by PC2 explaining 2% of the variation (**Figure S4A**). Furthermore, Euclidean clustering analyses of the methylation of 1,000 most variable CpGs, confirmed separation of hPACs, hCiCs, and hFiCs (**Figure S4B**). These data indicated that the methylome landscape of neo-cartilages, has also specific characteristics that reflect differences in the cell identity of the hCiC, hFiC or hPAC derived chondrocytes.

Transcriptome wide landscape of 3D neo-cartilage derived hiPSC sources relative to hPAC

To characterize and biologically interpret differences in the methylome wide landscapes of neo-cartilage deposited by hiPSC derived chondroprogenitors relative to hPACs, we next set out to prioritize on differences in the methylome that are most likely transcriptionally active. Hereto, we performed RNA sequencing of neo-cartilages derived from hPACs (N=7, day 14 of differentiation), hCiCs (N=9, day 35 of differentiation), and hFiCs (N=3, day 35 of differentiation). Prior to assessing the differences in methylation that are most likely transcriptionally active, we first performed an exploratory analyses of the normalized VST values of the transcriptome of hCiCs relative to hPACs. Expression data of hFiCs is provided in **Table S1**, however, because of the low number of samples (N=3) and the high similarity between the two hiPSCs cell sources, the hFiC transcriptomic data was not included in further downstream analyses. As shown in **Figure S5**, the average expression levels of all genes in hPAC (Ave Exp = 4.9) was slightly lower as compared to hCiC (Ave Exp = 5.4) while the range in expression levels in hPACs (between 1.8-16) was larger as compared to hCiC (between 2.0-14). Moreover, as shown by the histogram the hPACs showed a large number of genes that are very lowly expressed (**Figure S5**).

Upon subsequently assessing relative gene expression levels defined as the standard deviation of z-scores with SD1 being lowest expressed genes (<SD2) and SD4 being highest expressed genes (>SD2) we could explore genes that mark authentic human primary chondrocytes expression patterns in hPACs relative to that in hCiCs (**Table S1**). The STRING protein-protein interaction network of top 50 of 919 genes that were highly expressed in hPAC (>SD2) and hCiC (>SD2), showed a highly interconnected dense network with significant gene enrichment in pathways involved in collagen fibril organization (GO:0030199, P=9.3x10^{-10}), extracellular matrix organization (GO:0030198, P=2.7x10^{-8}), and cell adhesion molecular binding (GO: 0050839, P=6.3x10^{-8}), represented by well-known cartilaginous genes such as *FN1*, *COL6A1/2, TNC, COL2A1*, and *MALAT1* (146), confirming quality of neo-cartilage deposited by hiPSCs derived chondroprogenitors (**Figure S6**). On the other hand, the STRING protein-protein interaction network of the n=69 genes that were highly expressed in hPAC (>SD2) but lowly expressed in hCiC (<SD2) showed a network with notable significant gene enrichment in pathways involved in skeletal system development (GO:0001501, P=8.0x10^{-4}), and Heparin binding (GO: 0008201, P=5.9x10^{-5}), represented by relevant cartilaginous genes such as *GDF5*, *TGFBR3,* and *CILP* (**Figure S7A**). Finally, the STRING protein-protein interaction network of the n=124 genes that were highly expressed in hCiCs (>SD2) but lowly expressed in hPAC (>SD2) showed significant gene enrichment in pathways involved in nervous system development (GO:0007399, P=2.5x10^{-21}), represented by highly interconnected genes such as *SOX2*, *SOX11*, *KIF1A,* and *STMN4* (**Figure S7B**).

Discordant aspects of integrated methylome and transcriptome patterns between 3D neo-cartilage from hCiC relative to hPAC

To identify differences in the methylome that are most likely transcriptionally active, we integrated differential transcriptome wide RNA sequencing data of hPACs (N=7) and hCiCs (N=9) neo-cartilage to the differential methylome wide data. In total N=94,771 significant differentially methylated CpGs (DM, P<0.05 after Bonferroni correction) and N=7,251 differentially expressed genes (DEG, FDR<0.05) were identified. To identify herein CpG-gene pairs that, most likely, reflect aberrant methylation set-points of gene expression in hCiCs, we subsequently selected CpG sites that had a singular mapping to a UCSC reference gene and had a significant high correlation ($r > 0.5$ and FDR < 0.05) between methylation and gene expression (**Figure S8**). As such we prioritized 2,378 discordant CpG-gene pairs that were highly interconnected and marked potential discordant set-points of gene expression between hCiC and hPACs (**Table S2**). Among these discordant CpG-gene pairs, the direction of differential methylation at CpG sites and gene expression was positive in N=722 (30%) and inverse in N=1656 (70%; **Table S2**). In the circos plot in **Figure 4A** the distribution of all discordant CpG-gene pairs across the genome are plotted while highlighting the top 15 most significant ones such as CpGs near *COMP* encoding cartilage oligomeric protein, *MMP7* encoding metalloproteinase 7 and the long noncoding RNA *HOTAIR*. Moreover, to visualize extend of FDR significant levels of differential methylation (**Figure S9**) and expression (**Figure S10**) of discordant CpG-gene pairs across the genome, we plotted FDR significant levels as Manhattan plots. As shown in **Figure S9** and **Figure S10**, multiple highly significant CpG-gene pairs were recognized. Among them, notable CpGs near OA risk genes (147) such as *MGP* encoding Matrix-Gla protein or *TNC* encoding Tenacin, and notable genes marking OA pathophysiology (70, 148) such as *TNFRSF11B* encoding TNF Receptor Superfamily Member 11b, *P4HB* encoding prolyl 4-hydroxylase subunit beta, *SMOC2* encoding SPARC related modular calcium binding 2, and *MMP3* encoding metalloprotease 3 (**Table S2**). Additionally, we noted among the highly significant differential CpG-gene pairs *CTNNB1* encoding beta-catenin. Together, by prioritizing on highly significant differential methylation at likely transcriptionally active CpG, we showed that differences in cell identity between hPACs and hCiCs is characterized by notable genes involved in OA (etio)pathophysiology (**Table S2**).

Pathway enrichment analyses of 2,378 differential set-points of gene expression between 3D neo-cartilage from hCiC relative to hPAC

To obtain insight into the pathways in which the CpG-gene pairs act that mark differences in cell identify between hPACs and hCiC, we next performed gene enrichment pathway analyses for genes lower (N=1237, **Table S3**) and higher (N=1141, **Table S4**) expressed in hCiCs relative to hPACS. Among the significant pathways of the lower expressed

genes in hCiC relative to hPACs (**Figure 4B**), we recognized closely related GO-terms such as 'extracellular matrix organization' (GO:0030198, FDR=7.7x10^{-18}) with diverse cartilage component genes such as *ACAN, COMP, DCN, COL2A1,* and *COL11A1*, 'chondrocyte differentiation' (GO:0002062, FDR=3.9x10^{-7}) with *NPR2, GDF5, TWSG1,* and *SOX9.* The latter gene also represented in the enriched pathway 'cartilage development' (GO:0051216, FDR=3.1x10^{-4}) with *MGP*, *CHRDL2,* and *CHI3L1* as notable other genes. Also, in **Table S3** highly significant enriched KEGG pathway 'Focal adhesion' (hsa04510, FDR=1.1x10^{-9}), with *CHAD* and multiple integrin genes such as *ITGA5, ITGAV, ITGA9* and *ITGA10*, many of which were overlapping with the 'extracellular matrix' pathway. Together, these lower expressed genes in hCiC relative to hPACs likely mark immaturity and/or lower quality of hCiC neo-cartilage (**Figure 4B**, **Table S3**). On the other hand, some of the lower expressed genes in hCiC relative to hPACs such as *TNFRSF11B* and *COL1A2* concurrent with genes such as *CLEC3A*, and *OMD* acting in the ossification pathway (GO:0001503 'FDR=2.1x10^{-4}, **Table S4**) actually reflect an OA phenotypic state of the chondrocytes. As such, these could merely reflect pathologic changes in set-points of gene expression of the hPAC derived neo-cartilage relative to that of hCiCs. This, since hPACs were isolated from preserved cartilage of elderly patients. Among the enriched pathways of the higher expressed genes in hCiC relative to hPACs (**Figure 4C**, **Table S4**), we particularly recognized closely related GO-terms involved in neurogenesis being 'neuron differentiation' (GO: GO:0030182, FDR= 6.2x10^{-7}) with diverse neurogenic genes such as *NEUROG2, SOX11,* and *WNT* genes, the latter genes overlapping with the enriched pathway 'cell fate commitment' (GO:0045165, FDR=8.3x10^{-3}). Additionally, we observed a highly significant enrichment within the 'Axon guidance' pathway (GO:0007411, FDR=6.0x10^{-9}) with multiple semaphorin genes such as *SEMA3D, SEMA4C, SEMA5B,* and *SEMA6B.* Semaphorins represent a family of transmembrane and secreted neuron-guidance molecules. Furthermore, among the genes enriched in the Axon guidance pathway we recognized *NGFR,* encoding the nerve growth factor receptor, and *LHX2* encoding LIM Homeobox 2 that acts as a transcriptional activator of neural cell types. This may reflect off-target during differentiation of hiPSCs towards neurogenic lineage.

Prioritization of differential methylated CpGs that mark stable set-points of transcription and that could facilitate strategies to improve hCiC neo-cartilage formation

Next, we wanted to prioritize for most eligible differential set-points of the transcriptionally active methylation that could facilitate targeted modifications and improve hiPSC derived neo-cartilage formation. Hereto we prioritized among the 2,378 CpG-gene pairs those that had the largest fold change (FC) in gene expression, the largest FC in methylation, and had the largest correlation. Additionally, to select for sensitively modifiable target genes we removed 21 CpG-gene pairs that were discordant yet were highly expressed in both hPACs and hCiCs (>1 Standard deviation, **Table S5**). This resulted in 195 CpG-gene pairs

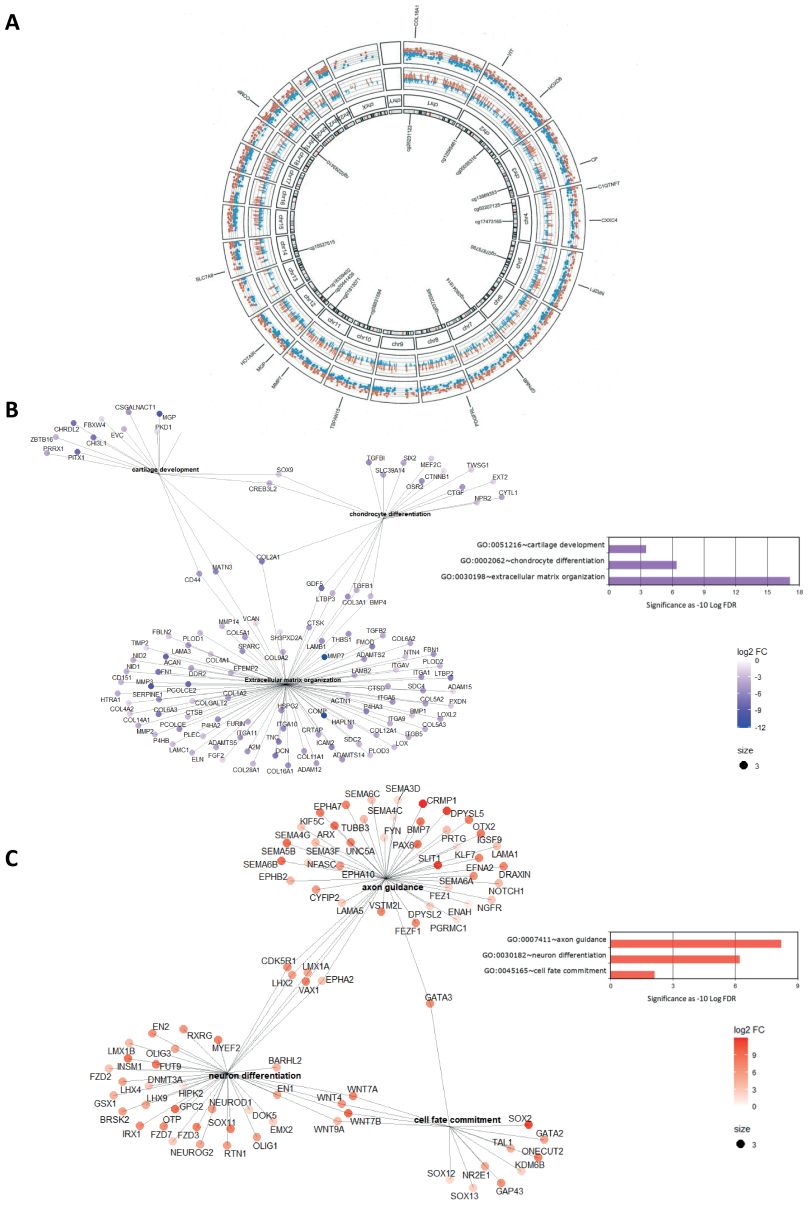


Figure 4 | Prioritized 2,378 discordant CpG-gene pairs marking discordant set-points of gene expression between hCiC and hPACs. (**A**) Circos plot showing the distribution of N=2378 differential CpG-gene pairs across the genome. Labeled are N=15 CpG-gene pairs with highest correlation between methylation and expression levels. The outer circle displays the gene expression. Blue shows downregulated genes, while red shows upregulated genes. The inside circle represents DNA methylation from CpGs that mapped in direct vicinity of DEGs, where blue is hypomethylated and red is hypermethylated. (**B**) Visualization of notable significant pathways enriched among 2,378 CpG-gene pairs lower and higher expressed in hCiCs relative to hPACs showing the links between genes and biological processes by using GO and KEGG networks. (**C**) Notable significant pathway enriched for CpG-gene pairs that were higher expressed in hCiCs relative to hPACs showing the links between genes and biological processes by using GO and KEGG networks. Size of the dots represents number of genes linked to each term. The color of the dots represents the fold change of the differentially expressed gene. P adjusted method= FDR depicted in the histograms.

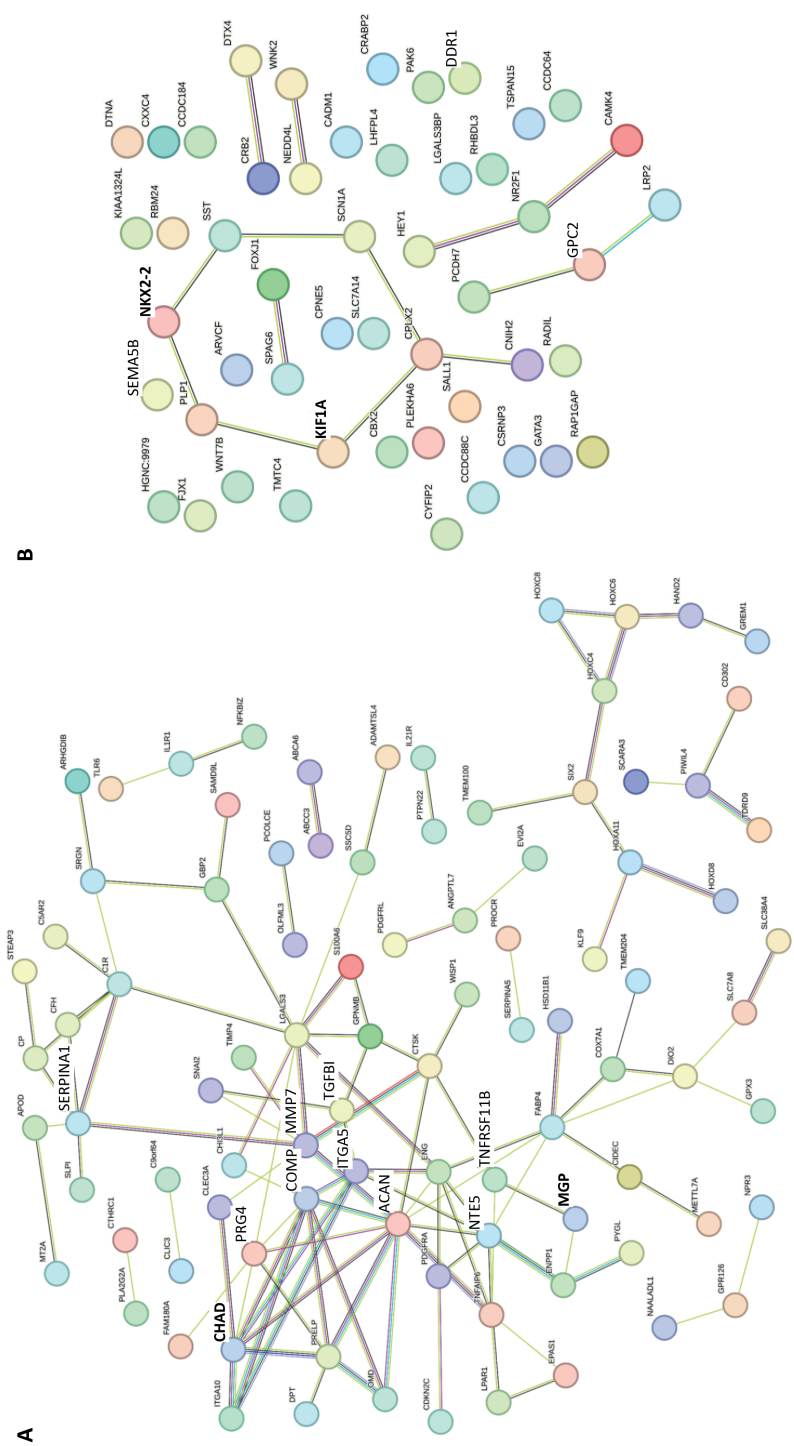


Figure 5 | STRING protein-protein interaction network of 195 eligible discordant set-points of gene expression that could facilitate strategies to improve hCiC neo-cartilage formation. (**A**) Significant STRING protein-protein interaction network ($P<1x10^{-16}$) among downregulated genes in hCiC derived neo-cartilage relative to hPAC (N=140). (**B**) Significant STRING protein-protein interaction network ($P=8.1x10^{-3}$) among of genes that were significantly upregulated in hCiC derived neo-cartilage relative to hPAC (N=55). The edges indicate both functional and physical protein associations, with selected genes of Figure 6 indicated in bold. Line thickness indicates the strength of data support. Minimum required interaction score = medium confidence (0.400).

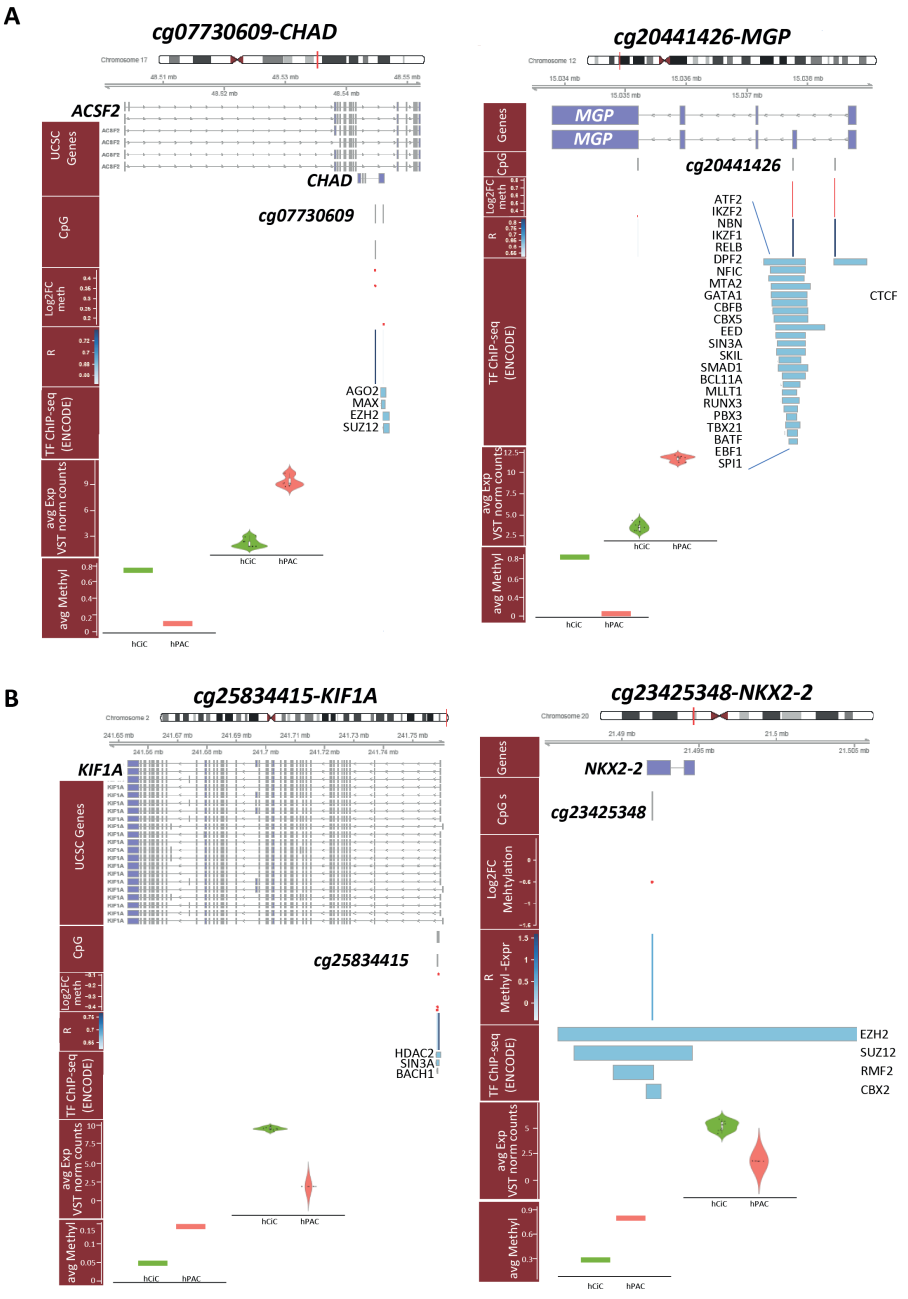


Figure 6 | Genomic plots of selected genes. Genomic plots of selected CpG-gene pairs eligible as epigenetically modifiable target genes with mapping of genes and CpG sites, Log2FC of methylation, correlation between methylation and expression, mapping of transcription factors binding sites (ChIP-seq ENCODE) as well as differences in expression and methylation between hCiCs and hPACs. (**A**) Genomic plots of cg07730609-CHAD and cg20441426-MGP with lower expressed in hCiCs relative to hPACs. (**B**) Genomic plots of cg25834415-KIF1A and cg23425348-NKX2-2 with higher expressed in hCiCs relative to hPACs.

with high differences in set-points of the transcriptionally active methylation that could reflect differences in phenotypic chondrocyte states (**Figure 5**, **Table S6**).

We recognized 140 CpG-gene pairs (72%) with lower and 55 CpG-gene pairs (28%) with higher gene expression in hCiCs relative to hPACs. To visually explore these potential epigenetically modifiable target genes, we generated STRING protein-protein interaction networks. **Figure 5A** represents lower expressed genes in hCiCs relative to hPACs with highly significant enriched protein-protein interactions (P=1.0x10^{-16}). Notable articular cartilage genes such as *PRG4, SERPINA1, ACAN, CHAD*, *NT5E,* and *MGP* that could mark cartilage immaturity were highly interconnected. Vice versa, **Figure 5B** represents higher expressed genes in hCiCs relative to hPACs, also with significant enriched protein-protein interaction (P=8.1x10^{-3}). Here, we recognize genes such as *SEMA5B, KIF1A, NKX2-2, DDR1* and *GPC2* involved in the generation and differentiation of neurons.

As proof of concept, we finally inspected the genomic region of four of these differential methylated CpGs that mark stable set-points of transcription. These genes were selected for being highly FDR significant, central and highly connected in the network (**Figure 5**), and being either low in hCiC derived neo cartilage relative to hPACs (skeletal development pathway i.e. *CHAD* and *MGP*) or high in hCiC derived neo-cartilage relative to hPACs (neurogenic pathway i.e. *KIF1A* and *NKX2-2*). As shown in **Figure 6A** the genomic regions of *CHAD*, *MGP* (lower expressed in hCiCs) and in **Figure 6B** the regions of *KIF1A*, *NKX2-2* (higher expressed in hCiCs) were plotted including the, location of the CpG sites, methylation-transcription correlation data, previously identified transcription factor binding sites (TFBS), RNA-seq expression levels, and CpG methylation levels. Notable is the overlap between the prioritized likely transcriptional active CpG sites and the mapping of TFBSs which underscore the validity of the applied prioritization scheme towards likely transcriptionally active CpG sites.

Discussion

To optimize generation of sustainable neo-cartilage constructs, we characterized molecular landscape of hiPSC derived neo-cartilages in comparison to those deposited by its primary counterpart, the human articular chondrocyte. Moreover, by using hiPSCs generated from two different tissue sources (skin fibroblasts and articular cartilage chondrocytes) we explored whether the epigenetic memory of articular chondrocytes could further facilitate *in vitro* chondrogenesis. However, the high similarity among neo-cartilages generated from different hiPSC lines, indicated that the articular cartilage epigenetic memory retained in hCiC, does not further improve the consistency and quality of the *in vitro* chondrogenesis (149-151). By subsequently prioritizing on

likely transcriptionally active methylation discordant between neo-cartilage from hCiC relative to hPAC, we identified relevant differences in phenotypic cell states between chondrocytes derived from hiPSCs and hPACs. Since this molecular level of information is known to be important in on/off target cell fate decisions (130), results of our study could be exploited to improve quality, purity, and maturity of hiPSC derived neo-cartilage matrix further. Ultimately to realize the introduction of sustainable, hiPSC derived neo-cartilage implantation into clinical practice.

Among the CpG-genes pairs lower expressed in hCiCs relative to hPACs, we identified genes that were enriched in pathways such as 'extracellular matrix organization', 'chondrocyte differentiation', and 'cartilage development' and likely mark immaturity and/or lower quality of hCiC neo-cartilage relative to hPACs. Among the identified discordant genes, we recognized many robust OA risk genes that have critical roles in articular cartilage maintenance such as *MGP, GDF5,* or *SERPINA1*. Moreover, other genes such as the identified integrins and *CHAD* are known to interact with structural ECM molecules, as well as, with cells and thus play a role in cartilage homeostasis (152). *CTNNB1*, among the highest significant differential CpG-gene pairs, is known to control, in interaction with *SOX9*, chondrocyte differentiation and could be an important marker of the reduced efficiency for hiPSC chondrogenesis (153). On a different note, *MGP* is known to regulate extracellular calcium levels via high affinity to its γ-carboxyglutamic acid (Gla) residues and is a critical marker of the chondrogenic cell lineage (154). More recently, *MGP* was recognized among the most robust and targetable OA risk genes by virtue of its vitamin K dependency (155, 156). In this respect it is tempting to speculate that enhancing MGP action by supplementation of vitamin K during hiPSC chondrogenic differentiation could be a potential alternative strategy to improve quality of deposited neo-cartilage.

We prioritized N=195 discordant genes that were highly significant differentially methylated and expressed between hCiC and hPACs, with large effect sizes, and with high correlation between methylation and expression, as eligible targets to improve hiPSC chondrogenic cell fate. The genomic regions plotted for compelling CpG-pairs highlighted that the mapping of these CpG sites indeed coincided with high confidence TFBSs. This adds to the validity that our prioritization scheme, although primarily based on association, has indeed identified differential set-points of transcriptionally active methylation hence potential epigenetically modifiable target genes. Nonetheless, a full exploration of the N=195 prioritized genes is required to implement strategies that improve quality, purity and maturity of hiPSC-derived neo-cartilages. We envision that such an exploration requires high throughput experimental validation by methodologies such as CRISPR-dCAS9 activation and interference (CRISPRi/a) during

hiPSC chondrogenic differentiation with single cell read-out and preferably followed by system biological approaches to model interactions. Eventually, methodologies such as dCas9-DNMT/TET methods that allow long term beneficial changes in set points of methylation, could be performed for ultimate validation.
Notable was also the lower expression of genes such as *TNFRS11B, P4HA2*, and *COL1A2* (157) in hCiC relative to hPAC neo-cartilage. These genes are also known to be consistently upregulated with OA pathophysiology (158). As such these differences in gene expression levels are likely due to the source of hPAC being harvested from aged preserved articular cartilage. On the other hand, it confirms that set-points of transcriptionally active methylation are stable throughout the harvesting and *in vitro* chondrogenesis steps of hPACs (110). We also identified discordant CpG-gene pairs that were upregulated in hCiC derived neo-cartilage relative to hPAC. These genes were enriched in pathways closely related to neurogenesis (**Table S4)**. These genes could be a reflection of off-target differentiation towards neurogenic lineage as outlined previously by Wu et al (159) using hiPSC derived from fibroblasts. Alternatively, these genes in the hiPSC derived neo-cartilage could be a reflection of the generation of a more fetal or immature cartilage like phenotype. Although we provided differential gene expression data between hPACs and hCiCs in **Table S2**, we reported an exploratory transcriptome analysis in the result section. This was done to circumvent the reporting of statistically significant yet subtle differences in expression between hPACs and hCiCs at the mature neo-cartilage stage. Such transcriptome wide differences per definition reflect the phenotypic cell state in mature hiPSC derived neo-cartilages and are actually not driving cell fate decisions, the primary focus of our study.

To identify differential methylation at CpG sites that likely mark aberrant set-points of gene expression, a key layer of molecular information that does facilitate critical lineage decisions hence cell fates (11), we integrated transcriptome data to the methylome data solely to extract biological relevant information to the differential methylated sites. Herein, we took the assumption that differences in methylation that affect expression of genes should follow specific rules of correlation between these two molecular levels of information.

Upon studying the discordant likely transcriptionally active methylation between neo-cartilage, we disregarded the hFiC datasets. This to assure that the comparative integrated methylome and transcriptome differential expression analyses was performed in powerful datasets generated from homogeneous cell populations. Additionally, the fact that our methylome wide data of neo-cartilages showed comparable high similarities within and between cell sources, we anticipated that the relative small hFiC dataset has likely no effect on the extent, nor content, of the conclusions of our study. Hence,

we are confident that the N=195 targets are robust across different cell sources used to generate hiPSCs.

A limitation of the current study is that we have used 3 hiPSC clones of chondrocytes of only one donor and 1 hiPSC clone of skin fibroblasts of 2 donors, that could have influenced the robustness of our results. Nonetheless, as reported, we have generated multiple differentiations of available clones and showed high methylome wide similarities within and among clones and cell sources confirming consistency of our differentiation protocol and validity of our conclusions. Moreover, our previous work supports the consistency of the chondrogenesis protocol across many different additional hiPSC lines (18, 128).

Conclusion

By applying an integrative multi-omics approach, we demonstrated high methylome wide similarity of neo-cartilages between and within cell sources underscoring the consistency and quality of the applied step-wise differentiation protocol (128, 129). Nonetheless, to realize hiPSC derived regenerative treatments into clinical practice, in-depth insight into the purity and quality of neo-cartilage is required. Hereto we also set out to identify discordant aspects of stable set-points of gene expression between neo-cartilage from hCiC relative to hPAC. Altogether, our findings provide important insights into discordant aspects of stable set-points of gene expression between neo-cartilage derived from hiPSCs relative to that of its primary counterpart, the human articular chondrocyte. These insights could be exploited to improve quality, purity, and maturity of hiPSC derived neo-cartilage matrix further, ultimately to realize the introduction of sustainable, hiPSC derived neo-cartilage implantation into clinical practice.

Abbreviations

hCiC	human cartilage-derived iPSCs differentiated to chondrocytes
hiPSC	human induced pluripotent stem cells
hFiC	human fibroblast-derived iPSCs differentiated to Chondrocytes
hPAC	human primary articular chondrocytes
OA	osteoarthritis

Ethics approval

Ethical permission for the described studies was obtained from the medical ethical committee (METC) of the Leiden University Medical Center:
1) “Research Artrotisch Articulair Kraakbeen (RAAK)” (P08.239 and P19.013, approved respectively d.d. December 2, 2008 and May 20, 2019)
2) “Parapluprotocol: hiPSCs” (P13.080 approved d.d. July 2, 2014)

Written informed consent was obtained from all participants.

Consent for publication
Not applicable.

Data availability
The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests
The authors declare that they have no competing interests.

Funding sources
This study is partially supported by the partners of Regenerative Medicine Crossing Borders (RegMed XB), a public-private partnership that uses regenerative medicine strategies to cure common chronic diseases and by the Dutch Scientific Research council NWO /ZonMW VICI scheme (91816631/528). The funding bodies played no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

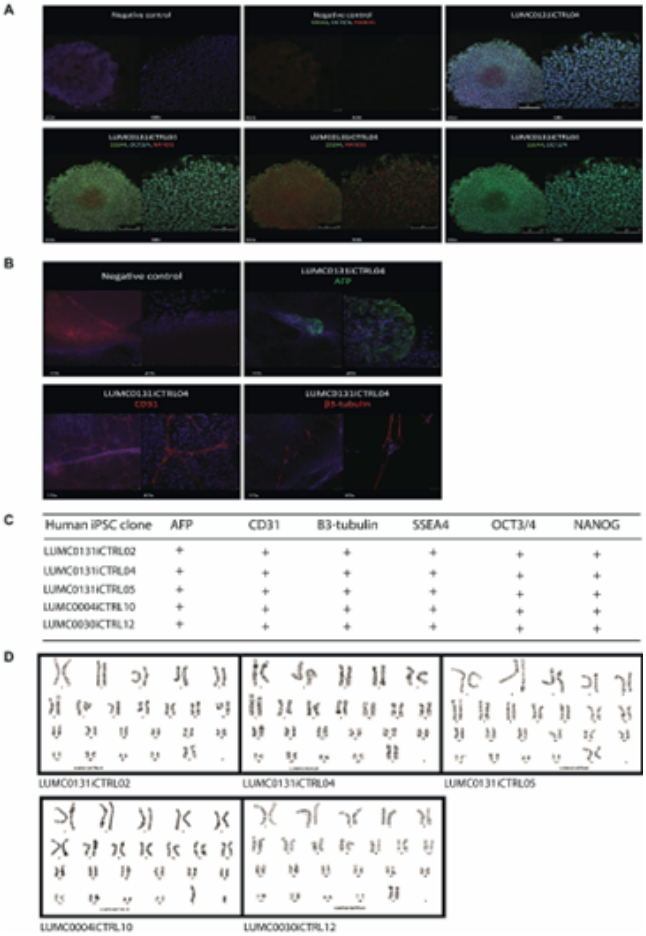
Author contributions
Conception and design: CF, YR, IM.
Collection and/or assembly of data: GH, ARR, NB, RN.
Data analysis and interpretation: GH, RCA, NB, RS, TK, KI, YR, IM.
Writing of the manuscript: GH, RCA, NB, YR, IM
Read and approved the final manuscript: all authors.

Acknowledgments
We thank Prof. Judith Bovee and all members of her team in the Dept. Pathology, Leiden University Medical Center, Leiden, The Netherlands.
No personal data are included in this study. All authors gave consent for publication.

References

1 Goldring, M. B. & Marcu, K. B. Cartilage homeostasis in health and rheumatic diseases. *Arthritis Res. Ther* **11**, 224 (2009).
2 Woolf, A. D., Erwin, J. & March, L. The need to address the burden of musculoskeletal conditions. *Best. Pract. Res. Clin. Rheumatol* **26**, 183-224 (2012).
3 Tuan, R. S., Chen, A. F. & Klatt, B. A. Cartilage regeneration. *J. Am. Acad. Orthop. Surg* **21**, 303-311, doi:21/5/303 [pii];10.5435/JAAOS-21-05-303 [doi] (2013).
4 Kimura, T., Yamashita, A., Ozono, K. & Tsumaki, N. Limited Immunogenicity of Human Induced Pluripotent Stem Cell-Derived Cartilages. *Tissue Eng Part A* **22**, 1367-1375, doi:10.1089/ten.TEA.2016.0189 (2016).
5 Bomer, N. *et al.* Neo-cartilage engineered from primary chondrocytes is epigenetically similar to autologous cartilage, in contrast to using mesenchymal stem cells. *Osteoarthritis. Cartilage* **24**, 1423-1430 (2016).
6 Kamaraj, A., Kyriacou, H., Seah, K. T. M. & Khan, W. S. Use of human induced pluripotent stem cells for cartilage regeneration in vitro and within chondral defect models of knee joint cartilage in vivo: a Preferred Reporting Items

for Systematic Reviews and Meta-Analyses systematic literature review. *Cytotherapy* **23**, 647-661, doi:10.1016/j.jcyt.2021.03.008 (2021).
7 Kretlow, J. D. *et al.* Donor age and cell passage affects differentiation potential of murine bone marrow-derived stem cells. *BMC Cell Biol* **9**, 60, doi:10.1186/1471-2121-9-60 (2008).
8 Qu, C. *et al.* Chondrogenic differentiation of human pluripotent stem cells in chondrocyte co-culture. *Int J Biochem Cell Biol* **45**, 1802-1812, doi:10.1016/j.biocel.2013.05.029 (2013).
9 Guha, P., Morgan, J. W., Mostoslavsky, G., Rodrigues, N. P. & Boyd, A. S. Lack of immune response to differentiated cells derived from syngeneic induced pluripotent stem cells. *Cell Stem Cell* **12**, 407-412, doi:10.1016/j.stem.2013.01.006 (2013).
10 Rodríguez Ruiz, A. *et al.* Cartilage from human-induced pluripotent stem cells: comparison with neo-cartilage from chondrocytes and bone marrow mesenchymal stromal cells. *Cell Tissue Res.*, doi:10.1007/s00441-021-03498-5 (2021).
11 Adkar, S. S. *et al.* Step-Wise Chondrogenesis of Human Induced Pluripotent Stem Cells and Purification Via a Reporter Allele Generated by CRISPR-Cas9 Genome Editing. *Stem Cells* **37**, 65-76, doi:https://doi.org/10.1002/stem.2931 (2019).
12 Parry, A., Rulands, S. & Reik, W. Active turnover of DNA methylation during cell fate decisions. *Nat Rev Genet* **22**, 59-66, doi:10.1038/s41576-020-00287-8 (2021).
13 Ziller, M. J. *et al.* Charting a dynamic DNA methylation landscape of the human genome. *Nature* **500**, 477-481, doi:10.1038/nature12433 (2013).
14 Doi, A. *et al.* Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nat Genet* **41**, 1350-1353, doi:10.1038/ng.471 (2009).
15 Kim, K. *et al.* Epigenetic memory in induced pluripotent stem cells. *Nature* **467**, 285-290, doi:10.1038/nature09342 (2010).
16 Lister, R. *et al.* Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* **471**, 68-73, doi:10.1038/nature09798 (2011).
17 Ohi, Y. *et al.* Incomplete DNA methylation underlies a transcriptional memory of somatic cells in human iPS cells. *Nat Cell Biol* **13**, 541-549, doi:10.1038/ncb2239 (2011).
18 Dambrot, C. *et al.* Polycistronic lentivirus induced pluripotent stem cells from skin biopsies after long term storage, blood outgrowth endothelial cells and cells from milk teeth. *Differentiation* **85**, 101-109, doi:10.1016/j.diff.2013.01.001 (2013).
19 Tuerlings, M. *et al.* RNA Sequencing Reveals Interacting Key Determinants of Osteoarthritis Acting in Subchondral Bone and Articular Cartilage: Identification of IL11 and CHADL as Attractive Treatment Targets. *Arthritis Rheumatol* **73**, 789-799, doi:10.1002/art.41600 (2021).
20 van Hoolwerff, M. *et al.* High-impact FN1 mutation decreases chondrogenic potential and affects cartilage deposition via decreased binding to collagen type II. *Sci Adv* **7**, eabg8583, doi:10.1126/sciadv.abg8583 (2021).
21 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21, doi:10.1093/bioinformatics/bts635 (2013).
22 Anders, S., Pyl, P. T. & Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166-169, doi:10.1093/bioinformatics/btu638 (2015).
23 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).
24 van Iterson, M. *et al.* MethylAid: visual and interactive quality control of large Illumina 450k datasets. *Bioinformatics* **30**, 3435-3437, doi:10.1093/bioinformatics/btu566 (2014).
25 Aryee, M. J. *et al.* Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* **30**, 1363-1369, doi:10.1093/bioinformatics/btu049 (2014).
26 Chen, Y. A. *et al.* Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics* **8**, 203-209, doi:10.4161/epi.23470 (2013).
27 Martin-Trujillo, A. *et al.* Rare genetic variation at transcription factor binding sites modulates local DNA methylation profiles. *PLoS Genet* **16**, e1009189, doi:10.1371/journal.pgen.1009189 (2020).
28 Fortin, J. P., Triche, T. J., Jr. & Hansen, K. D. Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. *Bioinformatics* **33**, 558-560, doi:10.1093/bioinformatics/btw691 (2017).
29 Zhang, D., Xue, J. & Peng, F. The regulatory activities of MALAT1 in the development of bone and cartilage diseases. *Front. Endocrinol. (Lausanne)* **13**, 1054827, doi:10.3389/fendo.2022.1054827 (2022).
30 Boer, C. G. *et al.* Deciphering osteoarthritis genetics across 826,690 individuals from 9 populations. *Cell* **184**, 4784-4818. e4717, doi:https://doi.org/10.1016/j.cell.2021.07.038 (2021).
31 Coutinho de Almeida, R. *et al.* Identification and characterization of two consistent osteoarthritis subtypes by transcriptome and clinical data integration. *Rheumatology*, doi:10.1093/rheumatology/keaa391 (2020).
32 Coutinho de Almeida, R. *et al.* RNA sequencing data integration reveals an miRNA interactome of osteoarthritis cartilage. *Ann Rheum Dis* **78**, 270-277, doi:10.1136/annrheumdis-2018-213882 (2019).
33 Kilpinen, H. *et al.* Common genetic variation drives molecular heterogeneity in human iPSCs. *Nature* **546**, 370-375, doi:10.1038/nature22403 (2017).
34 Kim, K. *et al.* Donor cell type can influence the epigenome and differentiation potential of human induced pluripotent stem cells. *Nat Biotechnol* **29**, 1117-1119, doi:10.1038/nbt.2052 (2011).
35 Nishizawa, M. *et al.* Epigenetic Variation between Human Induced Pluripotent Stem Cell Lines Is an Indicator of Differentiation Capacity. *Cell Stem Cell* **19**, 341-354, doi:10.1016/j.stem.2016.06.019 (2016).
36 Hessle, L. *et al.* The skeletal phenotype of chondroadherin deficient mice. *PLoS One* **8**, e63080, doi:10.1371/journal.pone.0063080 (2014).
37 Akiyama, H. *et al.* Interactions between Sox9 and beta-catenin control chondrocyte differentiation. *Genes Dev* **18**, 1072-1087, doi:10.1101/gad.1171104 (2004).
38 Luo, G., D'Souza, R., Hogue, D. & Karsenty, G. The matrix Gla protein gene is a marker of the chondrogenesis cell lineage during mouse development. *J. Bone Miner. Res.* **10**, 325-334, doi:10.1002/jbmr.5650100221 (1995).
39 Houtman, E. *et al.* Characterization of dynamic changes in Matrix Gla Protein (MGP) gene expression as function of genetic risk alleles, osteoarthritis relevant stimuli, and the vitamin K inhibitor warfarin. *Osteoarthritis and Cartilage*, doi:https://doi.org/10.1016/j.joca.2021.05.001 (2021).
40 Boer, C. G. *et al.* Vitamin K antagonist anticoagulant usage is associated with increased incidence and progression of osteoarthritis. *Ann. Rheum. Dis.* **80**, 598-604, doi:10.1136/annrheumdis-2020-219483 (2021).
41 Xu, H. *et al.* Suppression of CRLF1 promotes the chondrogenic differentiation of bone marrow-derived mesenchymal stem and protects cartilage tissue from damage in osteoarthritis via activation of miR-320. *Mol Med* **27**, 116, doi:10.1186/s10020-021-00369-1 (2021).
42 Rodríguez Ruiz, A. *et al.* The role of TNFRSF11B in development of osteoarthritic cartilage. *Rheumatology (Oxford)* **61**, 856-864, doi:10.1093/rheumatology/keab440 (2022).



Human iPSC clone	AFP	CD31	B3-tubulin	SSEA4	OCT3/4	NANOG
LUMC0131iCTRL02	+	+	+	+	+	+
LUMC0131iCTRL04	+	+	+	+	+	+
LUMC0131iCTRL05	+	+	+	+	+	+
LUMC0004iCTRL10	+	+	+	+	+	+
LUMC0030iCTRL12	+	+	+	+	+	+

Figure S1 | Characterization of hiPSC cell-lines. (A and B) representative pictures of hiPSCs confirming adequate pluripotency and spontaneous differentiation capacity of human skin fibroblast-derived iPSCs (2 cell-lines from healthy donors: LUMC0004iCTRL12-LUMC0030iCTRL12) and human cartilage-derived iPSCs (3 Clones from preserved cartilage of a patient who underwent joint replacement surgery due to OA: LUMC0131iCTRL02- LUMC0131iCTRL04- LUMC0131iCTRL05), respectively. **(C)** Presence of pluripotency and differentiation markers in all hiPSCs sources used in the study. **(D)** Confirming the normal karyotypes in all hiPSCs sources after reprogramming.

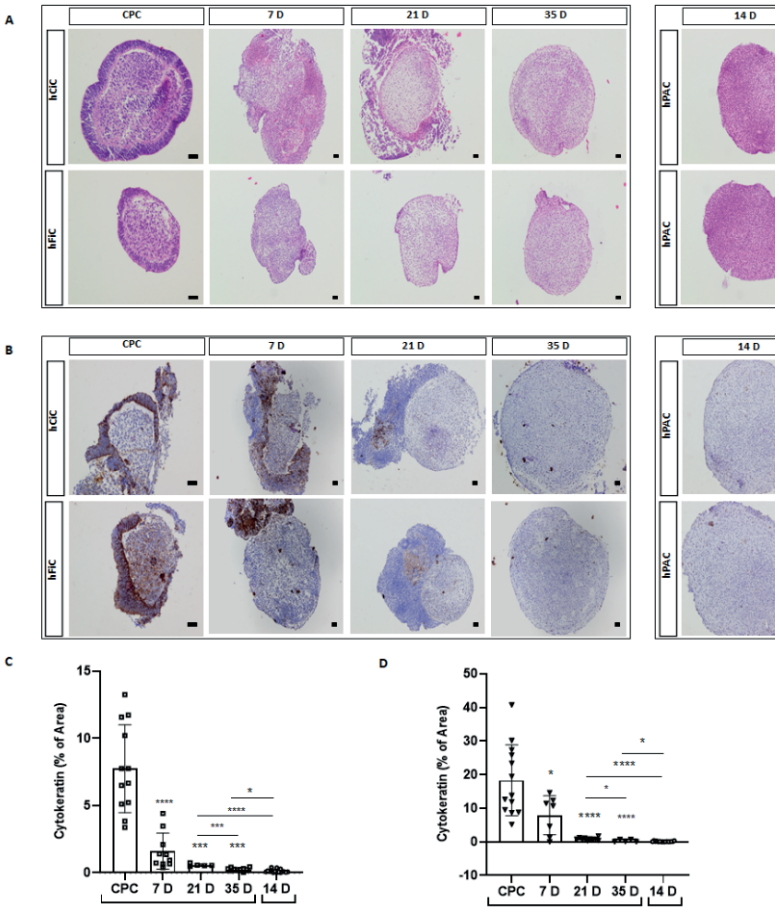


Figure S2 | Histology and immunohistochemistry of hiPSC derived chondroprogenitor cells towards neo-cartilage. (A) Histological evaluation of neo-cartilage from human chondrocyte-derived iPSCs (hCiC) and from human fibroblast-derived iPSCs (hFiC) by Haematoxylin-eosin (H&E) staining, compared to primary articular chondrocytes (hPACs; day 7, 14, 21, and 35 of chondrogenic differentiation). (B) Immunohistochemical staining (Anti-Cytokeratin 7 Antibody (CK)) showing the presence of keratinocytes distributed in a mesh-like form amongst the undifferentiated/undefined population around the 3D neo-cartilage during the first 21 days of chondrogenic differentiation. **(C)** Quantification of cytokeratin presence presented by percentage of area in hCiC compared to hPAC. (D) Quantification of cytokeratin immunohistochemistry presented by percentage of area in hFiC compared to hPAC. Scale bars in CPC: 50 µm, other scale bars: 100 µm. All data are presented as mean ± STD. Two-tailed Student's t tests were used to compare 1) each group relative to CPC (stars on top of the bars) and 2) each group relative to hPAC (indicated by lines). *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001 (N=11-14 pellets each).

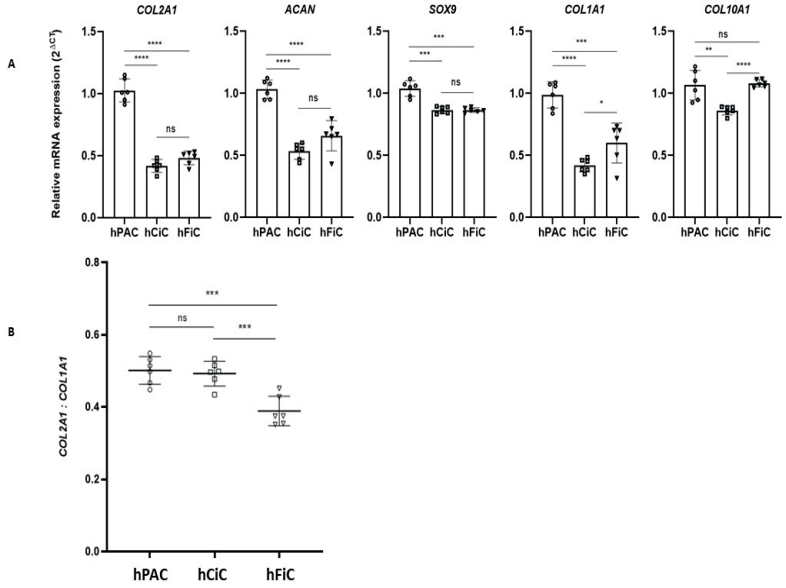


Figure S3 | mRNA expression levels of chondrogenic genes. (A) mRNA expression of chondrogenic genes (*COL2A, SOX9* and ACAN), fibrotic (*COL1A1*), and hypertrophic (*COL10A1*) cartilage markers at day 35 of differentiation (N=6) as determined by RT-qPCR. Representative of three independent experiments. **(B)** The ratio of the expression of COL2A1 to COL1A1 compared between hCiC, hFiC, and hPAC. The graph represents the chondrogenic potential of the neo-cartilages. Two-tailed Student's t tests were used to compare means between groups. All data are presented as mean ± STD. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

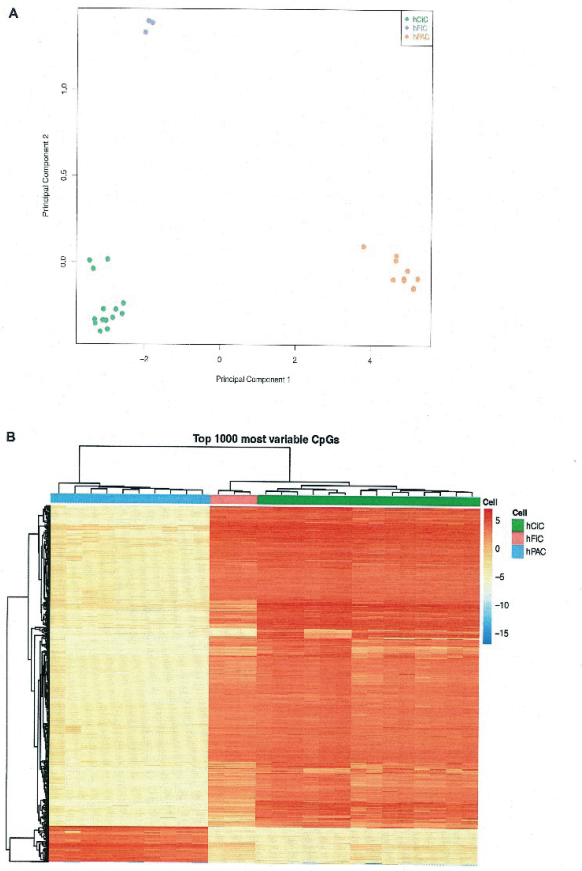


Figure S4 | (A) Principal component Analysis (PCA) of genome-wide methylation sites in human iPSCs and hPAC. The green dots are displaying hCiC, blue dots hFiC and orange dots hPAC. Component 1 (x-axis) and component 2 (y-axis). **(B)** Heatmap constructed from the 1000 most variable CpGs (y-axis) that showed the largest variation (Standard Deviation) over the samples (x-axis). Samples are hierarchically clustered (dendrogram on the top), showing three distinct clusters: hCiC (green), hFiC (red), and hPAC (blue) together.

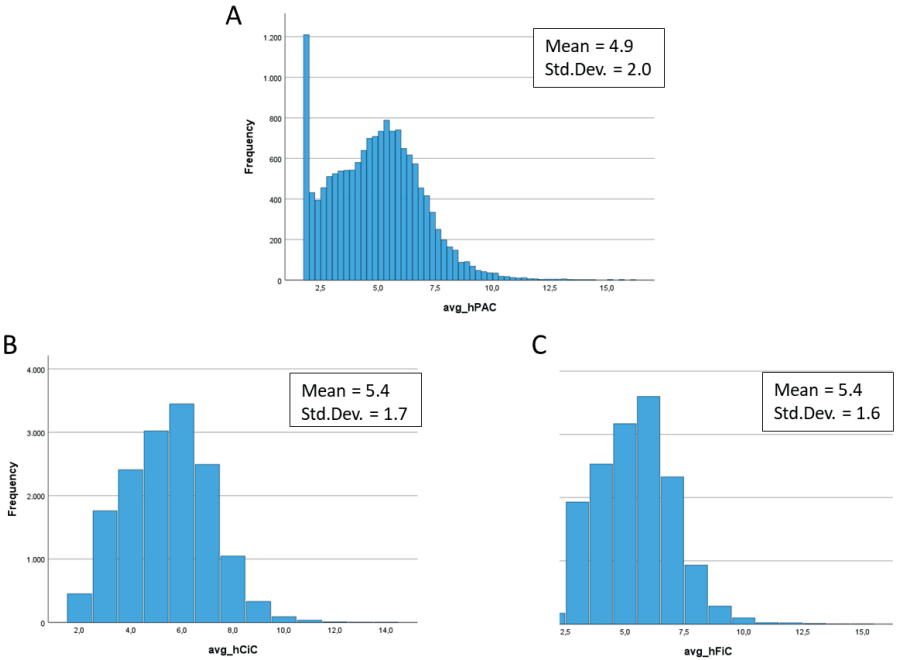
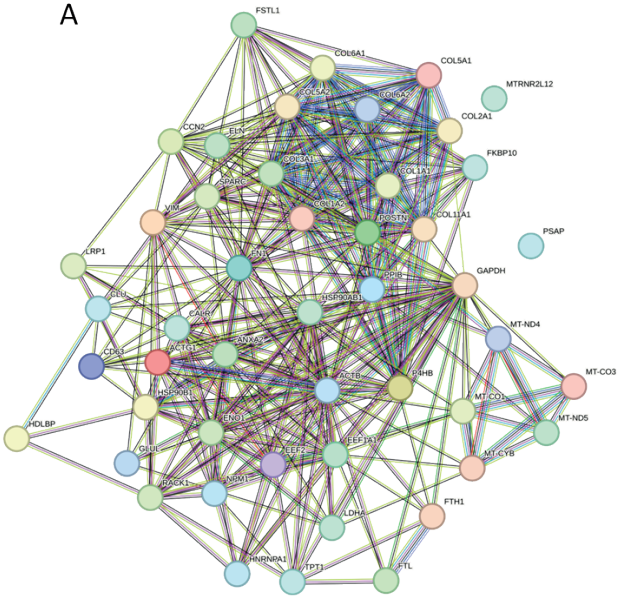


Figure S5 | Histogram of transcriptome wide gene expression (N=15249 genes) across (A) human primary chondrocytes (N=7) (B) human cartilage-derived iPSCs differentiated to chondrocyte (hCi-Chondrocyte: hCiC) and (C) human fibroblast-derived iPSCs differentiated to Chondrocyte (hFi-Chondrocyte: hFiC).

A



B

number of nodes: 48
number of edges: 345
average node degree: 14.4
avg. local clustering coefficient: 0.684

expected number of edges: 99
PPI enrichment p-value: < 1.0e-16
your network has significantly more interactions than expected (what does that mean?)

Functional enrichments in your network

explain columns

Biological Process (Gene Ontology)

GO-term	description	count in network	strength	false discovery rate
GO:0030199	Collagen fibril organization	9 of 59	1.8	9.29e-10
GO:0030198	Extracellular matrix organization	12 of 278	1.25	2.74e-08
GO:0097435	Supramolecular fiber organization	14 of 570	1.0	2.01e-07
GO:0010033	Response to organic substance	25 of 2692	0.58	8.49e-07
GO:0072359	Circulatory system development	15 of 901	0.83	5.15e-06

(more ...)

Molecular Function (Gene Ontology)

GO-term	description	count in network	strength	false discovery rate
GO:0030020	Extracellular matrix structural constituent conferring tensile ...	8 of 28	2.07	9.35e-11
GO:0005201	Extracellular matrix structural constituent	11 of 131	1.54	9.35e-11
GO:0048407	Platelet-derived growth factor binding	6 of 11	2.35	2.97e-09
GO:0050839	Cell adhesion molecule binding	14 of 560	1.01	6.29e-08
GO:0005198	Structural molecule activity	14 of 776	0.87	3.22e-06

(more ...)

Cellular Component (Gene Ontology)

GO-term	description	count in network	strength	false discovery rate
GO:0005615	Extracellular space	38 of 3247	0.68	4.88e-18
GO:0005576	Extracellular region	40 of 4175	0.59	7.78e-17
GO:0005788	Endoplasmic reticulum lumen	17 of 312	1.35	7.05e-16
GO:0062023	Collagen-containing extracellular matrix	18 of 407	1.26	1.42e-15
GO:0031012	Extracellular matrix	19 of 552	1.15	9.31e-15

(more ...)

Figure S6 | STRING protein-protein interaction network concordant genes between hPAC and hCiCs (A) Highly interconnected dense network representing top 50 of 919 genes highly expressed in hPAC (>SD2) and hCiC (>SD2). The edges indicate both functional and physical protein associations. Line thickness indicates the strength of data support. Minimum required interaction score = medium confidence (0.400). (B) Functional gene enriched pathway in the networ

number of nodes: 68
number of edges: 49
average node degree: 1.44
avg. local clustering coefficient: 0.436

expected number of edges: 10
PPI enrichment p-value: < 1.0e-16
your network has significantly more interactions than expected (what does that mean?)

Functional enrichments in your network

explain columns

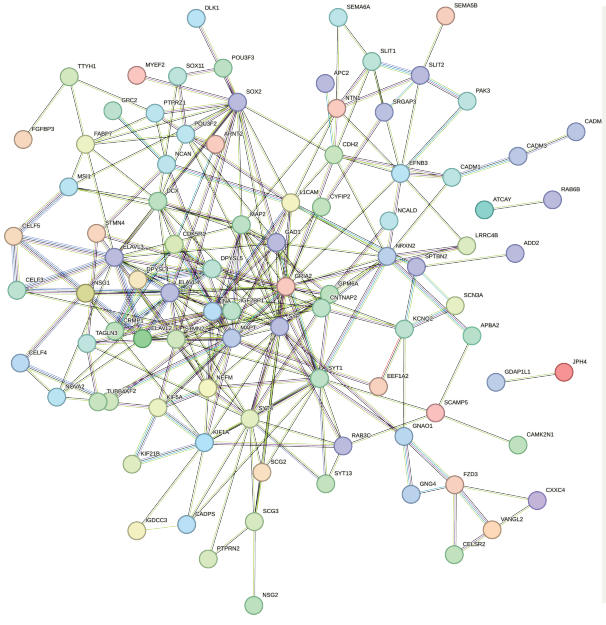
>	Biological Process (Gene Ontology)			
	GO-term	description	count in network	strength	▲ false discovery rate
	GO:2000026	Regulation of multicellular organismal development	20 of 1389	0.62	0.00047
	GO:0051239	Regulation of multicellular organismal process	28 of 2749	0.47	0.00047
	GO:0048519	Negative regulation of biological process	29 of 3313	0.33	0.00072
	GO:0001501	Skeletal system development	12 of 513	0.83	0.00080
	GO:0051241	Negative regulation of multicellular organismal process	16 of 1035	0.65	0.0012
					(more ...)

	Molecular Function (Gene Ontology)				
	GO-term	description	count in network	strength	▲ false discovery rate
	GO:0008201	Heparin binding	9 of 173	1.18	5.94e-05
	GO:0005539	Glycosaminoglycan binding	10 of 245	1.07	5.94e-05

	Cellular Component (Gene Ontology)				
	GO-term	description	count in network	strength	▲ false discovery rate
	GO:0005576	Extracellular region	47 of 4175	0.51	7.41e-14
	GO:0031012	Extracellular matrix	19 of 552	1.0	2.98e-11
	GO:0005615	Extracellular space	39 of 3247	0.54	2.98e-11
	GO:0062023	Collagen-containing extracellular matrix	11 of 407	0.89	7.05e-05
	GO:0071944	Cell periphery	40 of 6015	0.28	0.00046

A

B



number of nodes: 121
number of edges: 252
average node degree: 4.17
avg. local clustering coefficient: 0.393

expected number of edges: 35
PPI enrichment p-value: < 1.0e-16
your network has significantly more interactions than expected (what does that mean?)

Functional enrichments in your network

explain columns

Biological Process (Gene Ontology)

GO-term	description	count in network	strength	false discovery rate
GO:0007399	Nervous system development	60 of 2188	0.65	2.49e-21
GO:0048666	Neuron development	32 of 840	0.79	7.75e-13
GO:0031175	Neuron projection development	29 of 674	0.85	8.86e-13
GO:0030182	Neuron differentiation	35 of 1062	0.73	8.86e-13
GO:0048731	System development	64 of 3867	0.43	1.78e-12

(more ...)

Molecular Function (Gene Ontology)

GO-term	description	count in network	strength	false discovery rate
GO:0008092	Cytoskeletal protein binding	23 of 1002	0.57	0.00026
GO:0015631	Tubulin binding	14 of 380	0.78	0.00032
GO:0005515	Protein binding	71 of 7242	0.2	0.0014
GO:0005488	Binding	102 of 12838	0.11	0.0029
GO:0036002	pre-mRNA binding	5 of 35	1.37	0.0043

(more ...)

Cellular Component (Gene Ontology)

GO-term	description	count in network	strength	false discovery rate
GO:0045202	Synapse	44 of 1350	0.72	2.09e-17
GO:0043005	Neuron projection	41 of 1391	0.68	1.11e-14
GO:0030054	Cell junction	49 of 2115	0.58	2.32e-14
GO:0030424	Axon	29 of 651	0.86	3.60e-14
GO:0120025	Plasma membrane bounded cell projection	43 of 2268	0.49	2.40e-09

Figure S7 | STRING protein-protein interaction network discordant genes between hPAC and hCiCs (A) Left panel: Network representing N=69 genes highly expressed in hPAC (>SD2) and lowly expressed in hCiC (<SD2). The edges indicate both functional and physical protein associations. Line thickness indicates the strength of data support. Minimum required interaction score = medium confidence (0.400). Right panel Functional gene enriched pathway in the network. **(B)** Left panel: Network representing N=124 genes lowly expressed in hPAC (<SD2) and highly expressed in hCiC (>SD2). The edges indicate both functional and physical protein associations. Line thickness indicates the strength of data support. Minimum required interaction score = medium confidence (0.400). Right panel Functional gene enriched pathway in the network.

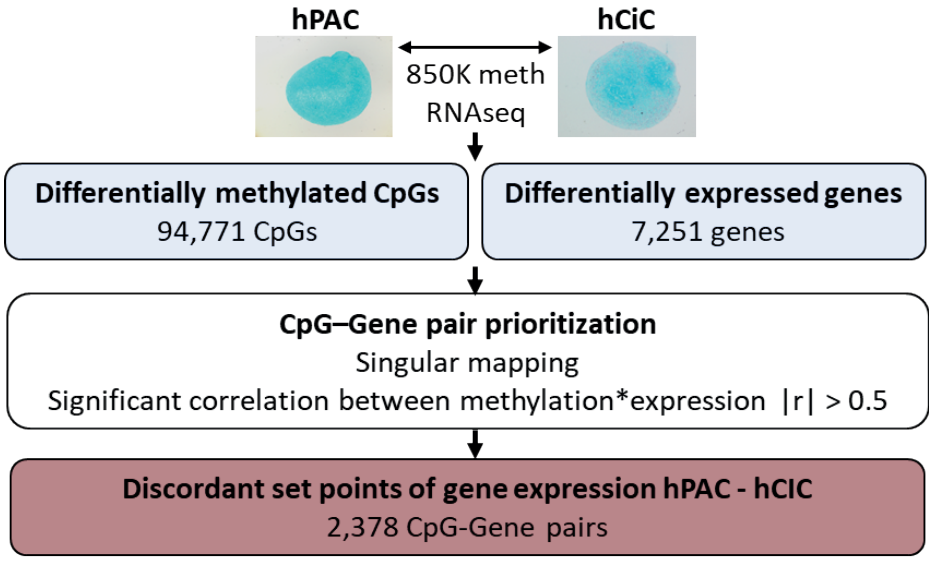


Figure S8 | Stepwise prioritization scheme for integrated methylation data and normalized gene expression 1) Integration of 94,771 differentially methylated CpGs and 7,251 differentially expressed genes between hCiC and hPAC 2) Selection of 2,378 unique CpG-gene pairs by prioritized among the 94,771 differential CpGs those that had an exact mapping on a reference gene that was also significant differential expression between hCiC and hPAC (FDR <0.05).

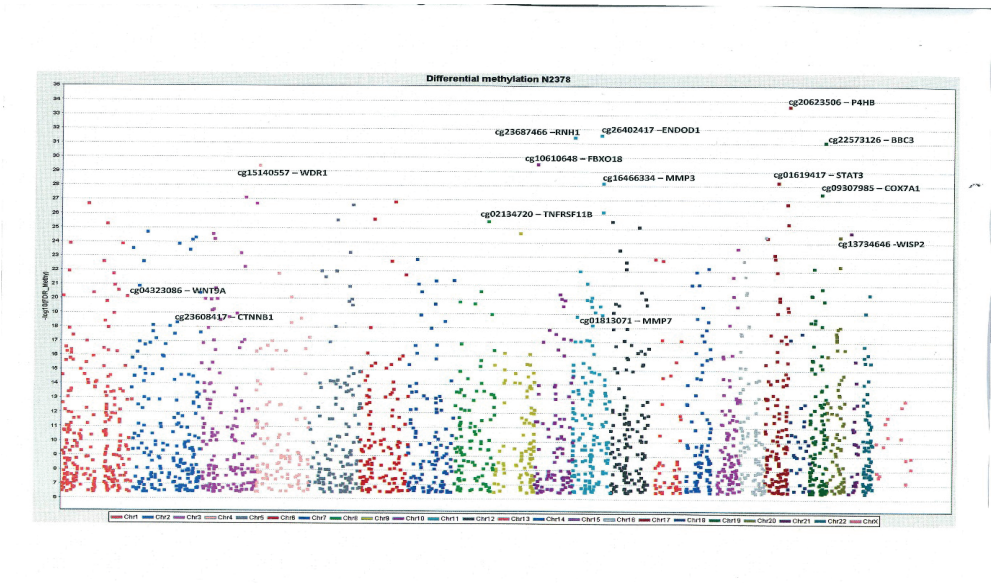


Figure S9 | Manhattan plot of FDR significant differential methylated CpG sites between hCiC and hPACs. Plotted are N=2379 prioritized CpG-gene pairs as determined by singular mapping, highest correlation between methylation and expression. See also Fig. 3 box 2. Y-axes indicates the log(FDR) value of the differential methylated CpGs.

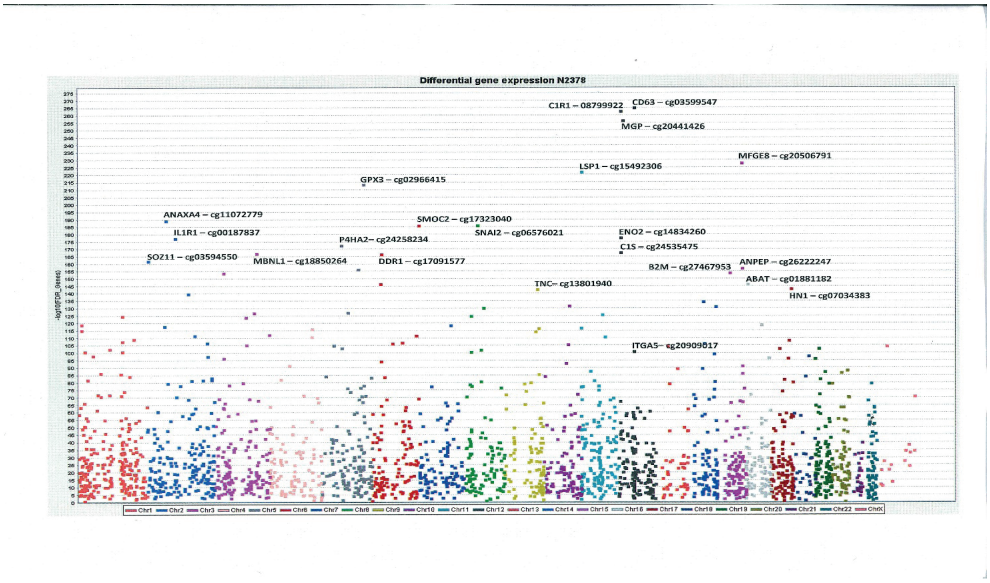


Figure S10 | Manhattan plot of FDR significant differential expressed genes between hCiC and hPACs. Plotted are the N=2379 prioritized CpG-gene pairs as determined by singular mapping, highest correlation between methylation and expression. See also Fig. 3 box 2. Y-axes indicates the log(FDR) value of the differential expressed genes.

Table S1. (partially) Normalized VST levels of genes expressed in 3D neo-cartilage derived from hPACs and from different hiPSC sources, showing top 50 similarly highly expressed 4th quartile genes.

Emsembl_ID	geneName	Z -score Average hPAC	Z-score Average hCiC	Z-score Average hFiC
ENSG00000156508	*EEF1A1*	5.17	4.99	5.91
ENSG00000251562	*MALAT1*	4.10	4.99	5.30
ENSG00000198804	*MT-CO1*	3.96	4.41	4.68
ENSG00000210082	*MT-RNR2*	3.83	4.41	4.68
ENSG00000075624	*ACTB*	3.07	3.83	4.07
ENSG00000184009	*ACTG1*	3.20	3.83	4.07
ENSG00000074800	*ENO1*	3.28	3.83	4.07
ENSG00000111640	*GAPDH*	3.70	3.83	4.07
ENSG00000026025	*VIM*	4.29	3.83	4.07
ENSG00000108821	*COL1A1*	4.36	3.24	4.07
ENSG00000168542	*COL3A1*	5.39	3.24	4.68
ENSG00000167658	*EEF2*	3.75	3.24	4.07
ENSG00000167996	*FTH1*	3.80	3.24	3.45
ENSG00000204628	*GNB2L1*	3.00	3.24	3.45
ENSG00000135486	*HNRNPA1*	2.99	3.24	3.45
ENSG00000096384	*HSP90AB1*	3.02	3.24	3.45
ENSG00000198938	*MT-CO3*	3.25	3.24	3.45
ENSG00000198886	*MT-ND4*	3.55	3.24	4.07
ENSG00000133112	*TPT1*	4.19	3.24	4.07
ENSG00000164692	*COL1A2*	4.74	2.66	4.07
ENSG00000196205	*EEF1A1P5*	3.08	2.66	2.83
ENSG00000049540	*ELN*	4.12	2.66	4.07
ENSG00000087086	*FTL*	3.64	2.66	3.45
ENSG00000166598	*HSP90B1*	3.28	2.66	2.83
ENSG00000198727	*MT-CYB*	3.03	2.66	2.83
ENSG00000198786	*MT-ND5*	3.09	2.66	3.45
ENSG00000181163	*NPM1*	2.95	2.66	3.45
ENSG00000133110	*POSTN*	3.43	2.66	4.68
ENSG00000197746	*PSAP*	3.11	2.66	2.83

Table S2. (partially) Differential CpGs (N=2378) that mapped in direct vicinity of differentially expressed genes (DEG) for which methylation and expression were correlated

CpG Site	Average Methylation hPAC	Average Methylation hCiC	logFC Methylation	FDR Methylation	Gene	Average Expression hPAC	Average Expression hCiC	logFC Genes	FDR Genes	R^2 Methylation Expression
cg02293410	0.53	0.14	-0.39	9.00E-11	*COMP*	15.18	3.87	-11.72	1,50E-083	0.92
cg07875786	0.2	0.03	-0.17	2.98E-08	*NR2F1*	3.73	9.31	5.26	3,59E-077	0.88
cg09431013	0.85	0.64	-0.21	1.27E-10	*LTBP2*	11.24	5.23	-6.47	3,67E-035	0.87
cg13384113	0.85	0.67	-0.18	1.16E-07	*ABI3BP*	11.56	4.75	-6.50	9,65E-017	0.87
cg20509092	0.21	0.04	-0.17	2.02E-08	*FAM169A*	1.78	6.10	7.64	2,36E-020	0.87
cg00407600	0.07	0.03	-0.04	2.32E-07	*CDKN2A*	5.58	1.80	-6.68	1,04E-032	0.86
cg03222811	0.27	0.07	-0.20	2.30E-07	*PTK2B*	6.58	4.48	-3.01	1,83E-033	0.86
cg15032098	0.23	0.03	-0.19	3.91E-08	*UCHL1*	4.90	10.37	4.99	4,14E-083	0.85
cg15640375	0.11	0.05	-0.06	2.79E-07	*LPPR3*	2.02	8.59	9.40	2,77E-033	0.85
cg25832796	0.21	0.05	-0.15	1.80E-10	*JPH4*	1.78	7.84	9.43	3,64E-032	0.85
cg03720545	0.63	0.16	-0.47	2.08E-10	*PDGFRL*	7.35	3.47	-4.65	8,49E-031	0.85
cg20800022	0.23	0.06	-0.16	2.68E-10	*GDF6*	4.87	1.50	-5.93	5,71E-022	0.85
cg16737670	0.1	0.05	-0.05	2.66E-07	*MCF2L*	2.08	7.66	7.78	3,86E-031	0.85
cg26366091	0.03	0.75	0.72	1.57E-12	*CHI3L2*	11.07	1.38	-12.80	1,37E-103	0.85
cg09397653	0.89	0.76	-0.13	3.44E-08	*ITGA11*	9.56	5.22	-4.68	5,49E-020	0.85
cg20569203	0.25	0.04	-0.20	8.78E-10	*MSI1*	2.15	8.27	7.88	1,62E-040	0.85
cg03173722	0.19	0.03	-0.16	3.89E-08	*RUFY3*	6.05	9.39	2.74	1,12E-092	0.84
cg25799864	0.14	0.05	-0.09	8.31E-08	*GDAP1L1*	1.78	7.99	9.63	8,20E-032	0.84
cg17153568	0.23	0.06	-0.17	1.14E-07	*SP9*	1.78	5.87	7.45	6,31E-020	0.84
cg13869353	0.56	0.10	-0.46	5.40E-13	*CP*	6.85	2.13	-6.01	2,89E-050	0.84
cg09967633	0.74	0.32	-0.41	7.66E-10	*CRISPLD2*	8.23	2.50	-7.55	1,11E-097	0.84
cg19142740	0.15	0.04	-0.12	3.06E-07	*CADM4*	2.16	8.21	8.06	9,07E-038	0.84
cg07207490	0.22	0.03	-0.19	1.66E-09	*PHLDB1*	9.23	8.01	-1.92	9,49E-042	0.84
cg15492306	0.82	0.67	-0.15	9.39E-12	*LSP1*	9.74	2.70	-8.82	6,91E-223	0.84

CpG Site	Average Methylation hPAC	Average Methylation hCiC	logFC Methylation	FDR Methylation	Gene	Average Expression hPAC	Average Expression hCiC	logFC Genes	FDR Genes	R^2 Methylation Expression
cg01813071	0.07	0.74	0.68	1.05E-19	*MMP7*	9.45	1.30	-12.19	1,43E-069	0.84
cg13882484	0.7	0.17	-0.53	1.09E-12	*MYO1D*	8.87	4.64	-4.98	4,57E-104	0.83
cg21889405	0.58	0.22	-0.36	2.42E-09	*IL6R*	6.82	2.40	-5.80	1,74E-033	0.83
cg04671541	0.39	0.07	-0.31	2.02E-07	*TAGLN3*	1.78	8.44	10.05	3,53E-036	0.83
cg04072597	0.66	0.17	-0.49	8.18E-10	*CRLF1*	8.69	3.57	-6.35	3,29E-060	0.83
cg13941993	0.11	0.05	-0.06	2.72E-07	*GAD2*	1.78	7.22	8.82	8,74E-028	0.83
cg07091412	0.19	0.02	-0.17	1.12E-07	*CAMK2N1*	2.85	8.05	5.58	1,71E-067	0.83
cg24132141	0.22	0.03	-0.19	3.52E-10	*IGFBP2*	3.79	10.72	6.97	1,63E-107	0.83
cg15127250	0.29	0.09	-0.20	1.50E-07	*GALNT6*	4.64	1.11	-7.66	2,48E-020	0.83
cg20441426	0.05	0.84	0.78	1.62E-20	*MGP*	11.84	3.57	-9.41	1,84E-257	0.83
cg00035316	0.53	0.04	-0.49	9.85E-12	*HOXD8*	4.81	1.19	-7.40	2,12E-022	0.83
cg00422935	0.15	0.04	-0.11	2.42E-08	*SEMA6A*	2.36	8.29	6.70	4,44E-060	0.82
cg00800512	0.14	0.07	-0.07	1.51E-07	*ZIK1*	3.33	5.39	2.44	1,57E-014	0.82
cg03228516	0.54	0.23	-0.31	8.57E-10	*CSNK1E*	5.95	8.34	1.86	1,26E-031	0.82
cg27078824	0.31	0.06	-0.25	3.06E-08	*SLC29A1*	9.07	5.88	-3.85	7,18E-085	0.82
cg26041814	0.07	0.69	0.62	1.60E-17	*GPNMB*	7.28	2.65	-6.39	1,34E-041	0.82
cg26227298	0.12	0.05	-0.06	4.39E-08	*SLC1A3*	2.19	7.74	6.98	8,55E-038	0.82
cg05094081	0.17	0.06	-0.10	1.92E-07	*DLX1*	1.87	7.52	8.82	2,60E-025	0.82
cg13734646	0.09	0.81	0.72	3.95E-23	*WISP2*	6.71	1.11	-9.93	9,68E-041	0.82
cg15527515	0.85	0.39	-0.45	8.31E-15	*SLC7A8*	7.85	3.71	-5.14	1,57E-065	0.82
cg23111830	0.13	0.05	-0.08	1.87E-08	*SOX21*	1.78	6.67	8.23	1,93E-024	0.82
cg04611203	0.17	0.07	-0.10	4.41E-08	*CPXM1*	1.78	7.47	9.05	1,03E-029	0.82
cg16110455	0.34	0.81	0.47	5.00E-17	*HSPB7*	8.12	1.62	-9.13	8,04E-032	0.82
cg25231123	0.79	0.60	-0.19	2.38E-07	*COL16A1*	10.91	5.82	-5.38	1,35E-025	0.82
cg08831594	0.32	0.03	-0.29	4.20E-12	TSPAN15	1.93	4.42	4.55	6,61E-009	0.82
cg02903822	0.02	0.51	0.49	4.23E-14	WISP1	7.39	1.82	-8.05	1,77E-077	0.82

Table S3. (Partially, selected top 5 of each Category) Pathway analyses genes (N=1237) down in hiCIC relative to hPACs

Category	Term	Fold Enrichment	Bonferroni	Count	Genes
Biological process	GO:0030198~extracellular matrix organization	4.85	7.85E-18	52	*VIT, COL16A1, ECM2, B4GALT1, ITGB5, COL14A1, ELN, CYR61, ADAMTS5, ADAMTS2, ADAMTSL4, ADAMTSL3, ITGAV, SOX9, CSGALNACT1, PDGFRA, MMP7, MMP2, MMP3, MMP14, COL2A1, CCDC80, COL4A2, COL4A1, CRISPLD2, MMP23B, PXDN, ITGA5, MATN3, MATN2, ITGA9, COL11A1, PAPLN, FURIN, TNFRSF11B, FBLN2, ADAMTS14, EGFLAM, ABI3BP, COL28A1, KAZALD1, SMOC2, COL3A1, COL1A2, COL5A1, ITGA10, COL5A3, SMOC1, COL5A2, TGFBI, COL9A2, RECK*
	GO:0030199~collagen fibril organization	7.32	1.46E-15	32	*FKBP10, COLGALT2, COL16A1, COL14A1, COL11A1, COL12A1, DPT, PLOD3, LOXL2, COMP, ACAN, ADAMTS14, ADAMTS2, CYP1B1, TGFB2, CRTAP, ANXA2, GREM1, COL3A1, COL2A1, BMP1, COL1A2, COL4A2, COL5A1, LOX, COL4A1, P4HA2, COL5A3, COL5A2, PXDN, FMOD, DDR2*
	GO:0001525~angiogenesis	3.54	3.15E-11	51	*NRP2, CLIC4, PLXND1, SERPINE1, CIB1, CTGF, ELK3, TMEM100, ANPEP, CYP1B1, ITGAV, HOXA7, HTATIP2, PDGFRB, ANXA2, MMP2, HSPG2, MMP14, COL4A2, ADAM15, HAND2, PXDN, MYH9, ANG, ITGA5, MFGE8, RAPGEF3, SHC1, EPAS1, NPR3, THY1, ARHGAP22, SRPX2, RNF213, HMOX1, APOD, WASF2, ANGPT1, CAV1, C1GALT1, FN1, TBX4, FMNL3, KLF5, FAP, CSPG4, TGFBI, AMOTL1, PLCD3, RAMP1*
	GO:0007155~cell adhesion	2.54	6.41E-11	82	*SRPX, NRP2, COL16A1, B4GALT1, TNFAIP6, ITGB5, TGFB1I1, COL12A1, WISP3, TNC, INPPL1, ICAM2, CIB1, LAMC1, CYR61, HAPLN1, LOXL2, CTGF, COMP, ISLR, STAB1, CYP1B1, SIRPA, ITGAV, SVEP1, EDIL3, ROM1, OMD, EMP2, ATP1B1, AZGP1, VCAN, MYOT, COL6A2, NINJ2, PXDN, ADAM12, CDH11, MXRA8, COL6A3, PKP1, ITGA5, MFGE8, TLN1, CD44, DDR2, ITGA9, ENG, FBLN7, CD151, PXN, LAMA3, DPT, NEXN, THY1, NID2, THBS1, THBS3, DPP4, ACAN, GPNMB, SUSD5, COL28A1, CD164, LAMB2, PCDHGC3, FN1, PARVA, LAMB1,*

Category	Term	Fold Enrichment	Bonferroni	Count	Genes
					SORBS3, PPFIBP1, COL5A1, FAP, FES, PRPH2, ITGA10, ITGA11, SSPN, FAT1, ITGBL1, CD9, CTNNB1, TGFBI
	GO:0001501~skeletal system development	4.09	1.22E-07	32	*HEXB, NPR3, PRELP, TNFRSF11B, HOXA13, FGFRL1, HOXC10, HAPLN1, HOXA11, HOXA10, COMP, CLEC3A, ACAN, ANKH, TRPS1, SOX9, PITX1, CMKLR1, TGFB2, VDR, MMP14, VCAN, COL2A1, EVC, BMP1, COL1A2, COL5A2, CDH11, COL9A2, MATN3, FGFR1, FBN1*
KEGG pathway	hsa04510:Focal adhesion	3.22	1.17E-09	45	*ITGB5, SHC1, PXN, LAMA3, TNC, LAMC1, PIK3R1, THBS1, MYL12A, EGFR, THBS3, MYLK, COMP, CCND3, CHAD, PDGFC, FLNB, ITGAV, PDGFRB, PDGFRA, LAMB2, CAV2, ACTN1, CAV1, ITGA1, FN1, LAMB1, PARVA, ACTN4, DIAPH1, COL2A1, COL1A2, COL4A2, COL4A1, ITGA10, COL6A2, ITGA11, COL6A3, CTNNB1, COL9A2, ITGA5, TLN1, MET, MYL9, ITGA9*
	hsa04512:ECM-receptor interaction	4.42	3.35E-08	27	*ITGB5, SDC4, LAMA3, TNC, LAMC1, THBS1, THBS3, COMP, CHAD, ITGAV, LAMB2, ITGA1, FN1, LAMB1, HSPG2, COL2A1, COL1A2, COL4A2, COL4A1, ITGA10, COL6A2, ITGA11, COL6A3, COL9A2, ITGA5, CD44, ITGA9*
	hsa04974:Protein digestion and absorption	3.49	3.01E-05	25	*KCNK5, COL16A1, COL14A1, COL11A1, ELN, COL12A1, ATP1A1, DPP4, KCNN4, SLC38A2, COL28A1, MME, ATP1B1, COL3A1, COL2A1, COL1A2, SLC7A8, COL4A2, COL5A1, COL4A1, COL6A2, COL5A3, COL5A2, COL6A3, COL9A2*
	hsa05205:Proteoglycans in cancer	2.53	1.59E-04	36	*CD63, ITGB5, SDC4, SDC2, PXN, ITPR1, ITPR3, PIK3R1, IQGAP1, FGF2, THBS1, EGFR, SLC9A1, PLCE1, FLNB, ITGAV, TGFB2, TGFB1, CAV2, MMP2, CAV1, STAT3, PLAUR, FN1, PTPN11, HSPG2, DCN, COL1A2, CTTN, CTNNB1, HCLS1, ARHGEF1, ITGA5, MET, CD44, FGFR1*
	hsa00532:Glycosaminoglycan biosynthesis - chondroitin sulfate / dermatan sulfate	6.48	0.009739797	9	*CSGALNACT1, CHSY1, CHPF2, DSE, UST, XYLT1, CHSY3, B4GALT7, CHST3*
Reactome pathway	R-HSA-1474244~Extracellular matrix organization	4.40	1.04E-32	93	*COLGALT2, SPARC, COL16A1, ITGB5, COL14A1, ELN, COL12A1, SERPINE1, TNC, ICAM2, PLOD3, PLOD2, LAMC1, PLOD1, FGF2, HAPLN1, LOXL2, COMP, ADAMTS5, ADAMTS2, EFEMP2, SH3PXD2A, CTSK, TIMP2, ITGAV, CTSD, CTSB,*

Category	Term	Fold Enrichment	Bonferroni	Count	Genes
					MMP7, ACTN1, MMP2, ITGA1, MMP3, PCOLCE, HSPG2, DCN, VCAN, MMP14, COL2A1, COL4A2, ADAM15, LOX, COL4A1, COL6A2, PXDN, ADAM12, COL6A3, ITGA5, MATN3, CD44, DDR2, PLEC, ITGA9, CD151, PCOLCE2, SDC4, COL11A1, SDC2, LAMA3, NTN4, HTRA1, FURIN, LTBP2, LTBP3, NID1, NID2, THBS1, FBLN2, ACAN, ADAMTS14, A2M, COL28A1, TGFB2, CRTAP, TGFB1, LAMB2, FN1, LAMB1, GDF5, BMP4, COL3A1, BMP1, COL1A2, COL5A1, P4HA2, ITGA10, P4HA3, COL5A3, COL5A2, ITGA11, COL9A2, P4HB, FMOD, FBN1
	R-HSA-1474290~Collagen formation	5.99	1.01E-16	38	*COLGALT2, CD151, COL16A1, PCOLCE2, COL14A1, COL11A1, LAMA3, COL12A1, PLOD3, PLOD2, PLOD1, LOXL2, ADAMTS14, ADAMTS2, CTSB, COL28A1, CRTAP, MMP7, MMP3, PCOLCE, COL3A1, COL2A1, BMP1, COL1A2, COL4A2, COL5A1, LOX, COL4A1, P4HA2, P4HA3, COL6A2, COL5A3, COL5A2, PXDN, COL6A3, COL9A2, P4HB, PLEC*
	R-HSA-3000178~ECM proteoglycans	5.97	1.74E-13	32	*SPARC, ITGB5, LAMA3, SERPINE1, TNC, LAMC1, HAPLN1, COMP, ACAN, ITGAV, TGFB2, TGFB1, LAMB2, FN1, LAMB1, HSPG2, DCN, COL3A1, VCAN, COL2A1, COL1A2, COL4A2, COL5A1, COL4A1, COL6A2, COL5A3, COL5A2, COL6A3, COL9A2, FMOD, MATN3, ITGA9*
	R-HSA-1650814~Collagen biosynthesis and modifying enzymes	6.14	2.26E-12	29	*COLGALT2, COL16A1, PCOLCE2, COL14A1, COL11A1, COL12A1, PLOD3, PLOD2, PLOD1, ADAMTS14, ADAMTS2, COL28A1, CRTAP, PCOLCE, COL3A1, COL2A1, BMP1, COL1A2, COL4A2, COL5A1, COL4A1, P4HA2, P4HA3, COL6A2, COL5A3, COL5A2, COL6A3, COL9A2, P4HB*
	R-HSA-3000171~Non-integrin membrane-ECM interactions	6.25	5.11E-11	26	*ITGB5, SDC4, COL11A1, SDC2, LAMA3, NTN4, TNC, LAMC1, FGF2, THBS1, ITGAV, TGFB1, LAMB2, ACTN1, FN1, LAMB1, HSPG2, COL3A1, COL2A1, COL1A2, COL4A2, COL5A1, COL4A1, COL5A3, COL5A2, DDR2*

Table S4. (Partially, selected top 5 of each Category) Pathway analyses genes (N=1141) up in hiCIC relative to hPACs

Category	Term	Fold Enrichment	Bonferroni	Count	Genes
GO:Biological pathway	GO:0007411~axon guidance	3.29	6.10E-09	45	CYFIP2, LAMA5, SEMA5B, DRAXIN, NOTCH1, EPHA10, SEMA3D, LAMA1, ARX, CRMP1, GATA3, SEMA3F, IGSF9, PRTG, PGRMC1, DPYSL5, TUBB3, KIF5C, DPYSL2, SLIT1, OTX2, FYN, EPHB2, FEZF1, SEMA6B, SEMA6C, NGFR, EPHA7, SEMA6A, UNC5A, SEMA4C, VSTM2L, PAX6, SEMA4G, VAX1, BMP7, ENAH, KLF7, NFASC, EFNA2, LMX1A, LHX2, FEZ1, CDK5R1, EPHA2
	GO:0007399~nervous system development	2.33	2.70E-07	68	SEMA5B, CNTFR, SMARCB1, NRSN1, STMN3, MSI1, NR2E1, BZW2, DPYSL5, MDK, DPYSL2, DPF1, ENC1, DPF3, EPHB2, PCDHAC2, SEMA6B, HES6, SRRM4, SMARCC1, SEMA6A, RBFOX2, DYRK1A, ISLR2, PAX3, GFRA2, OLFM1, DOK4, MAP1B, FEZ1, RUFY3, KCNQ2, ARHGEF7, CAMK2B, CHRNA3, DLX6, CRMP1, GDPD5, CPLX2, ACVR1B, BRINP2, IGF2BP1, ST8SIA4, IGF2BP3, IGF2BP2, ZNF423, N4BP3, SCN3B, SH2B2, HES4, PCDHA6, GPR56, LIMK1, SEMA4C, NR2F1, CABLES1, SMARCA4, TTLL7, CCDC88A, PCDHB2, DLG4, NOVA1, FABP7, UNC119, APBA1, ATXN10, SIM1, APBA2
	GO:0007268~chemical synaptic transmission	2.91	2.32E-06	42	GABRB3, GRIA2, CHRNA3, NLGN2, CHRNA4, CACNA1B, SLC1A2, NRXN2, SLC1A3, DOC2B, LRP8, CACNA1G, GRM3, GRM2, HRH3, SLITRK5, KCNMB4, NPTX1, TAC1, NPTX2, BSN, SLC12A7, SNCA, CHRNB4, DTNA, SLC12A5, GABRA5, GAD1, GAD2, SYN2, GJD2, CACNB3, CACNB4, PCDHB2, DLG4, SST, KCNQ2, PLP1, UNC119, APBA1, APBA2, KCNK3
	GO:0030182~neuron differentiation	3.16	2.51E-06	37	BARHL2, BRSK2, RTN1, OTP, NEUROD1, FUT9, GPC2, RXRG, WNT4, IRX1, EMX2, FZD3, FZD2, MYEF2, OLIG3, GSX1, FZD7, WNT7B, DNMT3A, EN2, EN1, WNT7A, SOX11, OLIG1, WNT9A, VAX1, HIPK2, LMX1A, DOK5, LHX2, LHX4, LMX1B, INSM1, NEUROG2, LHX9, CDK5R1, EPHA2

Category	Term	Fold Enrichment	Bonferroni	Count	Genes
	GO:0045944~positive regulation of transcription from RNA polymerase II promoter	1.64	2.74E-05	130	FOXA1, TCERG1, SMARCB1, CRTC1, RBPJ, SOX21, SOX2, SOX1, SALL1, SALL2, SALL4, DPF1, MYB, HOXA2, DPF3, KAT7, SMARCC1, LMO3, EBF2, SOX11, SOX12, EBF4, ISL1, POU3F3, ADCYAP1, MYCN, SFRP2, LMX1A, TFR2, KAT6A, RFX4, HOXB4, HOXB3, LMX1B, ZFPM1, DLX1, CASZ1, DLX2, NOTCH1, CTBP2, ARX, GATA3, GATA2, NEUROD1, NHLH2, ZNF827, NHLH1, DHX36, HLTF, OTX2, OTX1, NKX2-2, TFAP2A, ZBTB18, AKIRIN2, TFAP2B, ZFHX3, POU2F1, JUP, BCL11B, BCL11A, HMGA1, NR2F1, PBX3, NR2F2, BMP7, PBX1, PHOX2B, SMARCA4, BCL9, LHX2, FEV, MAFF, ID4, LHX4, TOP2A, BARHL2, CSRNP3, ONECUT2, PTPRN, CHD7, AKAP8L, PSIP1, TCF20, RSF1, ARRB1, NR2E1, FOXM1, HEY1, SIX3, ZIC1, HRAS, KDM6B, EDN1, EN2, PAX3, EN1, PAX6, SREBF2, RGMA, TAL1, MTF2, TET3, ITGA6, PTMA, KDM3A, HDAC2, DOT1L, DLL1, STOX2, HSF2, APBB1, MSX1, RXRG, ZNF585A, TBX1, ARNT2, GSX1, DAB2IP, WNT7A, FOXJ1, MEIS2, ACVR2A, HIPK2, KLF7, NR6A1, ZNF219, F2RL1, PFKM, SSBP3
GO:Molecular Function	GO:1990837~sequence-specific double-stranded DNA binding	2.30	1.76E-08	74	FOXA1, BARHL2, ONECUT2, NR2E1, BACH2, LMNB1, HEY1, DPF1, SIX3, ZIC1, HOXA2, HES6, EMX2, OLIG3, EN2, PAX3, EN1, PAX6, SOX12, POU3F1, ISL1, SREBF2, MYCN, LMX1A, PAX7, RFX4, HOXB4, LMX1B, HOXB8, DLX1, DLX2, DLX6, ARX, GATA3, GATA2, NEUROD1, SCRT2, NHLH2, NHLH1, HSF2, OTX2, MSX1, OTX1, RXRG, IRX1, ZBTB18, TFAP2A, TBX1, FOXB1, ARNT2, TFAP2B, SKOR1, TCF7L1, GSX1, BCL11B, GSX2, NR2F1, ZIC4, ZIC5, VAX1, MEIS2, PBX1, PHOX2B, LIN28B, NR6A1, TLX3, FEV, MAFF, SP8, LHX4, SP9, NEUROG2, LHX9, ZSCAN1
	GO:0000978~RNA polymerase II core promoter proximal region sequence-specific DNA binding	1.54	0.0013465	118	FOXA1, IKZF2, RBPJ, SOX21, BACH2, SOX2, SOX1, SALL1, SALL2, SALL3, SALL4, MYB, NKX1-2, HOXA2, HES6, KDM2B, SOX13, EBF2, SOX11, EBF3, SOX12,

Category	Term	Fold Enrichment	Bonferroni	Count	Genes
					EBF4, POU3F1, ISL1, POU3F3, HIC2, MYCN, RFX4, HOXB4, HOXB3, ZNF710, ZNF793, DLX1, DLX2, DLX6, ARX, FOXO6, GATA3, GATA2, NEUROD1, SCRT2, NHLH2, NHLH1, DHX36, OTX2, OTX1, PATZ1, ZNF423, ZSCAN18, NKX2-2, TFAP2A, ZBTB18, TFAP2B, ZFHX3, POU2F1, BCL11B, BCL11A, HMGA1, NR2F1, PBX3, NR2F2, ZFHX4, PBX1, PHOX2B, TLX3, LHX2, FEV, MAFF, NFE2L3, SP8, SP9, ZSCAN1, ZNF254, ONECUT2, CHD7, NR2E1, ZIC2, HEY1, SIX3, ZIC1, FEZF1, ZIK1, KDM6B, EMX2, OLIG3, DNMT3A, EN2, PAX3, EN1, PAX6, OLIG1, SREBF2, PAX7, TAL1, NTN1, MTA1, HSF2, E2F5, RXRG, ZNF585A, HES4, IRX1, TBX1, FOXB1, ARNT2, SKOR1, TCF7L1, GSX1, FOXJ1, ZIC4, ZIC5, VAX1, MEIS2, KLF7, NR6A1, ZNF219, INSM1, NEUROG2
	GO:0001228~transcriptional activator activity, RNA polymerase II transcription regulatory region sequence-specific binding	1.93	0.0035569	55	FOXA1, BARHL2, CSRNP3, ONECUT2, NR2E1, RBPJ, SOX21, SOX2, SOX1, SALL2, SIX3, ZIC1, MYB, EBF2, SOX11, EBF3, PAX6, SOX12, ISL1, MYCN, LMX1A, RFX4, HOXB3, DLX2, CASZ1, ARX, GATA3, GATA2, NEUROD1, STOX2, NHLH2, NHLH1, HSF2, OTX2, MSX1, OTX1, ZNF585A, NKX2-2, TFAP2A, ARNT2, TFAP2B, POU2F1, GSX1, BCL11B, PBX3, FOXJ1, MEIS2, PBX1, PHOX2B, KLF7, NR6A1, LHX2, FEV, MAFF, LHX4
	GO:0000981~RNA polymerase II transcription factor activity, sequence-specific DNA binding	1.51	0.0045513	116	FOXA1, OTP, RBPJ, SOX21, BACH2, SOX2, SOX1, SALL1, SALL2, SALL3, SALL4, MYB, NKX1-2, HOXA2, HES6, SOX13, EBF2, SOX11, EBF3, SOX12, EBF4, POU3F1, ISL1, POU3F3, MYCN, LMX1A, RFX4, HOXB4, HOXB3, LMX1B, HOXB8, ZNF793, DLX1, CASZ1, DLX2, DLX6, ARX, FOXO6, GATA3, GATA2, NEUROD1, SCRT2, NHLH2, NHLH1, OTX2, OTX1, ZNF423, ZSCAN18, NKX2-2, TFAP2A, ZBTB18, TFAP2B, ZFHX3, POU2F1, PBX3, ZFHX4, PBX1, PHOX2B, TLX3, LHX2, FEV, MAFF, NFE2L3, SP8, LHX4, SIM1, SP9, LHX9, ZSCAN1, ZNF254, HMX3, BARHL2, CSRNP3, ONECUT2, NR2E1, FOXM1,

Category	Term	Fold Enrichment	Bonferroni	Count	Genes
					ZIC2, HEY1, SIX3, ZIC1, ZIK1, EMX2, OLIG3, EN2, PAX3, EN1, PAX6, OLIG1, SREBF2, PAX7, TAL1, ZBTB46, NTN1, HSF2, MSX1, E2F5, RXRG, HES4, PURG, IRX1, TBX1, FOXB1, ARNT2, SKOR1, TCF7L1, GSX1, GSX2, FOXJ1, ZIC4, ZIC5, VAX1, MEIS2, KLF7, NR6A1, ZNF219, NEUROG2
	GO:0003700~transcription factor activity, sequence-specific DNA binding	1.79	0.0082482	62	FOXA1, CSRNP3, FOXM1, IKZF2, RBPJ, SOX21, BACH2, SOX2, SOX1, HEY1, ZIC2, ZIC1, FEZF1, HES6, SOX13, EBF2, PAX3, SOX11, PAX6, EBF4, POU3F1, SREBF2, POU3F3, MYCN, PAX7, TAL1, RFX4, HOXB3, LMX1B, ZNF710, DLX2, DLX6, FOXO6, GATA3, GATA2, NEUROD1, HSF2, OTX2, OTX1, E2F5, RXRG, TFAP2A, TBX1, ARNT2, TFAP2B, ZFHX3, TCF7L1, BCL11B, BCL11A, GSX2, PBX3, NR2F1, FOXJ1, NR2F2, PBX1, KLF7, NR6A1, ZNF219, MAFF, NFE2L3, INSM1, SIM1
KEGG pathway	hsa04360:Axon guidance	3.09	4.18E-09	44	CAMK2B, SEMA5B, SEMA3D, SRC, CAMK2A, LRRC4, CXCR4, SEMA3F, NTN1, GNAI1, RND1, PAK1, PARD6B, ABLIM1, ABLIM2, DPYSL5, DPYSL2, SLIT1, PAK6, FYN, PLXNC1, SRGAP3, EPHB2, HRAS, PAK2, WNT4, SEMA6B, SEMA6C, FZD3, EPHA7, SEMA6A, UNC5A, TRPC3, TRPC4, PTCH1, LIMK1, SEMA4C, SEMA4G, BMP7, RGMA, ENAH, EFNA2, EPHA2
	hsa04740:Olfactory transduction	5.26	5.31E-04	13	CAMK2B, GNAL, PDE1B, ADRBK2, GNG7, CAMK2A, CNGA3, ARRB1, CALM1, PRKACB, CNGB1
	hsa04916:Melanogenesis	3.15	0.0015816	21	CAMK2B, FZD3, TCF7L1, EDN1, FZD2, FZD7, WNT7B, CAMK2A, WNT7A, WNT9A, ADCY8, GNAI1, ADCY5, GNAO1, CALM1, HRAS, PLCB2, PRKACB, WNT4
	hsa04713:Circadian entrainment	3.15	0.0015816	21	CAMK2B, RYR1, GRIA2, KCNJ6, CAMK2A, ADCY8, CACNA1H, GNAI1, ADCY5, CACNA1G, GNAO1, ADCYAP1, GNG2, GNG7, NOS1AP, GNG8, CALM1, PLCB2, PRKACB
	hsa04725:Cholinergic synapse	2.90	0.0034742	22	CAMK2B, CHRNA3, KCNJ6, CHRNB4, CHRNA4, CAMK2A, CHAT, CACNA1B, ADCY8, GNAI1, ADCY5, GNAO1, GNG2,

Category	Term	Fold Enrichment	Bonferroni	Count	Genes
					GNG7, CAMK4, KCNQ2, BCL2, GNG8, FYN, HRAS, PLCB2, PRKACB
	hsa04310:Wnt signaling pathway	2.4144501	0.004529	29	CAMK2B, CTBP2, CAMK2A, NLK, CXXC4, MAPK8, ZNRF3, CCND2, PRKACB, WNT4, TLE3, FZD3, TCF7L1, TLE2, FZD2, FZD7, WNT7B, CTNNBIP1, WNT7A, CSNK1E, AXIN2, WNT9A, SFRP1, VANGL2, SFRP2, DAAM1, CCDC88C, PLCB2, LGR5
Reactome pathway	R-HSA-112316~Neuronal System	3.0491026	9.09E-17	79	GABRB3, KCNG1, KCNK9, KCNC2, GRIK5, RASGRF1, GRIK3, ABAT, AP2A2, KIF17, PTPRF, TUBB3, PRKACB, HRAS, KCNH2, KCNH3, CHRNB4, KCNK10, KCND3, ACTN2, CACNA2D2, SYN2, GJD2, CACNB3, TUBB2B, CACNB4, ALDH5A1, PRKAR1B, KCNQ2, ARHGEF7, PLCB2, VAMP2, SHANK1, CAMK2B, GRIA2, CHRNA3, NLGN2, SRC, CHRNA4, STXBP1, CAMK2A, CHAT, CACNA1B, SLC1A2, NRXN2, SLC1A3, ADCY8, PPM1E, GNAI1, ADCY5, BEGAIN, GNG2, PRKAR2B, GNG7, SLITRK3, SLITRK6, GNG8, SLITRK5, KCNMB4, KCNN2, KCNN3, CACNG2, CACNG4, LRRC4B, KCNJ6, GABRA5, GAD1, GAD2, SYT7, GNAL, DLG4, LRFN1, CAMK4, APBA1, CALM1, APBA2, KCNK3
	R-HSA-112315~Transmission across Chemical Synapses	3.1121489	7.52E-11	54	GABRB3, GRIK5, RASGRF1, GRIK3, ABAT, AP2A2, KIF17, TUBB3, PRKACB, HRAS, CHRNB4, ACTN2, CACNA2D2, SYN2, CACNB3, TUBB2B, CACNB4, ALDH5A1, PRKAR1B, ARHGEF7, PLCB2, VAMP2, CAMK2B, GRIA2, CHRNA3, SRC, CHRNA4, STXBP1, CAMK2A, CHAT, CACNA1B, SLC1A2, SLC1A3, ADCY8, PPM1E, GNAI1, ADCY5, GNG2, PRKAR2B, GNG7, GNG8, CACNG2, CACNG4, KCNJ6, GABRA5, GAD1, GAD2, GNAL, DLG4, CAMK4, APBA1, CALM1
	R-HSA-1266738~Developmental Biology	1.8218606	1.66E-07	112	FOXA1, MSI1, RBPJ, SOX2, SALL1, CDH2, DPYSL5, TUBB3, DPYSL2, SALL4, DPYSL3, MYB, HOXA2, PLXNC1, EPHB2, PRKACB, SCN1A, EPHA7, SEMA6A, UNC5A, POU3F1, ANK1, ISL1, STX1B, CACNB3, CACNB4, DOK4,

Category	Term	Fold Enrichment	Bonferroni	Count	Genes
					DOK5, KCNQ2, HOXB4, HOXB3, PFN2, EPHA2, SPTBN4, SHC3, NOTCH1, EPHA10, GATA2, RAP1GAP, CACNA1H, CACNA1G, NEUROD1, ST8SIA4, SRGAP3, SCN3B, MYH10, NKX2-2, JUP, NR2F2, PBX1, SMARCA4, EFNA2, LHX2, PTPRA, DLG4, RNPS1, LHX4, LHX9, NCAN, MAST1, PCSK6, AP2A2, RND1, SPTAN1, HRAS, HSP90AA1, TRPC3, TRPC4, GAB1, PAX6, RGMB, GFRA2, SREBF2, RGMA, VAV2, ENAH, GAP43, TUBB2B, DPPA4, TAL1, ARHGEF7, HDAC2, SRC, LAMA1, CRMP1, CXCR4, ACVR1B, DLL1, NTN1, ABLIM1, PAK1, MAPK8, ABLIM2, SLIT1, PAK6, FYN, CACNG2, PAK2, CACNG4, CLASP2, SPAG9, SCD5, LIMK1, LGI2, ACVR2A, NFASC, KRT18, NR6A1, INSM1, TRIM33, CDK5R1
	R-HSA-112314~Neurotransmitter receptors and postsynaptic signal transmission	3.1255847	1.97E-07	40	GABRB3, CAMK2B, GRIA2, CHRNA3, CHRNA4, SRC, GRIK5, RASGRF1, CAMK2A, GRIK3, ADCY8, AP2A2, KIF17, PPM1E, GNAI1, ADCY5, GNG2, PRKAR2B, TUBB3, GNG7, GNG8, CACNG2, HRAS, PRKACB, CACNG4, KCNJ6, CHRNB4, ACTN2, GABRA5, TUBB2B, GNAL, PRKAR1B, DLG4, CAMK4, APBA1, CALM1, ARHGEF7, PLCB2
	R-HSA-162582~Signal Transduction	1.4494456	1.98E-07	221	CYFIP2, VIPR2, WIPF3, STMN2, RBPJ, SOX2, STMN1, MYB, PRKACB, TMPO, RBFOX2, ACTN2, SOX13, CSNK1E, SFRP1, SFRP2, PRKAR1B, DAAM1, RRAGD, SPATA13, PFN2, EPHA2, SPTBN4, CRABP2, CTBP2, PDE1B, CRABP1, PDGFB, NEDD4L, GATA3, IQGAP2, NLK, SKA2, PRKAR2B, RDH13, RALGDS, ATP6V0A1, NGFR, FZD3, FZD2, JUP, MCAM, FZD7, GREB1, BRAF, PBX1, SMARCA4, CCDC88A, FAM135A, VANGL2, NMU, CCDC88C, ID4, CALM1, CDC42SE2, GABRB3, SHB, PCSK6, AP2A2, PTPRF, SMPD3, LFNG, HEY1, SPTAN1, HRAS, EDN1, GPR37, STRADB, GAB1, TACR1, WNT9A, DUSP8, IL17RD, GFRA2, VAV2, TUBB2B, CDC42EP4, ARHGEF7, PLCB2, PHLPP1, CXCR4, ADCY8, GNAI1, ADCY5, NEURL1B, PAK1, ADRBK2, CHN1, PAK6, CKB,

Category	Term	Fold Enrichment	Bonferroni	Count	Genes
					EPS15, RXRG, PAK2, WNT7B, DAB2IP, WNT7A, FABP7, CAMK4, F2RL1, LGR5, CDK5R1, FOXA1, DGKD, WWC1, IGF1R, TUBB3, BASP1, CDH1, SALL4, TLE3, USP7, TLE2, F2R, CTNNBIP1, DISP2, ADCYAP1, MYCN, MADD, MAPKAPK5, ABCG1, RTN4R, SHC3, NOTCH1, DTX1, RASAL2, DTX4, RHOBTB2, ARHGAP12, HRH3, RHPN1, SRGAP3, RHPN2, TAC1, MYH10, MARK3, WNT4, CBX8, POU2F1, CBX4, PTCH1, CBX2, GNAO1, ARHGAP33, BCL9, GNAL, PTPRA, DLG4, BCL2, YWHAE, RASGRF1, MAST1, ARHGEF10L, ARRB1, TRH, NR2E1, DUSP16, RND1, LMNB1, SPRED3, GRM3, GRM2, MDK, PPP2R5E, MCF2L, LBR, YWHAG, YWHAH, PLEKHG1, HSP90AA1, TRPC3, KSR1, ARHGEF16, AXIN2, OPRM1, SSTR2, ARL4C, VGF, DAGLB, CAMK2B, LAMA5, KDM3A, HDAC2, SRC, LAMA1, RGS16, CAMK2A, ACVR1B, FAM169A, DLL1, PARD6B, MAPK8, GNG2, ZNRF3, MTA1, PTPRZ1, GNG7, SLITRK3, IGF2BP1, GNG8, RGS20, SLITRK5, FYN, E2F5, SH2B2, MAPK4, CLASP2, PPP1R14A, LINGO1, TCF7L1, PCDH7, LIMK1, ACVR2A, TNRC6C, SST, SPRY1, RGS12, PDE7A, TRIM33