Osteoarthritis and Cartilage 32 (2024) 1419–1432

Osteoarthritis and Cartilage

Inverse genetics tracing the differentiation pathway of human chondrocytes

H.T. Do # †, M. Ono # ‡, Z. Wang #, W. Kitagawa # †, A.T. Dang # †, T. Yonezawa #, T. Kuboki † ‡, T. Oohashi #, S. Kubota § *

Department of Molecular Biology and Biochemistry, Okayama University Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama, Japan
† Department of Oral Rehabilitation and Regenerative Medicine, Okayama University Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama, Japan
‡ Department of Oral Rehabilitation and Implantology, Okayama University Hospital, Okayama, Japan
§ Department of Biochemistry and Molecular Dentistry, Okayama University Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama, Japan

ARTICLE INFO

Article history:
Received 21 December 2023
Accepted 18 June 2024

Keywords:
Cartilage
Chondrocyte
Differentiation
Reprogramming
CCN family

SUMMARY

Objective: Mammalian somatic cells can be reprogrammed into induced pluripotent stem cells (iPSCs) via the forced expression of Yamanaka reprogramming factors. However, only a limited population of the cells that pass through a particular pathway can metamorphose into iPSCs, while the others do not. This study aimed to clarify the pathways that chondrocytes follow during the reprogramming process.
Design: The fate of human articular chondrocytes under reprogramming was investigated through a time-coursed single-cell transcriptomic analysis, which we termed an inverse genetic approach. The iPS interference technique was also employed to verify that chondrocytes inversely return to pluripotency following the proper differentiation pathway.
Results: We confirmed that human chondrocytes could be converted into cells with an iPSC phenotype. Moreover, it was clarified that a limited population that underwent the silencing of *SOX9*, a master gene for chondrogenesis, at a specific point during the proper transcriptome transition pathway, could eventually become iPSCs. Interestingly, the other cells, which failed to be reprogrammed, followed a distinct pathway toward cells with a surface zone chondrocyte phenotype. The critical involvement of cellular communication network factors (CCNs) in this process was indicated. The idea that chondrocytes, when reprogrammed into iPSCs, follow the differentiation pathway backward was supported by the successful iPS interference using *SOX9*.
Conclusions: This inverse genetic strategy may be useful for seeking candidates for the master genes for the differentiation of various somatic cells. The utility of CCNs in articular cartilage regeneration is also supported.

Introduction

Differentiation from stem cells to somatic cells is eventually executed via the dynamic transition of the transcriptome, in which transcription factors play critical roles.[1] Cartilage is produced and maintained by chondrocytes, and the SRY-box transcription factor (SOX) 9 gene is widely recognized as a master gene of chondrogenesis.[2,3] However, it is not yet known what initiates *SOX9* at a particular point along the differentiation pathway toward chondrocytes. This is partly because of the difficulty in characterizing the transcriptomic transition that occurs in differentiating cells in vivo. We discovered several methods that force stem cells to

* Correspondence to: Department of Biochemistry and Molecular Dentistry, Okayama University Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, 2-5-1 Shikata-cho, Kita-ku, Okayama 700-8525, Japan.
E-mail addresses: pa62884x@s.okayama-u.ac.jp (H.T. Do), mitsuaki@md.okayama-u.ac.jp (M. Ono), wangziyi@s.okayama-u.ac.jp (Z. Wang), wakana.kitagawa@s.okayama-u.ac.jp (W. Kitagawa), ppjj5xuj@s.okayama-u.ac.jp (A.T. Dang), tomoy@okayama-u.ac.jp (T. Yonezawa), kuboki@md.okayama-u.ac.jp (T. Kuboki), oohashi@cc.okayama-u.ac.jp (T. Oohashi), kubota1@md.okayama-u.ac.jp (S. Kubota).

https://doi.org/10.1016/j.joca.2024.06.009

differentiate into cartilage-producing cells in vitro; however, the pathway of transcriptomic transition and even the final transcriptomic landscape of these cells could be different from those naturally occurring in vivo. This is an issue relevant to regenerative therapeutics with cells differentiated in vitro from pluripotent stem cells.[4,5]

In contrast to the complexity during differentiation into somatic cells, the reprogramming pathway of somatic cells back into induced pluripotent stem cells (iPSCs) appears quite simple. Indeed, the overexpression of only four genes (Yamanaka factors) encoding SOX2, octamer binding transcription factor (OCT)3/4, Küppel-like factor 4 (KLF4) and myelocytomatosis protooncogene product (MYC) is believed to convert most somatic cells into iPSCs.[6,7] However, we suspect that the reprogramming pathway is not a single process common to all cells, through which any cell-specific transcriptome is simply erased, but must be a cell-type-specific process. In particular, although Yamanaka factors are overexpressed in somatic cells, only a limited population is successfully converted into pluripotent cells, indicating the requirement of additional conditions.[8,9] One possible idea to account for this limited efficacy of reprogramming is that cells following the proper pathway of transcriptomic transition could only obtain pluripotency. In other words, under the driving force of reprogramming, a particular population of cells can follow the pathway leading to iPSCs, whereas others may take off-targeted pathways leading to dead ends. Regarding this pathway toward iPSCs, a promising hypothesis can be drawn from recent studies. It is known that the transient induction of transcripts from murine endogenous retrovirus L plays a critical role in the early development of mice.[10] In 2014, it was reported that the hyper-activation of human endogenous retrovirus (HERV) H was observed at a late stage in the reprogramming of human dermal fibroblasts into iPSCs.[11] Interestingly, the activation of HERV-H has been shown to be necessary for complete reprogramming into iPSCs. Conversely, HERV-H is highly expressed to create topologically associated domains in human pluripotent stem cells, which declines along with differentiation.[12] Indeed, the reprogramming of human dermal fibroblast into iPSCs inversely recapitulated the early differentiation process of pluripotent cells. Based on these findings, we hypothesized that somatic cells metamorphose into iPSCs by retracing their natural and proper differentiation pathway backward. If so, by tracing the transition pathway of the transcriptome during the reprogramming of the cells of interest, a differentiation master gene of those cells could be retrospectively discovered (Fig. 1A). In order to verify this hypothesis, we chased the transcriptomic trajectory from human chondrocytes toward iPSCs by time-coursed single-cell transcriptomic analysis. This inverse genetic approach may enable us to view the changing transcriptomic landscape from pluripotent stem cells to articular chondrocytes backward. At the same time, the fates of off-targeted cells were also pursued. Unexpectedly, those cells followed a single pathway toward cells with surface zone chondrocyte characteristics under the cell culture conditions employed (Fig. 1A). The critical involvement of cellular communication network factor (CCN) family members[13,14] in this terminal conversion to surface zone chondrocytes is suggested. Here, the utility of inverse genetics in identifying a master regulator of cytodifferentiation is indicated.

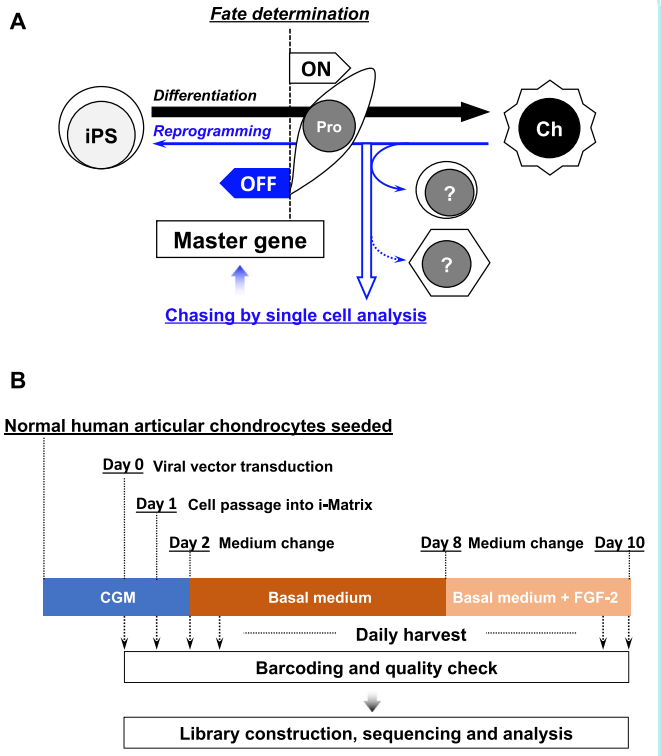


Fig. 1 Osteoarthritis and Cartilage

Research strategy of this study. (A) Overview of the inverse genetic approach. The transcriptome transition pathways (blue lines) of both the cells under reprogramming and those that failed to be reprogrammed (off-targeted) are chased, following a time course. The former may enable the discovery of a master gene of chondrogenesis, whereas the latter clarifies the fate of those cells. Abbreviations: iPS, induced pluripotent cell; Pro, chondrocyte progenitor; Ch, chondrocytes; ?, off-targeted cells. (B) Experimental time course of the single-cell transcriptome analysis for inverse genetics. CGM: chondrocyte growth medium.

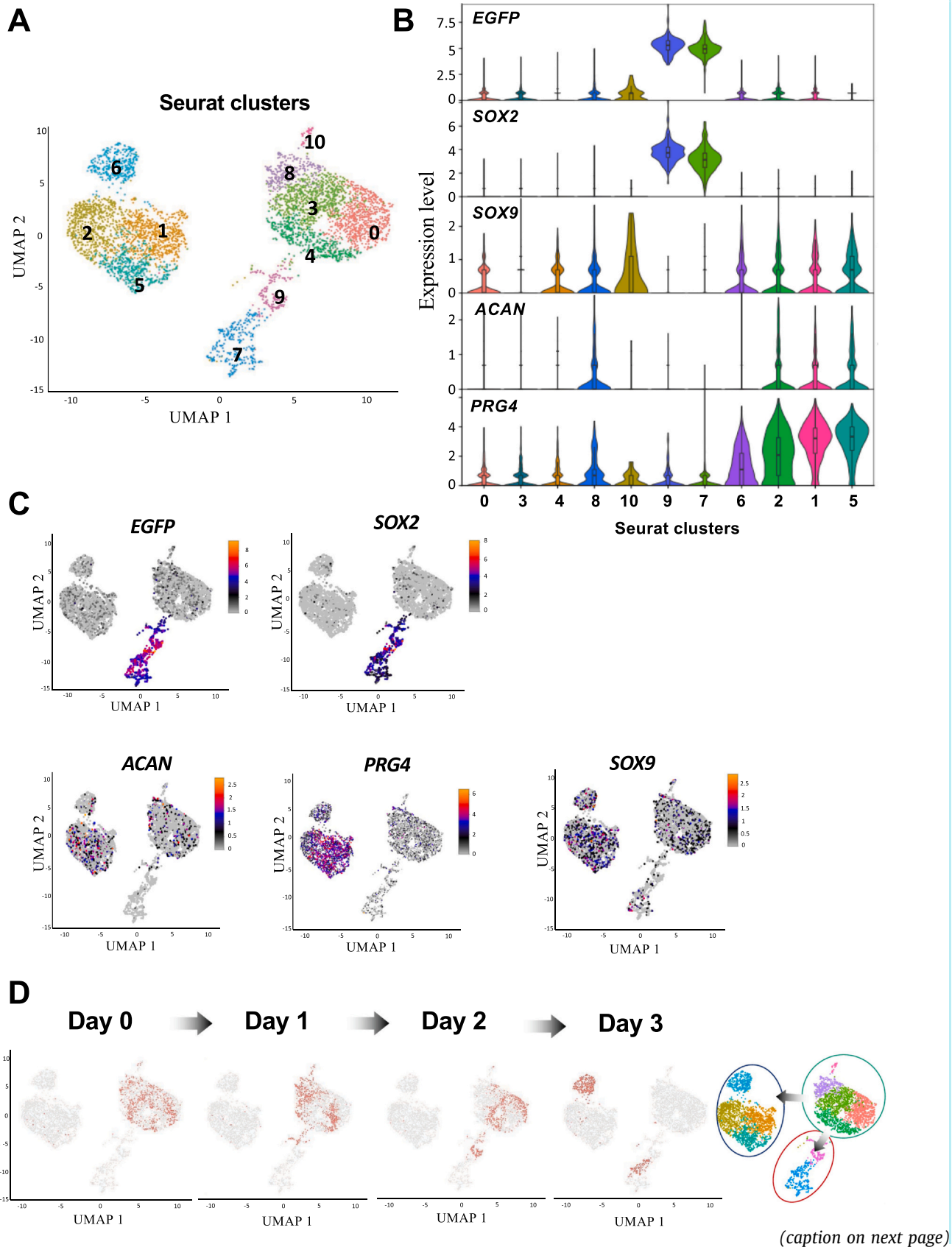
Method

Viral vectors used for genetic reprogramming and iPS interference

For the genetic reprogramming of human chondrocytes into iPSCs, the gene delivery of individually defined factors—*OCT4, KLF4, SOX2, c-MYC, NANOG* and *LIN28*—was performed using a stealth RNA virus (SRV) iPSC-3 vector that also expresses enhanced green fluorescent protein gene (*EGFP*) (Tokiwa-Bio, Tsukuba, Japan). This Sendai virus-based vector persistently expresses transgenes without being integrated into host genome and thus does not undergo epigenetic silencing upon genetic reprogramming.[15] For iPS interference, retroviral vectors expressing *SOX9* and *EGFP* were used.

Cell culture and genetic reprogramming

Human normal articular chondrocytes were commercially available (Lonza, Basel, Switzerland). The initial step of this study via single-cell transcriptomic analysis and primary iPS interference experiment was performed by the cells from a healthy donor (NHAC - kn 38588). Thereafter, the obtained finding was extensively confirmed by iPS interference experiments with those from another donor (NHAC - kn 42653, Lonza). The overview and details for reprogramming are described in Fig. 1B and the Supplemental Method. Numbers of the cells eventually collected on each day are 170,000 (day 0), 230,000 (day 1), 160,000 (day 12), 90,000 (day 3), 100,000 (day 5), 130,000 (day 6), 110,000 (day 7), 160,000 (day 8), 250,000 (day 9) and 140,000 (day 10).



(caption on next page)

Fig. 2 Osteoarthritis and Cartilage

Two distinct pathways for transcriptome transition revealed via scRNA-seq. (A) A UMAP plot of chondrocytes under genetic reprogramming conditions was analyzed using Seurat. Colors marked the 11 distinct clusters. Each dot represents each cell. (B) Violin plot for the expression levels of representative genes in different subclusters. (C) Distribution of *EGFP-, SOX2-, SOX9-,* and *ACAN*-positive cells in the total population on UMAP. Expression levels computed via scTransform R-package in Seurat are represented with colors. (D) Actual time course of the transcriptome transition in chondrocytes under reprogramming until day 3. Red dots and light-gray ones on UMAP represent the cells harvested on the indicated day and the total cells analyzed, respectively. The chondrocytes in Seurat clusters circled in green followed two distinct fates. Their directions of transition along the time course are indicated with arrows.

Single-cell RNA sequencing

Chondrocytes collected at each time course were labeled with a 10X Genomics 3′ CellPlex kit. After the cellular labeling, all cells were mixed, and cell sorting was performed on a FACSAria III Cell Sorter (BD Biosciences, San Diego, CA, USA) to remove dead cells and debris via forward scatter, side scatter and 4′,6-diamidino-2-phenylindole (DAPI)/Calcein-AM staining. The viability of the cells forwarded to the analysis was 95.2%. A total of 6513 cells were then processed with Chromium Next GEM Single Cell 3′ GEM kit v.3.1 (10X Genomics Inc., Pleasanton, CA, USA) according to the manufacturer's instructions. Complementary DNA libraries were sequenced using the NovaSeq platform (Illumina). In total, 409,314,286 reads were obtained (62,846 reads per cell). To remove doublet cells, Solo software was used.[16] Data processing detail is described in the Supplemental Method.

RNA velocity and trajectory analysis

RNA velocity analysis of single cells was performed using scVelo (version 0.2.5) software[17] with default parameters. Subsequently, trajectory analysis was performed by using the Single-cell Trajectory Reconstruction Exploration And Mapping (STREAM; version = 1.1) algorithm, which is a trajectory inference method that can robustly reconstruct developmental trajectories with relatively accurate pseudotime estimation from single-cell RNA sequencing (scRNA-seq) data.[18] Details are provided in the Supplemental Method.[19]

Evaluation of iPS interference

In order to verify if the constitutive expression of *SOX9* interferes with the reprogramming of chondrocytes into iPSCs, RNA interference experiments were carried out following an established protocol.[20] Experimental details are given in the Supplemental Method.

Statistical analysis

Statistical analyses of the two experimental groups were performed by using an unpaired Student's *t*-test with Prism 10 (GraphPad, MA. USA). Differentially expressed gene (DEG) analysis of the scRNA-seq data to identify cell-type-specific genes was performed using the nonparametric Wilcoxon rank sum test (detailed in Supplemental Method).

Data availability

The data presented in this study are available in the Source data as detailed in the Supplemental Method.

Results

Two distinct pathways for transcriptome transition in chondrocytes under genetic reprogramming conditions

In order to reprogram human articular chondrocytes into pluripotent stem cells, we infected the cells with SRV-iPSC3. Thereafter, the cells were harvested daily, and their transcriptomes were analyzed following a time course, as shown in Fig. 1B.

Single-cell transcriptomic analysis revealed that there are 11 clusters in the uniform manifold approximation and projection (UMAP) (Fig. 2A). Based on violin and dot plots of gene expression, we could classify the clusters into three distinct groups. These groups consist of clusters 0, 3, 4, 8 and 10, with no prominent transcriptomic characteristics; clusters 9 and 7, with high *EGFP* and *SOX2* expression and without *SOX9* expression; and clusters 6, 2, 1 and 5, presenting higher expression levels in *PRG4* that encodes lubricin (Fig. 2B and C). Interestingly, from days 0 to 3, the cells gathered separately from clusters 0, 3, 4, 8 and 10 into two distinct clusters, 6 and 7/9 (Fig. 2D). This indicates that cells in cluster 6 did not receive, or received but did not express, the transgenes for reprogramming and thus, stayed within a lineage of chondrocytes. Therefore, the cells moving toward cluster 7 through cluster 9 were those proceeding to pluripotent cells, whereas the others in cluster 6 failed to be reprogrammed and followed a single pathway into something else (Fig. 2D).

Spontaneous silencing of SOX9 occurred at an early stage of transcriptomic reprogramming: re-discovery of SOX9 as a master gene of chondrocytic differentiation

If our hypothesis is correct, chondrocytes being reprogrammed into iPSCs retrospectively follow the differentiation pathway. To verify this, the expression of *SOX9* in the reprogrammed chondrocytes was traced, along with the transition from cluster 4 to cluster 7. It was demonstrated that the cells migrated within 3 days from cluster 4 to clusters 9 and 7 (Fig. 2D), which were characterized by the high expression of *EGFP* and *SOX2* (Figs. 2B and 3A). After that, the cells expressing these genes gradually changed the transcriptomes to those of *NANOG*-positive pluripotent cells located at the bottom of cluster 7 on the UMAP (Fig. 3B and C). These data suggest that the cells under reprogramming followed a distinct transcriptomic trajectory toward iPSCs (Fig. 3D). Importantly, the expression of *SOX9* and *ACAN* was conversely turned off upon the transcriptomic transition from cluster 4 to cluster 9; Fig. 3A, which happened between day 1 and day 2 during the actual time course (Supplemental Fig. 1), recapitulating the cellular commitment toward chondrocytes in a retrospective manner. These findings again indicate that *SOX9* had to be silenced at a specific point, following the pathway of chondrocytic differentiation in an inverse manner during reprogramming. In other words, only the cells that retrospectively recapitulated the proper differentiation events could ultimately become iPSCs. Interestingly, marker genes of articular chondrocytes were silenced not before,

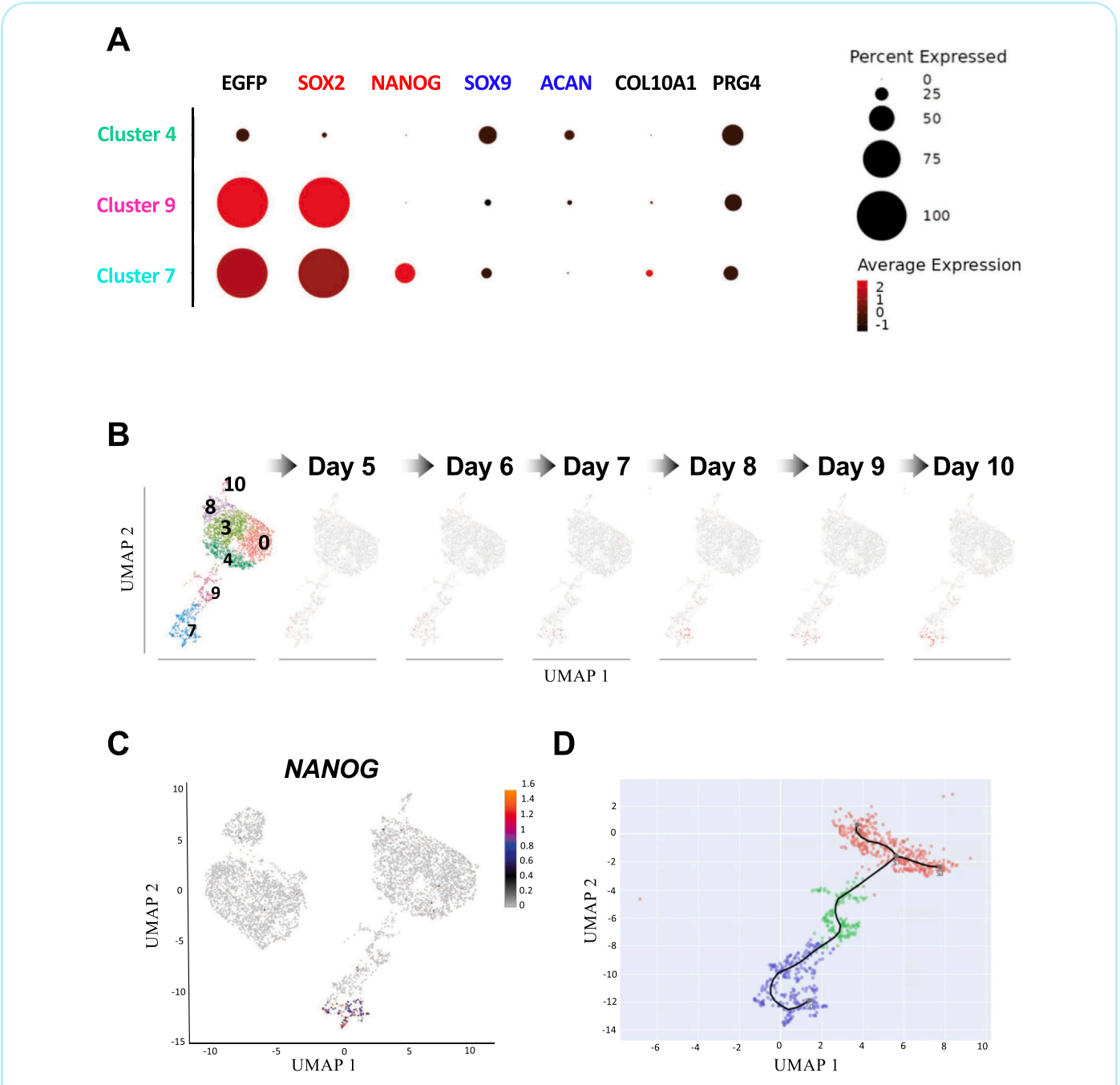
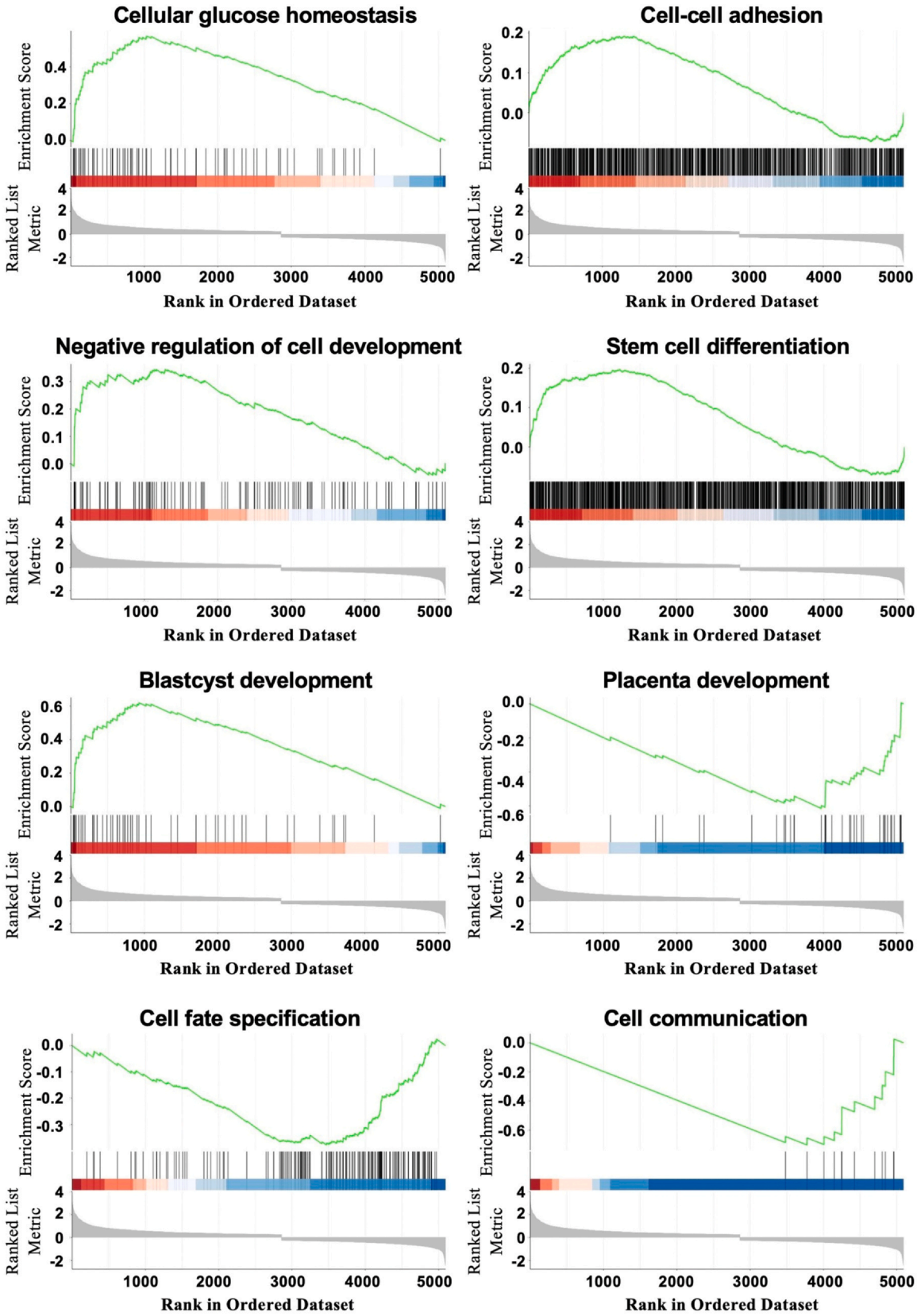


Fig. 3

Osteoarthritis and Cartilage

Transcriptomic transition pathway from chondrocytes to pluripotent cells. (A) Expression of the marker genes of pluripotent cells and chondrocytes among Seurat clusters 4, 9 and 7. Cluster 4 consists of the cells between day 0 and day 2, whereas cluster 9 is composed of the cells at days 1 and 2, and Day 3 cells in this pathway are mapped mostly in cluster 7 (Fig. 2D). Expression levels indicated in red and black represent z-scores. *SOX9* expression was turned off upon the transition from cluster 4 to cluster 7. (B) Actual time course of the transcriptome transition in chondrocytes toward iPSCs until day 10. Red and light-gray dots represent the cells harvested on the indicated day and total cells in the indicated clusters, respectively. (C) Distribution of *NANOG*-positive cells. Note that only a limited population of the cells in cluster 7 are *NANOG*-positive. (D) Transition trajectory of the states computed via STREAM, indicating that the cells in cluster 4 are directed toward becoming pluripotent cells.

but upon or after *SOX9* silencing along with actual time course (Supplemental Figure 1). These findings confirm that the expression of *SOX9* and these marker genes changes concullently either in differentiation or reprogramming processes. Additionally, a minor population with *COL10A1* expression is found also in cluster 7. Considering that these cells were not accompanied by strong *NANOG* expression (Supplemental Figure 2), such a cell was an incompletely reprogrammed one without reaching pluripotency.



(caption on next page)

Fig. 4

Osteoarthritis and Cartilage

GSEA of the cells in cluster 7 versus cluster 9. Transcriptomes of these clusters were analyzed with the gene sets under the gene ontologies closely related to energy metabolism, cell differentiation, embryonic development and cellular communication. Gene ontologies to which corresponding gene sets belong are displayed at the top.

Proper reprogramming pathway confirmed by gene set enrichment analysis

Regarding the transcriptomic transition pathway from cluster 4 to cluster 7 via cluster 9, genes that differentially expressed cluster 7 versus cluster 9 were extracted and were subjected to gene set enrichment analysis (GSEA) on gene ontology (Fig. 4). Genes critical for cellular glucose metabolism were induced toward the higher dependence of energy metabolism on glycolysis in pluripotent stem cells. Cell-to-cell adhesion potential was enhanced, which is consistent with the acquisition of the colony-forming ability of stem cells. In line with this finding, negative regulation took place during cell differentiation, whereas differentiation into stem cells was promoted. Strong and poor enrichment of the genes for blastocyst and placenta development, respectively, suggests that these cells progressed to pluripotency rather than totipotency. Finally, the results on cell fate specification and cell communication indicate that the cells lost a specific differentiated phenotype established through intercellular communications. Collectively, these results indicate that this transcriptome transition pathway is that of reprogramming into pluripotent cells.

Reprogramming force driving the cells against their apparent transcriptomic velocity

Next, we analyzed the dynamics of the transcriptome in the cells involved in clusters 4, 7 and 9 through RNA velocity algorithms (scVelo). Then, we aligned the cells in clusters on a pseudotime axis computed via STREAM (Fig. 5A). Overall, the alignment of each cell was consistent with that of the actual time, indicating the accuracy of this pipeline in describing population relationship. However, the direction of pseudotime was inverse to that of the actual time. PAGA velocity graph based on the RNA velocity analysis also showed the same direction as that given by STREAM (Fig. 5B–D). On the contrary, the results of cell cycle analysis support the actual time course, in which proliferating cells in cluster 4 gradually enter dormancy as stem cells during transition toward cluster 7 (Fig. 5E).

In order to account for this discrepancy, GSEA of DEGs in cluster 7 versus cluster 9 was performed. As a result, we found that RNA splicing was negatively regulated in a genome-wide manner (Fig. 5F), thus affecting the results of RNA velocity. Therefore, the accumulation of unspliced RNAs due to genome-wide repression of splicing may result in an apparent increase in computed RNA velocity based on such an algorithm, which may explain the observed discrepancy in the directions between actual transcriptomic transition and the computed RNA velocity (Supplemental Figure 3). These results suggest that the overexpression of the reprogramming sextet is a strong drive force that alters the transcriptome of the cells through the post-transcriptional regulation of mRNAs.

If this notion is correct, in the absence of the sextet, it is assumed that the computed velocity of the transcriptome in the cells should be consistent with the actual transition of the transcriptome along the time course. To confirm this idea, RNA velocity analysis was performed among the cells around cluster 2. As expected, the RNA velocities of the cells in cluster 6 show the same direction as that of the actual transcriptomic transition from cluster 6 to cluster 2 (Fig. 6A and Supplemental Figure 4).

The single pathway of transcriptomic transition found in the off-targeted cells toward PRG4-positive chondrocytes

Initially, we suspected that the cells that failed to be reprogrammed might have followed diverse differentiation pathways and resulted in multiple cell types with undefinable phenotypes (Fig. 1A; indicated with blue lines pointing to the cells at the right end). However, under the cell culture conditions employed, almost all of these cells eventually displayed a comparable transcriptome profile (clusters 1 and 5; Fig. 6A and B). In contrast to the cells in clusters 7 and 9, significant portions of those cells expressed *ACAN* and *SOX9*, even more prominently, indicating the retention of a chondrocytic phenotype. Importantly, the vast majority of *PRG4*-positive cells were accumulated in clusters 1, 2 and 5 (Fig. 6B; right panel). These cells are characterized by enhanced *PRG4* expression and silenced *COL10A1*, which represents the phenotype of surface zone chondrocytes that produce lubricin for the locomotive function of articular cartilage.[21] The fact that no positive correlation was observed between the expression levels of *PRG4* and reprogramming factors in those cells rules out a possibility that *PRG4* is a target gene of these factors (Supplemental Figure 5). This outstanding *PRG4* expression is in a sharp contrast with the chondrocytes in cluster 4, which also show silenced *COL10A1* (Fig. 3A), probably because of the onset of reprogramming. The cells harvested from day 7 to day 10 stayed in the same *PRG4*-highly positive clusters 1 and 5 without major changes in their transcriptomes (Fig. 6A). In addition, laminin genes that are known to be closely related to stemness[23] were induced in the cells in clusters 1 and 5 (Supplemental Figure 6). The top 10 induced genes in clusters 1 and 5 share *BST2, MX1*, *IF16* and *PRG4*. *BST2* encodes bone marrow stromal cell antigen 2 with an unknown function[22]; interestingly, both *MX1* and *IFI6* are interferon-inducible genes.[24,25]

Transient induction of CCN family members prior to the terminal transition

To characterize the precursors of *PRG4*-positive cells, DEGs in cluster 6 in comparison with the total population were extracted. As shown in Fig. 7A, *MGP, CCN2* and *SCRG1,* which encode matrix Gla protein,[26] CCN factor 2 and stimulator of chondrogenesis 1, respectively, were ranked as the top three upregulated genes. With regard to downregulated genes, three transgenes used for reprogramming, *EGFP, POU5F1 (OCT3/4)* and *SOX2*, were identified, confirming the reprogramming failure. Although *MYC* was also downregulated (logFC = −1.642941 at p value adj. = 1.49803E-09), the other reprogramming genes were not found downregulated, suggesting the involvement of partially reprogrammed cells.

Among six members of the CCN family,[27,28] positive roles in chondrogenesis have been reported for *CCN1, CCN2* and *CCN4*.[29,30] *CCN2* is the most differentially expressed among the clusters, followed by *CCN1* and *CCN4* (Supplemental Figure 7). All three members are parallelly upregulated in clusters 2 and 6 in contrast to the tenascin C gene (*TNC*), which is a marker of articular chondrocytes. On the contrary, these genes were silenced in clusters 7 and 9, showing the same pattern (Fig. 7B). It is also noted that the expression level of *CCN4* is globally lower than the others, and thus a different scale is used for the z-scores. No significant differences were observed in the other *CCN* family members. From this point of view, chondrocytes under reprogramming may be classified into

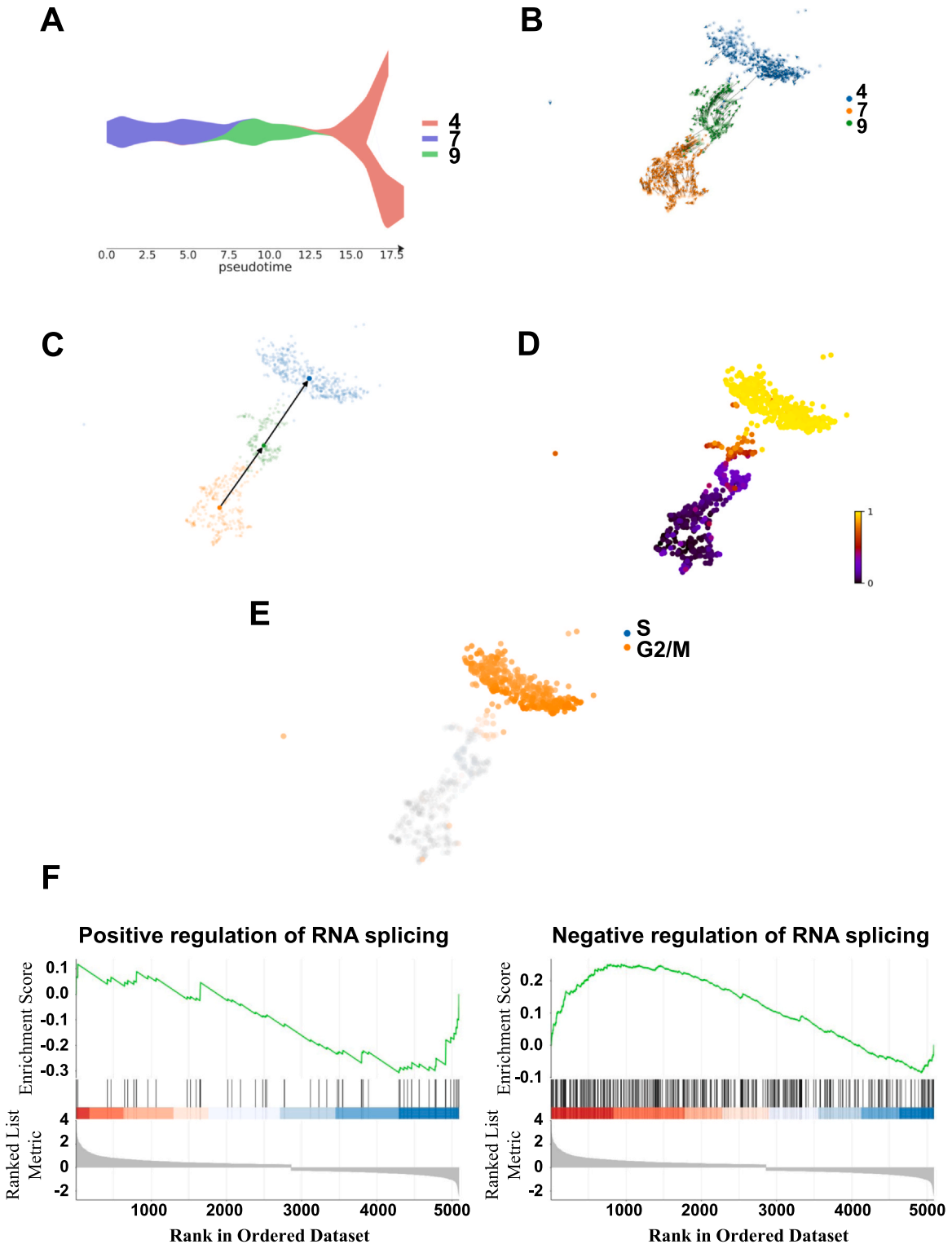


Fig. 5

Osteoarthritis and Cartilage

Forced transition of transcriptome against computed RNA velocity by the reprogramming factors. (A) Pseudotime trajectory from Seurat clusters from 7 to 4 through 9. Note that it is entirely inverse of the actual time course. (B) RNA velocity analysis of each cell in Seurat clusters 4, 9 and 7. (C) PAGA velocity graph among the clusters and (D) velocity pseudotime (D) of each cell, which again confirmed inverse RNA velocity direction against actual time. (E) Distribution of the cells in G2/M and S phases in those clusters. (F) GSEA of DEGs in cluster 7 versus cluster 9, suggesting global repression of RNA splicing in cluster 7.

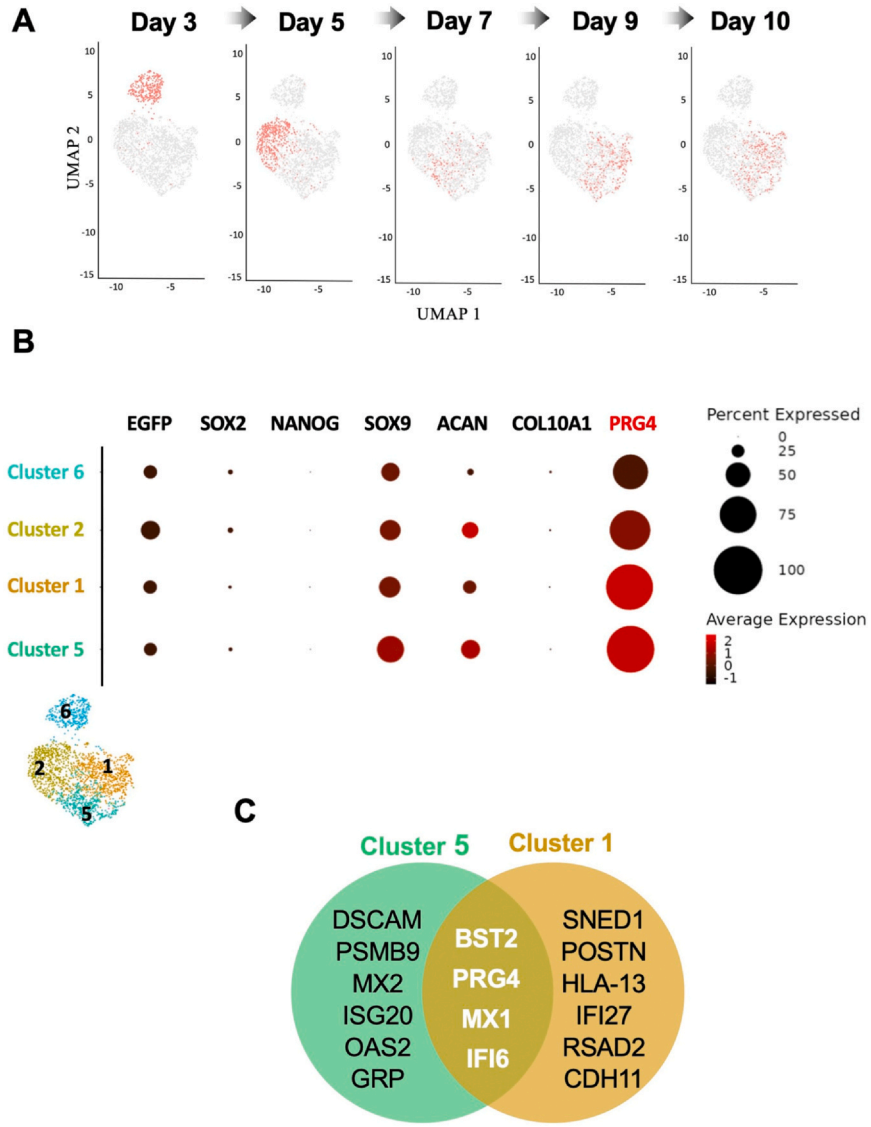


Fig. 6 Osteoarthritis and Cartilage

The off-targeted cells in cluster 6 follow a single transition pathway toward surface zone chondrocytes. (A) Actual time course of the transcriptome in chondrocytes that failed to be reprogrammed. The cells for each day (red) are plotted in the total cells (light gray) in clusters 6, 2, 1, and 5. (B) Expression of pluripotent and chondrocytic marker genes among clusters 6, 1, 2 and 5. Z-scores are represented with red/black gradation. (C) Top 10 DEGs in the cells in clusters 1 and 5. DEGs were extracted against the expression levels of the total population.

three groups (Fig. 7C): clusters 2 and 6 as CCN-induced cells and clusters 7 and 9 as CCN-silenced ones.

Phenotypic confirmation of reprogramming from chondrocytes and SOX9 as a master gene by iPS interference

To verify that chondrocytes were converted into pluripotent cell colonies and that *SOX9* expression needs to be turned off along with the reprogramming pathway, the iPS interference strategy was employed. One day after infection, *GFP*-positive cells were evenly distributed in both groups (Fig. 8A and B), indicating successful transfection of reprogramming factors. *SOX9* overexpression was also confirmed in the *SOX9* retrovirus transfected group (Supplemental Figure 8A). On day 6, the cells with *GFP* grew in the control without *SOX9*, whereas no such growth was evident in *SOX9*-overexpressing cells. Then, on day 8, in the control group, the number and size of the colonies increased, and large colonies with highly expressed *GFP* showed typical iPSC-like morphology on day 12 (Fig. 8B). In contrast, colony formation was not yet evident among the cells with *SOX9* on day 8, and few colonies were observed after 4 days (Fig. 8A and B).

Moreover, staining with fluorescently labeled rBC2LCN,[31] which recognizes human embryonic stem cell-specific sugar chains, confirmed that pluripotent cells constituted these colonies (Fig. 8C), thus indicating that human chondrocytes could be successfully reprogrammed. Furthermore, compared with the control, the group that overexpressed *SOX9* had far fewer rBC2LCN-positive colonies (Supplemental Figure 8B and Fig. 8D). These data also indicate a high

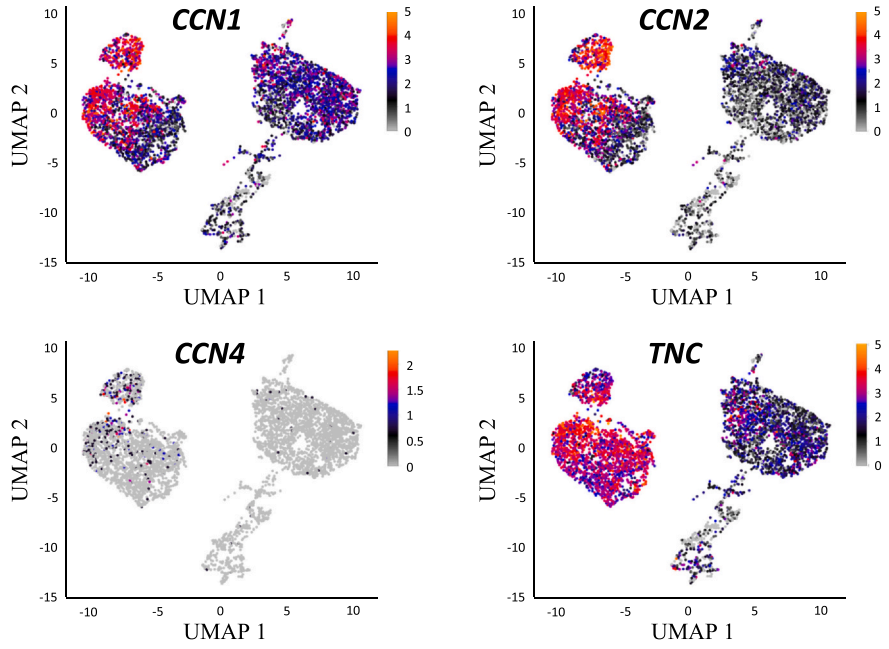
A

Seurat cluster 6 vs the others

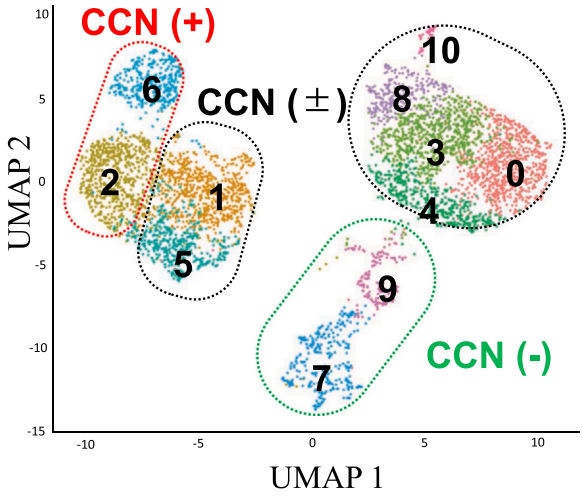
2.50941585	2.73E-107	*MGP*	-4.1370376	0.002902945	*EGFP*
2.40112723	4.2277E-127	*CCN2*	-3.3873987	0.000230222	*POU5F1 (OCT3/4)*
2.37800927	1.0827E-156	*SCRG1*	-2.6913026	3.572E-101	*CENPF*
2.22176644	1.4444E-109	*CYP1B1*	-2.6786217	2.4472E-104	*MKI67*
2.01935897	8.2667E-116	*FABP3*	-2.6354694	0.0063866	*SOX2*
2.0040675	4.0239E-139	*GXYLT2*	-2.6188408	4.47363E-56	*H4C3*
1.82234384	1.83797E-51	*KRT7*	-2.4888225	6.2734E-110	*RRM2*
1.71785076	6.3451E-137	*DAB2*	-2.4411478	1.14503E-91	*CCNB1*
1.70716477	1.9372E-103	*CLIC3*	-2.361667	2.06409E-82	*UBE2C*
1.7022903	8.0232E-146	*TRIB3*	-2.3256037	2.27441E-82	*ASPM*

log2FC

B



C



(caption on next page)

Fig. 7

Osteoarthritis and Cartilage

Involvement of CCNs in the transcriptome transition of the off-targeted cells. (A) Top 10 DEGs in the cells in cluster 6. Upregulated and downregulated ones are shown in red and green backgrounds, respectively. (B) Expression of three *CCN* family members and a marker of articular chondrocytes, tenascin C (*TNC*) among the clusters. (C) Summary of the result in B on two-dimensional UMAP. The cells can be classified into 3 groups based on the expression levels of CCNs, as indicated.

transfection efficiency (> 70%) of the retroviral vector expressing *SOX9*. This interference did not occur in the cells transduced with lentivirus expressing *EGFP*, ruling out the influence of lentiviral infection and protein overproduction (Fig. 8E). Therefore, turning off *SOX9* was confirmed to be the indispensable step during the reprogramming of chondrocytes into pluripotent cells, which inversely recapitulated the induction of *SOX9* that directs mesenchymal stem cells toward chondrocytic differentiation.

Discussion

Our time-coursed single-cell transcriptomic analysis enabled us to clarify the real-time trajectory of transcriptomic transition during the reprogramming of chondrocytes into pluripotent cells. Consequently, we confirmed that *SOX9* expression was turned off early among the cells proceeding to iPSCs within two days after the initiation of reprogramming. These cells constituting approximately 20% of cluster 9 eventually became *NANOG*-positive pluripotent cells (Fig. 3). However, among the cells in clusters 1, 2, 5 and 6, which failed to be reprogrammed, more than 50% of the cells were found to be *SOX9*-negative. Considering together with the presence of *SOX9*-negative dedifferentiated chondrocytes in the original population plotted in clusters 0, 3, 4, 8 and 10 (Fig. 2B and C), turning off *SOX9* is necessary but not sufficient for the conversion of chondrocytes into iPSCs. Conversely, certain cells at day 1 are already plotted in *SOX9*-negative cluster 9 that is committed to pluripotency, indicating that dedifferentiated cells could be reprogrammed as well. For successful reprogramming, this turning-off event must occur early within a restricted time frame. During the in vitro differentiation of mesenchymal stem cells into chondrocytes, *SOX9* is induced.[32,33] Therefore, turning off *SOX9* during reprogramming recapitulated turning on *SOX9* during differentiation, indicating that the reprogramming process followed the differentiation process backward. Namely, we could re-discover *SOX9* as a master gene of chondrogenesis by using this inverse genetic approach, confirming our initial hypothesis. Furthermore, the potential of *SOX9* as a regulator of differentiation was functionally proven using the iPS interference methodology.

In our time-coursed analysis, the transcriptomic transition pathway to iPSCs was clearly shown to begin at cluster 4 toward cluster 7 through cluster 9. Nevertheless, RNA velocity showed transition directions opposite to the actual ones. This apparently contradictory data represent the property of the Yamanaka quartet as a strong driving force that redirects the transcriptome against its intrinsic direction. The mechanism by which these factors altered the transcriptome against RNA velocity is unclear. However, these transcription factors may regulate target gene expression not only at transcriptional but also post-transcriptional levels. Indeed, the GSEA results clearly indicate the negative regulation of splicing upon the transition from cluster 9 to cluster 7. This unexpected regulation in splicing complicates the interpretation of the RNA velocity analysis results. In contrast, the RNA velocities of the cells that failed in the reprogramming process were, overall, consistent with the actual transition of the transcriptome.

According to our initial hypothesis, the off-targeted cells were expected to be sporadically plotted without forming a cluster on the UMAP. On day 3, all such cells, including GFP-positive ones, unexpectedly migrated from clusters 0, 3, 4, 8 and 10 to the single cluster 6, suggesting that these cells follow their intrinsic differentiation pathway. The actual time course of the transcriptomic transition on the UMAP indicates that the final destinations of these cells are clusters 1 and 5, which are characterized by a strong expression of *PRG4* and the retention of *SOX9* expression. Thus, the cells in these clusters show the phenotype of surface zone chondrocytes that can be defined only by *PRG4* expression.[26] In this context, the cells in clusters 2 and 6 are the precursors of those in clusters 1 and 5. Interestingly, the elevated expression of *CCN1, CCN2* and *CCN4* is noted in the cells in clusters 2 and 6. These findings suggest the critical role of CCN family members in the terminal differentiation of articular chondrocytes. CCN family members show diverse and context-dependent functions via the interaction with a variety of cofactors.[27,28,34,35] In particular, all CCN family members are expressed in growth plate cartilage.[36] Among them, *CCN1, CCN2* and *CCN4* play critical roles in chondrogenesis. *CCN1* regulates the early chondrogenesis of limb bud cells, whereas *CCN4* enhances the differentiation of human bone marrow stromal cells into chondrocytes.[29,30] *CCN2* is required for proper endochondral ossification and murine skeletal development.[37] A number of studies have indicated the solid roles of *CCN2* in chondrogenesis and skeletal development.[38] The proliferation and differentiation of auricular chondrocytes are also enhanced by *CCN2*.[39] More importantly, CCN2 promotes the proliferation and maturation of articular chondrocytes to regenerate damaged articular cartilage in vivo.[40,41] A bioengineered CCN2 derivative was also effective in repairing OA cartilage.[42] Consistent with these findings, the present data indicate that transient *CCN2* induction is a critical step in regenerating articular cartilage in an integrated manner.

Our novel strategy may be useful for identifying a master gene of differentiation. If the expression of a particular gene is turned off, accompanied by a decline in marker genes during reprogramming, this may indicate a master transcription factor of the differentiation process. Since this method required maximal transduction efficiency to obtain the cells sufficient for the analysis, the Sendai virus-based vector was used. We should note that the iPSC-like cells we obtained may not be used to generate somatic cells because of persistent expression of transgenes, which prevents cell differentiation. To complete the reprograming and use these cells for cell engineering, the RNA transgenes have to be eliminated by an siRNA as indicated by the manufacturer (Tokiwa-Bio).[15]

Recently, the transient induction of endogenous retroviral RNAs was found to provoke a major transition in the transcriptome, thereby initiating the differentiation of zygotes,[10] thus suggesting that RNAs regulate master transcription factors upstream. In the present study, several lncRNAs were found to be turned off during reprogramming. The function of these RNAs as master regulators of chondrocytic differentiation is currently under investigation. It is also of particular interest whether other master transcription factors of mesenchymal cell differentiation, such as runt-related

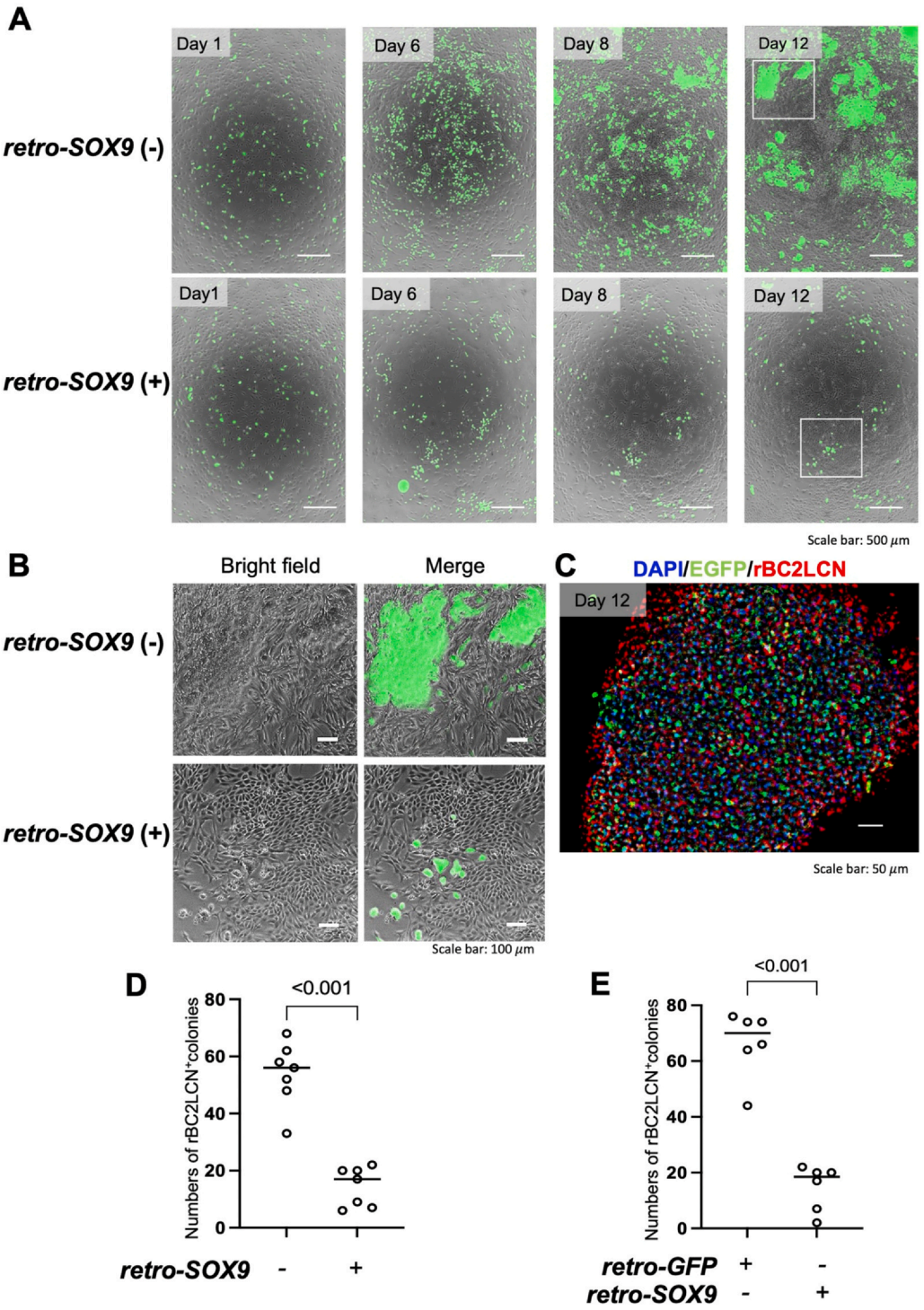


Fig. 8

Osteoarthritis and Cartilage

RNA interference analysis confirming the potential of chondrocytes to be reprogrammed via six transcription factors and the integrity of SOX9 as a master transcription factor of chondrogenesis. (A) Reprogramming of human articular chondrocytes with or without *SOX9* overexpression. Green represents GFP-positive cells with the reprogramming transgenes. Scale bars: 500 μm. (B) The magnified images of the portions indicated in panel A by rectangles are shown. Scale bars: 100 μm. (C) Staining of a representative colony formed on day 12 in the control culture with rBC2LCN and DAPI. GFP signals are also visualized. Scale bar: 50 μm. (D) Quantitative analysis of the colonies observed on Day 12 in 7 independent cell cultures with or without *SOX9* overexpression. (E) The result of the same analysis between *GFP*- and *SOX9*- overexpressed groups with 6 independent cell cultures, respectively. Chondrocytes were obtained from a donor distinct from the one whose cells were used for the single-cell transcriptomic analysis. Probability values between the indicated groups were computed via Student's *t*-test. Labels: *retro-GFP*, *GFP* was transfected by a retroviral vector; *retro-SOX9*: *SOX9* was transfected by a retroviral vector.

transcription factor 2 (*RUNX2*) and myogenic differentiation 1 (*MYOD1*) genes confer iPS-interference in chondrocytes, or not. Such investigation would further support the utility of inverse genetics and iPS-interference methodologies, or might suggest an unknown intersection of differentiation pathways of mesenchymal cells.

Funding sources

This study was supported by JSPS KAKENHI (grant numbers: JP21K19603 and JP23K17439). The role of the sponsor in this study is not specified.

CRediT authorship contribution statement

H.T.D.: Data curation, Investigation, Writing – original draft. **M.O.:** Conceptualization, Data curation, Funding acquisition, Project administration, Resources, Writing – original draft. **Z.W.:** Investigation. **W.K.:** Investigation. **A.T.D.:** Investigation. **T.K.:** Supervision, Writing – review & editing. **T.O.:** Supervision, Writing – review & editing. **S.K.:** Conceptualization, Funding acquisition, Writing – original draft.

Competing interest statement

The authors declare no conflicts of interest.

Acknowledgments

The authors are grateful for the secretarial assistance of Ms. Yoshiko Miyake. We also would like to express our gratitude to our respective families, colleagues and the leaders of Dental School of Haiphong University of Medicine and Pharmacy, Vietnam, who have supported us over the years.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.joca.2024.06.009.

References

1. Balsalobre A, Drouin J. Pioneer factors as master regulators of the epigenome and cell fate. Nat Rev Mol Cell Biol 2022;23:449–64. https://doi.org/10.1038/s41580-022-00464-z
2. Lefebvre V, Angelozzi M, Haseeb A. SOX9 in cartilage development and disease. Curr Opin Cell Biol 2019;61:39–47. https://doi.org/10.1016/j.ceb.2019.07.008
3. Ohba S. Genome-scale actions of master regulators directing skeletal development. Jpn Dent Sci Rev 2021;57:217–23. https://doi.org/10.1016/j.jdsr.2021.10.001
4. Attwood SW, Edel MJ. iPS-cell technology and the problem of genetic instability – can it ever be safe for clinical use? J Clin Med 2019;8:288. https://doi.org/10.3390/jcm8030288
5. Suman S, Domingues A, Ratajczak J, Ratajczak MZ. Potential clinical applications of stem cells in regenerative medicine. Adv Exp Med Biol 2019;1201:1–22. https://doi.org/10.1007/978-3-030-31206-0_1
6. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 2006;126:663–76. https://doi.org/10.1016/j.cell.2006.07.024
7. Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K, *et al*. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 2007;131:861–72. https://doi.org/10.1016/j.cell.2007.11.019
8. Okada M, Oka M, Yoneda Y. Effective culture conditions for the induction of pluripotent stem cells. Biochim Biophys Acta 2010;1800:956–63. https://doi.org/10.1016/j.bbagen.2010.04.004
9. Wang W, Yang J, Liu H, Lu D, Chen X, Zenonos Z, *et al*. Rapid and efficient reprogramming of somatic cells to induced pluripotent stem cells by retinoic acid receptor gamma and liver receptor homolog 1. Proc Natl Acad Sci U S A 2011;108:18283–8. https://doi.org/10.1073/pnas.1100893108
10. Sakashita A, Kitano T, Ishizu H, Guo Y, Masuda H, Ariura M, *et al*. Transcription of MERVL retrotransposons is required for preimplantation embryo development. Nat Genet 2023;55:484–95. https://doi.org/10.1038/s41588-023-01324-y
11. Ohnuki M, Tanabe K, Sutou K, Teramoto I, Sawamura Y, Narita M, *et al*. Dynamic regulation of human endogenous retroviruses mediates factor-induced reprogramming and differentiation potential. Proc Natl Acad Sci U S A 2014;111:12426–31. https://doi.org/10.1073/pnas.1413299111
12. Zhang Y, Li T, Preissl S, Amaral ML, Grinstein JD, Farah EN, *et al*. Transcriptionally active HERV-H retrotransposons demarcate topologically associating domains in human pluripotent stem cells. Nat Genet 2019;51:1380–8. https://doi.org/10.1038/s41588-019-0479-7
13. Bork P. The modular architecture of a new family of growth regulators related to connective tissue growth factor. FEBS Lett 1993;327:125–30. https://doi.org/10.1016/0014-5793(93)80155-n
14. Perbal B. The concept of the CCN protein family revisited: a centralized coordination network. J Cell Commun Signal 2018;12:3–12. https://doi.org/10.1007/s12079-018-0455-5
15. Nishimura K, Sano M, Ohtaka M, Furuta B, Umemura Y, Nakajima Y, *et al*. Development of defective and persistent Sendai virus vector: a unique gene delivery/expression system ideal for cell reprogramming. J Biol Chem 2011;286:4760–71. https://doi.org/10.1074/jbc.M110.183780
16. Bernstein NJ, Fong NL, Lam I, Roy MA, Hendrickson DG, Kelley DR. Solo: doublet identification in single-cell RNA-seq via semi-supervised deep learning. Cell Syst 2020;11:95–101.e5. https://doi.org/10.1016/j.cels.2020.05.010
17. Bergen V, Lange M, Peidli S, Wolf FA, Theis FJ. Generalizing RNA velocity to transient cell states through dynamical modeling. Nat Biotechnol 2020;38:1408–14. https://doi.org/10.1038/s41587-020-0591-3
18. Chen H, Albergante L, Hsu JY, Lareau CA, Lo Bosco G, Guan J, *et al*. Single-cell trajectories reconstruction, exploration and mapping of omics data with STREAM. Nat Commun 2019;10:1903. https://doi.org/10.1038/s41467-019-09670-4
19. Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol 2019;20:296. https://doi.org/10.1186/s13059-019-1874-1
20. Hikichi T, Matoba R, Ikeda T, Watanabe A, Yamamoto T, Yoshitake S, *et al*. Transcription factors interfering with dedifferentiation induce cell type-specific transcriptional profiles. Proc Natl Acad Sci U S A 2013;110:6412–7. https://doi.org/10.1073/pnas.1220200110
21. Kozhemyakina E, Zhang M, Ionescu A, Ayturk UM, Ono N, Kobayashi A, *et al*. Identification of a Prg4-expressing articular cartilage progenitor cell population in mice. Arthritis Rheumatol 2015;67:1261–73. https://doi.org/10.1002/art.39030
22. Ishikawa J, Kaisho T, Tomizawa H, Lee BO, Kobune Y, Inazawa J, *et al*. Molecular cloning and chromosomal mapping of a bone marrow stromal cell surface gene, BST2, that may be involved in

pre-B-cell growth. Genomics 1995;26:527–34. https://doi.org/10.1016/0888-7543(95)80171-h
23. Xu W, Wang W, Liu D, Liao D. Roles of cartilage-resident stem/progenitor cells in cartilage physiology, development, repair and osteoarthritis. Cells 2022;11:2305. https://doi.org/10.3390/cells11152305
24. Haller O, Staeheli P, Schwemmle M, Kochs G. Mx GTPases: dynamin-like antiviral machines of innate immunity. Trends Microbiol 2015;23:154–63. https://doi.org/10.1016/j.tim.2014.12.003
25. Pan C, Hu T, Liu P, Ma D, Cao S, Shang Q, *et al*. BM-MSCs display altered gene expression profiles in B-cell acute lymphoblastic leukemia niches and exert pro-proliferative effects via overexpression of IFI6. J Transl Med 2023;21:593. https://doi.org/10.1186/s12967-023-04464-1
26. Ochi K, Derfoul A, Tuan RS. A predominantly articular cartilage-associated gene, SCRG1, is induced by glucocorticoid and stimulates chondrogenesis in vitro. Osteoarthritis Cartilage 2006;14:30–8. https://doi.org/10.1016/j.joca.2005.07.015
27. Perbal B. CCN proteins: multifunctional signalling regulators. Lancet 2004;363:62–4. https://doi.org/10.1016/S0140-6736(03)15172-0
28. Takigawa M. The CCN proteins: an overview. Methods Mol Biol 2017;1489:1–8. https://doi.org/10.1007/978-1-4939-6430-7_1
29. Wong M, Kireeva ML, Kolesnikova TV, Lau LF. Cyr61, product of a growth factor-inducible immediate-early gene, regulates chondrogenesis in mouse limb bud mesenchymal cells. Dev Biol 1997;192:492–508. https://doi.org/10.1006/dbio.1997.8766
30. Yoshioka Y, Ono M, Maeda A, Kilts TM, Hara ES, Khattab H, *et al*. CCN4/WISP-1 positively regulates chondrogenesis by controlling TGF-β3 function. Bone 2016;83:162–70. https://doi.org/10.1016/j.bone.2015.11.007
31. Tateno H, Toyota M, Saito S, Onuma Y, Ito Y, Hiemori K, *et al*. Glycome diagnosis of human induced pluripotent stem cells using lectin microarray. J Biol Chem 2011;286:20345–53. https://doi.org/10.1074/jbc.M111.231274
32. de Crombrugghe B, Lefebvre V, Behringer RR, Bi W, Murakami S, Huang W. Transcriptional mechanisms of chondrocyte differentiation. Matrix Biol 2000;19:389–94. https://doi.org/10.1016/s0945-053x(00)00094-9
33. Ishikawa T, Nishida T, Ono M, Takarada T, Nguyen HT, Kurihara S, *et al*. Physiological role of urothelial cancer-associated one long noncoding RNA in human skeletogenic cell differentiation. J Cell Physiol 2018;233:4825–40. https://doi.org/10.1002/jcp.26285
34. Rachfal AW, Brigstock DR. Structural and functional properties of CCN proteins. Vitam Horm 2005;70:69–103. https://doi.org/10.1016/S0083-6729(05)70003-0
35. Kubota S, Takigawa M. The CCN family acting throughout the body: recent research developments. Biomol Concepts 2013;4:477–94. https://doi.org/10.1515/bmc-2013-0018
36. Kawaki H, Kubota S, Suzuki A, Lazar N, Yamada T, Matsumura T, *et al*. Cooperative regulation of chondrocyte differentiation by CCN2 and CCN3 shown by a comprehensive analysis of the CCN family proteins in cartilage. J Bone Miner Res 2008;23:1751–64. https://doi.org/10.1359/jbmr.080615
37. Ivkovic S, Yoon BS, Popoff SN, Safadi FF, Libuda DE, Stephenson RC, *et al*. Connective tissue growth factor coordinates chondrogenesis and angiogenesis during skeletal development. Development 2003;130:2779–91. https://doi.org/10.1242/dev.00505
38. Kubota S, Takigawa M. The role of CCN2 in cartilage and bone development. J Cell Commun Signal 2011;5:209–17. https://doi.org/10.1007/s12079-011-0123-5
39. Fujisawa T, Hattori T, Ono M, Uehara J, Kubota S, Kuboki T, *et al*. CCN family 2/connective tissue growth factor (CCN2/CTGF) stimulates proliferation and differentiation of auricular chondrocytes. Osteoarthritis Cartilage 2008;16:787–95. https://doi.org/10.1016/j.joca.2007.11.001
40. Nishida T, Kubota S, Nakanishi T, Kuboki T, Yosimichi G, Kondo S, *et al*. CTGF/Hcs24, a hypertrophic chondrocyte-specific gene product, stimulates proliferation and differentiation, but not hypertrophy of cultured articular chondrocytes. J Cell Physiol 2002;192:55–63. https://doi.org/10.1002/jcp.10113
41. Nishida T, Kubota S, Kojima S, Kuboki T, Nakao K, Kushibiki T, *et al*. Regeneration of defects in articular cartilage in rat knee joints by CCN2 (connective tissue growth factor). J Bone Miner Res 2004;19:1308–19. https://doi.org/10.1359/JBMR.040322
42. Abd El Kader T, Kubota S, Nishida T, Hattori T, Aoyama E, Janune D, *et al*. The regenerative effects of CCN2 independent modules on chondrocytes in vitro and osteoarthritis models in vivo. Bone 2014;59:180–8. https://doi.org/10.1016/j.bone.2013.11.010