RESEARCH ARTICLE

JOR *Spine* OPEN ACCESS

# Developmental morphogens direct human induced pluripotent stem cells toward an annulus fibrosus-like cell phenotype

Ana P. Peredo [1,2,3]    |    Tonia K. Tsinman [1,2,3]    |    Edward D. Bonnevie [2,3]    |    Xi Jiang [2]    |    Harvey E. Smith [2,3]    |    Sarah E. Gullbrand [2,3] iD    |    Nathaniel A. Dyment [1,2]    |    Robert L. Mauck [1,2,3] iD

[1] Department of Bioengineering, University of Pennsylvania, Philadelphia, Pennsylvania, USA

[2] Department of Orthopaedic Surgery, University of Pennsylvania, Philadelphia, Pennsylvania, USA

[3] Corporal Michael J. Crescenz VA Medical Center, Translational Musculoskeletal Research Center, Philadelphia, Pennsylvania, USA

**Correspondence**
Robert L. Mauck, Department of Bioengineering, University of Pennsylvania, Philadelphia, PA 19104, USA.
Email: lemauck@pennmedicine.upenn.edu

## Abstract

**Introduction:** Therapeutic interventions for intervertebral disc herniation remain scarce due to the inability of endogenous annulus fibrosus (AF) cells to respond to injury and drive tissue regeneration. Unlike other orthopedic tissues, such as cartilage, delivery of exogenous cells to the site of annular injury remains underdeveloped, largely due to a lack of an ideal cell source and the invasive nature of cell isolation. Human induced pluripotent stem cells (iPSCs) can be differentiated to specific cell fates using biochemical factors and are, therefore, an invaluable tool for cell therapy approaches. While differentiation protocols have been developed for cartilage and fibrous connective tissues (e.g., tendon), the signals that regulate the induction and differentiation of human iPSCs toward the AF fate remain unknown.

**Methods:** iPSC-derived sclerotome cells were treated with various combinations of developmental signals including transforming growth factor beta 3 (TGF-β3), connective tissue growth factor (CTGF), platelet derived growth factor BB (PDGF-BB), insulin-like growth factor 1 (IGF-1), or the Hedgehog pathway activator, Purmorphamine, and gene expression changes in major AF-associated ECM genes were assessed. The top performing combination treatments were further validated by using three distinct iPSC lines and by assessing the production of upregulated ECM proteins of interest. To conduct a broader analysis of the transcriptomic shifts elicited by each factor combination, and to compare genetic profiles of treated cells to mature human AF cells, a 96.96 Fluidigm gene expression array was applied, and principal component analysis was employed to identify the transcriptional signatures of each cell population and treatment group in comparison to native AF cells.

**Results:** TGF-β3, in combination with PDGF-BB, CTGF, or IGF-1, induced an upregulation of key AF ECM genes in iPSC-derived sclerotome cells. In particular, treatment with a combination of TGF-β3 with PDGF-BB for 14 days significantly increased gene expression of collagen II and aggrecan and increased protein deposition of collagen I and elastin compared to other treatment groups. Assessment of genes uniquely

highly expressed by AF cells or SCL cells, respectively, revealed a shift toward the genetic profile of AF cells with the addition of TGF-β3 and PDGF-BB for 14 days.

**Discussion:** These findings represent an initial approach to guide human induced pluripotent stem cells toward an AF-like fate for cellular delivery strategies.

**KEYWORDS**
culture systems, regenerative medicine, stem cell

## 1 | INTRODUCTION

The intervertebral disc (IVD) is a soft tissue that resides between the bony segments of the spine and is essential for bearing the high loads that arise with activities of daily living. This tissue achieves its function through its highly specialized microarchitecture and biochemical constitution.[1] The disc is a composite tissue that enables motion in six degrees of freedom through its compression-resistant nucleus pulposus (NP) core and its surrounding elastic, high tensile strength annulus fibrosus (AF). Given its demanding loadbearing role, the IVD is often injured, with IVD herniations affecting 2%–3% of the world population.[2,3] These insults accumulate over time due to the tissue's inability to self-repair, leading to severe pathologies such as disc degeneration and herniations later in life.

Disc herniations arise when AF bulging impinges on the adjacent neural elements or when AF lesions become large enough for migration of NP tissue beyond the anatomical boundaries of the disc. This can result in the compression of adjacent nerves, resulting in debilitating pain and numbness in symptomatic cases.[4] Current clinical strategies for the treatment of symptomatic disc herniations involves resection of the herniated tissue to relieve nerve impingement. However, in such procedures, the AF tear is not repaired, leaving behind a compromised tissue structure that will likely fail to heal and that is predisposed to progressive and repeated herniation at that site. Indeed, the incidence of re-herniation can reach 25% in some instances.[5] Additionally, individuals with lumbar disc herniations, regardless of surgical or nonoperative treatment of the herniation, are more likely to experience long-term low back pain compared to the general population.[6]

Strategies to repair the AF, including the delivery of growth factors, drugs, and biomaterials, have been explored but have shown limited benefits due to rapid drug clearance, limited molecule diffusion, and a lack of endogenous cell engagement and recruitment.[7–9] The AF is a dense fibrocartilaginous tissue that is formed early during embryonic development by highly anabolic cells that secrete and organize specialized matrix into a functional structure.[10,11] The inner region of the AF accumulates glycosaminoglycans and type II collagen while the outer region is rich in type I collagen and elastin. Once the tissue matures, resident AF cells decrease protein production and only a few cells are left inhabiting the fully formed dense matrix.[12,13] These cells have limited biosynthetic abilities, and are further diminished with injury-induced apoptosis.[14] Therefore, cell therapies that deliver specialized exogenous cells could enhance the restoration of the AF post-herniation.

Cellular delivery strategies for the AF remain underdeveloped due to a lack of AF-associated markers and an optimal cell type for tissue repair. Prior work has demonstrated the potential of AF cell-based approaches to repair the injured or degenerative AF in small animal models,[15,16] however, the translational value of this approach is limited due to the difficulties in obtaining healthy autologous AF cells in human patients. While mesenchymal stromal cells and adipose-derived stem cells have been used in previous attempts, the ability of these cells to take on an AF-associated phenotype is not yet proven. Furthermore, some reports of in vivo delivery of these cells resulted in the formation of osteophytes,[17] highlighting the importance of cell priming and specialization before in vivo delivery.

Human induced pluripotent stem cells (iPSCs) have gained widespread use for cellular therapies due to their infinite cell-renewal ability and their potential to differentiate into all three embryonic germ layers through sequential developmentally guided specification. To date, the differentiation of human iPSCs toward the AF fate has not been attempted. While differentiation approaches for NP cells have been reported,[18–20] these have limited application for the differentiation of AF cells due to differences in embryonic lineage (notochord vs. paraxial mesoderm, respectively).[11] Indeed, marked differences in transcriptomic signatures have been established between the healthy adult NP and AF cells populations.[21] AF cells are specified from mesenchymal progenitors that, also give rise to cartilage, ligament, and tendon during embryonic development.[22–24] Although the iPSC differentiation strategies for cartilage and tendon can be used as guiding principles, the molecular cues that drive human AF specification have not been defined.

During embryonic development, the paraxial mesoderm undergoes segmentation into somites that are dorsoventrally patterned into the dorsal somite (dermomyotome) and the ventral somite (sclerotome).[11,25] The sclerotome gives rise to the axial skeleton, including bone, cartilage, and fibrocartilaginous tissues. Sclerotome-derived mesenchymal progenitors undergo condensation to form the AF around the NP that is concomitantly forming through notochord segmentation and expansion.[11] A number of molecular factors have been implicated in AF formation. This includes transforming growth factor β (TGF-β), which is indispensable for the formation, maintenance, and growth of the IVD.[26,27] Insulin-like growth factors (IGFs) have also been implicated in skeletal growth and development.[28] Although less is known about their precise role during IVD development, they do stimulate proliferation and extracellular matrix (ECM) synthesis in IVD cells.[29,30] When combined with platelet-derived growth factor (PDGF), IGF reduces apoptosis in human AF cells[29] and,

in concert with TGF-β, increases the production of collagen and other ECM proteins.[31] Moreover, the combination of PDGF and TGF-β directs sclerotome cells toward a fibroblast fate resembling the AF morphology.[32] Another factor, connective tissue growth factor (CTGF), is involved in several cellular functions and is important in skeletal development. During early embryogenesis, CTGF is highly expressed in the somites and notochord, and remains expressed in mature tissues that arise from these embryonic structures.[33] Similarly, sonic hedgehog (SHH) has been implicated in regulating notochord patterning and IVD development.[34] Although these signaling factors have been implicated in IVD development, the effect of these factors (individually or in combination) for the differentiation of AF cells remains unknown.

In this study, we explored the changes induced by developmental signals, independently or in concert, in the differentiation of human iPSCs toward AF cell-like cells. Due to the lack of molecular markers unique to the AF cell population, AF "fate" was assessed by the expression and production of proteins characteristic of AF tissue. Developmental signals were screened by treating iPSC-derived sclerotome cells and assessing gene expression changes at different time points of induction. TGF-β3, in combination with PDGF-BB, CTGF, or IGF-1, induced an upregulation of key AF ECM genes. The synergistic effects observed were validated by using three distinct iPSC lines and by assessing the production of upregulated ECM proteins of interest. Finally, to conduct a broader analysis of the transcriptomic shifts elicited by each factor combination, and to compare genetic profiles of treated cells to mature human AF cells, a 96.96 Fluidigm gene expression array was applied. Genes spanning ECM interactions and proteins, differentiation-specific markers, tissue-specific markers, and genes involved in several cellular processes, were carefully selected from literature available at the time. Principal component analysis was employed to identify the transcriptional signatures of each cell population and treatment group in comparison to native AF cells. Together, these studies represent the first step toward developing a differentiation method for iPSC-derived AF cells for future use in cell-based repair of the AF following disc herniation.

# 2 | METHODS

## 2.1 | iPSC maintenance and expansion

The human iPSC line AICS-0061-036 was purchased from the Coriell Institute for Medical Research. The CHOP WT10.3 and CHOP WT4.2 human iPSC cell lines were donated by the Human Pluripotent Stem Cell Core at the Children's Hospital of Philadelphia (CHOP). For expansion and maintenance, human iPSCs were cultured in tissue culture plates coated with hESC qualified Matrigel (BD Biosciences, 352277) with feeder-free mTeSR™1 maintenance medium (Stemcell Technologies, 85850). The enzyme-free Gentle Cell Dissociation Reagent (Stemcell Technologies, 100-0485) was used to dissociate iPSCs into cell aggregates for routine passaging.

## 2.2 | iPSC differentiation

### 2.2.1 | Differentiation of human iPSCs toward sclerotome

Human iPSCs were passaged as cell clusters and seeded on tissue culture plates coated with Growth Factor Reduced Matrigel Matrix (Corning, 354 230). At ∼70% confluence, cells were fed with mTESR™1 approximately 3 h before the start of the differentiation protocol. Washing Medium, composed of 50% IMDM, GlutaMAX™ (ThermoFisher, 31980030) and 50% Ham's F12 Nutrient Mix (Corning, MT10-080-CV) volume/volume (v/v), was used to wash the cells two times before the start of differentiation. Cells were washed with Washing Medium before each differentiation step. iPSCs were maintained and expanded in hypoxic conditions at 5% $O_2$, 5% $CO_2$, and 90% $N_2$. Initial stages of differentiation were also conducted in hypoxic conditions up to the end of the sclerotome stage, after which cells were kept in normoxic conditions at ∼21% $O_2$, 5% $CO_2$, and 74% $N_2$.

Human iPSCs were induced to differentiate toward the sclerotome fate following a previously established protocol, with slight modifications.[35] The base medium used did not include penicillin/streptomycin and polyvinyl alcohol as reported by Adkar and colleagues. Briefly, iPSCs were guided through a stepwise differentiation approach wherein cells were directed toward the following fates in the respective order: anterior primitive streak (APS) for 24 h, paraxial mesoderm (PM) for 24 h, early somite (ES) for 24 h, and sclerotome (SCL) for 72 h (Figure S1). Mesoderm differentiation medium, composed of washing medium supplemented with 1% v/v chemically defined lipid concentrate (ThermoFisher, 11905031), 1% v/v Corning ITS™ Premix (Fisher, CB-40350), and 450 μM of 1-thioglycerol (Sigma, M6145), was used for all differentiation steps, with media changes every day. For each differentiation step, mesoderm differentiation medium supplemented with stage-specific factors, was added to cell cultures as depicted in Figure S1 (Table S1). At the end of each differentiation stage, cell samples ($n = 4$/cell line) were collected and frozen for gene expression analysis. Experiments were repeated for each cell line ($n = 3$ lines).

### 2.2.2 | Factor screening for induction of annulus fibrosus fate

After the 72 h of sclerotome induction, cells were washed twice with washing medium. Control (Ctrl) wells were fed mesoderm differentiation medium. To screen factors for their ability to promote the AF fate, transforming growth factor beta 3 (TGF-β3) (+T, 10 ng/mL), connective tissue growth factor (CTGF) (+C, 100 ng/mL), platelet derived growth factor BB (PDGF-BB) (+P, 2 ng/mL), insulin-like growth factor 1 (IGF-1) (+I, 5 ng/mL), or the Hedgehog pathway activator, purmorphamine (+Pu, 2 μM) were added independently or in combination (Figure 2A) (Table S1) for 7 or 14 days, with media changed every 3 days. These factor concentrations were derived from

prior literature.[31,32,35–37] Induction of the AF fate after sclerotome induction was performed using cell culture incubator conditions of 5% $CO_2$ and atmospheric $O_2$ levels. At the timepoint of choice, samples were either fixed for immunofluorescence staining or collected, pelleted, and frozen at $-80°C$ for gene expression analysis. Large-scale screening was performed with one iPSC cell line (CHOP WT10.3) and was repeated 3 times (Figure 2). The effects were confirmed using three iPSC cell lines (AICS-0061-036, CHOP WT10.3, and CHOP WT4.2) for the factors identified as having an inductive effect (Figure 3).

## 2.3 | Gene expression analysis via quantitative RT-PCR

RNA isolation of digested samples was performed using the Directzol RNA Miniprep kit with DNAse-I treatment to remove trace DNA before RNA elution (Zymo Research, R2050). RNA was quantified via Nanodrop spectrophotometry. cDNA was synthesized using the SuperScript™ IV VILO Master Mix (Invitrogen, 11756050) according to the manufacturer's protocol. Human genomic DNA, extracted from human cells using the GeneJET Genomic DNA Purification kit (ThermoFisher, K0721), was used to generate a standard curve (3000000 copies). Absolute quantitative RT-PCR was run using Fast SYBR™ Green Master Mix (ThermoFisher, 4385618) for 40 cycles with validated custom designed primers (Table S2), using a genomic DNA standard curve. Changes in gene expression, reported as relative expression, were quantified by normalizing copy numbers for each gene of interest in each sample to the copy numbers of the TBP housekeeping gene in the respective sample (Table S2).

## 2.4 | Immunofluorescence staining and imaging

### 2.4.1 | Staining

For immunofluorescence staining, iPSCs were seeded on μ-Slide 8 well polymer coverslip chamber slides (Ibidi, #80826). Differentiation steps were followed as described, including coating wells with growth factor reduced Matrigel before cell seeding. At the defined time points, cells were washed $3\times$ with PBS, after which they were fixed with 4% paraformaldehyde for 20 min at room temperature (RT). After fixation, the cells were washed $3\times$ with PBS. For intracellular staining, cells were permeabilized with 0.1% v/v Triton X-100 solution in PBS for 10 min at 4°C, after which they were rinsed $3\times$ with PBS. For both intracellular and extracellular staining, blocking was performed by adding blocking solution composed of 3% weight/volume bovine serum albumin in PBS for 1 h at RT. After blocking, cells were washed with PBS $3\times$, and primary antibodies were added overnight at 4°C (Table S3). Cells were then washed with PBS $3\times$ and secondary antibodies, phalloidin, and/or DAPI were added for 1 h. at RT, after which cells were washed $3\times$ with PBS and imaged.

### 2.4.2 | Confocal imaging

Confocal imaging was performed using a Nikon A1R+ confocal microscope. For SRY-box 9 (SOX9) intranuclear staining, cells counterstained with DAPI, and for cells co-stained for F-actin, stage-specific embryonic antigen-4 (SSEA-4), and E-cadherin, images were obtained at the mid-plane using $60\times$ magnification. For visualization of type I collagen and elastin, and for imaging of cells co-stained for F-actin, images were obtained at $20\times$ magnification.

### 2.4.3 | Matrix staining intensity measurements

Type I collagen and elastin deposition were evaluated through staining intensity measurements for $n = 4$ wells for the Ctrl, +TP, +TC, and +TI conditions after 14 days of induction post-sclerotome differentiation. Wells were imaged and divided intro quartiles, from which mean gray value was obtained using ImageJ. Mean gray value for each well was reported (Figure S2).

### 2.4.4 | Nuclear morphology and Sox9 intranuclear intensity measurements

Nuclear aspect ratio and levels of SOX9 intranuclear fluorescence intensity were assessed using CellProfiler, ver 3.1.9. Nuclei were viewed using DAPI and were segmented based on a minimum cross entropy algorithm with nuclei excluded below 20 and above 60 pixels. Nuclei that were touching the image boundary were also excluded. Major and minor axes were measured, and their ratio (long to short axis) was calculated and is presented as the nuclear aspect ratio ($n$ >1000 cells/sample). SOX9 average fluorescent intensity was measured on a per-nucleus basis using nuclei identified by minimum cross entropy algorithm ($n$ >1000 cells/sample).

## 2.5 | Fluidigm 96.96 gene expression assay

### 2.5.1 | Sample sourcing and preparation

One sample for each of the three iPSC lines at SCL stage and day 14 of culture for the +TP, +TC, +TI, and +TPCI groups (TP14, TC14, TI14, and TPCI14, respectively) were collected, pelleted, and frozen. Human AF cells obtained from human cadaveric donors (within 48 h of death) with no medical history of spinal degeneration or disease were purchased from Articular Engineering. AF cell lines AF26 (30-year-old donor), AF1356 (26-year-old donor), and AF1363 (59-year-old donor) were used at passage 1. AF cells were received at passage 0 and expanded to passage 1 in basal media composed of Dulbecco's Modified Eagle Medium (11965-084, Gibco) supplemented with 10% v/v fetal bovine serum (R&D Systems, S11150) and 1% v/v Penicillin/Streptomycin/Fungizone (Gibco, 15240-062).

## 2.5.2 | RNA isolation, cDNA preparation, and preamplification

RNA was isolated from the 45 samples (see above) as previously described with DNase-I treatment before RNA elution (see above). cDNA was synthesized using the SuperScript™ IV VILO Master Mix with ezDNase™ Enzyme (Invitrogen, 11766050), with a second DNase digestion to remove gDNA according to manufacturer's protocol. No reverse transcriptase controls were generated from pooled iPSC, SCL or AF cells by not including reverse transcriptase for cDNA synthesis. Generated cDNA for the 48 samples was pre-amplified for 14 cycles using the specific 96 gene targets. Gene targets were chosen to encompass a wide array of cellular phenotype, signaling, and activity. Tissue-specific markers including ECM proteins for cartilage, AF, and NP tissue as well as stage-specific markers for differentiation toward from a pluripotent state toward sclerotome were included. Importantly, genes associated with downstream signaling for the induction factors were also assessed, in addition to genes spanning important cell–cell and cell-ECM processes. Adipogenic, osteogenic, and mesenchymal markers were also included for analysis. For cDNA pre-amplification, a primer pool containing the 20X TaqMan primers for the gene targets was used (Table S4), consisting of the primers diluted 1:100 (0.2X) in DNA suspension buffer (10 mM Tris, pH 8.0, 0.1 mM EDTA). Two microliter of each cDNA sample was combined with 1.25 μL of pooled diluted TaqMan probes, 1 μL of Fluidigm Preamp Master Mix (Fluidigm, 100–5580), and 0.75 μL of water for a combined 5 μL reaction volume. The resultant pre-amplified cDNA was diluted 1:5 in the DNA suspension buffer.

## 2.5.3 | Fluidigm gene expression chip and data postprocessing

qPCR of samples ($n = 3$ cell lines or donors/group) was performed on the Fluidigm Biomark HD at the Penn Genomic Analysis Core. A Fluidigm Dynamic Array IFC (BMK-M-96.96) was loaded with pre-amplified cDNA for the 48 samples in duplicate and 20X TaqMan probes for the 96 genes, including 3 housekeeping genes (Table S4). The probe for TBXT was custom made using the Custom Assay Design tool by ThermoFisher. A brightfield image of the loaded Fluidigm chip was assessed to detect bubbles in the reaction wells. If a bubble was detected, the reaction well was excluded from analysis. Housekeeping genes (TBP, 18S, and RPS17) were chosen based on prior experimentation and validation. To account for variations in total RNA between samples, $\Delta C_T$ for a sample was calculated by subtracting the duplicate $C_T$ average for a gene of interest from the housekeeping gene average (TBP, 18S, and RPS17) for that sample.

## 2.5.4 | Principal component analysis (PCA)

PCA and hierarchical clustering were performed using ClustVis (open-source web tool, beta version).[38] $\Delta C_T$ for all genes except housekeeping genes were entered into ClustVis. For PCA of SCL and AF samples to identify signature genes for each cell type, only data from these groups were used in the PCA. Plots were generated using custom R scripts with the "ggplot2," "rgl," and "Rcmdr" packages (R, Posit, version 4.1.2 (2021-11-01)).

## 2.6 | Illustrations

Illustrations were generated using Inkscape (open-source, version 1.0.1 (c497b03c, 2020-09-10)) or BioRender.com (BioRender).

## 2.7 | Statistical analysis

Statistical analyses were conducted using GraphPad PRISM (version 10, GraphPad Software). Data are reported as mean ± standard deviation (SD). The Shapiro–Wilk normality test was used to determine the need for nonparametric testing (alpha = 0.05). For gene expression comparisons at the stages of differentiation leading up to SCL, gene expression was compared via a one-way ANOVA with Tukey's multiple comparisons post hoc analysis. For gene expression, nuclear aspect ratio, and SOX9 nuclear intensity comparisons between iPSC and SCL groups, an unpaired t-test with Welch's correction was used. Gene expression between groups treated with factors for AF cell-like differentiation of SCL cells was compared to Ctrl groups using a one-way ANOVA with Tukey's multiple comparisons post hoc analysis. Matrix staining intensity mean gray values for collagen I and elastin were compared using a one-way ANOVA with Tukey's multiple comparisons post hoc analysis. For comparison of principal component scores and gene expression between SCL and AF groups, an unpaired t-test with Welch's correction was used. To compare principal component scores and gene expression between SCL, AF, and other treatment groups, a one-way ANOVA was used followed by Tukey's multiple comparisons post hoc analysis.

## 3 | RESULTS

### 3.1 | Differentiation of human iPSCs toward sclerotome

To begin this process, we first differentiated human iPSCs toward a sclerotome fate, using a protocol slightly modified from Adkar et al. Successful sclerotome phenotype was verified in three distinct cell lines (AICS-0061-036, CHOP WT10.2, and CHOP WT4.3) (Figure S1)[35] via gene expression analysis for key stage-specific transcription factors and markers via quantitative RT-PCR (Figure S1). As expected, over the course of differentiation, there was a gradual decrease in the expression of the pluripotency marker POU domain, class 5, transcription factor 1 (POU5F1), also known as Oct-4 (Figure 1A). After 24 h of APS induction, the expression of one APS marker mix paired-like homeobox (MIXL1) gene surged and rapidly
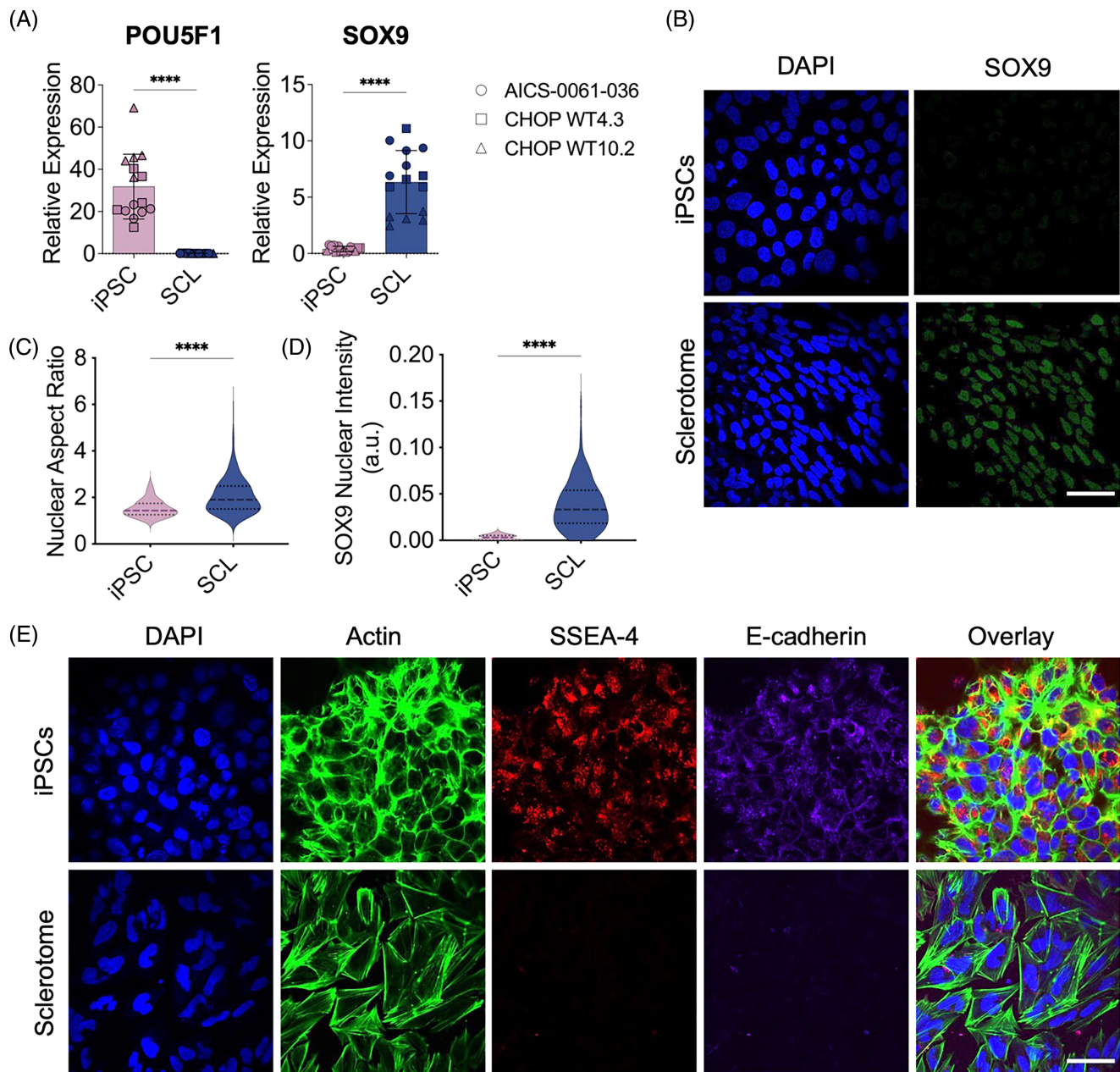
**FIGURE 1** Directed differentiation of human induced pluripotent stem cells (iPSCs) toward the sclerotome (SCL) fate. (A) POU5F1 and SOX9 gene expression at the iPSC and SCL stages of differentiation for three iPSC cell lines normalized to housekeeping gene expression ($n = 5$/cell line; mean ± SD). (B) Representative immunofluorescence imaging for Sox9 intranuclear staining at the iPSC and SCL stages of differentiation (scale: 50 μm). (C) Cell nucleus aspect ratio and (D) Sox9 nuclear intensity measured from immunofluorescence images for iPSC and SCL cells ($n$ >1000 cells). (E) Representative immunofluorescence imaging for cell nuclear morphology (DAPI), actin filaments (phalloidin), Ssea-4 pluripotency marker, and E-cadherin for iPSC and SCL stages of differentiation (scale: 50 μm). Statistical significance denoted by *$p$-value <0.05, **$p$ < 0.01, ***$p$ < 0.001, ****$p$ < 0.0001.

decreased with the start of paraxial mesoderm (PM) induction. The t-box transcription factor T (TBXT) gene, associated with primitive streak and mesoderm differentiation, was upregulated during the APS and PM stages. Platelet derived growth factor receptor α (PDGFRα), an early mesoderm marker (mainly for paraxial mesoderm), showed a sustained gradual increase from PM to SCL stages. Finally, the sclerotome marker gene SRY-box 9 (SOX9) showed peak expression at the SCL stage after 72 h of sclerotome induction (Figure 1A, Figure S1).

To further validate, the generation of sclerotome cells from human iPSCs, the presence of intranuclear Sox9 was verified via immunofluorescence staining. After 72 h of sclerotome induction, there was an increased intensity of nuclear Sox9 compared to uninduced iPSCs (Figure 1B,D). Sclerotome cells also developed an elongated nuclear morphology with greater nuclear aspect ratios compared to iPSCs, characteristic of cellular differentiation (Figure 1C). These changes in nuclear morphology were accompanied

by a shift from cortical F-actin to the development of robust F-actin stress fibers (Figure 1E). Furthermore, the sclerotome cells lost expression of the pluripotency marker SSEA-4 and showed a reduction in cell–cell contacts as evidenced by a decrease in E-cadherin staining, corroborating the shift from a pluripotent to a differentiated state (Figure 1E).

## 3.2 | Candidate screening reveals synergistic role of factors in combination with TGF-β to drive annulus fibrosus-like fate

After validating the iPSC-derived sclerotome cells, factors present during the embryonic development of the IVD and AF were investigated with respect to their ability to promote induction of an AF cell-like fate. TGF-β3, CTGF, PDGF-BB, IGF-1, or the Hedgehog pathway activator, Purmorphamine (Pu) were added independently or in combination with TGF-β3 (due to its instrumental role in IVD

morphogenesis) (Figure 2A). The expression of genes for major AF extracellular matrix (ECM) proteins including type I collagen (COL1A1), type II collagen (COL2A1), elastin (ELN), and aggrecan (ACAN) was compared among groups after 7 or 14 days of treatment with different combinations. All genes assessed showed peak expression after 14 days of treatment compared to 7 days (Figure 2B,C). After 7 days of induction, the group treated with TGF-β3 + PDGF-BB (+TP) showed significantly higher expression of COL1A1, ELN, and ACAN compared to the untreated control cells (Ctrl) and other treatment groups. By 14 days of exposure to factors, the dual-factor combination groups that included TGF-β3 (i.e., +TP, +TC, and +TI), all showed significantly greater expression of COL1A1, ELN, and ACAN compared to the Ctrl group, and trended toward greater expression of COL2A1. Interestingly, combining TGF-β3 with Purmorphamine (+TPu) did not show this effect. Furthermore, the addition of the Hedgehog pathway activator in the triple factor combinations (+TPPu, +TCPu, and +TIPu) decreased the expression of these ECM genes by 14 days compared to the double factor combinations that
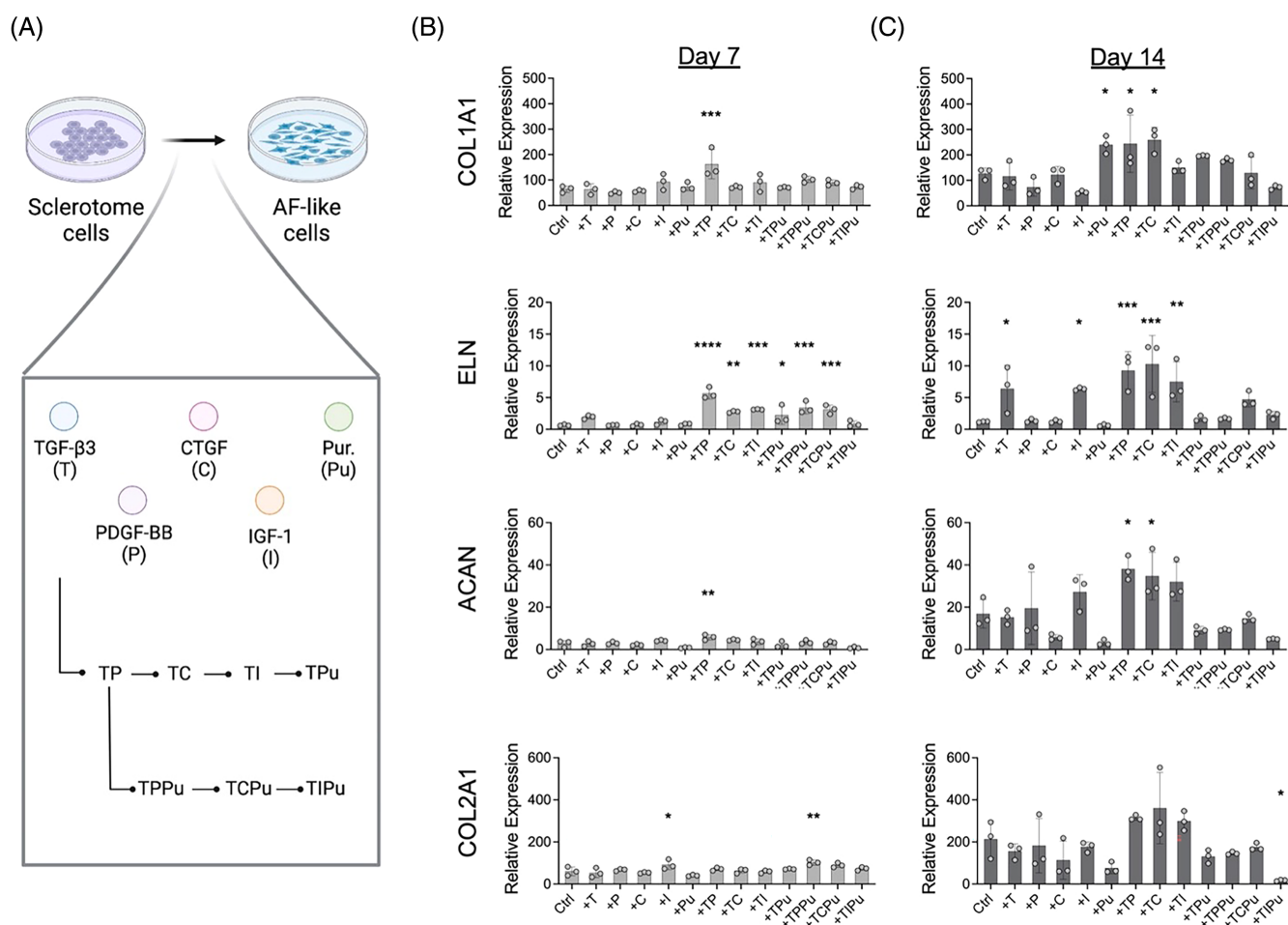


**FIGURE 2** TGF-β in combination with PDGF-BB, CTG, or IGF-1 drives the upregulation of major AF ECM genes. (A) Schematic of the factors tested to drive sclerotome cells toward the annulus fibrosus fate. (B) Gene expression for major annulus fibrosus extracellular matrix protein genes, COL1A1, ELN, ACAN, and COL2A1, after 7 or 14 days of exposure to differentiation factors, normalized to housekeeping gene expression (n = 3; mean ± SD) using cell line CHOP WT10.3. Statistical significance compared to Ctrl denoted by *p-value <0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

did not include Purmorphamine (+TP, +TC, and +TI) (Figure 2C). Based on these findings, the Hedgehog pathway activation via the addition of Purmorphamine was not pursued further.

To verify that the double factor treatments led to comparable effects across iPSC cell lines, +TP, +TC, and +TI treatments were assessed using three distinct lines and compared to Ctrl. An additional group in which these factors were combined (+TPCI) was added to

this analysis. After 14 days of treatment, expression of COL1A1, COL2A1, ELN, and ACAN was greater for all treatment groups compared to Ctrl, across all cell lines (Figure 3A). The +TP group resulted in significantly higher expression of COL2A1 and ACAN, while ELN was significantly higher than Ctrl with +TC or +TI treatment after 14 days (Figure 3A). The combination of TGF-β3, PDGF-BB, CTGF, and IGF-1 did not have a synergistic effect on the expression of any
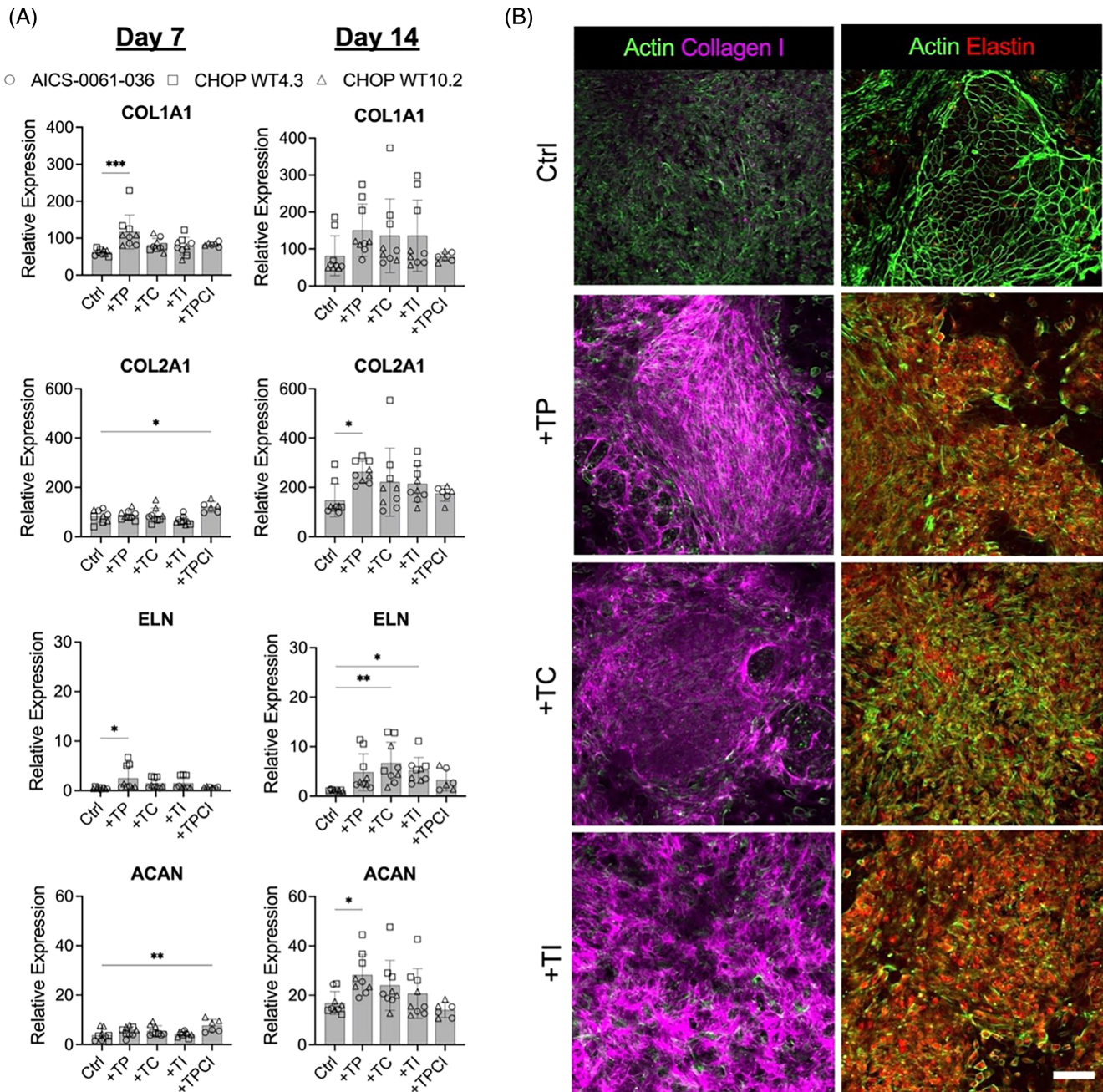


**FIGURE 3** Upregulated expression and production of major annulus fibrosus extracellular matrix proteins with top performing factor combinations. (A) Gene expression for major annulus fibrosus extracellular matrix protein genes COL1A1, ELN, COL2A1, and ACAN after 7 or 14 days of exposure to top performing differentiation factor combinations, normalized to housekeeping gene expression for three different iPSC cell lines ($n = 3$/cell line; mean ± SD). (B) Immunofluorescence staining for type I collagen and elastin, co-stained for F-Actin (phalloidin) after 14 days of exposure to factor combinations (scale: 100 μm). Statistical significance compared to Control (Ctrl) denoted by *$p$-value <0.05, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.

these key AF ECM genes. To assess whether changes in gene expression translated to the deposition of matrix, immunofluorescence staining for type I collagen and elastin – the major structural proteins of the outer AF—was performed after 14 days of treatment (Figure 3B). Compared to Ctrl, induced cells showed a greater distribution and density of collagen and elastin deposition, providing evidence of AF-like matrix elaboration (Figure 3B, Figure S2).

## 3.3 | Fluidigm gene expression array for the comparison of factor treatment effects

To conduct a broader analysis of the gene expression differences between human AF cells and iPSC derived SCL cells that were induced toward AF differentiation, a Fluidigm 96.96 gene expression array was carried out. This enabled simultaneous comparison of adult AF, SCL, and +TP, +TC, +TI, +TPCI groups after 14 days of treatment (Figure S3). Ninety-three genes, spanning collagens, proteoglycans and glycoproteins, ECM-remodeling factors, differentiation specific markers, cell–cell and cell-ECM interaction proteins, and mechanotransduction mediators, were included in the qPCR array (Table S4). TBP, 18S, and RPS17 were used as housekeeping genes.

From these data, we first carried out principal component analysis (PCA) of $\Delta C_t$ values for human adult AF versus SCL cells to identify the genetic signatures of each cell population (Figure 4A, Figure S4). PC1 captured 74.5% of the total variance while PC2 captured 9.8%, together representing 84.3% of the variance between the groups (Figure 4B). PC1 scores were significantly different between groups, with AF scores being higher than SCL scores, while no differences in PC2 scores were observed (Figure 4C). Several genes involved in the establishment and maintenance of AF ECM (TIMP1, FN1, COL6A1, SPARC, DCN) were among the top 8 genes with highest PC1 loading values (Figure 4D,E). Other top scoring PC1 loading genes included THY1, ACAN, LUM, CA12, TGFBR2, COMP, ITGA5, COL12A1, GDF5, PAX1, PXN, LGALS3, MYH9, COL1A1, ROCK2, IGFBP7, PIEZO1, COL1A2, CTGF, TGFB1, and ITGB5 (Table S5). Meanwhile, several of the lowest PC1 loading value genes (TBXT, TBX6, MSGN1) represented established mesoderm markers, characteristic of SCL cells (Figure 4D,E). This analysis thus provided a clear set of markers to define genetic profiles for each cell population.

Next, to assesses the genetic differences between different treatment groups, PCA was conducted using $\Delta C_t$ values for AF, SCL, and SCL cells treated for 14 days with factor combinations (TP14, TC14, TI14, and TPCI14) across the 93 genes analyzed with the Fluidigm gene expression array (Figure S5). A heatmap of the processed data revealed first-level clustering of the day 14 treatment groups with SCL cells, separated from AF cells. This indicates that, after 14 days of induction, treated cells continued to show greater gene expression similarities to SCL cells compared to adult human AF cells. Second-level clustering revealed differences between SCL cells and all day 14 treatment groups.

Interestingly, the iPSC cell line AICS-0061-036 (AICS) showed cluster separation from the CHOP WT4.2 (WT4.2) and CHOP

WT10.3 (WT10.3) cell lines. Further assessment of the global gene expression patterns showed marked differences across several genes between AICS and the other two cell lines. Specifically, all AICS cells, regardless of treatment type, showed lower expression of key genes that were highly upregulated by adult AF cells compared to WT4.2 and WT10.3-treated cells. However, this marked difference in gene expression was not apparent at the SCL stage, indicating the cell line-dependent differences arose in later stages of differentiation with extended factor treatment (Figure S5).

From the overall analysis, PC1 and PC2 scores captured 62.1% of the variance between the groups assessed. Plotting of PC1 and PC2 scores for each sample showed overlap between all day 14 treatment groups (TP14, TC14, TI14, and TPCI14), with all populations lying between SCL and adult AF cells (Figure S6). Further inspection of the PC1 and PC2 score differences among treatment groups revealed no significant differences after 14 days of induction with the different factor combinations (Figure S6), suggesting that the differences between treatment groups did not induce drastic differences in gene expression. Assessment of the eight highest and eight lowest scoring genes identified from the PCA of adult AF and SCL cells showed increased expression of genes characteristic of AF cells with 14 days of induction and a reduction in expression of genes highly expressed at the SCL stage across all treatments (Figure S6). However, the AICS cell line showed lower expression of AF-associated genes compared to WT4.2 and WT10.3, regardless of factor combination.

To further explore the effects of treatment, we compared genetic profiles of the TP14 group to those of SCL and adult AF cells. TP14 was chosen due to the significantly increased expression of COL2A1 and ACAN compared to other treatment groups, in addition to the physical manifestation of COL1A1 and ELN in IF staining (Figure 3A). PCA of AF, SCL, and TP14 is shown in Figure 5A. PC1 score comparisons, which captured 52.8% of the variance, revealed significant differences between all groups (Figure 5B). PC1 scores gradually increased as a function of differentiation progression, increasing from SCL to the adult AF stage. Meanwhile, the PC2 score, which captured 17.7% of the variance, only revealed differences between the SCL and TP14 group. TP14 and AF cells did not have significantly different PC2 scores (Figure 5C).

Assessment of the eight genes uniquely highly expressed by AF cells or SCL cells (therefore contributing most to the PC1 scores), respectively, revealed a shift toward the genetic profile of AF cells with the addition of TGF-β3 and PDGF-BB for 14 days (TP14) (Figure 5D,E). Specifically, addition of these factors increased the expression of FN1, MKX, a key transcription factor involved in AF development, and SPARC to levels expressed by mature adult AF cells (Figure 5E). Other genes highly expressed by mature AF cells such as TIMP1, NT5E, ENG, COL6A1, and DCN also showed increases in expression with treatment. Concomitantly, genes highly expressed by SCL cells were downregulated with treatment, with TBX6 and TBXT reaching expression levels apparent in mature AF cells (Figure S7). Other genes highly expressed by SCL, such as MSGN1, SP7, KDR, and KRT8, were reduced to a lesser extent by treatment.

Expression of ECM-related genes such as COL1A1, COL12A1, ACAN, PRG4, LUM, COMP, and FMOD was dramatically upregulated
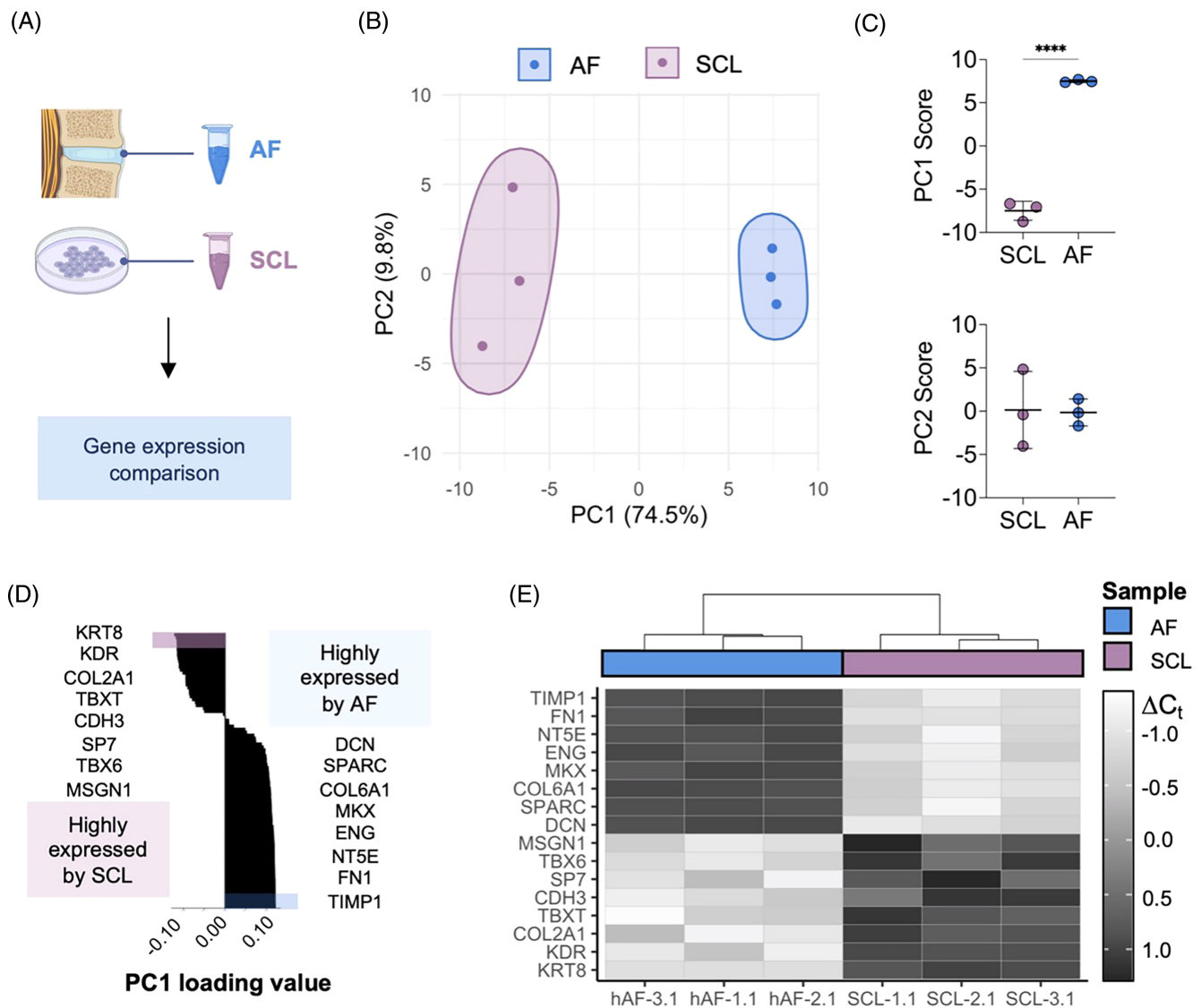
**FIGURE 4** Principal component analysis of sclerotome and AF cells reveal differences in Fluidigm gene expression array profiles. (A) The gene expression patterns of AF cells from three human donors and sclerotome (SCL) cells differentiated from three human iPSC cell lines were compared. (B) Principal component analysis (PCA) of the data represented by mapping PC1 and PC2 scores for SCL and AF groups. Percent of described total variance by each PC is listed in parentheses. Concentration ellipses for each group are demarcated via clustering. (C) (Top) PC1 scores and (Bottom) PC2 scores for SCL and AF groups ($n = 3$/group; mean ± SD). (D) Ranked PC1 loading values for genes included in the Fluidigm gene expression array, listing the eight highest and eight lowest valued genes. For a full list of ranked PC1 loading values, see Table S5. (E) Heatmap of $\Delta C_t$ values for genes listed in (D) with hierarchical clustering based on the 93 genes included in the gene expression array with color coding by sample type. For the heatmap showing all 93 genes used in the cluster analysis, see Figure S5. Statistical significance denoted by *$p$-value <0.05, **$p$ < 0.01, ***$p$ < 0.001, ****$p$ < 0.0001.

with treatment, in some cases reaching levels comparable to mature AF cells (Figure 6A). Other genes involved in the TGFB signaling pathway such as TGFB3, and IGF-related genes (IGF-1, IGFBP5, and IGFBP7) showed expression changes toward levels measured in adult AF cells with treatment (Figure 6B). The increase in expression for AF ECM proteins in treated cells was accompanied by a reduction in the expression of genes involved in early embryonic development, including PODXL, FOXC2, and proteins involved in cell-to-cell adhesion, namely CDH4, and CDH5 (Figure 6C). Other genes involved in diverse cellular processes, such as PIEZO2 and GDF5, showed expression

shifts toward AF levels with treatment (Figure 6D). Interestingly, expression of the Scleraxis transcription factor (Scx) gene, SCX, required for AF and tendon development, was markedly upregulated with treatment, surpassing expression levels measured in the adult AF (Figure 6E). Similarly, the expression of TNMD and ADAMTS6 surpassed that of mature AF cells with treatment, while PAX1 expression decreased with treatment (Figure 6E). These results demonstrated that treatment of SCL cells with a combination of TGF-β3 and PDGF-BB for 14 days differentiated cells away from SCL toward an anabolic state characteristic of fibrocartilaginous cells.
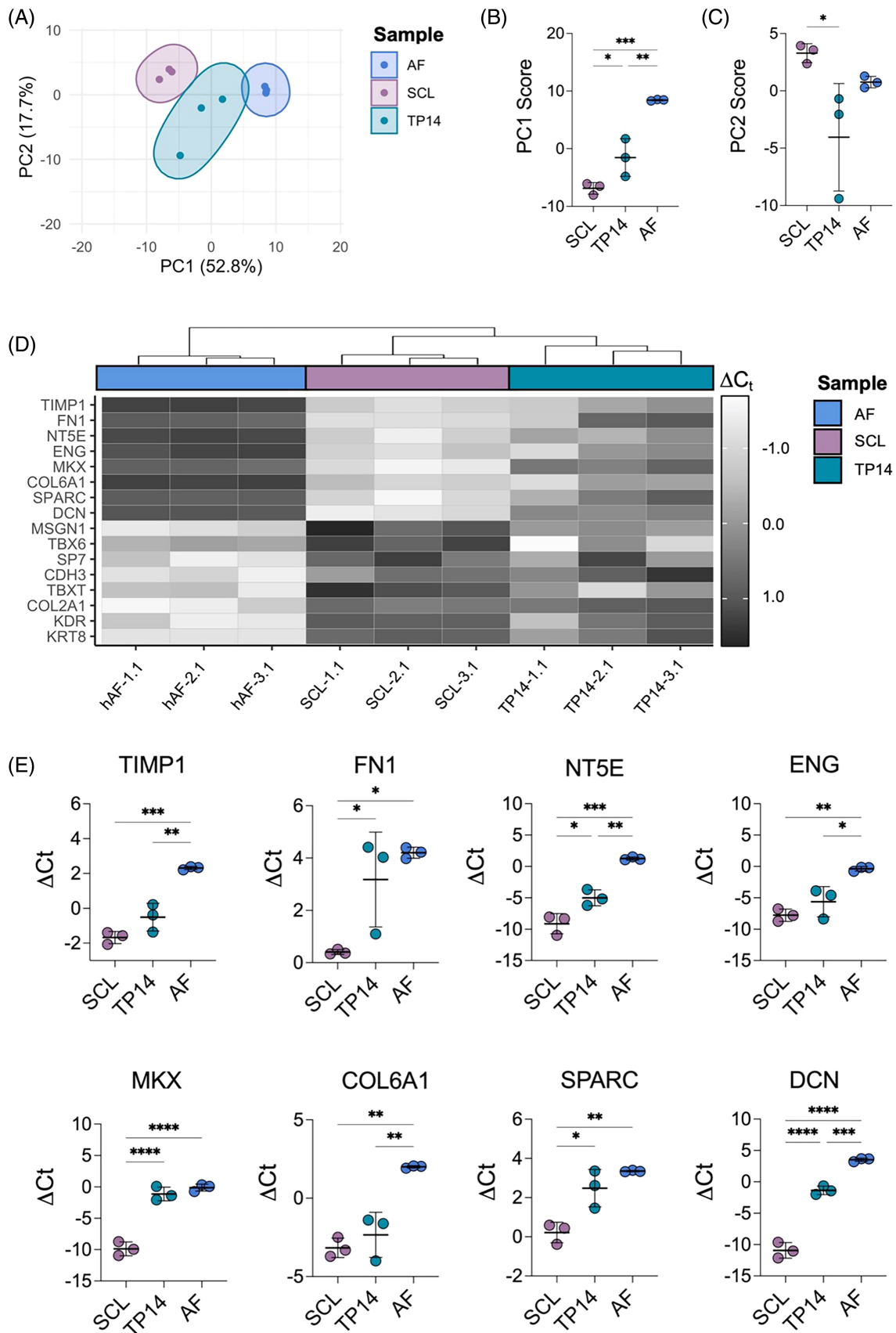
**FIGURE 5** Legend on next page.

# 4 | DISCUSSION

In this study, we screened developmental molecular signals involved in axial skeletal and IVD embryonic development to establish their effects on the induction of an AF-like phenotype in human iPSCs. The impact of TGF-β, PDGF, CTGF, IGF-1, and Hedgehog signaling were investigated individually or in combination with TGF-β, which is known to be indispensable for disc development, maintenance, and maturation.[26,27] Using this screen, the synergy of TGF-β, when combined with other factors, was evidenced at the transcriptomic level. This was marked by a dramatic increase in expression of AF-associated ECM genes that resulted in increased translation of AF-associated ECM proteins. Using a 96.96 Fluidigm gene expression array, the transcriptional signatures of cells before and after treatment were compared to that of mature AF cells. This established that combinations of developmental signals resulted in iPSC-derived sclerotome (SCL) cells to partially shift away from their primitive SCL fate and toward the early pro-anabolic AF-like phenotype.

To date, AF healing has only been shown during fetal and neonatal stages, indicating that early AF cells have a unique reparative potential ideal for tissue replenishment and healing.[39,40] Recreating this early AF phenotype from human iPSCs has never been attempted but may represent a promising approach as a cell therapy for disc herniation interventions. Although TGF-β has been identified as a key regulator of IVD formation, the effect of other signals present during AF embryonic development remains unknown. In this study, we demonstrate that TGF-β alone is insufficient to drive the upregulation of AF-associated ECM genes. Similarly, other signals such as PDGF-BB, CTGF, IGF-1, or SHH activation had limited effects when applied independently. However, TGF-β combinations with PDGF-BB, CTGF, or IGF-1 dramatically increased the expression of major AF ECM proteins, which resulted in increased matrix deposition. Surprisingly, broader examination of transcriptomic changes resulting from these signal combinations through a 96.96 Fluidigm gene expression array did not show significant differences between treatment combinations. This indicates that although TGF-β synergizes with PDGF-BB, CTGF, or IGF-1, the synergy remains relatively comparable for the genes assessed. Future studies using RNA sequencing will provide a greater appreciation of the transcriptomic signatures of each cell population. Furthermore, investigating cells from donors of different ages will enable the characterization of the genetic shifts associated with differentiation from an early embryonic AF cell phenotype, established in our study, to the more mature phenotype that has been further characterized.[21]

Even though our analysis did not yield differences between the TGF-β combinations, all treatments drove differences in gene expression toward the AF fate and away from the SCL fate. Analysis of the TP14 group showed a significant pro-anabolic shift in cells, where 14 days of treatment induced the upregulation of major and minor AF-associated ECM proteins involved in embryonic collagen fibrillogenesis.[41] Strikingly, Scx and Mkx, two transcription factors highly involved and necessary for early embryonic AF development, were significantly upregulated by treatment. SCX expression in treated cells reached levels greater than that of the mature AF cell expression level, in line with the high SCX expression observed in the embryonic AF that is attenuated with maturation.[23,42] Scx-lineage cells have also been implicated in the only reported instances of AF healing and in the proper attachment of ligamentous tissues.[23,40,43] AF healing in neonates is directed by Scx-lineage cells that proliferate and deposit type I collagen to repair AF lesions.[40] Scx-lineage cells that express TNMD and SOX9, two genes upregulated by treatment, have also been implicated in proper AF formation.[33,35] Similar to Scx, Mkx was upregulated with treatment to levels observed in mature AF cells.[44,45] Together, these results highlight the strong inductive potential of the signal combinations used to direct the AF specification of SCL progenitors. Future studies will compare these genetic shifts to the transcriptomic profiles of adult chondrocytes and NP cells.

Although several AF-associated genes were upregulated upon treatment, PCA analysis revealed that treated cells remained closer to SCL cells than to mature AF cells. This may be due the comparison of early embryonic cells to fully formed AF cells obtained from adult donors. Obtaining human embryonic AF tissue would provide the most accurate comparison, but this remains a challenging endeavor. Furthermore, the adult AF cells used in this study were acquired from a commercial vendor that did not discriminate cells based on region of origin in the native tissue. As such, the transcriptomic profiles identified represent human AF cells from multiple regions in the tissue. The segregation of AF cells based on anatomical region could further enhance our understanding of cellular heterogeneity present in the native tissue. Ultimately, driving cells toward an embryonic AF fate rather than a mature AF phenotype may yield a cell population more amenable to stimulating tissue repair or regeneration.[40]

Another potential reason for the incomplete shift away from the SCL transcriptomic profile is the lack of additional morphogenetic signals that may complete the differentiation process. Our study explored limited signaling players and did not investigate the induction effects of different concentrations, timing of initial exposure, and exposure duration. In addition to investigating these different aspects

**FIGURE 5** Treatment of sclerotome cells with factors involved in AF embryonic development drive changes in gene expression toward the expression patterns of adult AF cells. (A) PCA of the data represented by mapping PC1 and PC2 scores for sclerotome (SCL), annulus fibrosus (AF), and TP14 (SCL + TGFβ3 + PDGF-BB for 14 days) groups, derived using three different cell lines for each group ($n = 3$/group). Percent of described total variance by each principal component (PC) is listed in parentheses. Concentration ellipses for each group are demarcated via clustering. (B) PC1 scores and (C) PC2 scores for SCL, AF, and TP14 groups ($n = 3$/group; mean ± SD). (D) Heatmap of $\Delta C_t$ values for the 8 highest and 8 lowest valued genes identified from the AF versus SCL PCA, with hierarchical clustering based on the 93 genes included in the gene expression array with color coding by sample type. For the heatmap showing all 93 genes used in the cluster analysis, see Figure S5. (E) $\Delta C_t$ values for genes with highest PC1 score in AF versus SCL PCA (highest expression shown by AF cells) ($n = 3$/group; mean ± SD). Statistical significance denoted by *$p$-value <0.05, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.

**FIGURE 6** Factor treatment drives changes in gene expression toward an anabolic AF cell-like state. (A) $\Delta C_t$ values for ECM-related, (B) signaling pathways, (C) differentiation-related, (D) miscellaneous genes, and (E) transcription factors involved in a variety of cellular processes for sclerotome (SCL), annulus fibrosus (AF), and TP14 (SCL + TGFβ3 + PDGF-BB for 14 days) cells ($n = 3$/group; mean ± SD). Statistical significance denoted by *$p$-value <0.05, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.
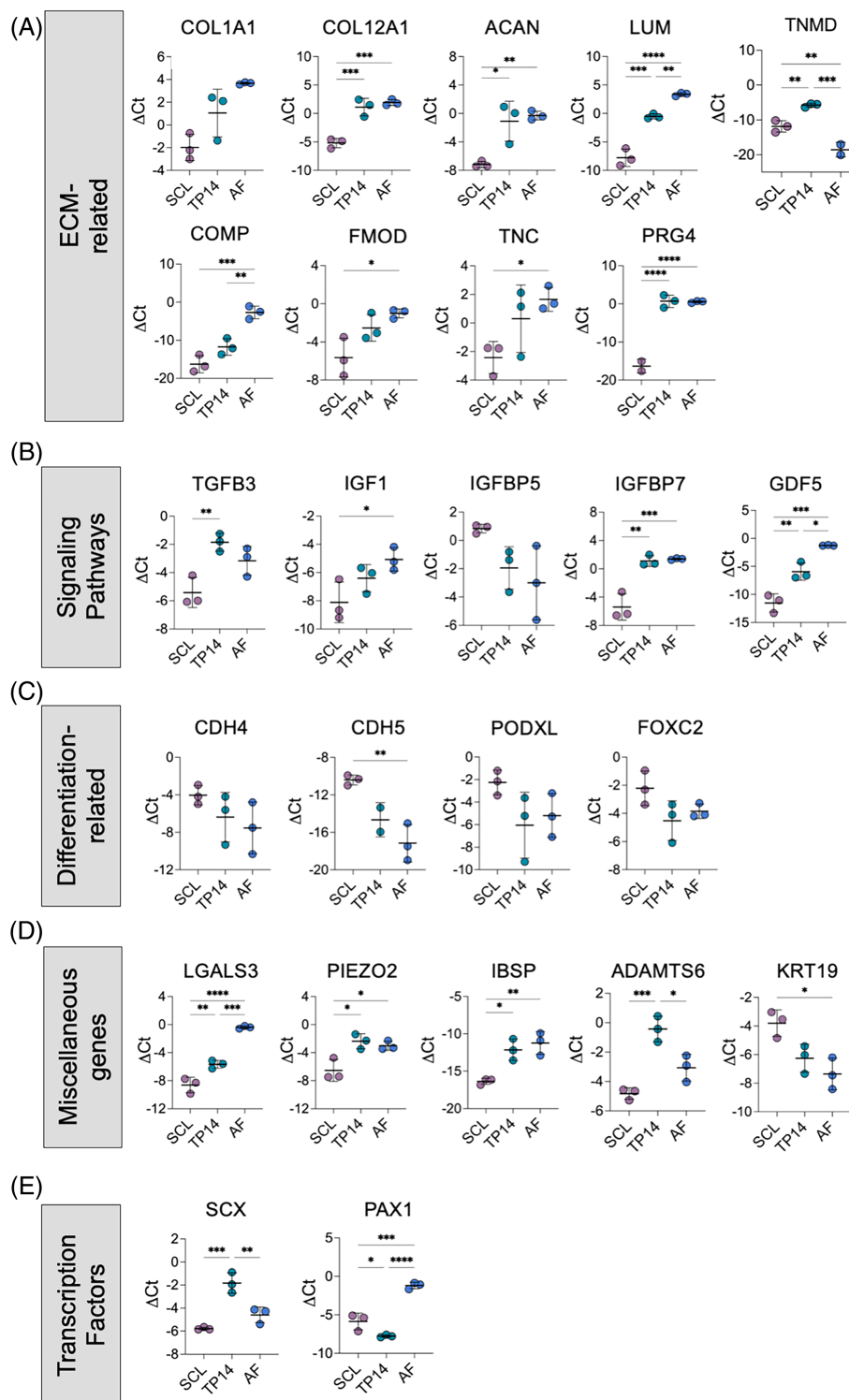


of treatment and other morphogen signals, the biophysical cues present during AF development could enhance AF specification. During embryonic development, the AF and surrounding tissues undergo continual growth and stiffening in unison with morphogenetic events.[46] Growth-mediated stresses and strains can dictate tissue segmentation patterns, activate molecular signaling events, and provide structural integrity to tissue structures during key morphogenetic events, playing an indispensable role throughout embryonic development.[47,48]

Future studies will focus on exploring the interplay between morphogens and biophysical cues to further improve the differentiation strategies employed for the generation of AF cells.

The AF is a tissue with numerous endogenous impediments for repair.[7] While this work sets the stage for the development of iPSC-based cellular therapies for AF repair, it is not without limitations. Given that we compared our treatments to SCL cells and mature AF cells only, further work is needed to compare expression profiles to

other fibrous cell populations (tendon, ligament), in addition to chondrocytes and nucleus pulposus cells. Such comparisons may also guide refinement of timing and doses of factors to more closely approximate the AF-associated phenotype. Additionally, functional aspects of treatment beyond gene expression will also need to be investigated in future work, including changes in proliferation rate, cell morphology, collagen, and proteoglycan production. Future investigations exploring more complex signals and their interplay with morphogens that guide AF fate induction will enable the development of a reparative AF cell population for the restoration of the AF structure and function. Combining the delivery of iPSC-based cellular therapies with the surgical removal of herniated tissue represents a promising approach to potentially improve repair outcomes and lower recurrent herniation rates after surgical intervention.

## AUTHOR CONTRIBUTIONS

Ana P. Peredo and Robert L. Mauck conceived the experiments. Ana P. Peredo conducted the experiments, performed statistical analysis, and generated figures, schematics, and illustrations. Edward D. Bonnevie conducted image analysis using CellProfiler. Ana P. Peredo, Tonia K. Tsinman, Xi Jiang, and Nathaniel A. Dyment were involved in the preparation and analysis of the Fluidigm GE 96.96 Dynamic Array. Harvey E. Smith, Sarah E. Gullbrand, Nathaniel A. Dyment, and Robert L. Mauck provided scientific and clinical guidance for the design of experiments. Ana P. Peredo and Robert L. Mauck wrote the manuscript, and all authors reviewed the manuscript and approved the final submission.

## CONFLICT OF INTEREST STATEMENT

Robert L. Mauck is Co-editor in Chief of JOR Spine and Sarah Gullbrand is an Editorial Board member of JOR Spine; both are co-authors of this article. They were excluded from editorial decision-making related to the acceptance of this article for publication in the journal.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author, Robert L. Mauck, upon reasonable request.

## ORCID

*Sarah E. Gullbrand* https://orcid.org/0000-0001-7806-6606
*Robert L. Mauck* https://orcid.org/0000-0002-9537-603X

## REFERENCES

1. Raj PP. Intervertebral disc: anatomy-physiology-pathophysiology-treatment. *Pain Pract*. 2008;8(1):18-44.
2. Stadnik TW, Lee RR, Coen HL, Neirynck EC, Buisseret TS, Osteaux MJ. Annular tears and disk herniation: prevalence and contrast enhancement on MR images in the absence of low back pain or sciatica. *Radiology*. 1998;206(1):49-55.
3. Jönsson B, Strömqvist B. Influence of age on symptoms and signs in lumbar disc herniation. *Eur Spine J*. 1995;4(4):202-205.
4. Schroeder GD, Guyre CA, Vaccaro AR. The epidemiology and pathophysiology of lumbar disc herniations. *Semin Spine Surg*. 2016;28(1):2-7.
5. Martens F, Vajkoczy P, Jadik S, Hegewald A, Stieber J, Hes R. Patients at the highest risk for reherniation following lumbar discectomy in a multicenter randomized controlled trial. *JBJS Open Access*. 2018;3(2): e0037.
6. Wong T, Patel A, Golub D, et al. Prevalence of long-term low Back pain after symptomatic lumbar disc herniation. *World Neurosurg*. 2023;170:163-173.e1.
7. Peredo AP, Gullbrand SE, Mauck RL, Smith HE. A challenging playing field: identifying the endogenous impediments to annulus fibrosus repair. *JOR Spine*. 2021;4(1):e1133.
8. Peredo AP, Gullbrand S, Smith HE, Mauck RL. Putting the pieces in place: mobilizing cellular players to improve annulus fibrosus repair. *Tissue Eng Part B Rev*. 2020;27(4):295-312.
9. Sloan SR, Lintz M, Hussain I, et al. Biologic annulus fibrosus repair: a review of preclinical in vivo investigations. *Tissue Eng Part B Rev*. 2018;24(3):179-190.
10. Hayes AJ, Benjamin M, Ralphs JR. Role of Actin stress fibres in the development of the intervertebral disc: cytoskeletal control of extracellular matrix assembly. *Dev Dyn*. 1999;215(3):179-189.
11. Sivakamasundari V, Lufkin T. Bridging the gap: understanding embryonic intervertebral disc development. *Cell Dev Biol*. 2012;1(2):103.
12. Maroudas A, Stockwell RA, Nachemson A, Urban J. Factors involved in the nutrition of the human lumbar intervertebral disc: cellularity and diffusion of glucose in vitro. *J Anat*. 1975;120(Pt 1):113-130.
13. Antoniou J, Steffen T, Nelson F, et al. The human lumbar intervertebral disc: evidence for changes in the biosynthesis and denaturation of the extracellular matrix with growth, maturation, ageing, and degeneration. *J Clin Investig*. 1996;98(4):996-1003.
14. Bonnevie ED, Gullbrand SE, Ashinsky BG, et al. Aberrant mechanosensing in injured intervertebral discs as a result of boundary-constraint disruption and residual-strain loss. *Nat Biomed Eng*. 2019;3 (12):998-1008.
15. Sato M, Asazuma T, Ishihara M, et al. An experimental study of the regeneration of the intervertebral disc with an allograft of cultured annulus fibrosus cells using a tissue-engineering method. *Spine (Phila Pa 1976)*. 2003;28(6):548-553.
16. Nukaga T, Sakai D, Schol J, Sato M, Watanabe M. Annulus fibrosus cell sheets limit disc degeneration in a rat annulus fibrosus injury model. *JOR Spine*. 2019;2(2):e1050.
17. Vadalà G, Sowa G, Hubert M, Gilbertson LG, Denaro V, Kang JD. Mesenchymal stem cells injection in degenerated intervertebral disc: cell leakage may induce osteophyte formation. *J Tissue Eng Regen Med*. 2012;6(5):348-355.
18. Tang R, Jing L, Willard VP, et al. Differentiation of human induced pluripotent stem cells into nucleus pulposus-like cells. *Stem Cell Res Ther*. 2018;9:61.
19. Xia K, Zhu J, Hua J, et al. Intradiscal injection of induced pluripotent stem cell-derived nucleus pulposus-like cell-seeded polymeric microspheres promotes rat disc regeneration. *Stem Cells Int*. 2019;2019: 6806540.

20. Zhang Y, Zhang Z, Chen P, et al. Directed differentiation of notochord-like and nucleus pulposus-like cells using human pluripotent stem cells. *Cell Rep*. 2020;30(8):2791-2806.e5.

21. Fernandes LM, Khan NM, Trochez CM, et al. Single-cell RNA-seq identifies unique transcriptional landscapes of human nucleus pulposus and annulus fibrosus cells. *Sci Rep*. 2020;10(1):15263.

22. Ferretti E, Hadjantonakis AK. Mesoderm specification and diversification: from single cells to emergent tissues. *Curr Opin Cell Biol*. 2019;61:110-116.

23. Sugimoto Y, Takimoto A, Akiyama H, et al. Scx+/Sox9+ progenitors contribute to the establishment of the junction between cartilage and tendon/ligament. *Development*. 2013;140(11):2280-2288.

24. Kaji DA, Montero AM, Patel R, Huang AH. Transcriptional profiling of mESC-derived tendon and fibrocartilage cell fate switch. *Nature Commun*. 2021;12(1):1-16.

25. Nakajima T, Shibata M, Nishio M, et al. Modeling human somite development and fibrodysplasia ossificans progressiva with induced pluripotent stem cells. *Development (Cambridge)*. 2018;145:(16):dev165431.

26. Baffi MO, Slattery E, Sohn P, Moses HL, Chytil A, Serra R. Conditional deletion of the TGF-β type II receptor in Col2a expressing cells results in defects in the axial skeleton without alterations in chondrocyte differentiation or embryonic development of long bones. *Dev Biol*. 2004;276(1):124-142.

27. Jin H, Shen J, Wang B, Wang M, Shu B, Chen D. TGF-β signaling plays an essential role in the growth and maintenance of intervertebral disc tissue. *FEBS Lett*. 2011;585(8):1209-1215.

28. Yakar S, Werner H, Rosen CJ. 40 years of IGF1: insulin-like growth factors: actions on the skeleton. *J Mol Endocrinol*. 2018;61(1):T115-T137.

29. Gruber HE, Norton HJ, Hanley EN. Anti-apoptotic effects of IGF-1 and PDGF on human intervertebral disc cells in vitro. *Spine*. 2000;25(17):2153-2157.

30. Gruber HE, Hoelscher GL, Ingram JA, Bethea S, Hanley EN. IGF-1 rescues human intervertebral annulus cells from in vitro stress-induced premature senescence. *Growth Factors*. 2008;26(4):220-225.

31. Hayes AJ, Ralphs JR. The response of foetal annulus fibrosus cells to growth factors: modulation of matrix synthesis by TGF-β1 and IGF-1. *Histochem Cell Biol*. 2011;136(2):163-175.

32. Loh KMM, Chen A, Koh PWW, et al. Mapping the pairwise choices leading from pluripotency to human bone, heart, and other mesoderm cell types. *Cell*. 2016;166(2):451-467.

33. Tran CM, Shapiro IM, Risbud MV. Molecular regulation of CCN2 in the intervertebral disc: lessons learned from other connective tissues. *Matrix Biol*. 2013;32(6):298-306.

34. Rajesh D, Dahia CL. Role of sonic hedgehog signaling pathway in intervertebral disc formation and maintenance. *Curr Mol Biol Rep*. 2018;4(4):173-179.

35. Adkar S, Wu C-L, Willard VP, 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*. 2014;32(6):1459-1467.

36. Kim M, Erickson IE, Choudhury M, Pleshko N, Mauck RL. Transient exposure to TGF-β3 improves the functional chondrogenesis of MSC-laden hyaluronic acid hydrogels. *J Mech Behav Biomed Mater*. 2012;11:92-101.

37. Matta A, Karim MZ, Isenman DE, Erwin WM. Molecular therapy for degenerative disc disease: clues from secretome analysis of the notochordal cell-rich nucleus pulposus. *Sci Rep*. 2017;7:45623.

38. Metsalu T, Vilo J. ClustVis: a web tool for visualizing clustering of multivariate data using principal component analysis and heatmap. *Nucleic Acids Res*. 2015;43(W1):W566-W570.

39. Torre OM, Das R, Berenblum RE, et al. Neonatal mouse intervertebral discs heal with restored function following herniation injury. *FASEB J*. 2018;32(9):4753-4762.

40. Torre OM, Mroz V, Benitez ARM, Huang AH, Iatridis JC. Neonatal annulus fibrosus regeneration occurs via recruitment and proliferation of Scleraxis-lineage cells. *npj Regen Med*. 2019;4(1):23.

41. Hayes AJ, Isaacs MD, Hughes C, et al. Collagen fibrillogenesis in the development of the annulus fibrosus of the intervertebral disc. *Eur Cell Mater*. 2011;22:226-241.

42. Yoshimoto Y, Takimoto A, Watanabe H, Hiraki Y, Kondoh G, Shukunami C. Scleraxis is required for maturation of tissue domains for proper integration of the musculoskeletal system. *Sci Rep*. 2017;7:45010.

43. Shukunami C, Takimoto A, Nishizaki Y, et al. Scleraxis is a transcriptional activator that regulates the expression of Tenomodulin, a marker of mature tenocytes and ligamentocytes. *Sci Rep*. 2018;8(1):3155.

44. Nakamichi R, Ito Y, Inui M, et al. Mohawk promotes the maintenance and regeneration of the outer annulus fibrosus of intervertebral discs. *Nat Commun*. 2016;7(1):1-14.

45. Nakamichi R, Kataoka K, Asahara H. Essential role of Mohawk for tenogenic tissue homeostasis including spinal disc and periodontal ligament. *Mod Rheumatol*. 2018;28(6):933-940.

46. van Uitert EM, Exalto N, Burton GJ, et al. Human embryonic growth trajectories and associations with fetal growth and birthweight. *Hum Reprod*. 2013;28(7):1753-1761.

47. Davidson LA. Mechanical design in embryos: mechanical signalling, robustness and developmental defects. *Philos Trans Royal Soc B: Biol Sci*. 2017;372(1720):20150516.

48. Chan CJ, Heisenberg C-P, Hiiragi T. Current biology review coordination of morphogenesis and cell-fate specification in development. *Curr Biol*. 2017;27:R1024-R1035.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.