

University of Kentucky

UKnowledge

---

Theses and Dissertations--Veterinary Science

Veterinary Science

---


2020

## COMPARATIVE CHONDROGENESIS OF INTERZONE AND ANLAGEN CELLS IN EQUINE SKELETAL DEVELOPMENT

ChanHee Mok

*University of Kentucky*, [chanhee.mok@uky.edu](mailto:chanhee.mok@uky.edu)

Author ORCID Identifier:

 <https://orcid.org/0000-0003-4159-7125>

Digital Object Identifier: <https://doi.org/10.13023/etd.2020.328>

[Right click to open a feedback form in a new tab to let us know how this document benefits you.](#)

### Recommended Citation

Mok, ChanHee, "COMPARATIVE CHONDROGENESIS OF INTERZONE AND ANLAGEN CELLS IN EQUINE SKELETAL DEVELOPMENT" (2020). *Theses and Dissertations--Veterinary Science*. 48.

[https://uknowledge.uky.edu/gluck\\_etds/48](https://uknowledge.uky.edu/gluck_etds/48)

This Doctoral Dissertation is brought to you for free and open access by the Veterinary Science at UKnowledge. It has been accepted for inclusion in Theses and Dissertations--Veterinary Science by an authorized administrator of UKnowledge. For more information, please contact [UKnowledge@sv.uky.edu](mailto:UKnowledge@sv.uky.edu).

## **STUDENT AGREEMENT:**

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

## **REVIEW, APPROVAL AND ACCEPTANCE**

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

ChanHee Mok, Student

Dr. James N. MacLeod, Major Professor

Dr. Daniel K. Howe, Director of Graduate Studies

COMPARATIVE CHONDROGENESIS OF INTERZONE AND ANLAGEN CELLS  
IN EQUINE SKELETAL DEVELOPMENT

---

DISSERTATION

---

A dissertation submitted in partial fulfillment of the  
requirements for the degree of Doctor of Philosophy  
in the College of Agriculture, Food and Environment  
at the University of Kentucky

By

ChanHee Mok

Lexington, Kentucky

Director: Dr. James N. MacLeod, Professor of Veterinary Science

Lexington, Kentucky

2020

Copyright © ChanHee Mok 2020

<https://orcid.org/0000-0003-4159-7125>

## ABSTRACT OF DISSERTATION

### COMPARATIVE CHONDROGENESIS OF INTERZONE AND ANLAGEN CELLS IN EQUINE SKELETAL DEVELOPMENT

At the presumptive sites of future synovial joints during mammalian skeletogenesis, articular cartilage develops from interzone located between the cartilaginous anlagen of bones. Thus, two types of cartilaginous tissues differentiate in close proximity. While anlagen cartilage is transient, progressing through endochondral ossification to form bones, articular cartilage is stable and functions throughout life to facilitate both low friction movement and load distribution. Despite important life-long functional properties, articular cartilage has a very limited intrinsic ability to repair structural defects. On the other hand, structural lesions in bones generally heal well by forming a cartilaginous callus and recapitulating endochondral ossification to repair fractures and other defects. Therefore, understanding the comparative aspects of interzone and anlagen cell differentiation may provide novel insights into emergent cell-based therapies to support articular cartilage regeneration. The objective of this dissertation research was to compare patterns of gene expression between equine interzone and anlagen cells across multiple post-induction time points to test the hypothesis that chondrogenic differentiation of these two cell lines is directed to articular and hypertrophic developmental pathways, respectively. The first part of the study was conducted using microfluidic RT-qPCR to analyze a selected panel of 93 genes. The data provided evidence that genes involved in transcriptional regulation and signaling transduction are differentially expressed as early as 1.5 hours after the start of chondrogenic induction, followed at later time points by effector genes such as those encoding cartilage matrix proteins. Then, RNA sequencing was used to expand the analyses at selected time points to a whole transcriptome level. A pilot single cell RNA sequencing experiment further described the two chondrogenic pathways characterizing subpopulations of these skeletal cell lines. Taken together, the results demonstrated that interzone and analgen cells respond very quickly but in different ways to the same inductive signals. Important regulatory mechanisms are likely activated almost immediately, within a few hours, after chondrogenic induction. These differential regulatory responses progress to cell type-specific profiles of effector genes that result in the two different cartilaginous tissues.

Key words: Arthritis, Bone, Cartilage, Horse, Synovial joint

ChanHee Mok

June 18, 2020

COMPARATIVE CHONDROGENESIS OF INTERZONE AND ANLAGEN CELLS  
IN EQUINE SKELETAL DEVELOPMENT

By

ChanHee Mok

Dr. James N. MacLeod

*Director of Dissertation*

Dr. Daniel K. Howe

*Director of Graduate Studies*

June 18, 2020

## ACKNOWLEDGEMENTS

First, I would like to express my sincere gratitude to my primary advisor, Dr. James MacLeod. He gave me the amazing opportunity to conduct this intriguing research project as well as to extend my scientific education. With enthusiasm and patient guidance, he trained me to be an independent scientist and has continued to believe in me. He has been incredibly inspiring, encouraging, and supportive throughout my entire journey through the PhD program.

My long academic training started 8 years ago. At the beginning of this long endeavor, Dr. Beob Gyun Kim introduced me to scientific research and encouraged me to pursue further graduate study, and over the years, he has always been by my side, guiding me to be a better scientist. Also, I cannot thank Dr. Kristine Urschel enough for giving me the opportunity to start equine research in Kentucky, the Horse Capital of the World, and to earn my academic training in the U.S. She has always been supportive of me developing my academic career over the years.

I would like to thank my PhD committee members, Drs. Kathryn Graves, Douglas Harrison, and Daniel Howe, for their mentorship. They have guided me in the right direction and provided priceless mentorship throughout my program. I am grateful to have had these amazing, influential mentors on my academic journey.

Also, Dr. James Matthews generously agreed to provide his expertise for my final examination as an outside examiner. I greatly appreciate his time and service.

My research has been built on all the knowledge that has been accumulated over multiple years in our laboratory, the MacLeod lab, and I know I could not make it this far without all the help and valuable teamwork from the lab. Therefore, I would like to thank all the current and former members of the MacLeod lab.

The most challenging part of my dissertation project was dealing with the enormous RNA-seq data sets. Dr. Ted Kalbfleisch helped me understand bioinformatics by answering so many beginner questions. Also, my data analyses would have been very challenging without Dr. Hossam Ali's help. He spent a lot of time with me and taught me how to sail on the sea of transcriptomic data.

I am also grateful for all the support and help that I could get from our department—the Maxwell H. Gluck Equine Research Center. The faculty members, staff, and fellow graduate students are always willing to help each other's research—being cooperative, professional, and very knowledgeable. I am deeply honored to be a part of this group.

I have met so many kind friends during this journey—fellow students and post-doc scholars in the veterinary science department and my old friends in the animal science department from my master's program. They have been by my side during this time of joy, frustration, happiness, and depression, as well as during the ordinary days. They made

Lexington, which is 7,000 miles away from my home, feel like home. I genuinely value their friendship.

Last but not least, I would like to thank my family and old friends in my home country. I spent seven years in the United States pursuing my master's and PhD. Over the years, they have never forgotten me and always have been there for me. I wanted to make them proud, and I hope this might be a moment where they smile.

I will not forget this precious experience throughout my future career. Again, I deeply appreciate every individual who helped me to get here and be prepared for the next step.



## Table of contents

ACKNOWLEDGEMENTS.....	iii
Table of contents .....	vi
List of Tables .....	ix
List of Figures .....	xii
Chapter 1. Introduction .....	1
Articular cartilage and joint health .....	1
Fetal limb skeletal development.....	5
Molecular mechanisms involved in limb skeletal development.....	8
Advances in technologies for gene expression evaluation.....	12
Overview of the dissertation .....	13
Chapter 2. Kinetics of gene expression changes in equine fetal interzone and anlagen cells during 14 days of <i>in vitro</i> chondrogenesis .....	15
Introduction .....	15
Materials and Methods.....	17
<i>Cell culture and sample collection</i> .....	17
<i>Total RNA isolation</i> .....	18
<i>Quantitative polymerase chain reaction (qPCR)</i> .....	19
<i>Data analyses and statistics</i> .....	27
Results.....	29
<i>Genes processed in data analyses</i> .....	29
<i>Kinetics of gene expression changes by cell type</i> .....	29
<i>Relative differences between cell types at each time point</i> .....	39
Discussion .....	48
<i>Relationship between early and delayed gene expression changes</i> .....	48
<i>Differential gene regulation between cell types</i> .....	50
<i>Future direction</i> .....	54
Conclusion.....	55
Chapter 3. Transcriptomic divergence between equine fetal interzone and anlagen cells during <i>in vitro</i> chondrogenic induction.....	56
Introduction .....	56
Materials and Methods.....	58

<i>Cell culture</i> .....	58
<i>RNA sample preparation</i> .....	59
<i>Bulk RNA sequencing</i> .....	60
<i>Data analysis pipeline</i> .....	60
Results.....	63
<i>Kinetics of gene expression profiles within a cell type</i> .....	65
<i>Common chondrogenic characteristics of interzone and analgen cell cultures</i> .....	90
<i>Distinct traits between interzone and analgen cell cultures</i> .....	97
Discussion .....	110
<i>Kinetics of gene expression patterns during 96 hours of in vitro chondrogenesis</i> .....	110
<i>Chondrogenic characteristics observed in both interzone and analgen cell cultures</i> .....	113
<i>Differential transcriptomic signatures between interzone and analgen cells</i> .....	120
<i>Future direction</i> .....	128
Conclusion.....	129
Chapter 4. A pilot single cell RNA-seq study: Evaluation of chondrogenic divergence between equine fetal interzone and analgen cell cultures at the single cell level .....	130
Introduction .....	130
Materials and Methods.....	132
<i>Cell culture</i> .....	132
<i>Single cell suspension preparation</i> .....	133
<i>Single cell cDNA library construction</i> .....	134
<i>RNA-sequencing</i> .....	137
<i>Data analysis pipeline</i> .....	137
Results.....	139
<i>Cell subpopulations within a sample</i> .....	140
<i>Aggregation of the entire sample set</i> .....	154
Discussion .....	169
<i>Subpopulations of cells in each sample</i> .....	170
<i>Shared features and unique properties between interzone and analgen cells</i> .....	171
<i>Future direction</i> .....	174
Conclusion.....	175
Chapter 5. Overall summary and future directions .....	176

References .....	185
Vita .....	194

## List of Tables

Table 2.1. The panel of 96 equine specific TaqMan® primer-probe sets used in the RT-qPCR analysis.....	21
Table 2.2. Timing of the first response to the chondrogenic induction protocol in new gene × cell type combinations within each of the three functional annotation groups .....	36
Table 2.3. Timing of the first response to the chondrogenic induction protocol within functional annotation group .....	38
Table 2.4. Four patterns of differential gene expression before and after inducing <i>in vitro</i> chondrogenesis.....	39
Table 3.1. Sequencing depth of all samples.....	64
Table 3.2. Mapping efficiency of all samples.....	64
Table 3.3. Number of differentially expressed genes (DEGs) within cell type when compared to previous time point.....	65
Table 3.4. Top five biological process gene ontology (GO) terms of differentially upregulated genes in interzone cell cultures during time courses.....	69
Table 3.5. Top five biological process gene ontology (GO) terms of differentially downregulated genes in interzone cell cultures during time courses .....	70
Table 3.6. Ten most highly correlated hub genes with time course in interzone cell cultures during entire experimental period.....	71
Table 3.7. Top five biological process gene ontology (GO) terms of differentially upregulated genes in anlagen cell cultures during time courses .....	77
Table 3.8. Top five biological process gene ontology (GO) terms of differentially downregulated genes in anlagen cell cultures during time courses .....	78
Table 3.9. Ten most highly correlated genes with time course in anlagen cell culture during entire experimental period .....	79
Table 3.10. Top five biological process gene ontology (GO) terms of differentially upregulated genes in fibroblast cultures during time courses.....	85
Table 3.11. Top five biological process gene ontology (GO) terms of differentially downregulated genes in fibroblast cultures during time courses.....	86
Table 3.12. Ten most highly correlated genes with time course in anlagen cell culture during entire experimental period .....	87
Table 3.13. Number of differentially expressed genes (DEGs) in interzone and anlagen cell cultures compared to the negative control at each time point .....	91
Table 3.14. Number of overrepresented biological process gene ontology (GO) terms identified in the chondrogenic cell cultures compared to fibroblast cultures.....	91

Table 3.15. Number of differentially expressed genes (DEGs) and candidate molecular switches in the chondrogenic cell cultures compared to fibroblast cultures.....	91
Table 3.16. Genes that were exclusively expressed in both interzone and anlagen cell cultures at the same time points while not expressed in fibroblast cultures .....	92
Table 3.17. Genes that were not expressed exclusively in both interzone and anlagen cell cultures at the same time points while expressed in fibroblast cultures.....	93
Table 3.18. Upregulated overrepresented biological processes in the chondrogenic cell lines compared to fibroblast cultures at all time points .....	95
Table 3.19. Downregulated overrepresented biological processes in the chondrogenic cell lines compared to fibroblast cultures at all time points .....	96
Table 3.20. Number of differentially expressed genes (DEGs) between interzone and anlagen cell cultures at each time point.....	97
Table 3.21. Exclusively expressed genes in either interzone or anlagen cell cultures .....	100
Table 3.22. Ten most upregulated differentially expressed genes in interzone cell cultures compared to anlagen cell cultures at each time point .....	102
Table 3.23. Ten most upregulated differentially expressed genes in anlagen cell cultures relative to interzone cell cultures at each time point.....	104
Table 3.24. Gene ontology (GO) IDs and corresponding terms present in Figure 3.15.....	107
Table 4.1. 10X Genomics' sample indexes.....	136
Table 4.2. cDNA library concentration and average size of fragments .....	136
Table 4.3. Sequencing results .....	140
Table 4.4. Number of cells in each cluster within sample .....	141
Table 4.5. Number of differentially expressed genes in each cluster within sample.....	141
Table 4.6. Top five upregulated and downregulated differentially expressed genes in each cluster of interzone cell pellet sample at 24h .....	142
Table 4.7. Top five upregulated and downregulated differentially expressed genes in each cluster of interzone cell pellet sample at 48h .....	143
Table 4.8. Top five upregulated and downregulated differentially expressed genes in each cluster of anlagen cell pellet sample at 24h .....	144
Table 4.9. Top five upregulated and downregulated differentially expressed genes in each cluster of anlagen cell pellet sample at 48h .....	145
Table 4.10. Number of differentially expressed genes in each cluster of the aggregate of all four samples .....	155
Table 4.11. Five most upregulated and downregulated differentially expressed genes in each cluster of the aggregate of all four samples .....	156

Table 4.12. Overrepresented biological processes from upregulated differentially expressed gene profile in Cluster 9 of the aggregate of all four samples ..... 160

Table 4.13. Overrepresented biological processes from upregulated differentially expressed gene profile in Cluster 13 of the aggregate of all four samples ..... 161

Table 4.14. Top ten overrepresented biological processes from downregulated differentially expressed gene profile in Cluster 6 of the aggregate of all four samples ..... 165

Table 4.15. Top ten overrepresented biological processes from differentially expressed gene profile in Cluster 2 of the aggregate of all four samples ..... 167

Table 4.16. Top ten overrepresented biological processes from differentially expressed gene profile in Cluster 4 of the aggregate of all four samples ..... 168

## List of Figures

Figure 1.1. Structure of articular cartilage.....	2
Figure 1.2. Synovial joint formation.....	6
Figure 1.3. Endochondral ossification.....	7
Figure 1.4. TGF- $\beta$ signaling pathways in cartilage formation and maintenance .....	11
Figure 2.1. Steady state mRNA levels (Ct, mean $\pm$ SEM) of three prospective endogenous controls (B2M, GUSB, and RPLP0) across the 336-hour experimental period .....	28
Figure 2.2. Changes in steady state mRNA levels in response to the chondrogenic stimulation relative to the 0h time point.....	33
Figure 2.3. Genes with significantly different steady state mRNA levels between passage 4 monolayer cultures and passage 5 cell pellet cultures at 0h.....	34
Figure 2.4. Histogram of percentages of the first response to the chondrogenic stimulation in new gene $\times$ cell type combinations within each of the three functional annotation groups .....	36
Figure 2.5. Four patterns of differential gene expression before and after inducing <i>in vitro</i> chondrogenesis.....	43
Figure 2.6. Steady state mRNA levels of classic cartilaginous biomarkers across the experimental period.....	46
Figure 2.7. Steady state mRNA levels of example genes that were differentially regulated in interzone cell cultures compared to the other cell types across the experimental period .....	47
Figure 2.8. Steady state mRNA levels of an example gene that was differentially regulated in anlagen cell cultures compared to the other cell types across the experimental period .....	47
Figure 3.1. Phred scores of the 90 samples assessed by FastQC v0.11.9 .....	63
Figure 3.2. Differentially expressed genes (DEGs; n= 2,822) over time within interzone cell cultures .....	68
Figure 3.3. Module eigengene (ME) groups with module–trait relationship scores in interzone cell cultures over time.....	71
Figure 3.4. Hub gene co-expression network in interzone cell cultures during 96 hours of <i>in vitro</i> chondrogenesis.....	72
Figure 3.5. Upstream regulators predicted by Ingenuity Pathway Analysis (IPA) between subsequent time points in interzone cell cultures.....	73
Figure 3.6. Differentially expressed genes (DEGs; n=3,784) over time within anlagen cell cultures .....	76
Figure 3.7. Module eigengene (ME) groups with module–trait relationship scores in anlagen cell cultures over time.....	79

Figure 3.8. Hub gene co-expression network in anlagen cell cultures during 96 hours of <i>in vitro</i> chondrogenesis.....	80
Figure 3.9. Upstream regulators predicted by Ingenuity Pathway Analysis (IPA) between subsequent time points in anlagen cell cultures .....	81
Figure 3.10. Differentially expressed genes (DEGs; n=3,212) over time within fibroblast cultures .....	84
Figure 3.11. Module eigengene (ME) groups with module–trait relationship scores in fibroblast cultures over time.....	87
Figure 3.12. Hub gene co-expression network in fibroblast cultures during 96 hours of <i>in vitro</i> chondrogenesis.....	88
Figure 3.13. Upstream regulators predicted by Ingenuity Pathway Analysis (IPA) between subsequent time points in fibroblast cultures.....	89
Figure 3.14. Differentially expressed genes (DEGs) between interzone and anlagen cell cultures at each time point plotted in volcano plots.....	98
Figure 3.16. Prediction of upstream regulators that differentially regulate the chondrogenic pathways between interzone and analgen cultures by Ingenuity Pathway Analysis (IPA) .....	109
Figure 4.1 Cell pellet proteoglycan staining.....	131
Figure 4.2. Automated electrophoresis analyses conducted for amplified library quality assessment using a Bioanalyzer 2100 (Agilent Technologies).....	136
Figure 4.3. Chondrocyte biomarker gene (COL2A1, COMP, and ACAN) expression in each sample, plotted by uniform manifold approximation and projection (UMAP) analysis .....	149
Figure 4.4. Interzone biomarker gene (GDF5, WNT9A, and ENPP2) expression in each sample, plotted by uniform manifold approximation and projection (UMAP) analysis .....	151
Figure 4.5. Chondrocyte hypertrophy biomarker gene (MMP13 and DLX5) expression in each sample, plotted by uniform manifold approximation and projection (UMAP) analysis.....	153
Figure 4.6. Non-linear dimensionality reduction analysis (uniform manifold approximation and projection; UMAP) conducted on the aggregate of all four samples .....	158
Figure 4.7. Biomarker gene expression in the aggregate of samples, plotted by uniform manifold approximation and projection (UMAP) analysis, with a separated view by sample .....	163
Figure 5.1. Molecular regulatory model in interzone and analgen cell cultures during <i>in vitro</i> chondrogenesis.....	181



## **Chapter 1. Introduction**

### **Articular cartilage and joint health**

Articular cartilage is a hyaline tissue lining the ends of bones on opposing surfaces in diarthrodial joints. Unlike other hyaline cartilage in the body, such as cartilage in the nose, larynx, trachea, or ribcage, articular cartilage is not covered by perichondrium; instead, it is encapsulated by a synovial membrane and surrounded by synovial fluid, from which nutrients are delivered to the tissue. The fibrous synovial joint capsule connects the perimeter articular surfaces on adjoining bone surfaces, anatomically delineating the total diarthrodial joint structure.

On a broad component level, articular cartilage consists of 65 – 80% water, 10 – 25% extracellular matrix (ECM), and 10% cells—articular chondrocytes, which produce the ECM (Fisher et al., 2019). The composition of the ECM in dry matter is 50% collagens (primarily type II and type IX), and the other 50% is non-collagenous proteins such as proteoglycans (e.g., aggrecan) and glycosaminoglycans (e.g., cartilage oligomeric matrix protein and hyaluronan; Lane and Weiss, 1975). The cells and ECM in articular cartilage are organized in a zonal structure: 1) superficial tangential zone, 2) middle zone, and 3) deep zone in order from the surface towards the subchondral bone. The border between articular cartilage and the subchondral bone is demarcated by the “tidemark,” a calcified cartilaginous junction.

In the superficial zone, articular chondrocytes are flattened, somewhat smaller in size, and relatively densely arranged. Cells become round in shape and are sparsely distributed in the middle zone. Towards the deep zone, chondrocytes are stacked together, making short columns, and arranged perpendicular to the surface. Approaching the subchondral bone, chondrocytes become hypertrophic and calcify the ECM (Figure 1.1).

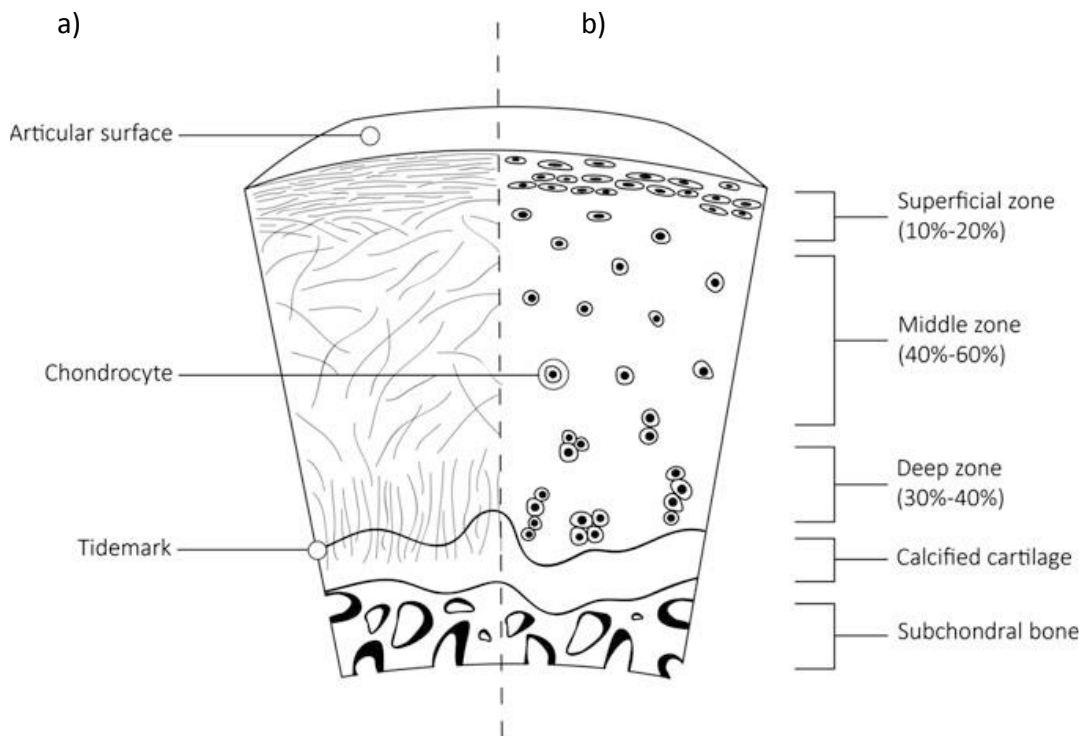


Figure 1.1. Structure of articular cartilage. a) Collagen fiber architecture; b) Cellular organization across the zones of articular cartilage (Copyright, Ondr sik et al., 2017, reproduced with permission)

The ECM is also arranged differently across the zones. Collagen fibrils, which make up a major part of the ECM content, are oriented parallel to the surface in the superficial zone and become isometrically distributed in the middle zone. Parallel to chondrocyte columns, collagen fibrils orient to a more vertical arrangement in the deep zone (Figure 1.1). At the same time, proteoglycan content becomes greater in the middle zone and the deep zone

compared to the superficial zone. Aggrecan is the major proteoglycan in articular cartilage and is highly sulfated and heavily glycosylated, which gives it a net negative electrostatic charge. This protein binds to hyaluronan and captures water, creating osmotic swelling pressure, and its interaction with collagens endows colloidal properties. Altogether, the characteristics of the ECM provide tensile strength and compressive stiffness of articular cartilage. This smooth, tough tissue facilitates low friction movement, shock absorption, and weight load distribution. Articular cartilage is a stable tissue, which maintains its structure and functions throughout life.

Despite its important biomechanical functions, traumatic injuries in articular cartilage often manifest as chronic arthritis and are the most prevalent joint diseases in various animals including humans, livestock animals, companion animals, and horses. Among the US equine population, chronic joint problems are the most common reason for lameness (20.9% of lame resident horses; USDA, 2017). Although the occurrence of articular cartilage degenerative diseases generally increases with age (USDA, 2017), these joint injuries are even more problematic in younger horses because their “product” usually centers on athletic performance. Horses, the primary patient population of interest in our laboratory, have several advantages as a model animal for these studies. In addition to the fact that the thickness of articular cartilage is comparable between horses and human (Sophia et al., 2009), several logistical challenges involving technical issues that would be present with small rodent models can be avoided with horses. The fetal limb buds are only a couple of millimeters in length even in horses, and isolating interzone and analgen cells from this small tissue is extremely difficult with smaller animal models and typically requires laser dissection

techniques and very limited cell yields. Also, horses are athletic individuals and therefore aspects of the data obtained from equine samples will be more relevant to the further translational experiments that address specific biomedical questions related to sport medicine.

While articular cartilage degeneration is frequently seen in synovial joints, the tissue's intrinsic ability to restore structural defects is very limited in mature mammals (Alford and Cole, 2005). Located in a hypoxic environment, this tissue is aneural, avascular, and alymphatic; these are the reasons why its degeneration is more troublesome. Because the tissue does not have any nerves, even though it is injured, the afflicted individual does not perceive pain directly from the articular cartilage. Nociceptive neurons would be located in the surrounding tissues such as the subchondral bone area or joint capsule. Furthermore, because this tissue does not have blood and lymph vessels, materials and factors needed for the tissue regeneration need to diffuse across large distances, which hinders the intrinsic tissue repair. Therefore, clinical interventions are challenged and efforts to support articular cartilage restoration met with limited success.

For this purpose, emergent cell-based therapies in articular cartilage regenerative medicine are being commonly applied; mesenchymal stem cells with multipotent differentiation potential are treated with chondrogenic induction factors and transplanted into articular cartilage lesions. Clinical outcomes to date, however, continue to have frustratingly limited success. The repaired tissue is often fibrous (fibrocartilage), and chondrocytes in repaired cartilage may undergo hypertrophy, followed by calcification of the

ECM (Beris et al., 2005, Caldwell and Wang, 2015). Thus, repaired cartilage has inferior biomechanical function and durability compared to normal articular cartilage, and therefore, the performance of the animal would be deteriorated. Further research is still required to improve clinical approaches for supporting articular cartilage regeneration.

### **Fetal limb skeletal development**

To advance current articular cartilage regenerative medicine, consideration of the normal developmental processes that generate limb skeletal elements may provide novel insights. In early embryonic stages, mesenchymal cells derived from the paraxial/lateral mesoderm aggregate together at a presumptive site of a limb. This mesenchymal condensation then undergoes chondrogenic differentiation forming a continuous, uninterrupted cartilaginous limb bud. Then, this cartilaginous tissue becomes properly segmented, resulting in several cartilaginous anlagen, which serve as templates for limb bones (Pitsillides and Ashhurst, 2008, Decker et al., 2014).

Between those cartilaginous tissues, “interzone” tissue develops, and cells in this region change their morphology, becoming flattened and densely packed. This tissue is characterized by paused chondrogenesis. At early stages in fetal development, multiple synovial joint elements—ligament, joint capsule, synovial membrane as well as articular cartilage—develop from this layer of interzone cells. As synovial joint formation proceeds in mammals, the space between adjoining bone surfaces becomes cavitated. During these processes, a portion of interzone cells resume chondrogenic differentiation for articular cartilage formation (Figure

1.2). Once articular chondrocytes are differentiated, these cells are stable, maintaining articular cartilage on the joint surface location for life.

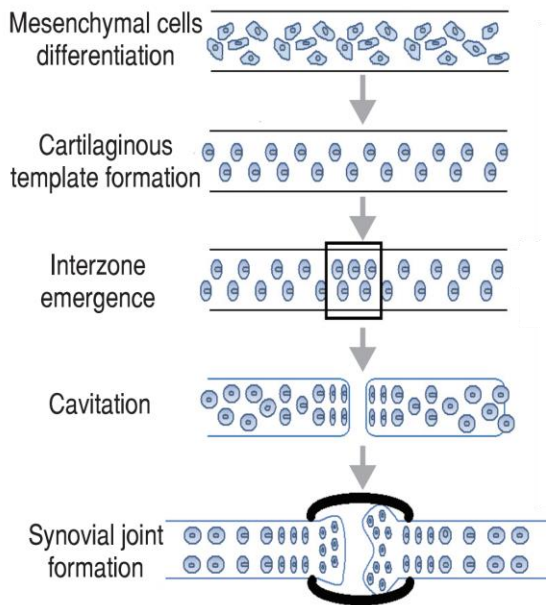


Figure 1.2. Synovial joint formation (Copyright: Moskalewski et al., 2013, reproduced with permission)

At the same time, cartilaginous anlagen undergo terminal hypertrophic differentiation. Starting from the center of each anlage, chondrocytes rapidly proliferate, mature, and start expressing hypertrophic ECM such as collagen type X. Then, the tissue becomes calcified, preventing the chondrocytes from approaching nutrients and in turn, resulting in apoptosis (Ham, 1952, Cameron, 1963). The cell death generates vacancy in the anlagen and allows blood vessels to invade the hypertrophic regions. The blood brings osteogenic factors and cells into the anlagen, establishing the primary ossification centers. Finally, the medullary cavity becomes enlarged, and hypertrophic cartilage is replaced by bone tissue in the normal process of bone formation, which is termed endochondral ossification (Figure 1.3). Thus, anlagen chondrocytes are transient.

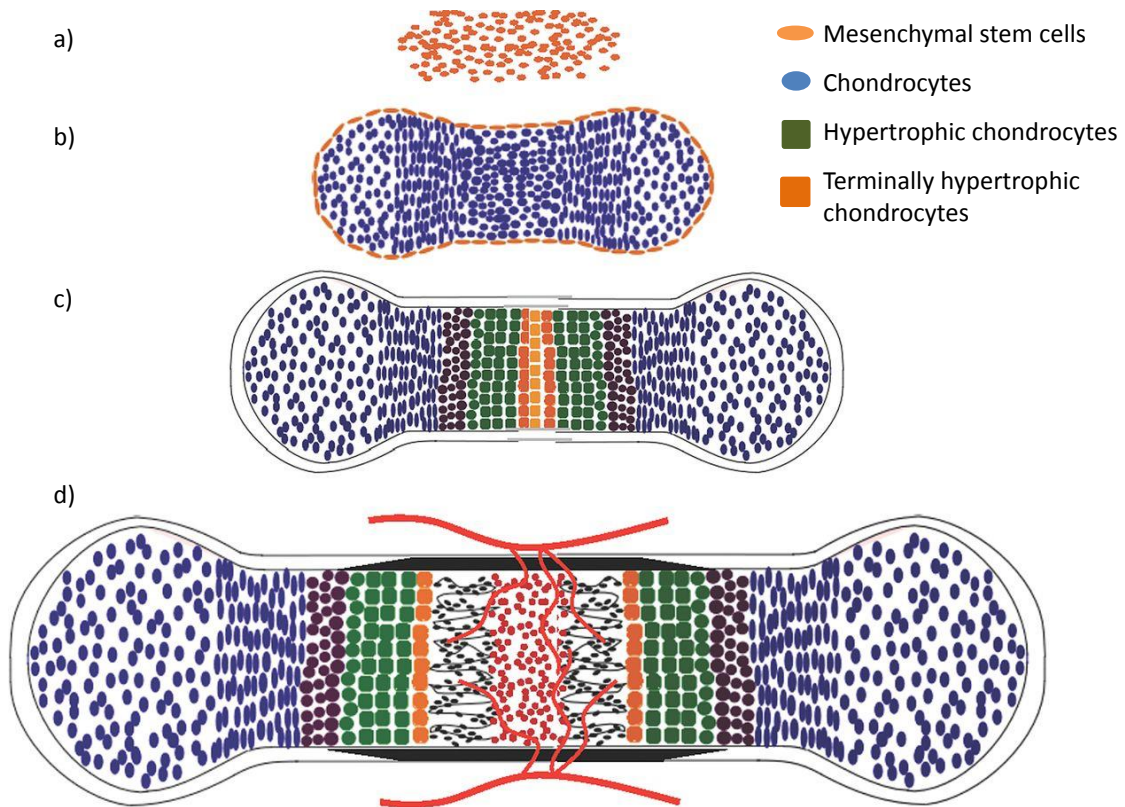


Figure 1.3. Endochondral ossification. a) Mesenchymal condensation; b) Chondrocyte differentiation; from the center of the mesenchymal condensation, cells differentiate into chondrocytes; c) Chondrocyte maturation; chondrocytes at the center of the cartilaginous anlage undergo hypertrophy until terminal stages; d) Cartilage vascularization; blood vessels (red lines) invade the center of the hypertrophic zone, and vascular invasion leads to resorption of cartilaginous matrix and deposition of bone (black) within the medullary cavity (Copyright: Long and Ornitz, 2013, reproduced with permission).

These two different developmental chondrogenic processes—one directed towards articular cartilage development and the other directed towards hypertrophic differentiation leading to bone formation—occur simultaneously and in close proximity. As a result of recapitulating developmental processes, fractured bones actually repair quite well, provided that the fracture is reduced, stabilized, and not compromised by infection or loss of blood supply. In stark contrast mammalian articular cartilage has limited potential for tissue regeneration.

Interestingly, as demonstrated roughly a decade ago in our laboratory, some vertebrate species such as mature axolotl salamanders retain interzone-like tissue in their distal limb joints which provides this amphibia with the ability to restore even large articular cartilage defects (Cosden et al., 2011). Furthermore, this axolotl interzone-like tissue has the potential to generate an entirely new diarthrodial joint *de novo* within a skeletal microenvironment (Cosden-Decker et al., 2012). These axolotl studies demonstrated the potential of interzone cells, which do not undergo hypertrophic differentiation, to regenerate articular cartilage tissue. Taken together, this developmental biology and previous literature suggest that understanding developmental processes of limb skeletal elements and the comparative aspect of interzone and anlagen cells may well provide important information to advance therapeutic approaches for mammalian articular cartilage regeneration.

### **Molecular mechanisms involved in limb skeletal development**

The homeobox (HOX) gene family of transcription factors is well-conserved across the species. These genes are widely involved in embryonic developmental processes and participate in axial patterning. Mammalian HOX paralogs (HOXA-D) are located on four separate chromosomes and their functional annotations are more similar to their own paralogs on the other chromosomes compared to their neighboring HOX genes on the same chromosome. The relative location within a chromosome defines their groups: anterior (HOX1-5), central (HOX6-8), and posterior (HOX9-13) clusters from the 3' end to the 5' end. While the HOX genes generally establish the anterior-posterior axis in the body, the posterior HOX genes (HOX9-13) are also involved in limb fields specification and appendicular skeletal



(limb) patterning (Mohanty-Hejmadi et al., 1992, Nelson et al., 1996, Pineault and Wellik, 2014). Along a limb bud, from proximal to distal, the posterior HOX genes are collinearly expressed as their relative location on chromosomes; HOX9 and 10 patterns the stylopod, HOX11 patterns the zeugopod, and HOX13 patterns the autopod. In the earlier developmental stages, their expression gradually overlaps while it is collinear. However, the expression patterns become restricted in the specific regions within the limb as development progresses.

At the presumptive sites of limbs, mesenchymal cells derived from the mesoderm come together in an aggregate structure to form limb buds, a process induced by FGF10 (Sekine et al., 1999). In an *in vitro* study, transforming growth factor beta 1 (TGF- $\beta$ 1) treatment resulted in chondrogenic differentiation in limb mesenchymal cells preceding the condensation, suggesting TGF- $\beta$ 1 may stimulate chondrogenesis during cartilage pattern formation (Leonard et al., 1991). The mesenchymal condensation and initial chondrogenesis are regulated by SOX9 expression and result in the production of a cartilaginous ECM, such as collagen type II and aggrecan core protein (Bi et al., 1999, Akiyama et al., 2002). This undisturbed cartilaginous tissue becomes subdivided into several anlagen, and each segment is strongly correlated with the spatially discrete domains of HOX9-13 expression (Nelson et al., 1996). Between these cartilaginous anlagen—at the future joint sites, cells begin expressing interzone marker genes (GDF5, WNT9A, and ENPP2) and stop expressing cartilaginous genes (Karsenty and Wagner, 2002, Pacifici et al., 2005). A portion of these interzone cells resume chondrogenesis and differentiate into articular chondrocyte, however, the molecular and cellular mechanisms have not been fully understood.

Some of the major cytokines that promote both chondrogenesis and hypertrophy in various cell lines are members of the transforming growth factor beta (TGF- $\beta$ ) family (Goldsmith et al., 2006, Dobaczewski et al., 2011, Wang et al., 2014). While multiple TGF- $\beta$  ligands can bind to multiple TGF- $\beta$  receptors, the downstream events are mediated by various combinations of receptors including, TGF- $\beta$  receptor type I (ALK1, and ALK5), type II (TGFBR2), and type III (TGFBR3; Figure 1.4). Also, these downstream signals can have different effects on limb skeletal development (Wang et al., 2014). The two canonical SMAD-dependent pathways are transduced by TGFBR2 and ALK5 and by TGFBR2 and ALK1. On the other hand, the noncanonical SMAD-independent pathway is transduced by TGFBR3 and ALK5 (Iwata et al., 2012). TGFBR3 not only transduces the noncanonical TGF- $\beta$  pathway, but also facilitates the TGFBR2 and ALK5 mediated canonical TGF- $\beta$  pathway by providing stable ligands (Shi and Massagué, 2003). The canonical TGF- $\beta$  pathway mediated by ALK5 activates SMAD2 and SMAD3, which are transcription factors promoting production of collagen type II and aggrecan core protein. Articular cartilage phenotypes are induced and maintained by SMAD2/3 signaling, which represses RUNX2-inducible MMP13 expression (Chen et al., 2012). Also, the noncanonical TGF- $\beta$  pathway has been reported to interact with SMAD2/3 signaling (Watanabe et al., 2001). Yet, the canonical pathway mediated by ALK1 induces hypertrophic differentiation by activating Smad1/5/8 (Nishida et al., 2013). However, the roles of the noncanonical TGF- $\beta$  pathway in chondrocyte hypertrophy are still unclear (Wang et al., 2014).

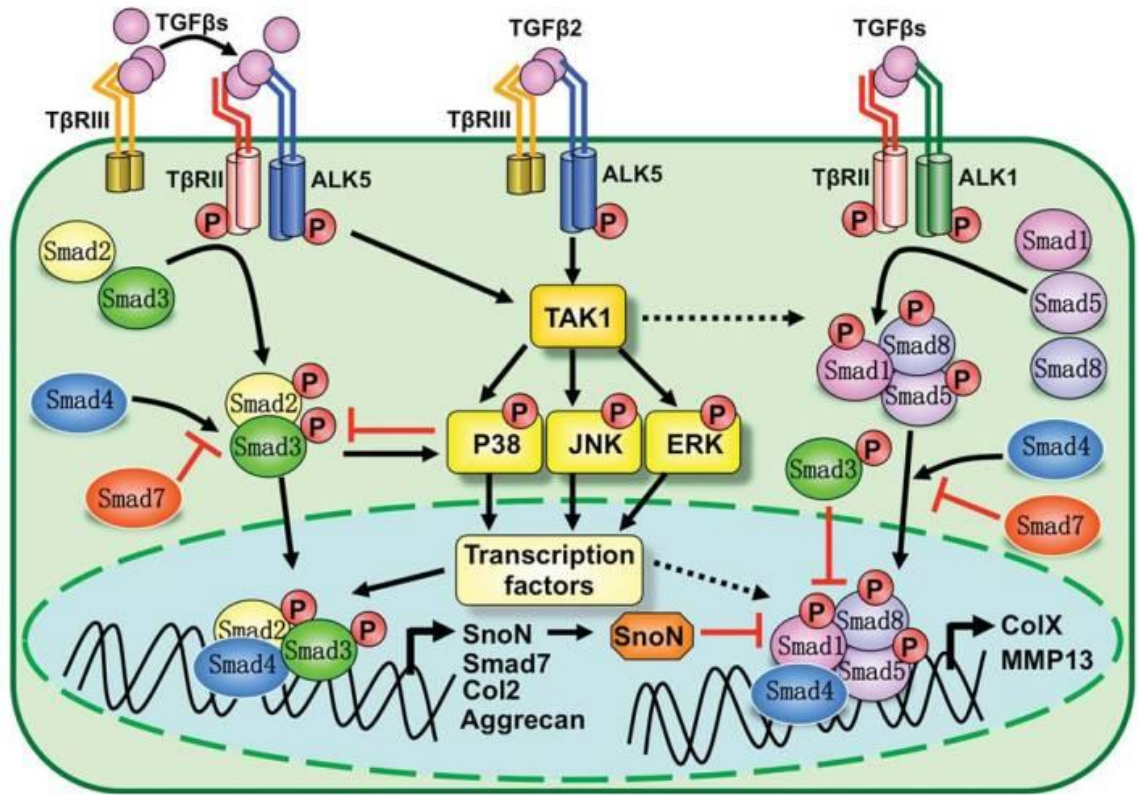


Figure 1.4. TGF- $\beta$  signaling pathways in cartilage formation and maintenance (Copyright: Wang et al., 2014, reproduced with permission).

In summary, various genes and signaling pathways are involved in molecular and cellular mechanisms that regulate the process and sequence of limb skeletal development. Thus, investigating kinetics or interactions of gene expression during these processes should help to elucidate the molecular mechanisms of the binary decision that is made by chondrocytes within a fetal limb bud, in which one population of cells take a non-hypertrophic program to stable articular cartilage, while others progress through terminal hypertrophic differentiation leading to osteogenesis. An important knowledge gap is whether interzone and anlagen cells are intrinsically determined to become articular cartilage and hypertrophic cartilage, respectively. If so, critical questions arise regarding the identity of the molecular regulators and the plasticity in their commitment to these two developmental pathways.

## **Advances in technologies for gene expression evaluation**

As the intricacy of biology have become more revealed, molecular approaches investigating the expression of a greater number of genes with a larger sample numbers have been developed. Today, “high throughput” capabilities have become a priority for studying gene expression leading to a diverse set of powerful technologies.

While conventional reverse transcription-quantitative polymerase chain reaction (RT-qPCR) has enabled sensitive and reproducible measurements in gene expression, this method is quite labor-intensive and expensive, requiring a greater volume of samples and reagents. Overcoming the disadvantages of conventional RT-qPCR systems, microfluidic RT-qPCR technology has allowed gene expression measurements from a number of targeted gene loci with a smaller amount of starting materials (both reagents and samples). By channeling reagents and cDNA samples within a microfluidic chip, thousands of gene expression reactions are simultaneously measured, and the reaction volume is scaled down to a nanoliter scale compared to traditional systems requiring microliters of reaction volumes. The microfluidic RT-qPCR systems work well for experiments investigating a defined set of target transcripts of interest.

On the other hand, gene expression can be assessed at the whole transcriptome level using next generation sequencing platforms. By profiling transcriptomic signatures, RNA-sequencing (RNA-seq) evaluates expression of not only genes that have been conventionally considered biomarkers or biologically relevant to the research subjects, but also genes that

have not been identified in their roles or have received less scientific attention. While traditional, bulk RNA-seq systems measure overall gene expression from all cells existing in a sample, single cell RNA-seq can separately profile the transcriptome from individual cells within a sample. Since single cell RNA-seq is still emerging and is a relatively novel technology, the high expense can be a barrier to broad use. However, single cell methods are rapidly bringing to consideration new and novel opportunities in transcriptome research.

### **Overview of the dissertation**

In this dissertation, chondrogenic divergence between equine fetal interzone and anlagen cell cultures was studied using various, advanced technologies for evaluating gene expression. The overall hypothesis tested in the present research project was that chondrogenic differentiation of interzone and anlagen cells is directed to articular and hypertrophic developmental pathways, respectively. In Chapter 2, using microfluidic RT-qPCR, the expression kinetics of 93 selected genes was evaluated at ten different time points during the 336-hour *in vitro* chondrogenesis. The hypothesis tested in this first study was that chondrogenic divergence between the two fetal skeletal cell lines will become evident within an earlier time frame—within the first 24 hours—after initiating the chondrogenic induction. Then, five time points prioritized based on the data from Chapter 2, and the different chondrogenic pathways between interzone and anlagen cells were further investigated in Chapter 3 using traditional, bulk RNA-seq. This study tested the hypothesis that regulatory genes will differentially respond to the chondrogenic stimulation between the two skeletal cell cultures within the first 1.5 hours, and more distinctive transcriptomic characteristics will

accumulate as time passes during the 96-hour experimental period. In addition, a pilot single cell RNA-seq study is reported in Chapter 4. The hypotheses of this pilot study were that interzone and anlagen cell pellet cultures will develop different levels of heterogeneity in cell subpopulations at 24h and 48h after inducing chondrogenesis, and these fetal cell lines will present cell type-unique traits as well as common chondrogenic characteristics. In closing, Chapter 5 summarizes the studies conducted for this dissertation project and indicates future research directions. Altogether, the new kinetics information on gene expression in interzone and anlagen cell cultures will not only enhance our understanding of these two developmental cell types, but also lay a foundation for future studies investigating novel therapeutic approaches to enhance articular cartilage repair.

## **Chapter 2. Kinetics of gene expression changes in equine fetal interzone and anlagen cells during 14 days of *in vitro* chondrogenesis**

### **Introduction**

During the early stages of limb formation, articular cartilage develops from interzone located at the presumptive sites of future synovial joints within the cartilaginous anlagen of bones. Therefore, two different types of cartilage differentiate in close proximity. While anlagen cartilage is transient, progressing through endochondral ossification to form bones, articular cartilage remains stable and functions throughout life to facilitate biomechanical load distribution and low friction movement between adjoining bone surfaces. Despite the important functional properties of articular cartilage, its intrinsic ability to restore structural defects is limited in mature mammals (Alford and Cole, 2005). Almost the polar opposite is true regarding the potential for bone tissue regeneration. Fractured bones repair quite well by recapitulating endochondral ossification; provided the fracture ends are brought together, stabilized, and not compromised by infection or loss of blood supply. Thus, research on the comparative cell biology between interzone and anlagen cells, as well as their developmental chondrogenic pathways may provide novel information relevant to improving mammalian articular cartilage regenerative treatments.

In an effort to understand the biology of fetal interzone and anlagen cells, their chondrogenic potential was measured and compared after 21 days in three dimensional pellet culture and continuous stimulation with a chondrogenic induction medium containing transforming growth factor beta 1 (TGF- $\beta$ 1; Adam et al., 2019). The results demonstrated that

interzone and anlagen cultures respond differently to chondrogenic stimulation based on expression of cartilaginous marker genes, such as aggrecan core protein (ACAN) and collagen type II alpha 1 chain (COL2A1). Also, the cell pellets showed distinguishing histological characteristics, including proteoglycan amount and distribution, as well as cellular morphology and arrangement. In other research studies, the protocol for *in vitro* chondrogenic differentiation has also been reported after 21 days, with the expression of extracellular matrix (ECM) genes measured at the mRNA or protein level used as targeted functional outcomes (McCarthy et al., 2012, Rakic et al., 2018). However, gene expression changes induced by TGF- $\beta$ , a well-established chondrogenic factor, start as early as 30 minutes to 1 hour after treating TGF- $\beta$  in culture (Franco et al., 2010, Aomatsu et al., 2011). Molecular details of these early responses are not fully understood, so an important gap of knowledge is whether there are qualitative or quantitative differences of gene expression kinetics over time in these two chondrogenic cell cultures.

By comparing a timed sequence of the cellular response to TGF- $\beta$ 1 induced chondrogenesis between interzone and anlagen cell cultures, the present study was designed to answer the following questions: 1) how do gene expression patterns change over time and 2) when do the differential pathways of the two cell types start to diverge in this *in vitro* chondrogenesis model? The hypothesis tested in this study is that divergent chondrogenic differentiation pathways in interzone and anlagen cultures will be evident within the first 24 hours after *in vitro* chondrogenic induction.



## Materials and Methods

### *Cell culture and sample collection*

Equine fetal interzone cells, anlagen cells, and dermal fibroblasts (a negative control) were previously harvested from seven 45-days-old fetuses (Adam et al., 2019), and the cells were frozen and stored at passage 2 (P2). Using standard protocols, frozen cells were thawed and cultured in T-75 polystyrene flasks with high glucose Dulbecco's modified Eagles medium (DMEM; cat No. 10569044; Gibco) supplemented by 10% (v/v) fetal bovine serum (heat inactivated; cat No. S11150H; Atlanta Biologicals) and 1% (v/v) penicillin/streptomycin (P/S; cat No. 15070063; Gibco) and termed 'complete medium' in this study. When the cell monolayers reached approximately 80% confluence, the adherent cells were lifted by 0.25% Trypsin-EDTA solution (cat No. 25200056; Gibco) and split into new flasks (seeding density of 10,000 cells/cm<sup>2</sup>). Cell viability in the suspension of >95% was confirmed by trypan blue dye exclusion test.

When P4 monolayers reached >80% confluence, a portion of the cells were used to collect total RNA. The cells were harvested in a guanidinium thiocyanate solution (1ml/T-75 flask; QIAzol Lysis Reagent; cat No. 79306; Qiagen), immediately snap-frozen, and stored at -80°C until total RNA isolation. The rest of P4 monolayers were lifted by trypsin digestion, and chondrogenic cell pellets were established at P5 as previously described (Adam et al., 2019). Each cell pellet was comprised of 500,000 viable cells and maintained in chondrogenic induction medium (high glucose DMEM + 1% P/S + bovine serum albumin, 12.5 mg/ml + ascorbic-2-phosphate, 50 µg/ml + TGF-β1 (Transforming Growth Factor-β1 human; cat No. T7039-50UG; Sigma-Aldrich), 10 ng/ml + 1% insulin-transferrin-selenium-sodium pyruvate +

Dexamethasone, 100 nM + 1% nonessential amino acid) for the full culture period.

Aliquots of pellet cultures were collected at ten different time points: baseline (0h), 1.5, 3, 6, 12, 24, 48, 96, 168, and 336 hours after the initiation of the chondrogenic induction. These time points were selected based on a literature review (Ranganathan et al., 2007, Rudini et al., 2008, Franco et al., 2010, Aomatsu et al., 2011, Koyama et al., 2013, Nejadnik et al., 2015, Yokota et al., 2014, Yamazaki et al., 2015) to assess the kinetics of gene expression over time after the TGF- $\beta$ 1 treatment in the cell pellet cultures. At each time point, the collected pellets were washed with phosphate-buffered saline (PBS), snap-frozen (3 pellets/vial), and stored at  $-80^{\circ}\text{C}$  until total RNA isolation.

#### ***Total RNA isolation***

Thawed cell monolayers and pellets (3 pellets per 1 ml of the guanidinium thiocyanate solution) were homogenized using a PowerGen homogenizer (Model 125; Fisher Scientific). Total RNA was then extracted and purified using a spin-column based RNeasy Mini kit (cat No. 74106; Qiagen) followed by ethanol precipitation. The quantity of RNA was determined using a Qubit™ RNA Broad Range Assay Kit (cat No. Q10211; Life Technologies) with a Qubit® 3.0 Fluorometer (cat No. Q33216; Life Technologies), and the 260/280 and 260/230 ratios were measured by a Nano Drop ND 1000 spectrophotometer (Thermo Fisher Scientific). Finally, RNA integrity numbers (RIN) were determined using a Bioanalyzer 2100 (Agilent Technologies) with an Agilent RNA 6000 Pico kit (cat No. 5067-1513; Agilent Technologies) to assess the quality of the RNA samples. The monolayer RNA samples resulted in 260/280 ratios of 2.0 – 2.1, 260/230 ratios of 2.3 – 2.7, and RINs of 8.7 – 10, except for one sample that showed a RIN

of 6.8. A substantial majority of cell pellet RNA samples had 260/280 ratios of 1.8 – 2.1, 260/230 ratios of 1.8 – 2.7, and RINs of 6.4 – 10. Out of 210 cell pellet RNA samples, ten samples had 260/230 ratios outside of this range, but gene expression patterns were consistent with experimental group averages so the data were retained. Any potential genomic DNA contamination was removed with dsDNase (cat No. K1672; Thermo Fisher Scientific) during reverse-transcription protocol using the Maxima First Strand cDNA Synthesis Kit for RT-qPCR (cat No. K1672; Thermo Fisher Scientific). The cDNA samples were then stored at –20°C pending qPCR analysis.

### ***Quantitative polymerase chain reaction (qPCR)***

#### **1) Targeted gene loci**

Three prospective endogenous control genes, B2M, GUSB, and RPLP0 (Mienaltowski et al., 2008) were evaluated in a preliminary RT-qPCR analysis conducted on a subset (n=2) of the entire sample set. Ninety-three genes (Table 2.1) of interest were selected for analysis based on equine cartilaginous tissue RNA-seq data generated in the MacLeod lab (Adam et al., in preparation) and a literature review. The earlier data or published reports from these genes were either 1) differentially expressed between interzone and anlagen tissue samples at three developmental ages (day-45 fetuses, day-60 fetuses and neonatal foals), 2) functionally annotated to include fetal developmental processes, 3) known to be involved in chondrogenic differentiation or 4) established components or regulators of TGF- $\beta$  signaling pathways.

These 96 gene loci were studied with commercially available (59 assays) and customized (37 assays) equine-specific TaqMan® primer-probe sets (Thermo Fisher Scientific; Table 2.1). Where possible, the primer-probe sets were designed to span two exons (88 assays), with eight assays designed within a single exon.

Table 2.1. The panel of 96 equine specific TaqMan® primer-probe sets used in the RT-qPCR analysis

Gene symbol	Gene name	ThermoFisher cat ID <sup>a</sup>	Predicted amplicon length, nt	Spanning exon design	EquCab 3.0 amplicon coordinates
ABCC9	ATP Binding Cassette Subfamily C Member 9	ARZTE9G	89	Spanning 2 exons	6:49224101-49222665
ABI3BP	ABI Family Member 3 Binding Protein	Ec06625599_m1	57	Spanning 2 exons	19:57461128-57461991
ADAMT5	ADAM Metalloproteinase with Thrombospondin Type 1 Motif 5	Ec03470666_m1	73	Spanning 2 exons	26:25148986-25137605
ADGRG1	Adhesion G Protein-Coupled Receptor G1	ARPRK7Z	75	Spanning 2 exons	3:10541647-10542471
ADGRG2	Adhesion G Protein-Coupled Receptor G2	ARU63XT	93	Spanning 2 exons	X:14858813-14858548
ALPK3	Alpha Kinase 3	Ec07042890_g1	60	Spanning 2 exons	1:93211011-93210198
ALPL	Alkaline Phosphatase, Biom mineralization Associated	ARYMKNG	120	Spanning 2 exons	2:33121046-33118823
ANGPTL4	Angiopoietin Like 4	Ec06997549_m1	62	Spanning 2 exons	7:6065657-6066055
APLNR	Apelin Receptor	Ec07019415_s1	82	Single exon	12:20444024-20443928
AQP1	Aquaporin 1 (Colton Blood Group)	AR2W9UC	82	Spanning 2 exons	4:61601355-61601398
ARHGEF15	Rho Guanine Nucleotide Exchange Factor 15	Ec07058476_g1	68	Spanning 2 exons	11:51540160-51540207
ASS1	Argininosuccinate Synthase 1	Ec06982992_g1	66	Spanning 2 exons	25:33865022-33867929
B2M <sup>b</sup>	Beta-2-Microglobulin	Ec03468699_m1	70	Spanning 2 exons	1:145964634-145961332
BMP2	Bone Morphogenetic Protein 2	ARRWFTX	55	Spanning 2 exons	22:16917593-16910100
BMPRI1A	Bone Morphogenetic Protein Receptor Type 1A	ARWCXHP	88	Spanning 2 exons	1:84431108-84427100
BOC	Boc Cell Adhesion Associated, Oncogene Regulated	ARZTE9E	120	Spanning 2 exons	19:47322846-47322396
CDH13	Cadherin 13	Ec03469102_m1	58	Spanning 2 exons	3:31762619-31825090

CDON	Cell Adhesion Associated, Oncogene Regulated	ARWCXHM	85	Spanning 2 exons	7:36010465-36005823
CHODL	Chondrolectin	Ec06984006_m1	66	Spanning 2 exons	26:17451675-17451892
CLU	Clusterin	Ec03468570_m1	133	Spanning 2 exons	2:56559215-56557872
COL10A1	Collagen Type X Alpha 1 Chain	ARXGR3J	73	Spanning 2 exons	10:65872577-65868970
COL1A1	Collagen Type I Alpha 1 Chain	Ec03469676_m1	154	Spanning 2 exons	11:26002328-26002234
COL2A1	Collagen Type II Alpha 1 Chain	Ec03467411_m1	81	Spanning 2 exons	6:66519131-66518656
COL5A3	Collagen Type V Alpha 3 Chain	Ec06999559_g1	56	Spanning 2 exons	7:51679496-51679953
COMP	Cartilage Oligomeric Matrix Protein	Ec03468072_m1	111	Spanning 2 exons	21:3988105-3988046
CREB5	Camp Responsive Element Binding Protein 5	AR323D9	77	Spanning 2 exons	4:59761418-59766406
CTGF	Connective Tissue Growth Factor	ARDJYMK	105	Spanning 2 exons	10:79344899-79344651
DCN	Decorin	Ec03468474_m1	102	Spanning 2 exons	28:18029682-18025682
DIO2	Iodothyronine Deiodinase Type II	Ec04320470_m1	86	Spanning 2 exons	24:25082308-25074448
DLX5	Distal-Less Homeobox 5	ARCE42N	147	Spanning 2 exons	4:40255725-40254648
ENTPD1	Ectonucleoside Triphosphate Diphosphohydrolase 1	Ec07040532_m1	81	Spanning 2 exons	1:33474308-33473045
ENTPD2	Ectonucleoside Triphosphate Diphosphohydrolase 2	Ec06983692_g1	62	Spanning 2 exons	25:39221285-39220720
FAM132A	C1q and TNF Related 12	AR323EC	90	Spanning 2 exons	2:48636431-48636589
FAM20A	Golgi Associated Secretory Pathway Pseudokinase	Ec07054339_m1	59	Spanning 2 exons	11:12154761-12155583
FGF1	Fibroblast Growth Factor 1	Ec01092738_m1	104	Spanning 2 exons	14:33844480-33856602
FGF18	Fibroblast Growth Factor 18	Ec03248217_g1	59	Spanning 2 exons	14:9070314-9056898
FGFR3	Fibroblast Growth Factor Receptor 3	Ec03470545_m1	118	Spanning 2 exons	3:120236405-120235774

FRZB	Frizzled Related Protein	ARYMKNJ	79	Spanning 2 exons	18:60567720 - 60563791
FZD1	Frizzled Class Receptor 1	ARFVMTG	82	Single exon	4:35193201-35193296
GALNT14	Polypeptide N-Acetyl Galactosaminyl Transferase 14	Ec06950408_m1	62	Spanning 2 exons	15:66761932-66787897
GDF5	Growth Differentiation Factor 5	Ec04321108_s1	89	Single exon	22:27084232-27084125
GDF6	Growth Differentiation Factor 6	Ec07097112_m1	102	Spanning 2 exons	9:44216931-44202151
GLI3	GLI Family Zinc Finger 3	ARFVMTF	128	Spanning 2 exons	4:13054722-13053538
GUSB <sup>b</sup>	Glucuronidase Beta	Ec03470630_m1	73	Spanning 2 exons	13:18879647-18882340
IBSP	Integrin Binding Sialoprotein	ARGZGDD	96	Spanning 2 exons	3:51431941-51431805
IGF2	Insulin-Like Growth Factor 2	Ec03469397_m1	158	Spanning 2 exons	12:34442429-34440361
IGFBP5	Insulin-Like Growth Factor Binding Protein 5	Ec03470296_m1	62	Spanning 2 exons	6:6400371-6386883
IGFBP7	Insulin-Like Growth Factor Binding Protein 7	Ec03469608_m1	85	Spanning 2 exons	3:77681966-77682663
IHH	Indian Hedgehog	Ec03470108_m1	61	Spanning 2 exons	6:8392633-8390869
ITGA7	Integrin Subunit Alpha 7	Ec06982346_m1	96	Spanning 2 exons	6:74384753-74382727
ITGAV	Integrin Subunit Alpha V	Ec03469608_m1	125	Spanning 2 exons	18:63632097-63636322
KCNJ8	Potassium Inwardly Rectifying Channel Subfamily J Member 8	AR7DRH4	94	Spanning 2 exons	6:49196363-49192450
LEF1	Lymphoid Enhancer Binding Factor	AR7DRH3	90	Spanning 2 exons	2:117834151-117834210
LOC100630171	Regakine-1	Ec07014483_s1	86	Single exon	11:37560494-37560588
MASP1	Mannan Binding Lectin Serine Peptidase 1	Ec06960466_m1	73	Spanning 2 exons	19:27583128-27580854
MET	MET Proto-Oncogene, Receptor Tyrosine Kinase	Ec02622441_m1	68	Spanning 2 exons	4:74113698-74115265
MGP	N-Methylpurine DNA Glycosylase	ARKA4G9	78	Spanning 2 exons	6:43505581-43504668

MMP2	Matrix Metalloproteinase 2	Ec03469994_m1	71	Spanning 2 exons	3:8344084-8345058
NEFL	Neurofilament Light	Ec06966469_m1	76	Spanning 2 exons	2:53936630-53936597
NPY	Neuropeptide Y	Ec06946514_m1	75	Spanning 2 exons	4:55899498-55901429
NTRK2	Neurotrophic Receptor Tyrosine Kinase 2	Ec07025737_s1	104	Single exon	23:5298376 -5298260
OMD	Osteomodulin	AR47WX6	118	Spanning 2 exons	23:55155149-55155748
OSR2	Odd-Skipped Related Transcription Factor 2	Ec07007011_m1	128	Spanning 2 exons	9:46596059 -46596942
PANX3	Pannexin 3	ARAACGR	134	Spanning 2 exons	7:34195152-34197741
PDLIM1	PDZ and LIM Domain 1	Ec07040588_m1	65	Spanning 2 exons	1:33957840-33958350
PLAT	Plasminogen Activator, Tissue Type	Ec06985220_g1	58	Spanning 2 exons	27:3277326-3278021
PLVAP	Plasmalemma Vesicle Associated Protein	Ec06971069_m1	101	Spanning 2 exons	21:2965701-2963843
PRKG2	Protein Kinase cGMP-Dependent 2	AR9HJ3Z	90	Spanning 2 exons	3:57147313-57150172
PTCH2	Patched 2	AREPT7H	61	Spanning 2 exons	2:13639004-13639168
RELN	Reelin	ARH6AXC	73	Spanning 2 exons	4:4288013-4287009
RET	Ret Proto-Oncogene	Ec03468172_m1	130	Spanning 2 exons	13:42671164-42670174
RPLP0 <sup>b</sup>	Ribosomal Protein Lateral Stalk Subunit P0	Ec04947733_g1	74	Spanning 2 exons	8:16270469-16271594
RUNX2	RUNX Family Transcription Factor 2	Ec03469741_m1	65	Spanning 2 exons	20:45071187-45090120
RUNX3	RUNX Family Transcription Factor 3	AR9HJ33	85	Spanning 2 exons	2:30378587-30389531
S100A1	S100 Calcium Binding Protein A1	ARU633W	107	Spanning 2 exons	5:40513357-40512077
S100A4	S100 Calcium Binding Protein A4	Ec07038302_m1	105	Spanning 2 exons	5:40581137-40582018
S1PR3	Sphingosine-1-Phosphate Receptor 3	AR9HJ32	57	Single exon	23:53100775-53100700
SERPINE1	Serpin Family E Member 1	Ec03469902_m1	56	Spanning 2 exons	13:9392324-9393407
SGMS2	Sphingomyelin	AR47WX9	74	Spanning 2 exons	2:117974497-117969273
SHC3	SHC Adaptor Protein 3	Ec06977865_m1	58	Spanning 2	23:53055119



				exons	-53062725
SLC38A1	Solute Carrier Family 38 Member 1	Ec06973498_m1	62	Spanning 2 exons	6:65158674-65157159
SMPD3	Sphingomyelin Phosphodiesterase 3	ARNKTM3	107	Spanning 2 exons	3:19318165-19316482
SNAI1	Snail Family Transcriptional Repressor 1	AR47WX7	139	Spanning 2 exons	22:39029820-39033420
SNAI2	Snail Family Transcriptional Repressor 2	ART2ADV	87	Spanning 2 exons	9:35268124-35269118
SP7	Sp7 Transcription Factor	AREPT7J	138	Single exon	6:71198223-71198072
SPARCL1	SPARC Like 1	Ec06992392_m1	73	Spanning 2 exons	3:51668155-51667472
STAB1	Stabilin 1	Ec06952812_g1	55	Spanning 2 exons	16:36722312-36721809
TGFBI	Transforming Growth Factor Beta Induced	ARMTX26	120	Spanning 2 exons	14:39336075-39335345
THBS4	Thrombospondin 4	Ec06947284_g1	58	Spanning 2 exons	14:86001213-86000560
TIMP2	TIMP Metalloproteinase Inhibitor 2	Ec03470558_m1	72	Spanning 2 exons	11:3846324-3848466
TLR2	Toll-Like Receptor 2	Ec03818334_s1	87	Single exon	2:80314660-80314562
TLR4	Toll-Like Receptor 4	Ec03468994_m1	99	Spanning 2 exons	25:22400336-22403588
TNFRSF21	TNF Receptor Superfamily Member 21	Ec06970391_m1	63	Spanning 2 exons	20:46704716-46683406
TNFSF11	TNF Superfamily Member 11	ARAACGT	71	Spanning 2 exons	17:27570379-27570216
TSPAN15	Tetraspanin 15	Ec07041353_m1	78	Spanning 2 exons	1:58097361-58125779
WNT9A	Wnt Family Member 9A	ARDJYMM	81	Spanning 2 exons	14:94400510-94401267

<sup>a</sup>Commercially available primer-probe sets. ThermoFisher catalogue IDs start with 'Ec' and their catalogue number is 4448892. Catalogue IDs for custom designed primer-probe sets start with 'A' and their catalogue number is 4441114.

<sup>b</sup>Three putative endogenous control genes

## 2) Positive control RT-qPCR assessments

Prior to conducting the main microfluidic RT-qPCR analysis, a preliminary assessment was conducted using a robotic ViiA™ 7 Real-Time PCR System (Thermo Fisher Scientific) in order to verify 1) amplification of two endogenous controls (GUSB and RPLP0) in all 231 cDNA samples (7 biological replicates × 3 cell lines × 11 time points, plus P4 monolayer samples) and 2) amplification of all 96 targeted gene loci by the TaqMan® primer-probe sets in a positive control sample. The positive control sample was prepared by pooling equal parts of 1) a pooled total RNA sample composed of 43 different equine tissue/cell sources (Hestand et al., 2015), and 2) a 35-days-old equine fetus homogenate. In these test analyses, negative controls included a no-reverse transcription sample and a no-template sample to confirm the absence of contaminating genomic DNA or RNA in individual samples or the system. The entire 231 samples (10 ng of cDNA/reaction) expressed GUSB (cycle threshold (Ct) values of  $21.44 \pm 0.06$ ) and RPLP0 (Ct values of  $18.21 \pm 0.06$ ). In addition, the positive control sample demonstrated amplification in 95 of the targeted gene loci (Ct values of 17.68 – 30.85). The one exception was primers specific for COL10A1 (Ct value of 35.92).

## 3) Microfluidic RT-qPCR

The 231 cDNA samples were prepared at a concentration of 10 ng/ul and divided into three 96 well-microfluidic chips (96 × 96 Fluidigm Dynamic Array; Fluidigm). Seven 3-fold dilution series (125, 41.67, 13.89, 4.63, 1.54, 0.51, and 0.17 ng/ul) of the positive control sample were added onto each chip to evaluate PCR efficiency and to also function as an inter-plate control. The plates were then shipped to the Roy J. Carver Biotechnology Center, a genomics core at the University of Illinois at Urbana-Champaign (Urbana, IL, USA). To

quantitate steady state mRNA levels for the targeted 96 gene loci in all experimental samples, a microfluidic RT-qPCR system (Biomark HD high throughput amplification system, Fluidigm) was utilized and operated with manufacturer-recommended protocols (Fluidigm Corporation, 2018). After 14-cycles of pre-amplification, steady state mRNA levels were measured with this microfluidic RT-qPCR system and the data processed using the Fluidigm Real-Time PCR Analysis software.

### ***Data analyses and statistics***

The two most stable endogenous controls (GUSB and RPLP0) across the sample set (Figure 2.1) were used for the gene expression normalization within a sample ( $\Delta Ct = Ct$  of a gene of interest – average  $Ct$  of GUSB and RPLP0). Then, the  $\Delta Ct$  of each target gene was calibrated with  $\Delta Ct$  of the same target gene in the positive control sample ( $\Delta\Delta Ct = \Delta Ct$  in a sample –  $\Delta Ct$  in the positive control). Finally,  $\Delta\Delta Ct$  values were converted to relative quantity ( $RQ = 2^{-\Delta\Delta Ct}$ ; fold changes based on expression in the positive control; Livak and Schmittgen, 2001). To determine statistical differences between the data points (targeted gene  $\times$  cell type  $\times$  time point), the fold change data were log-transformed and analyzed using SAS statistical software, version 9.3 (SAS Institute Inc., Cary, NC). One-way multivariate analysis of variance was conducted with Tukey's honest significance test for multiple comparison adjustments. The significance threshold was defined as  $P < 0.05$ .

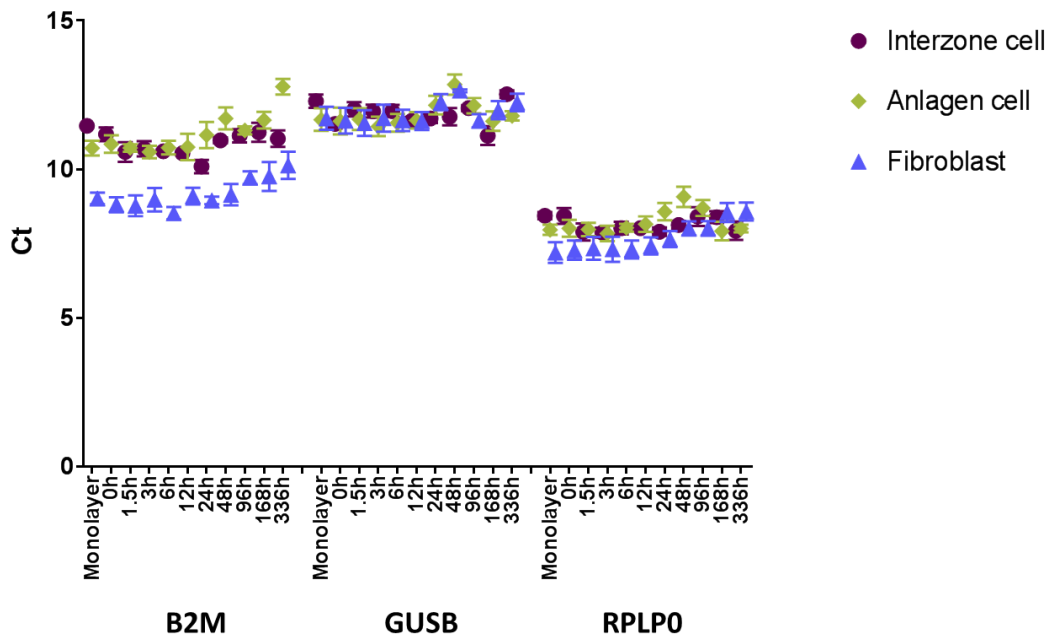


Figure 2.1. Steady state mRNA levels (Ct, mean  $\pm$  SEM) of three prospective endogenous controls (B2M, GUSB, and RPLP0) across the 336-hour experimental period. GUSB and RPLP0 displayed more stability across the three cell types and time course.

## Results

### *Genes processed in data analyses*

Of the 93 targeted genes of interest, six loci were not processed for further data analyses; five genes (ARHGEF15, CHODL, NPY, RET, and STAB1) had little or no relative expression based on the control sample (RQs of <0.02), and IGF2 had low fluorescent intensity in two out of three microfluidic chips which resulted in loss of 5 – 6 biological replicates. Thus, data from 87 genes were analyzed further.

These 87 genes were categorized into three groups based on their established functional annotation: 1) 15 genes regulating transcription, 2) 51 genes involved in signal transduction, and 3) 23 genes involved in ECM biology. Two genes (MASP1 and NEFL) were not categorized into any of the three annotation groups. Four genes were included in two of the three annotation groups; ENTPD1, ENTPD2, and LEF1 in both the transcription and signaling groups, and THBS4 in both the signaling and ECM groups.

### *Kinetics of gene expression changes by cell type*

Time point differences in steady state levels of mRNA for individual gene loci were assessed by comparing values to baseline at 0h within a cell type. Significant upregulation and downregulation events were noted (Figure 2.2). At the first time point, 1.5 hours, the significant changes observed were all upregulation. The time points from 3 hours on had instances of both upregulation and downregulation. Among the total of 261 gene × cell type combinations (87 genes × 3 cell types), 110 combinations showed upregulation (Figure 2.2.a), 83 combinations showed downregulation (Figure 2.2.b), 22 combinations showed mixed

patterns of upregulation and downregulation (Figure 2.2.c). Steady state mRNA levels were stable across all time points in 46 cell type x gene loci combinations (Figure 2.1.d).

Figure 2.2.a) Upregulated genes

(Time point, h)

Genes	Cell type	1.5	3	6	12	24	48	96	168	336
SNAI1	FB	***	***	***	***	***	***			
IHH	ANL	***	***	***	***	***	***	***	***	***
IHH	FB	*	**			**	***	***	***	***
S1PR3	ANL	**	***	***	**	***	***	***	***	***
S1PR3	FB	**	***	***	***	***	***	***	***	***
ALPK3	IZ	***	***							
RUNX3	IZ	***	**							
RUNX3	ANL	***								
RUNX3	FB	***	***							
ANGPTL4	IZ	***	***	***	***	***	***	*		
ANGPTL4	FB	***	***	***	***	***	***	***	***	***
AQP1	IZ	*	**	**	**	**	**	**	**	**
AQP1	ANL	***	***	***	***	***	***	***	***	***
AQP1	FB	**	***	***	***	***	***	***	***	***
FAM132A	IZ	*	**	**	**	**	**	**	**	**
FGF1	ANL	**	***	***	***	***	***	***	***	***
IHH	IZ	*	***	***	***	***	***	***	***	***
S1PR3	IZ	***	**	*	***	***	***	***	***	***
WNT9A	IZ	**	***							
WNT9A	FB	*	**	***	**	*	***	***	***	***
COL5A3	IZ	***	***	***	***	***	***	***	***	***
COL5A3	ANL	***	***	***	***	***	**			
COL5A3	FB	***	***	***	***	***	***	***	***	***
COMP	ANL	*	***	***	***	***	***	***	***	***
SGMS2	IZ	***								
SGMS2	ANL	*						**	*	
THBS4	IZ	*	***	***	***	***	***	***	***	***
THBS4	ANL	***	***	***	***	***	***	***	***	***
THBS4	FB	**	***	***	***	***	***	***	***	***
ASS1	FB	***	***	***	***	***	*	*	*	*
LEF1	IZ	**			*	***	**	***	**	**
BMPRI1A	ANL		**	***						
BOC	ANL	**	**	**			**	**	**	**
CDON	IZ	*	*	***						
CDON	ANL	***	***	**	**		***	***	***	***
FAM132A	ANL		*	***	***	***	*	***	**	**
GDF5	ANL	**	***	***	***	***	***	***	***	***
KCNJ8	ANL	***	***	***	***	***	***			
LOC100630171	ANL	*								
LOC100630171	FB	**						**		
PRKG2	FB	***	***	***	***	***	***	***	***	***
ALPL	FB	***	***	***	***	**	**	***	***	***
OMD	ANL	***	***	***	***	***	***	***	***	***
OMD	FB	**	***	***	***	***	***	***	***	***
RELN	ANL	*	*	***	***	***	***	***	***	***
TGFBI	IZ	***	***	***	***	***	***	***	***	***
ASS1	IZ			***	***	***		*	**	
ASS1	ANL	***	***	***	***	***	***	***	***	***
ENTPD2	ANL			*	**	***	*	***	***	***
ENTPD2	FB	***	***	***	***	***	***	***	***	***
RUNX2	IZ	**	***	***	***	***	***	***	***	***
RUNX2	ANL	**	***	***	***	***	***	***	***	***
RUNX2	FB	***	***	***	***	***	***	***	***	***
BOC	FB	*	**	***	***	***	***	*		
GDF5	IZ	***	***	***	***	***	***	***	***	***
ALPL	IZ	*	*							
ALPL	ANL	**	**	*						
DCN	IZ	***	***	***	***	***	***	***	***	***
DCN	ANL	*	***	***	***	***	***	***	***	***
FAM20A	ANL	**	***	***	***	***	*	*	*	*
MMP2	ANL	*	***	***	***	***	***	***	***	***
OMD	IZ	***	***	***	***	***	***	***	***	***
SMPD3	ANL	***	***	***	***	***	***	***	***	***

(continued Figure 2.2.a)

(Time point, h)

Genes	Cell type	1.5	3	6	12	24	48	96	168	336
ENTPD2	IZ					**	**			
LEF1	ANL					*	***	***		
SP7	IZ					**	**	***	**	**
SP7	ANL					**	***	***	***	***
BOC	IZ					**				
S100A1	IZ					***	***	***	***	***
S100A1	ANL					***	***	***	***	***
S100A1	FB					**	***	***	***	***
TNFSF11	FB					***	***	***	***	***
COL2A1	ANL					*	***	***	***	***
COMP	IZ					***	***	***	***	***
COMP	FB					***	***	***	***	***
MMP2	IZ					**	***	***	**	*
MMP2	FB					***	***	***	***	***
SPARCL1	IZ					**	***	***	***	***
SPARCL1	FB					***	***	***	***	***
ABCC9	ANL						**	**		
BMP2	ANL						**	***	**	***
FGF1	IZ						**	*		
FGF1	FB					***				
ITGAV	FB					***	**			
PLVAP	FB					***	***	**		
PRKG2	ANL					***	***	***	***	***
TNFSF11	ANL					***	***	**		
MGP	ANL					*	***	***	***	***
RELN	IZ					***	***	***	***	***
SPARCL1	ANL					*	***	***	***	***
APLNR	FB							**	***	***
BMP2	IZ							*		
FGFR3	IZ							***	***	***
FGFR3	ANL							***	***	***
FRZB	ANL							***	***	***
GDF5	FB							**	***	***
IBSP	FB							**	*	
PANX3	ANL							**	***	**
PLVAP	IZ							**	***	**
DCN	FB							*	**	**
FAM20A	FB							*	*	*
SMPD3	FB							*	***	***
TLR2	FB								***	***
COL10A1	ANL								**	*
FAM20A	IZ								***	*
RELN	FB								**	***
SMPD3	IZ								**	***
BMP2	FB									**
NTRK2	FB									**
COL2A1	FB									**

Annotation group
Transcription regulation
Signaling transduction
Extracellular matrix
Not categorized

Upregulated based on 0h	***, P<0.0001
	** , P<0.01
	* , P<0.05

Figure 2.2.b) Downregulated genes

Genes	Cell type	(Time point, h)																			
		1.5	3	6	12	24	48	96	168	336											
OSR2	IZ		*					*													
OSR2	ANL		***	***	*	***	***														
OSR2	FB		***	**			**	**			**	***									
SNAI2	FB		***	***																	
CTGF	ANL		***	***	***	***	***	***	**	**	***										
CTGF	IZ		***	***	***	***	***	***													
CTGF	FB		**	***	***	***	***	***	***	**	**	**	**	**	**	**	**	**	**	**	**
DIO2	ANL		**	*	***																
GLI3	IZ			***	*															*	
GLI3	ANL			***		*															***
GLI3	FB			***		**	*			**	***										***
CREB5	IZ			**	**					***	***	***	***	***	***	***	***	***	***	***	***
CREB5	ANL			**						***	***	***	***	***	***	***	***	***	***	***	***
ADGRG2	IZ			**	***	***	***	***	*	***											
GDF6	IZ			***				*													*
GDF6	ANL			***	**	***	**	**	***	***	***	***	***	***	***	***	***	***	***	***	***
GDF6	FB			***	**	***	***	*													*
MET	IZ			**	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
MET	ANL			**	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
MET	FB			***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
PTCH2	IZ			***	***	***	***	***	***	*											
PTCH2	ANL			***	***	***	***	***	***												**
PTCH2	FB			***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
SLC38A1	IZ			*	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
SLC38A1	ANL			*	***	***	*	***													
SLC38A1	FB			***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
TLR4	FB			**	**																***
TNFRSF21	ANL			*	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
ADAMT5	FB			*	*																
PDLIM1	IZ				***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
ENTPD1	IZ				***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
ENTPD1	ANL				**	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
ENTPD1	FB				**	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
ADGRG2	ANL				**	***	***	**													
TLR4	ANL				*	**															***
SHC3	ANL				*	**															***
FGF18	ANL				**					**	***	***	***	***	***	***	***	***	***	***	***
ADAMT5	IZ				***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
ADAMT5	ANL				**	***	**	***	***	***	***	***	***	***	***	***	***	***	***	***	***
MASP1	IZ				**	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
MASP1	ANL				**	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
MASP1	FB				**	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
PDLIM1	ANL				***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***
IGFBP7	FB						*	***	***	***	***	***	***	***	***	***	***	***	***	***	***
SERPINE1	IZ						**	***	***	***	***	***	***	***	***	***	***	***	***	***	***
TLR4	IZ						**														
TSPAN15	ANL						**	***	***	***	***	***	***	***	***	***	***	***	***	***	***

(continued Figure 2.2.b)

Genes	Cell type	(Time point, h)																			
		1.5	3	6	12	24	48	96	168	336											
PDLIM1	FB										***	***	***	***	***	***	***	***	***	***	***
FRZB	FB											***	**	***	***	**					
IGFBP7	IZ											**	***	**							
IGFBP7	ANL											***	*								
PLAT	IZ											*	***	***	***	***	***	***	***	***	***
SERPINE1	ANL											**	***	***	***	***	***	***	***	***	***
TSPAN15	IZ											***	***	***	***	***	***	***	***	***	***
ITGA7	IZ											**	***	**	***	***	***	***	***	***	***
ITGA7	FB											**	***	**	***	***	***	***	***	***	***
NEFL	IZ											***	***	***	***	***	***	***	***	***	***
CREB5	FB																				***
ADGRG1	FB																				***
CDH13	ANL																				***
CDH13	FB																				***
IGFBP5	IZ																				**
PLAT	ANL																				***
SERPINE1	FB																				***
ABI3BP	IZ																				***
ITGA7	ANL																			*	**
TIMP2	ANL																				**
TIMP2	FB																			**	*
NEFL	ANL																				***
CDH13	IZ																				***
IGFBP5	FB																				***
S100A4	ANL																				***
FZD1	IZ																				***
FZD1	ANL																				***
ITGAV	IZ																				*
PLAT	FB																				**
S100A4	IZ																				***
S100A4	FB																				**
TNFRSF21	FB																				**
TSPAN15	FB																				**
CLU	ANL																				***
CLU	FB																				***
COL1A1	ANL																				***

Annotation group
Transcription regulation
Signaling transduction
Extracellular matrix
Not categorized

Downregulated based on 0h	***, P<0.0001
	** , P<0.01
	* , P<0.05



Figure 2.2.c) Genes with mixed upregulation and downregulation patterns

Genes	Cell type	(Time point, h)										
		1.5	3	6	12	24	48	96	168	336		
SNAI1	IZ	***	***	***	***	***	***	***			**	
SNAI1	ANL	***	***	***	***	***	***	***			*	
ALPK3	ANL		***	***		***	***				***	
ALPK3	FB		***	***		**	*		***	**		
SNAI2	IZ		**	**				*	**			
SNAI2	ANL		***	***				***	***	***		
DLX5	IZ		*					*	***	***	**	
DLX5	FB		**						***	***	***	
SP7	FB		**					*	**		*	
ADGRG1	IZ		**	**				*	***	***	***	
ADGRG1	ANL		**	**					***	***	***	
ANGPTL4	ANL		***	***	***	***	***	***			***	
WNT9A	ANL		**	***	*					**	***	
TNFSF11	IZ		**					*	***	***	***	
COL2A1	IZ		*	**						***	***	
TGFBI	ANL			***	***	***	***	***			*	
NEFL	FB			**	***				***	***	***	
IGFBP5	ANL			**	***	***					***	
TGFBI	FB				***	***	***				***	
FAM132A	FB					*	**				***	
PLVAP	ANL						**	*			*	
ITGAV	ANL						**				***	

Figure 2.2.d) Genes with no response to the chondrogenic stimulation

Genes	Cell type	(Time point, h)										
		1.5	3	6	12	24	48	96	168	336		
DLX5	ANL											
LEF1	FB											
ABCC9	IZ											
ABCC9	FB											
ADGRG2	FB											
APLNR	IZ											
APLNR	ANL											
BMPR1A	IZ											
BMPR1A	FB											
CDON	FB											
DIO2	IZ											
DIO2	FB											
FGF18	IZ											
FGF18	FB											
FGFR3	FB											
FRZB	IZ											
FZD1	FB											
IBSP	IZ											
IBSP	ANL											
KCNJ8	IZ											
KCNJ8	FB											
LOC100630171	IZ											
NTRK2	IZ											
NTRK2	ANL											
PANX3	IZ											
PANX3	FB											
PRKG2	IZ											
SHC3	IZ											
SHC3	FB											
TLR2	IZ											
TLR2	ANL											
TNFRSF21	IZ											
ABI3BP	ANL											
ABI3BP	FB											
CLU	IZ											
COL10A1	IZ											
COL10A1	FB											
COL1A1	IZ											
COL1A1	FB											
GALNT14	IZ											
GALNT14	ANL											
GALNT14	FB											
MGP	IZ											
MGP	FB											
SGMS2	FB											
TIMP2	IZ											

Annotation group
Transcription regulation
Signaling transduction
Extracellular matrix
Not categorized

Upregulated based on 0h	***, P<0.0001
Downregulated based on 0h	** , P<0.01
	* , P<0.05

Figure 2.2. Changes in steady state mRNA levels in response to the chondrogenic stimulation relative to the 0h time point. IZ = interzone cell; ANL = anlagen cell; FB = fibroblast. a) Upregulated genes, b) downregulated genes and c) genes with mixed patterns of upregulation and downregulation after inducing chondrogenesis. d) Genes with no response to the chondrogenic stimulation.

## 1) Monolayer vs. 0h

To assess the effect of trypsin digestion and centrifugation required for establishing cell pellets from monolayers, steady state mRNA levels in P4 cell monolayers were compared to that of P5 0h cell pellets. Among the 87 targeted loci, only four genes (APLNR, GDF5, S1PR3, and TLR2) had significantly lower expression levels in monolayer cultures compared to 0h cell pellets in one or more cell types ( $P < 0.05$ ; Figure 2.3). These genes are all categorized in the signal transduction. Interestingly, GDF5 and S1PR3 subsequently displayed consistent up-regulation following the onset of chondrogenic induction.

Genes	Cell type	Monolayer	(Time point, h)											
			0	1.5	3	6	12	24	48	96	168	336		
APLNR	IZ	***	Ref. <sup>a</sup>											
APLNR	ANL	***	Ref.											
APLNR	FB	***	Ref.							**	***	***		
GDF5	IZ	**	Ref.				***	***	***	***	***	***	***	***
GDF5	ANL	*	Ref.			**	***	***	***	***	***	***	***	***
GDF5	FB	***	Ref.							**	***	***	***	***
S1PR3	IZ	***	Ref.		***	**	*	***	***	***	***	***	***	***
S1PR3	ANL		Ref.	**	***	***	**	***	***	***	***	***	***	***
S1PR3	FB		Ref.	**	***	***	***	***	***	***	***	***	***	***
TLR2	IZ	*	Ref.											
TLR2	ANL	***	Ref.											
TLR2	FB		Ref.									***	***	

Upregulated based on 0h	***, $P < 0.0001$
Downregulated based on 0h	** , $P < 0.01$
	* , $P < 0.05$

Figure 2.3. Genes with significantly different steady state mRNA levels between passage 4 monolayer cultures and passage 5 cell pellet cultures at 0h. IZ = interzone cell; ANL = anlagen cell; FB = fibroblast.

<sup>a</sup>The reference point (baseline) for comparison was 0h for the same sample.

## 2) Timing of initial differential expression relative to 0h within a cell type

Eighty-six out of the 87 targeted genes displayed a significant change with at least one timepoint in response to the chondrogenic induction protocol. The lone exception was

GALNT14, in which steady state mRNA levels did not change significantly at any time point in either of the three cell types (Figure 2.2.d). Comparing the three functional annotation categories, percentages of loci displaying their first onset of change (gene × cell type combinations) were calculated at each time point (Table 2.2; Figure 2.4).

Table 2.2. Timing of the first response to the chondrogenic induction protocol in new gene × cell type combinations within each of the three functional annotation groups. Data are reported as the percent (%) of gene loci within the annotation group with significant changes.

Annotation group	Time point								
	1.5h	3h	6h	12h	24h	48h	96h	168h	336h
Transcription regulation	7.0	34.9	16.3	25.6	11.6	2.3	2.3	0.0	0.0
Signal transduction	3.3	18.9	21.3	9.8	9.8	13.1	12.3	3.3	8.2
Extracellular matrix	0.0	18.2	12.7	20.0	12.7	9.1	12.7	7.3	7.3
Total	3.3	20.9	18.6	14.9	10.2	10.7	11.2	3.7	6.5

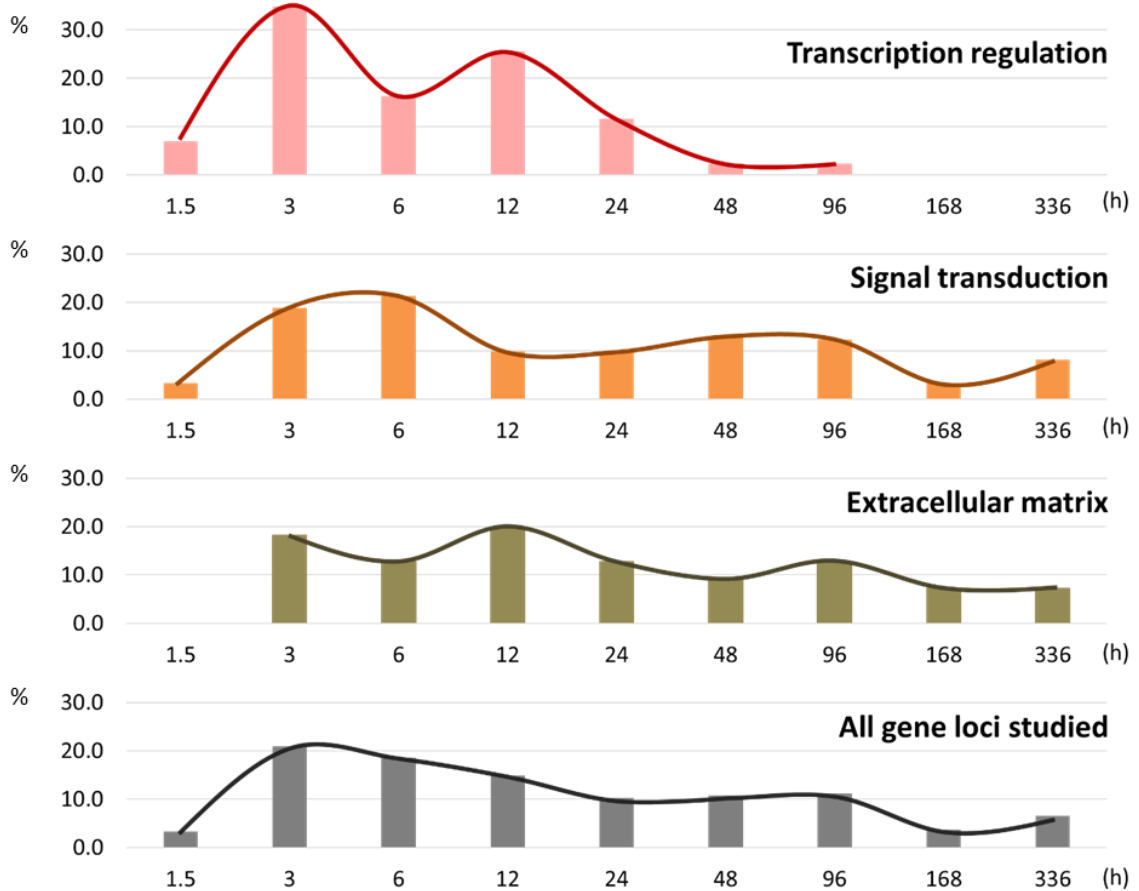


Figure 2.4. Histogram of percentages of the first response to the chondrogenic stimulation in new gene × cell type combinations within each of the three functional annotation groups.

In the transcription regulation group, the first half (58.1% of the total) of gene × cell type combinations showed their first chondrogenic responses within the first 6 hours, with the mode observed at 3h (34.9% of the total). The other 41.9% of gene × cell type combinations first responded to the chondrogenic stimulation between 12 – 96h. Only 4.6% of gene × cell type combinations responded between 48 – 96h, and these were observed solely in fibroblast cultures. It is interesting to note that in interzone and anlagen cell cultures, steady state mRNA levels for all of the genes with transcription regulation functional annotation had changed significantly within the first 24 hours following chondrogenic induction (Table 2.3). No transcription regulating genes displayed their first change later than 96h in any cell type.

By comparison to the transcription regulation group, genes involved in signaling cascades showed slightly delayed responses. Roughly half (53.3%) of the gene × cell type combinations changed significantly within the first 12 hours. Genes in the signaling group also showed relatively slower responses in fibroblast cultures.

Response to the chondrogenic induction protocol was more delayed for genes encoding ECM proteins or involved in ECM metabolism. Indeed, no differences were observed in this functional annotation group at the 1.5h time point. The first half (50.9%) of gene × cell type combinations were observed between 3 – 12h (Table 2.3).

In the negative control dermal fibroblasts, responses to chondrogenic induction were delayed compared to interzone and anlagen cells. Transcriptional regulatory genes all showed their first responses within 24 hours in the chondrogenic cells, compared to some initial

changes delayed until 96h in fibroblast cultures. In the signal transduction group, 8.3% or 2.1% of the first responses were recorded at the last time point (336h) in interzone cell cultures and anlagen cell cultures, respectively, while 15.4% of those was recorded at 336h in fibroblast cultures. Also, the majority of first reactions in the ECM group were at 3h (23.5%) and 12h (23.5%) in interzone cultures, 12h (28.6%) in anlagen cultures, and 96h (23.5%) in fibroblast cultures.

Table 2.3. Timing of the first response to the chondrogenic induction protocol within functional annotation group

(% of gene loci with significant changes)

Annotation group	Time point								
	1.5h	3h	6h	12h	24h	48h	96h	168h	336h
Transcription regulation									
Interzone cell	6.7	33.3	20.0	26.7	13.3	0.0	0.0	0.0	0.0
Anlagen cell	7.1	28.6	14.3	28.6	21.4	0.0	0.0	0.0	0.0
Fibroblast	7.1	42.9	14.3	21.4	0.0	7.1	7.1	0.0	0.0
Signal transduction									
Interzone cell	0.0	27.8	19.4	5.6	13.9	11.1	11.1	2.8	8.3
Anlagen cell	4.3	17.0	25.5	14.9	6.4	17.0	10.6	2.1	2.1
Fibroblast	5.1	12.8	17.9	7.7	10.3	10.3	15.4	5.1	15.4
Extracellular matrix									
Interzone cell	0.0	23.5	5.9	23.5	17.6	11.8	5.9	11.8	0.0
Anlagen cell	0.0	19.0	14.3	28.6	4.8	9.5	9.5	4.8	9.5
Fibroblast	0.0	11.8	17.6	5.9	17.6	5.9	23.5	5.9	11.8

**Relative differences between cell types at each time point**

Steady state mRNA levels at each time point were compared in pairwise comparisons between cell types, and the results categorized into four groups (Table 2.4; Figure 2.5). While 14 – 18 genes were already differentially expressed at 0h in the comparisons between cell types, most genes (69 – 73 genes among the 87 targeted loci) did not show differences initially. The focus of this study is on relative differences between interzone and anlagen cells.

Table 2.4. Four patterns of differential gene expression before and after inducing *in vitro* chondrogenesis

Cell type comparison	Already different at baseline (0h)		Not different at baseline (0h)	
	Retained differences after chondrogenesis	Lost differences after chondrogenesis	Became different after chondrogenesis	Remained similar after chondrogenesis
IZ vs. ANL	11 genes	3 genes	47 genes	26 genes
IZ vs. FB	14 genes	0 gene	41 genes	32 genes
ANL vs. FB	18 genes	0 gene	47 genes	22 genes
<b>Common genes in all comparisons</b>	3 genes	0 gene	15 genes	7 genes
<b>Specific to the comparison between IZ and ANL cultures</b>	5 genes: GDF6, MGP, OMD, PDLIM1, and RUNX2	3 genes: COL2A1, COMP, and DIO2	9 genes: ALPK3, ASS1, AQP1, CTGF, FAM20A, FGF1, ITGA7, PLVAP, and TLR4	12 genes: ADGRG1, ALPL, APLNR, DCN, ENTPD1, FAM132A, FZD1, GDF5, IGFBP5, IGFBP7, MMP2, and PLAT

IZ = interzone cell; ANL = anlagen cell; FB = fibroblast.

Figure 2.5.a) Genes that were differentially expressed in the indicated pairwise cell type comparison at baseline (0h) and also retained differences after inducing chondrogenesis

Comparison between interzone and anlagen cell cultures

Genes	(Time point, h)											
	0	1.5	3	6	12	24	48	96	168	336	336	
<b>RUNX3</b>	***	***	***	***	***	***	***	***	***	***	***	***
<b>ABI3BP</b>	***	***	***	***	***	***	***	**	***	**	***	
<b>MGP</b>	***	***	***	***	***	***	***					
<b>LEF1</b>	***	***	***	**	***	***	***	***	***			
<b>BOC</b>	***	***	***	**	***	***	***	***	**	***	**	
<b>SHC3</b>	***	***	**	*		*	**	*				
<b>DLX5</b>	**	***	***	***	***	***	***					
<b>GDF6</b>	**	**	**	**	***	***	**	***	***	***	***	
<b>RUNX2</b>	**	**			**		*	*				
<b>OMD</b>	**							**	**			
<b>PDLIM1</b>	*								***	***		

Comparison between interzone cell and fibroblast cultures

Genes	(Time point, h)											
	0	1.5	3	6	12	24	48	96	168	336	336	
<b>ASS1</b>	***	***	***	***	***	***	***	***	***	***	***	***
<b>LEF1</b>	***	***	***	***	***	***	***	***	***	***	***	***
<b>COMP</b>	***	***	***	***	***	***	***	***	***	***	***	***
<b>TLR4</b>	***	***	***	***	***	***	***	***	***	***	*	
<b>DLX5</b>	***	***	***	***	***	***	***	***	***	***		
<b>DCN</b>	***	***	***	***	***	***	***	***	***	**	***	
<b>ENTPD1</b>	***	***	***	***	***	***	***	**	**	***	**	
<b>ABI3BP</b>	***	***	***	***	**	*		*				
<b>SHC3</b>	***	***	**	***	***	***	***	***	***	***	**	
<b>FGF1</b>	***	***	*		***	***	***	***				
<b>AQP1</b>	*	*	**	***	***							
<b>CTGF</b>	*	*	*		*	***	***	***	***	***	***	
<b>RUNX3</b>	*	*		**	**	*	**					
<b>FAM20A</b>	***		***	***	***	**		***	**	**	**	

Comparison between anlagen cell and fibroblast cultures

Genes	(Time point, h)											
	0	1.5	3	6	12	24	48	96	168	336	336	
<b>ASS1</b>	***	***	***	***	***	***	***	***	***	***	***	***
<b>ENTPD1</b>	***	***	***	***	***	***	***	***	***	***	***	***
<b>TLR4</b>	***	***	***	***	***	***	***	***	***	***	***	***
<b>DLX5</b>	***	***	***	***	***	***	***	***	***	***	**	**
<b>DCN</b>	***	***	***	***	**	*						
<b>TNFSF11</b>	***	***	***	***								***
<b>AQP1</b>	***	***	**	***	***	**				**	***	
<b>LEF1</b>	***	***	**	**	***						***	
<b>FGF1</b>	***	***										
<b>BOC</b>	**	***	**		*	***	***	***	***	*		
<b>FAM20A</b>	**	**	**							**	***	
<b>IGFBP5</b>	**	**	*								**	
<b>ITGA7</b>	**	**		*	*	*					***	
<b>DIO2</b>	**	**										**
<b>CDON</b>	**	**										
<b>CTGF</b>	**	*	**	**			**	***	***	**		
<b>CLU</b>	*	***	*		*	*					**	
<b>RUNX3</b>	*							*	***	***	***	

Figure 2.5.b) Genes that were differentially expressed in the indicated pairwise cell type comparison at baseline (0h) but lost differences after inducing chondrogenesis

Comparison between interzone and anlagen cell cultures

Genes	(Time point, h)											
	0	1.5	3	6	12	24	48	96	168	336	336	
<b>DIO2</b>	**											
<b>COL2A1</b>	**											
<b>COMP</b>	**											

Annoation group	
Upregulated in interzone cell culture	***, P<0.0001
Upregulated in anlagen cell culture	** , P<0.01
Upregulated in fibroblast culture	* , P<0.05
<b>Transcription regulation</b>	
<b>Signal transduction</b>	
<b>Extracellular matrix</b>	
<b>Not categorized</b>	



Figure 2.5.c) Genes that were not differentially expressed in the indicated pairwise cell type comparison at baseline (0h) but responded differently to the chondrogenic stimulation

Comparison between interzone and anlagen cell cultures

(Time point, h)

Genes	0	1.5	3	6	12	24	48	96	168	336
S1PR3		**	**	**						
RELN		*	**	***	***	***	***	***	**	
TNFSF11		**	***	***	**		**	*		
FGF1				**	***	***	**			
SP7				**			*			
BMPR1A				*	*					
SMPD3					***	***	***	***	***	***
FAM20A					***	**		**		
FGF18					**	**		***	***	***
IHH					**			*	***	**
TNFRSF21						**				***
MASP1					*			**	***	**
PRKG2							**	***	***	**
ABCC9							**	**		
SERPINE1							**			
THBS4								***	***	***
FGFR3								***	***	***
S100A1								***	***	***
MET								**	**	***
IBSP								**	*	
COL5A3								**		
PANX3								*		

<b>Annoation group</b>
<b>Trascription regulation</b>
<b>Signal transduction</b>
<b>Extracellular matrix</b>
<b>Not categorized</b>

Upregulated in interzone cell culture	***, P<0.0001
Upregulated in anlagen cell culture	** , P<0.01
Upregulated in fibroblast culture	* , P<0.05

Comparison between interzone cell and fibroblast cultures

(Time point, h)

Genes	0	1.5	3	6	12	24	48	96	168	336
MGP		**	**	**	**			*		
SNAI1		**								
S1PR3		*	***	***	***			**	**	**
SLC38A1		*								
ADGRG1			**					*	**	
SGMS2			*							
PRKG2				***	***	***	***	***	***	***
GDF5				**	***	***	***	***	*	
MET				**	***	**	*	***	***	***
ALPL					*	**	**			**
FZD1					*	**				
IHH					*			***	**	*
COL5A3					*				**	**
ADAMTS5						***	***	***	***	***
COL2A1						**	***	***	***	***
SERPINE1						**	***	**	***	***
NEFL						**	***			
GLI3						**	**	***	***	***
TNFSF11						*				*
PLAT							**	***	***	***
S100A4						**	**		***	
IBSP							*	***	**	
CLU							*			**
ENTPD2							*			
ABCC9							*			
FAM132A								***	***	***
WNT9A								***	***	***
THBS4								***	***	***
TGFB1								***	***	**
MMP2								**	**	*
IGFBP5								**	**	
RELN								*		

Comparison between anlagen cell and fibroblast cultures

(Time point, h)

Genes	0	1.5	3	6	12	24	48	96	168	336
COMP		**	***	***	***	***	***	***	***	***
OSR2		*				*	**			
SLC38A1		*								*
SNAI1		*								
SGMS2			*	*					***	*
GDF5				***	***	***	***	***	***	***
IHH				***	***	***	**	***	***	***
COL5A3				**	***	***	***	***	***	***
FZD1				**	***	***	**			
TNFRSF21				**	***	***	*			*
FGF18					***	***	**	**	***	**
SMPD3					**	***	***	***	***	**
RELN					*	***	***	***	***	**
ALPL					*	**				***
ADAMTS5						**	**	***	***	***
COL2A1						**	***	***	***	***
SP7						**	*			
MASP1						**		**	***	***
S100A4						*	**	**	***	***
FGFR3								***	***	***
S100A1								***	***	***
WNT9A								***	***	***
PLAT								**	***	***
SERPINE1								**	***	***
SHC3								**	***	
MMP2								**		**
FAM132A								*	***	***
MGP								*		

(continued Figure 2.5.c) Genes that were not differentially expressed in the indicated pairwise cell type comparison at baseline (0h) but responded differently to the chondrogenic stimulation

Comparison between interzone and anlagen cell cultures

(Time point, h)

Genes	0	1.5	3	6	12	24	48	96	168	336
SNAI2									***	***
CLU									***	***
CDH13									***	***
ENTPD2									***	**
ASS1									**	***
TGFBI									**	***
S100A4									**	***
SGMS2									**	
WNT9A									*	***
SNAI1									*	
ALPK3										***
CREB5										***
ANGPTL4										***
AQP1										***
TSPAN15										***
COL1A1										***
TIMP2										***
CTGF										**
PLVAP										**
ITGA7										**
GLI3										*
SLC38A1										*
ADAMTS5										*
TLR2										*
TLR4										*

Comparison between interzone cell and fibroblast cultures

(Time point, h)

Genes	0	1.5	3	6	12	24	48	96	168	336
GALNT14									***	
FGFR3									**	***
RUNX2									**	**
FRZB									*	*
IGFBP7										***
ANGPTL4										**
APLNR										**
FGF18										*
TLR2										*

Comparison between anlagen cell and fibroblast cultures

(Time point, h)

Genes	0	1.5	3	6	12	24	48	96	168	336
ENTPD2									***	***
PDLIM1									***	***
SNAI2									***	***
ANGPTL4									***	***
CDH13									***	***
GDF6									***	***
PANX3									***	*
APLNR									**	***
COL1A1									**	***
BMPR1A									**	
ABI3BP									**	
ADGRG1									*	**
CREB5										***
NTRK2										***
TLR2										***
IGFBP7										**
ITGAV										**
TSPAN15										**
TIMP2										**

Annoation group	
Upregulated in interzone cell culture	***, P<0.0001
Upregulated in anlagen cell culture	** , P<0.01
Upregulated in fibroblast culture	* , P<0.05
Trascription regulation	
Signal transduction	
Extracellular matrix	
Not categorized	

Figure 2.5.d. Genes that were not differentially expressed in the indicated pairwise cell type comparison at baseline (0h) and also were not different after the chondrogenic stimulation

Comparison between interzone and anlagen cell cultures	Comparison between interzone cell and fibroblast cultures	Comparison between anlagen cell and fibroblast cultures
<b>Genes</b>	<b>Genes</b>	<b>Genes</b>
ENTPD1	ALPK3	ALPK3
OSR2	CREB5	GLI3
ADGRG1	OSR2	RUNX2
APLNR	PDLIM1	ABCC9
FAM132A	SNAI2	ADGRG2
GDF5	SP7	BMP2
IGFBP5	ADGRG2	FRZB
IGFBP7	BMP2	IBSP
PLAT	BMPR1A	KCNJ8
ADGRG2	BOC	LOC100630171
BMP2	CDH13	MET
CDON	CDON	PLVAP
FRZB	DIO2	PRKG2
FZD1	GDF6	PTCH2
ITGAV	ITGAV	S1PR3
KCNJ8	KCNJ8	COL10A1
LOC100630171	LOC100630171	GALNT14
NTRK2	NTRK2	OMD
PTCH2	PANX3	SPARCL1
ALPL	PLVAP	THBS4
DCN	PTCH2	TGFB1
MMP2	S100A1	NEFL
SPARCL1	TNFRSF21	
COL10A1	TSPAN15	
GALNT14	COL10A1	
NEFL	COL1A1	
	ITGA7	
	OMD	
	SMPD3	
	SPARCL1	
	TIMP2	
	MASP1	

Annoation group
Trascription regulation
Signal transduction
Extracellular matrix
Not categorized

Figure 2.5. Four patterns of differential gene expression before and after inducing *in vitro* chondrogenesis. Each pattern is shown in a), b), c), and d), respectively.

### 1) Genes with differential expression levels prior to the chondrogenic stimulation

Among the cell type pairwise comparisons, there were three common genes (DLX5, LEF1, and RUNX3) that had different initial expression levels at 0h and showed unique cell type specific properties to the chondrogenic induction protocol. Specific to the comparison between interzone and anlagen cultures, five genes (GDF6, MGP, OMD, PDLIM1, and RUNX2) started with different expression levels and also retained differential profiles following the chondrogenic induction (Table 2.4; Figure 2.5.a).

If initial mRNA levels of a gene were different in the comparisons to fibroblast cultures, the gene also showed different expression levels after inducing chondrogenesis. However, between interzone and anlagen cell cultures, three genes (COL2A1, COMP, and DIO2) with different initial mRNA levels lost differences across all post-chondrogenic induction time points (Table 2.4; Figure 2.5.b).

### 2) Genes with no differential expression levels prior to the chondrogenic stimulation

Among the 69 – 73 genes that initially had not different expression levels in pairwise cell-type comparisons, 41 – 47 genes developed significant steady state mRNA differences after inducing chondrogenesis. Between interzone and anlagen cultures, there were 47 genes in this category, and only nine (ALPK3, ASS1, AQP1, CTGF, FAM20A, FGF1, ITGA7, PLVAP, and TLR4) that were specific to the comparison between these two chondrogenic cell lines (Table 2.4; Figure 2.5.c).

On the other hand, 22 – 32 genes still did not change in pairwise cell-type comparisons,

seven of which showed no difference in any cell type comparison: ADGRG2, BMP2, COL10A1, KCNJ8, LOC100630171, PTCH2, and SPARCL1. Among these seven genes, there was no gene involved in transcription regulation. In addition, twelve genes were commonly regulated under the chondrogenic stimulation specifically between interzone and anlagen cultures: ADGRG1, ALPL, APLNR, DCN, ENTPD1, FAM132A, FZD1, GDF5, IGFBP5, IGFBP7, MMP2, and PLAT (Table 2.4; Figure 2.5.d).

### 3) Individual gene loci demonstrate clear examples of cell type specific differences

Cartilage biomarkers (COL2A1 and COMP) were upregulated in the chondrogenic cell lines at later time points (Figure 2.6). While their steady state mRNA levels were also increased in fibroblast cultures towards the end of the experimental period, the relative levels compared to the chondrogenic cell cultures were consistently lower in non-chondrogenic cell cultures (COL2A1, from 24h to 336h; COMP, at all post-chondrogenic induction time points; Figure 2.5.a and c).

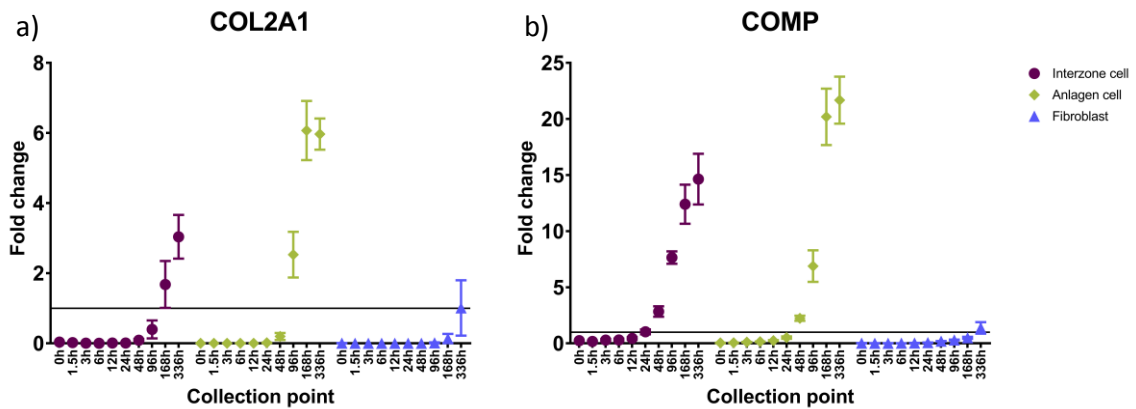


Figure 2.6. Steady state mRNA levels of classic cartilaginous biomarkers across the experimental period. Fold changes were calculated based on the positive control sample (pooled equine adult tissue and fetus RNA). Mean  $\pm$  SEM (n=7); a) COL2A1; b) COMP

There were examples of genes that resulted in unique regulation patterns in either interzone or anlagen cell cultures compared to the other two cell types over the time course. Changes in steady state mRNA levels of ABI3BP showed a clear example of downregulation (Figure 2.7.a), and those of PRKG2 showed an example of minimal expression (Figure 2.7.b) specifically in interzone cell cultures during the experimental period. In contrast, PANX3 was upregulated only in anlagen cell cultures toward the later time points while its mRNA levels

were consistently lower or not detected in the other two cell cultures across the time points (Figure 2.8).

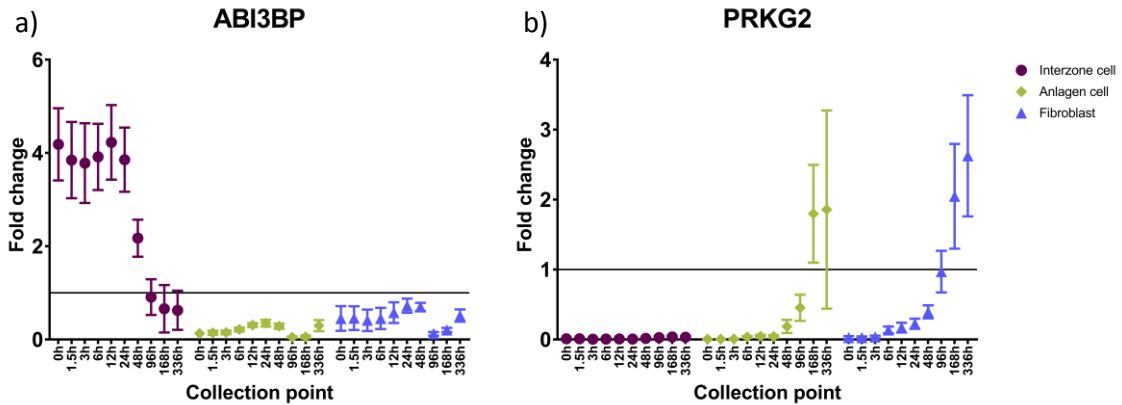


Figure 2.7. Steady state mRNA levels of example genes that were differentially regulated in interzone cell cultures compared to the other cell types across the experimental period. Fold changes were calculated based on the positive control sample (pooled equine adult tissue and fetus RNA). Mean  $\pm$  SEM ( $n=7$ ); a) ABI3BP was downregulated and b) PRKG2 was not expressed only in interzone cell cultures.

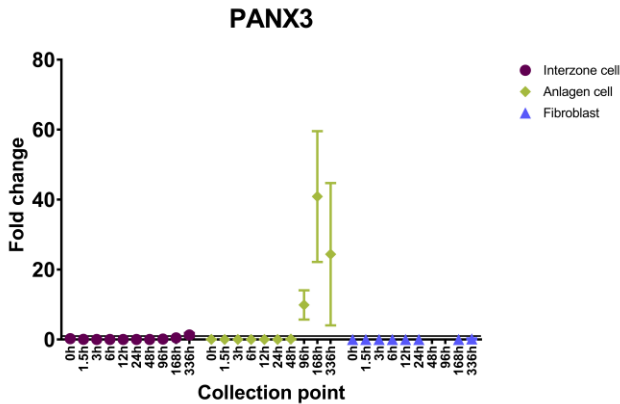


Figure 2.8. Steady state mRNA levels of an example gene that was differentially regulated in anlagen cell cultures compared to the other cell types across the experimental period. Fold changes were calculated based on the positive control sample (pooled equine adult tissue and fetus RNA). Mean  $\pm$  SEM ( $n=7$ ); PANX3 was upregulated only in anlagen cell cultures towards later time points.

## Discussion

The present study was conducted to test the hypothesis that divergent chondrogenic pathways in interzone and anlagen cell cultures will be evident within the earlier time frame—within the first 24 hours after inducing *in vitro* chondrogenesis. In all three cell types including the negative control fibroblasts, changes in steady state mRNA levels in a subset of the targeted genes started as early as 1.5 hours after initiating the chondrogenic induction. All genes that showed responses at 1.5h had functional annotation categorized in regulatory events such as transcription or signal transduction. Moreover, initial responses in transcription regulatory events in the two chondrogenic cell lines occurred within the first 24 hours, while delayed responses were observed in the negative control (Table 2.3). The earlier changes in these regulatory genes likely diverge chondrogenic pathways in interzone and anlagen cell cultures, and differences may provide insight into divergent aspects of their chondrogenic fate. From the pairwise comparisons between interzone and anlagen cell cultures at each time point, starting mRNA levels of 73 genes were not significantly different. For the 47 genes that differentially responded to the TGF- $\beta$ 1 chondrogenic induction at some point in the 336-hour of experimental period, 12 significantly changed within the first 24 hours (Table 2.4; Figure 2.5.c). Thus, the results support the hypothesis tested in this study.

### ***Relationship between early and delayed gene expression changes***

The first responses to the chondrogenic induction measured at 1.5h were involved in either transcription regulation or signal transduction while no ECM related genes changed their mRNA levels at the first collection point. The majority of first responses were observed



within the first 24 hours in all three annotation groups: at 3h in the transcription group, 6h in the signaling group, and 12h in the ECM group. Most of the initial changes in transcription regulatory category occurred within a day (0 – 24h), while about 40% of the first responses in the signaling group and the ECM group occurred after 24 hours. In fact, no gene loci in the ECM related genes had no response at the first collection point in any cell cultures (Table 2.2; Figure 2.4). Taken together, the results demonstrate a sequence whereby genes encoding proteins functionally annotated in transcription, signaling events, and ECM are expressed in roughly that order: altered transcriptional events leading to subsequent changes in signaling cascades and finally the effector genes involved in ECM biology.

The model derived from this observation is that transcription regulating genes might be primary genes required for secondary changes in downstream signaling events and ECM accumulation. To define “(early) primary” and “(delayed) secondary” genes from a perspective of cell biology, an experiment designed to include protein synthesis inhibitors must be performed (Yamamoto and Alberts, 1976). If a gene changes its expression levels under the presence of protein synthesis inhibitors, it can be called “primary” response because it does not require *de novo* protein synthesis to change its expression levels. Therefore, to further elucidate these relationships and identify if the responses were either primary or secondary, a mechanistic experiment will be required.

It is interesting to note that, all changes in mRNA levels observed at 1.5h (Table 2.2) were upregulation compared to the baseline (Figure 2.2.a and c). On the other hand, downregulation started to be detected by 3h. This may confirm that degradation of

transcripts requires more time than *de novo* synthesis (Eser et al., 2014). Regulatory genes may have relatively faster decay rates in order to rapidly alter signal transduction; all the genes that showed their first responses as downregulation between 3 – 6h were involved in either transcription or signaling events, except for two cases (Figure 2.2.b and c).

There were only a few genes that showed changes at the 0h time point compared to P4 monolayer cell cultures, indicating little subsidiary effects of the chondrogenic induction protocol besides the chondrogenic stimuli treatment (Figure 2.3). These genes are generally involved in cell migration and adhesion (Hashimoto et al., 2005, Scott et al., 2007, Zeng et al., 2007, Shwartz et al., 2016, Ogle et al., 2017), and thus the results may suggest that being lifted from monolayers and spun down to form pellets might require the cells to alter gene expression regulating these biological processes regardless of cell types.

### ***Differential gene regulation between cell types***

#### Differences between chondrogenic and non-chondrogenic cell cultures

From the observations made in individual cell types over the time course, delayed responses in all functional annotation categories were observed in the fibroblasts compared to the either interzone or anlagen cells (Table 2.3). While this may well be attributed to chondrogenic potential of the cell types and relative sensitivity to the induction medium, it is important to note that the targeted gene loci were selected based in part on established annotations related to chondrogenesis and expression in skeletal tissues. As such, there is a selection bias and a full transcriptome analyses would help to sort this out.

Consistent with their established functional annotations, COL2A1 and COMP, encoding major cartilaginous ECM proteins, were significantly upregulated in the interzone and anlagen cell pellets at later time points. The results with these biomarker effector genes confirmed their chondrogenic potential (Figure 2.5.a and c; Figure 2.6).

#### Differences between interzone and anlagen cell cultures

Interzone and anlagen tissues in equine fetuses at day-45 of gestation are readily distinguished morphologically with a dissecting microscope. Along with the morphogenic distinction, 14 out of the 87 gene loci targeted started with different mRNA levels at 0h. Among the 14 genes, 11 genes still expressed different profiles after inducing chondrogenesis (Table 2.4; Figure 2.5.a) and five of them showed the differential patterns only between interzone and anlagen cultures but not in the negative control fibroblasts. OMD and RUNX2 were upregulated in anlagen cell cultures after the chondrogenic induction, and their known annotation is related to ECM production in the bone (Ninomiya et al., 2007, Mevel et al., 2019). Thus, the results may confirm that anlagen cultures under chondrogenic stimulation progressed towards the pathways leading to bone formation. On the other hand, MGP and GDF6 were upregulated in interzone cultures, and their reported functional annotations are inhibiting ectopic tissue calcification (Luo et al., 1997) and diarthrodial joint formation (Settle et al., 2003), respectively. In addition, mRNA levels of PDLIM1, a transcription coactivator gene, were greater in interzone cell cultures at the latest time points. From a previous RNA-seq dataset generated in the MacLeod lab, this gene was not differentially expressed between interzone and anlagen tissue lineages from 45-day-old equine fetuses and neonatal foals, but the expression of PDLIM1 was 3.41 fold greater ( $P < 0.0001$ ) in interzone tissue compared to

cartilagenous anlagen from 60-day-old fetuses (Adam et al., in preparation). The present results and the previous data may suggest that this transcription regulatory gene might become upregulated in interzone at a later stage during articular cartilage development. Taken together, expression patterns of these genes in the culture model are consistent with the articular and hypertrophic cartilagenous tissue outcomes of interzone and anlagen cells respectively.

Interestingly, three genes that started with different steady state mRNA levels in interzone and anlagen cells, lost those differences after inducing chondrogenesis across the entire experimental period (Table 2.4; Figure 2.5.b). Two of the three, COL2A1 and COMP, are major cartilagenous ECM genes. These data are consistent with the induction medium driving both cell lines to produce common cartilagenous ECM under the same chondrogenic stimulation.

As noted above, steady state mRNA levels for 73 of the 87 targeted gene loci were not significantly different at the initial 0h time point. In comparing cell types, gene expression either diverged or changed in a similar way. Forty seven genes differentially responded to the chondrogenic induction (Table 2.4; Figure 2.5.c), but only nine of these were specific to the comparison between interzone and anlagen pellets. FGF1 was upregulated early (6h) in anlagen cultures and justifies further investigation. For the other 26 genes, the changes displayed profile similarities after the chondrogenic induction (Table 2.4; Figure 2.5.d). GDF5 is an established interzone biomarker (Bi et al., 1999, Hyde et al., 2007), but in this *in vitro* model using primary cells derived from day 45 equine fetuses, its mRNA profiles were not different between interzone and anlagen cultures at both pre- and post-chondrogenic

induction time points. The discrepancy between the current data and the previous findings in other species may be due to differences in developmental age. Similarly, while ALPL is known to be involved in bone mineralization (Golub and Boesze-Battaglia, 2007), its mRNA levels were not different between the two chondrogenic cell cultures. This may not be surprising given that the culture system was not intended to model osteogenesis.

Several gene loci displayed clear cell type-specific expression profiles. Example of unique steady state mRNA levels in interzone cell cultures were ABI3BP and PRKG2 (Figure 2.7). ABI3BP is a novel ECM gene and is required to switch the cellular status from proliferation to differentiation in mesenchymal stem cells (Hodgkinson et al., 2013). The mRNA levels of this gene were initially greater in interzone cell cultures compared to the other cell lines (Figure 2.5.a), started to decrease by 96h (Figure 2.2.b), and became as low as the other cell lines at the last time point. In literature, this gene was differentially upregulated in articular cartilage compared to hypertrophic growth plate both *in vivo* and *in vitro* (Hissnauer et al., 2010). In contrast, mRNA levels of PRKG2 were minimal and remained unchanged over time in interzone cell cultures (Figure 2.2.d) while this gene became upregulated in the other two cell lines at later time points (Figure 2.2.a). PRKG2 is known to be involved in mammalian skeletal development, and its null mutation resulted in 23 – 30% decreased length of limb bones (Pfeifer et al., 1996) due to impaired chondrocyte hypertrophy (Kawasaki et al., 2008). Thus, the present data indicate that interzone cell cultures had less hypertrophic potential compared to the other two cell types.

PANX3 was a locus specifically upregulated in anlagen cell cultures compared to the other

cell cultures (Figure 2.8). This gene is expressed in cartilaginous anlagen of a developing limb (Bond et al., 2011), and when PANX3 was knocked down, hypertrophic differentiation was delayed and attenuated (Oh et al., 2015). Consistent with its functional annotation, the patterns of its mRNA levels across the experimental period confirmed the greater hypertrophic potential in anlagen cell cultures.

In summary, results demonstrated that while interzone and anlagen cells are both chondrogenic, they display some clear differences in response to the same TGF- $\beta$ 1 chondrogenic induction signal. Data from this *in vitro* model had several gene expression profiles broadly consistent with well-established developmental fates of interzone and anlagen cells within limb buds.

### ***Future direction***

Although the current study reports interesting time point differences and examples of genes that may render interzone and anlagen cell cultures unique, there are important questions remaining: 1) how would the kinetics of gene expression change at a transcriptome level and 2) what would be the key regulator(s) of differential chondrogenic pathways between these cell lines?

The present study focused on the 93 selected genes that have differential gene expression levels between interzone and anlagen tissues or functional annotations related to fetal skeletal development. However, there are more than 20,000 structurally annotated protein coding gene loci in mammals. Therefore, whole-transcriptome analyses at critical time points

may also reveal important regulator or effector genes that have conventionally received less attention. By conducting pathway analyses on the transcriptomic data, critical signaling mechanisms that regulate divergent chondrogenic differentiation pathways may be identified.

### **Conclusion**

In conclusion, the results confirmed the chondrogenic potential of interzone and anlagen cells, but they also documented distinct functional responses to the same chondrogenic stimulation. The new information elucidated by this study centers on expression kinetics of the targeted genes with established functional annotation relevant to chondrogenesis, cellular responses to TGF- $\beta$ 1, and the regulation of cellular differentiation. Overall, transcription regulatory responses preceded the other responses in signal transduction or ECM maintenance, and effector genes involved in ECM biology showed relatively delayed responses compared to regulatory genes. The data demonstrated that '(early) primary' regulatory gene expression changes start as early as 1.5 hours after inducing chondrogenesis, followed by changes in expression profiles of '(delayed) secondary' effector genes at later time points. Further investigation is required at a transcriptome level to identify key regulators and pathways that drive differential chondrogenesis in interzone and anlagen cells in different directions.

### **Chapter 3. Transcriptomic divergence between equine fetal interzone and anlagen cells during *in vitro* chondrogenic induction**

#### **Introduction**

Different chondrogenic pathways are involved in developmental processes of limb skeletal elements in mammals starting early in fetal development and continuing through gestation and even into the postnatal period (Shimizu et al., 2007, Pitsillides and Ashhurst, 2008). The initial chondrogenesis occurs in the mesenchymal condensation of a limb bud resulting in cartilaginous tissue formation. In developing limb tissues, SOX9 is expressed from the proliferating chondrocytes (Ng et al., 1997) and promotes expression of cartilaginous marker genes, COL2A1, ACAN, and COMP (Bi et al., 1999, Akiyama et al., 2002). Then, this continuous—uninterrupted—cartilaginous anlage becomes segmented into several units, which serve as templates of limb bones, separated by a different developmental tissue called the interzone which develops into all of the structures and tissue types present in synovial joints. Early developmental events of interzone are characterized by the decreased expression of chondrogenic biomarkers such as SOX9, COL2A1, and ACAN, and increased expression of interzone markers such as GDF5, WNT9A, CHRD, and ENPP2 (Bi et al., 1999, Karsenty and Wagner, 2002, Pacifici et al., 2005). While chondrocytes in anlagen progress through hypertrophic differentiation during endochondral ossification, a subset of the cells in interzone tissue that paused chondrogenic differentiation resume chondrogenesis and differentiate into articular chondrocytes. Even though several key marker genes in these processes have been reported in the literature, details related to the kinetics of gene expression have not been fully understood.



Beyond understanding developmental processes of skeletogenesis, the importance of comprehending comparative cell biology of interzone and anlagen cells likely has relevance to stark differences in the intrinsic ability of articular cartilage and bone tissue to repair structural defects. While bone fractures repair well by recapitulating the developmental process, articular cartilage has a very limited capacity to heal lesions. To study the cell biology of equine interzone and anlagen, primary cell cultures derived from fetal limb tissues collected at day 45 of gestation were studied before and 21 days after the *in vitro* induction of chondrogenesis (Adam, in preparation). In this previous RNA-seq study, transcriptomic profiles of interzone and anlagen cell cultures were relatively similar at the start of the experiment, but substantially diverged 21 days after being grown under the exact same chondrogenic induction conditions.

In a follow-up set of experiments to understand the biology of these chondrogenic cell lines, as reported in Chapter 2 of this thesis, the kinetics of gene expression profiles of 93 selected genes in equine fetal interzone and anlagen cell pellets grown in a chondrogenic medium were evaluated by RT-qPCR at 10 time points, specifically 0, 1.5, 3, 6, 12, 24, 48, 96, 168, and 336 hours after inducing *in vitro* chondrogenesis. The results confirmed that both cell lines are chondrogenic, but showed clear differences in expression profiles of certain regulatory genes starting even with the first 1.5h sample, as well as those encoding ECM components at subsequent time points. The new information obtained from the experiments in Chapter 2 further elucidated the kinetics of gene expression in these cell cultures during the 336-hour experimental period, although the findings were based on a finite number of the targeted genes. Thus, comparing steady state RNA levels between interzone and anlagen

cells on a fully transcriptomic scale would provide additional knowledge of individual gene loci and functional ontologies while providing the opportunity for a much more rigorous assessment of important chondrogenic and other cell biology pathways.

Therefore, the objective of the present study was to elaborate chondrogenic divergence between interzone and anlagen cell cultures at a whole transcriptome level under the same chondrogenic induction conditions. By analyzing transcriptomic data, potential candidate regulators that diverge the chondrogenic pathways between these two skeletal cell lines might be identified. The hypothesis tested in this chapter was that chondrogenic divergence between interzone and anlagen cell cultures will be initiated by differentially expressed regulatory genes within the first 1.5 hours, and more distinctive characteristics represented by the activation of different pathways and acquisition of distinguishing ECM profiles will accumulate as time passes in culture.

## **Materials and Methods**

### ***Cell culture***

Interzone cells, anlagen cells, and dermal fibroblasts (negative control) from day-45 equine fetuses were previously harvested (Adam et al., 2019) and used for chondrogenic differentiation experiments as reported in Chapter 2 of this thesis. A subset of these RNA samples were used for the present transcriptomic analysis. Briefly, three-dimensional cell pellet cultures were established at passage 5 and treated with a chondrogenic induction medium containing TGF- $\beta$ 1 (10 ng/ml). Then, cell pellets were collected at 10 different time

points during the 336-hour experimental period, followed by total RNA extraction for gene expression analyses. The subset of these samples were analyzed further in the current chapter, specifically six biological replicates at each of five collection points (0, 1.5, 3, 12, and 96h). Including the baseline, these time points were prioritized based on the study results reported in Chapter 2. By 1.5h, regulatory genes started to respond to the chondrogenic stimulation. At 3h, the major peak of overall first gene expression responses was observed. A second major peak of first responses of transcription regulatory genes was observed at 12h. Finally, most effector genes involved in ECM maintenance were exhibiting changes in steady state RNA levels by 96h after initiation of chondrogenic induction. More detailed procedures are described in the Materials and Methods section in Chapter 2.

### ***RNA sample preparation***

For bulk RNA-seq, any potential genomic DNA contamination in the total RNA isolates was removed using an RNase-Free DNase kit (cat No. 79254; Qiagen). The post-DNase step concentration of RNA was measured by a Qubit® 3.0 Fluorometer (cat No. Q33216; Life Technologies) with a Qubit® RNA Broad Range Assay Kit (catalog No. Q10211; Life Technologies). Both purity and RNA structural integrity was assessed using a Bioanalyzer 2100 (Agilent Technologies) with an Agilent RNA 6000 Pico kit (catalog No. 5067-1513; Agilent Technologies). All samples resulted in RIN values >8. The samples were diluted with nuclease-free water to 20 ng/ul and frozen at –80 °C. Then, 1,000 ng of each sample was shipped to the Roy J. Carver Biotechnology Center (University of Illinois at Urbana-Champaign, Urbana, IL, USA) for mRNA sequencing.

### ***Bulk RNA sequencing***

Construction of cDNA libraries from the RNA samples was conducted using a TruSeq Stranded mRNA kit (cat #. 20020595, Illumina). In the protocol, the samples were purified using poly-T oligo-attached magnetic beads and fragmented into smaller pieces (average library fragment size of 440 bp) by divalent cations under increasing temperature. The strand orientation was distinguished using Actinomycin D. In turn, Illumina Read 1 adaptors (P5) were ligated to the antisense strands, and Read 2 adaptors (P7) including sample-specific barcoding sequences (i7 index) were added to the sense strands. After the adaptor ligation step, the libraries were amplified by PCR for 14 cycles and quantitated by qPCR. The amplified individual libraries (100 ng/each) were pooled. The pool was diluted to 5nM, and 9 ul of the pool was loaded into an S4 lane on a NovaSeq 6000 (Illumina) for the sequencing. The cDNA libraries were sequenced for 151 cycles from each end of the fragments (total 302 cycles) with NovaSeq S4 reagents (cat. # 20027466; Illumina). The sequence Fastq files were generated and demultiplexed with the bcl2fastq v2.20 Conversion Software (Illumina). Quality scores were accessed by FastQC v0.11.9.

### ***Data analysis pipeline***

Once the resulting Fastq files were transferred to the University of Kentucky from the Roy J. Carver Biotechnology Center, adaptor sequences were trimmed out by Trim Galore v0.6.5. The trimmed reads were mapped onto the latest equine reference genome (EquCab 3.0, GCA\_002863925.1; Kalbfleisch et al., 2018) using Tophat (v2.0.8) and quantified with an ENSEMBL equine gene structure annotation database (v98) through the Cufflinks (v 2.2.1) pipeline (Trapnell et al., 2012). Read counts were normalized to the fragments per kilobase of

transcript per million mapped reads (FPKM) unit.

#### Differential expression analysis

Using the CuffDiff function in Cufflinks (v 2.2.1), differentially expressed genes (DEGs) were evaluated with focuses on 1) profiling kinetics of steady state mRNA changes within a cell type, 2) identifying common chondrogenic traits of interzone and anlagen cells relative to fibroblasts, and 3) characterizing differences between the two chondrogenic cell lines during *in vitro* chondrogenesis. The thresholds defining differential expression were  $\log_2$  fold change (FC)  $> |1|$  and the false discovery rate (FDR) adjusted P-value  $< 0.05$ . The DEGs were visualized in a heatmap or a volcano plot using R packages (gplot, Warnes et al., 2015; EnhancedVolcano, Blighe et al., 2019).

#### Gene ontology enrichment analysis

Gene ontology (GO) enrichment analyses were conducted with DEGs by the Functional Annotation Tool from DAVID Bioinformatics Resources (v6.8; <https://david.ncifcrf.gov>; Huang et al., 2009)) to systemically describe the DEGs with a focus on biological processes. The results were either listed in tables or visualized using GOplot (Walter et al., 2015).

#### Hub gene analysis

Hub gene analyses were conducted using an R package, WGCNA (Weighted gene co-expression network analysis, v1.66), to identify modules of highly correlated genes with the time sequence within a cell type. Gene significance (GS) of a gene was assessed based on the correlation of the gene with the trait (in this case, the time course). Module eigengene (ME)

groups were defined by module-trait relationship scores ranged from -1 to 1, and the module-trait score closest to 1 represents the most correlated features with the trait. Module membership (MM) of a gene was determined by the correlation of its expression profile with each ME group. Genes that have  $GS > 0.8$  and  $MM > 0.9$  with the most significant ME group were used to build a gene co-expression network within a cell type (Cytoscape, v2.8.3).

#### Upstream pathway prediction

Upstream regulators identified computationally as having an increased potential of being responsible for differential gene expression profiles 1) between two consecutive time points within a cell type and 2) between interzone and anlagen cell types at each time point were predicted using the Ingenuity Pathway Analysis (IPA) software (Qiagen). The differential expression data input cutoffs for this analysis were  $\log_2FC > |0.3|$  and FDR adjusted P-values  $< 0.05$ . From the analysis results, when a P-value of overlap was less than 0.05 and an activation z-score was greater than or equal to  $|2|$ , the predicted genes were likely to be upstream regulators which distinguished between two conditions being compared: 1) a time point vs. its previous time point within a cell type or 2) interzone cell cultures vs. anlagen cell cultures at each time point.

## Results

Sequencing depth of all samples (90 samples total, reflecting 3 cell types × 5 time points × 6 biological replicates) ranged from 29 – 45 million paired-reads per sample (average of 35 million paired-reads/sample; Table 3.1). Phred scores were greater than 30 at all base positions in all 90 samples (base call accuracy > 99.9%; Figure 3.1). With an average mapping efficiency of 95.0%, the sequencing of the entire sample set resulted in well-balanced mapping rates onto the latest version of the equine reference genome (EquCab 3.0; Kalbfleisch et al., 2018; Table 3.2).

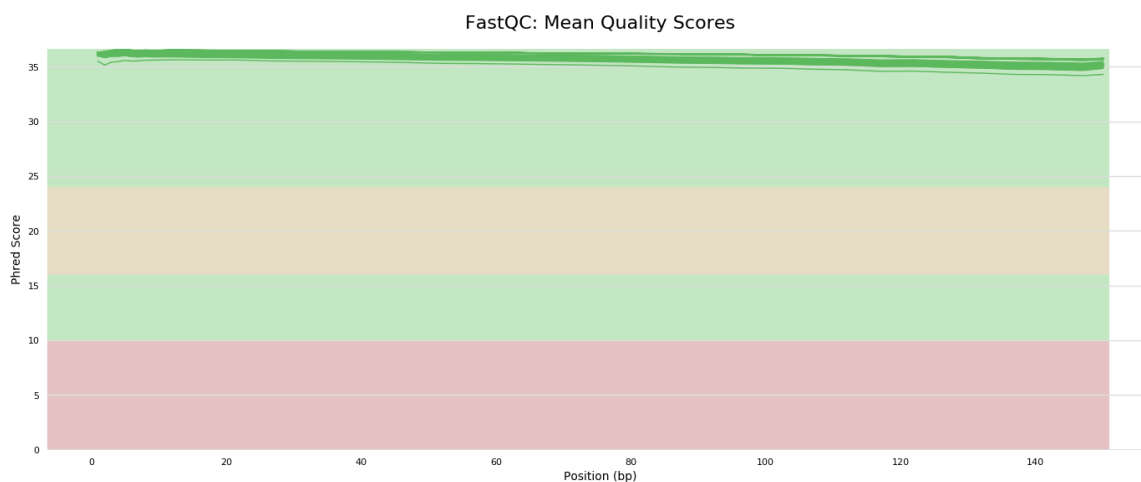


Figure 3.1. Phred scores of the 90 samples assessed by FastQC v0.11.9. Each green line represents each sample.

Table 3.1. Sequencing depth of all samples, million paired-reads/sample

Time point Biological replicate	Interzone cell cultures					Anlagen cell cultures					Fibroblast cultures				
	0h	1.5h	3h	12h	96h	0h	1.5h	3h	12h	96h	0h	1.5h	3h	12h	96h
H1	31	40	36	34	35	39	34	36	34	33	29	33	45	35	40
H2	34	35	35	37	37	35	40	30	34	35	32	38	32	36	38
H5	32	34	33	38	33	38	40	37	37	32	33	35	39	40	35
H6	35	35	32	42	36	34	43	37	33	33	34	35	33	34	32
H9	32	39	38	34	32	37	37	32	34	37	36	40	35	39	37
H11	39	38	43	39	40	34	36	35	32	34	34	38	31	30	36

Table 3.2. Mapping efficiency of all samples, %

Time point Biological replicate	Interzone cell cultures					Anlagen cell cultures					Fibroblast cultures				
	0h	1.5h	3h	12h	96h	0h	1.5h	3h	12h	96h	0h	1.5h	3h	12h	96h
H1	95.2	95.1	95.0	95.1	95.1	95.1	94.8	94.9	95.0	95.1	95.6	95.3	95.3	94.9	94.6
H2	95.2	95.4	95.4	95.0	95.2	95.2	95.4	95.5	95.4	95.0	95.2	95.4	95.2	95.0	94.7
H5	95.5	94.9	95.1	95.4	94.7	95.4	95.2	95.4	95.4	95.0	94.4	95.3	95.3	94.7	94.9
H6	94.6	94.7	94.9	95.0	94.4	94.0	95.0	94.9	94.7	94.9	95.2	95.1	94.5	94.4	94.6
H9	95.4	95.3	95.4	95.2	95.1	95.3	95.1	94.5	95.3	95.0	95.1	95.3	94.1	94.8	94.8
H11	95.2	95.2	95.1	95.0	94.9	93.2	94.8	95.0	94.9	94.5	95.1	94.7	95.4	94.6	94.7



### ***Kinetics of gene expression profiles within a cell type***

Differentially expressed genes were evaluated between paired time point comparisons (0h vs. 1.5h; 1.5h vs. 3h; 3h vs. 12h; 12h vs. 96h) within a cell type. The numbers of DEGs determined at a given time point compared to a previous time point are shown in Table 3.3.

The following subsections describe the kinetics of gene expression profiles in each cell type evaluated by diverse analyses. First, expression patterns of total DEGs across the entire experimental period were visualized. Also, GO enrichment analysis was conducted, and the five most overrepresented biological process terms in each subsequent time comparison are presented. Hub genes and their co-expression network were defined based on gene expression patterns over time. Finally, upstream regulators were predicted in each time sequence comparison.

Table 3.3. Number of differentially expressed genes (DEGs) within cell type when compared to previous time point

Cell type	Interzone cell		Anlagen cell		Fibroblast	
Total unique DEG entities across the entire experimental period <sup>a</sup>	2,822		3,784		3,212	
Regulation pattern	UP	Down	UP	Down	UP	Down
Time point comparison						
1.5h based on 0h	141	29	110	16	96	10
3h based on 1.5h	304	207	234	133	289	118
12h based on 3h	858	696	1,031	965	988	838
96h based on 12h	691	742	1,091	1,304	830	1,027
Total unique DEG entities across the entire experimental period in each regulation pattern <sup>b</sup>	1,861	1,532	2,242	2,252	2,026	1,861

<sup>a</sup>The numbers include both upregulated and downregulated DEGs

<sup>b</sup>The numbers include DEGs within each regulation pattern, either upregulated or downregulated.

## 1) Interzone cell analyses

### 1-a) Expression patterns of DEGs during the experimental period

Along the time course, interzone cell cultures resulted in 2,822 DEGs: 1,816 genes were upregulated, 1,532 genes were downregulated, and 538 genes changed relative directional orientation (up- or downregulation) over the experimental period. While the gene expression pattern changes between the baseline and the first collection point (1.5h) were not separated, a clear distinction of transcriptomic patterns started to be observed by 3h (Figure 3.2.a). Some genes that were expressed at a higher level during earlier time points became downregulated at later time points, and the opposite was also observed. By 96h, the differential gene expression profiles became almost reversed from the baseline. Towards the later time points, a greater number of genes were differentially expressed at a time point compared to its previous time point (Figure 3.2.b and c). Chronologically closer time comparisons shared more common DEGs (Figure 3.2.b and c).

### 1-b) Gene enrichment analysis

The five most upregulated and downregulated overrepresented biological processes for each time comparison are listed in Table 3.4 and 3.5, respectively. At the earlier time sequences, transcription regulatory events were most significantly represented by the profiles of both upregulated and downregulated DEGs. Towards the later time points, biological processes involved in chondrogenic differentiation became distinctive among upregulated DEGs.

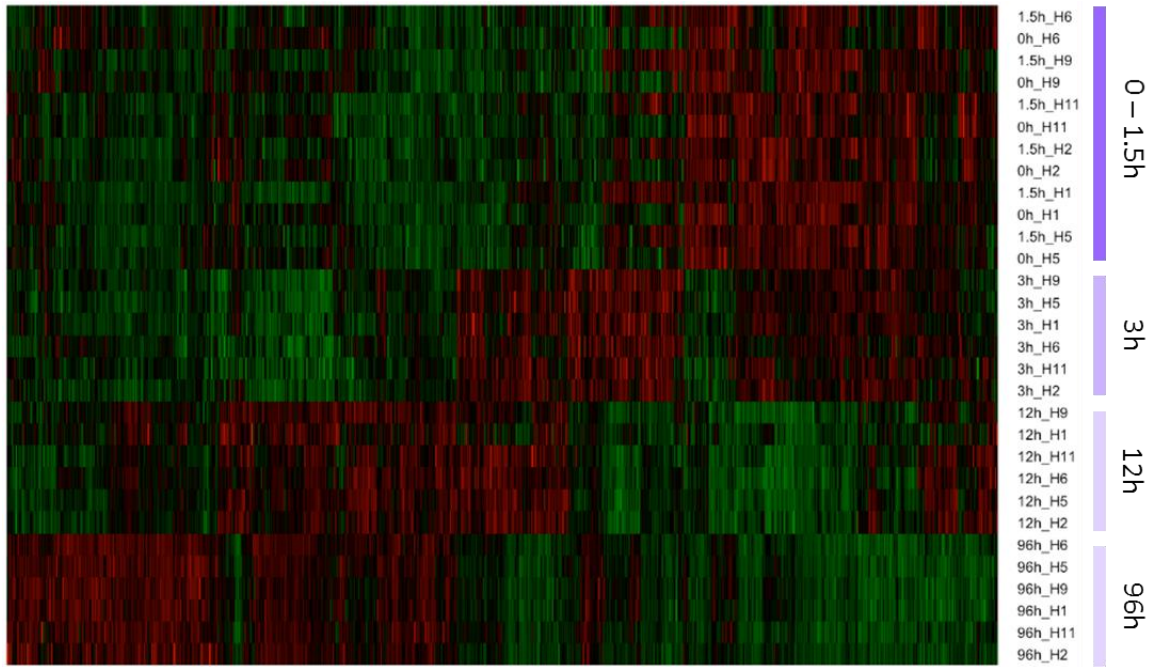
#### 1-c) Hub genes and co-expression network

The most significant gene correlated with the time course was FOLR2 (GS=0.99; P-value=2.2E-24), and the top ten genes that were highly correlated with the time sequence are listed in Table 3.6. All of the ten most significant genes were categorized into the Turquoise group by ME identification. Among the 27 ME groups evaluated by WGCNA, the Turquoise group had the greatest module-trait relationship (0.96; P-value=1E-17) with the time course in interzone cell cultures (Figure 3.3). The co-expression network of the genes, which were highly correlated with the Turquoise group, is exhibited in Figure 3.4. The gene that had the highest MM with the Turquoise group was OLFML3 (MM=0.98; P-value=7.1E-22). Based on existing knowledge, however, interactions of OLFML3 with the other members in the Turquoise group were not demonstrated, therefore not shown in Figure 3.4.

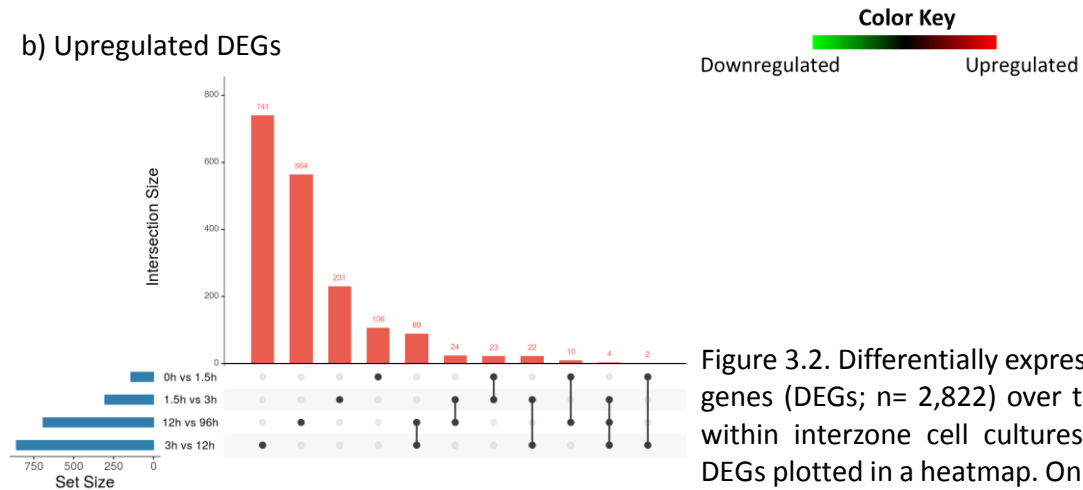
#### 1-d) Upstream regulator prediction

During the first 1.5 hours, a smaller number of upstream regulators were predicted in both activation and inhibition status (Figure 3.5). The greatest number of computationally predicted upstream regulators were observed between 3h and 12h. It is interesting to note that TGFB1 was predicted to be an upstream regulator from the baseline to 1.5h (z-score=4.95; P-value=2.8E-39) while the major chondrogenic factor that was used in the experiment was the protein encoded by TGFB1.

a) Heatmap



b) Upregulated DEGs



c) Downregulated DEGs

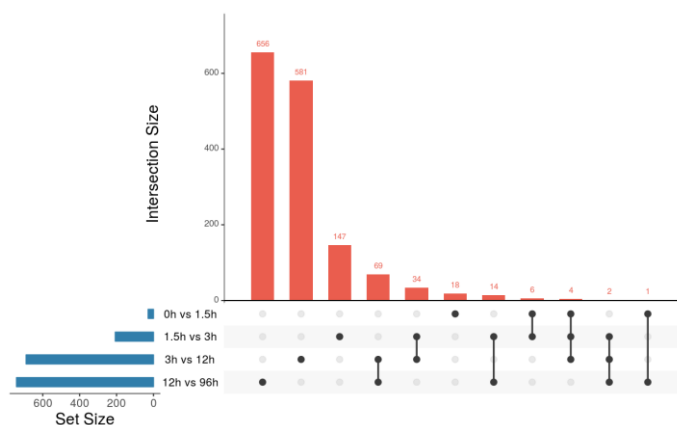


Figure 3.2. Differentially expressed genes (DEGs; n= 2,822) over time within interzone cell cultures. a) DEGs plotted in a heatmap. On the right side, time points (#h) and biological replicates (H#) are labeled. b) UpSet figure of upregulated DEGs (n=1,816) between time points. c) UpSet figure of downregulated DEGs (n=1,532) between time points. The set size graphs of the UpSet figures represent the number of DEGs in each time point comparison. Note that each comparison has a different set size.

Table 3.4. Top five biological process gene ontology (GO) terms of differentially upregulated genes in interzone cell cultures during time courses

Time course	GO ID	Term	P-value <sup>a</sup>	Differentially expressed genes from the dataset
0h vs 1.5h	GO:0000122	Negative regulation of transcription from RNA polymerase II promoter	7.51E-07	DLX2, PLK3, EZR, IRF2BPL, NRARP, SMAD7, EFNA1, RELB, NR4A2, LMCD1, PER1, FOXC2, ID3, SOX9, JUNB, HIC1
	GO:0045944	Positive regulation of transcription from RNA polymerase II promoter	1.92E-06	IL6, RELB, F2RL1, CCNL1, NFKBIA, NR4A1, DLL1, FADD, SOX9, JUNB, IL11, LIF, ADRB2, IRF2BPL, VEGFA, ZC3H12A, PER1, SERTAD1
	GO:0043410	Positive regulation of MAPK cascade	4.11E-06	LIF, TNFRSF1B, ADRB2, IL6, RELT, PDGFB, IL11
	GO:0045892	Negative regulation of transcription, DNA-templated	4.91E-05	GCLC, CRY2, CEBPB, PDGFB, CCDC85B, KLF10, RELB, IRF1, ID3, SOX9
	GO:1900745	Positive regulation of p38mapk cascade	9.71E-05	GADD45G, ZC3H12A, GADD45B, GADD45A
1.5h vs 3h	GO:0006954	Inflammatory response	4.67E-05	TLR10, CCL2, ELF3, TNFRSF25, CSF1, RELB, AFAP1L2, NFKB1, TGFB1, CALCB, S1PR3, TNFRSF1B, CCL20, CYP26B1
	GO:0060412	Ventricular septum morphogenesis	1.69E-04	NOTCH1, HEYL, WNT11, ZFPM1, GJA5
	GO:0007219	Notch signaling pathway	1.22E-03	S1PR3, NRARP, DTX1, HEYL, TMEM100, TGFB1, ANGPTL4
	GO:0003151	Outflow tract morphogenesis	1.66E-03	HEYL, WNT11, ZFPM1, PLXND1, GJA5
	GO:0071407	Cellular response to organic cyclic compound	1.87E-03	ALPL, RGS20, CCL2, GLI2, TGFB1
3h vs 12h	GO:0060021	Palate development	5.51E-04	MEF2C, DHRS3, SATB2, WDPCP, MEOX2, HAND2, DLX5, ARID5B, INSIG1, IFT172, SNAI2
	GO:0050680	Negative regulation of epithelial cell proliferation	1.03E-03	MEF2C, MTSS1, IFT122, CDKN2B, IFT172, GDF5, MCC, TINF2
	GO:0071498	Cellular response to fluid shear stress	3.47E-03	MEF2C, MTSS1, PTGS2, HAS2
	GO:0007155	Cell adhesion	6.13E-03	HAPLN2, HAPLN1, HAPLN4, ICAM5, POSTN, NCAM1, ITGA9, COMP, TGFB1, CNTN2, ACAN, HAS2, HRC, SPP1
	GO:0019221	Cytokine-mediated signaling pathway	9.30E-03	FLRT3, ASPN, STAT4, IL22RA1, LRRC4B, EPOR, DCN, LRRC15, GHR
12h vs 96h	GO:0001958	Endochondral ossification	5.88E-07	BMP4, CSGALNACT1, FGF18, DLX5, COL2A1, MMP16, MMP13, SCX
	GO:0070374	Positive regulation of ERK1 and ERK2 cascade	1.47E-06	BMP4, FGF18, BMP2, CCL2, F2RL1, CHI3L1, CCL19, FGF10, CD74, CCL11, SPRY2, ARRB1, PLA2G2A, ANGPT1, PDGFD, PLA2G5, F2R
	GO:0002062	Chondrocyte differentiation	8.56E-06	BMP4, BMP2, WNT5B, OSR2, GDF5, COL2A1, COL11A2, SCX
	GO:0032331	Negative regulation of chondrocyte differentiation	1.61E-05	BMP4, RARG, GDF5, NKX3-2, SNAI2, GREM1
	GO:0090090	Negative regulation of canonical Wnt signaling pathway	2.03E-05	NKD1, NOG, AMER1, BMP2, WNT5B, DACT1, PRICKLE1, SNAI2, FRZB, WWTR1, GREM1, GLI3, GLI1

<sup>a</sup>Modified Fisher Exact P-Value was used to evaluate enrichment of genes in GO annotation terms.

Table 3.5. Top five biological process gene ontology (GO) terms of differentially downregulated genes in interzone cell cultures during time courses

Time course	GO ID	Term	P-value <sup>a</sup>	Differentially expressed genes from the dataset
0h vs 1.5h	GO:0001759	Organ induction	0.016	SPRY1, SIX1
	GO:0006355	Regulation of transcription, DNA-templated	0.018	EGR1, SIX1, CITED2, ZNF2
	GO:0042310	Vasoconstriction	0.023	HTR1B, EDN1
	GO:0001657	Ureteric bud development	0.044	SPRY1, SIX1
	GO:0043407	Negative regulation of MAP kinase activity	0.049	SPRY1, RGS2
1.5h vs 3h	GO:0035914	Skeletal muscle cell differentiation	3.20E-08	MEF2C, FOS, EGR2, ATF3, BTG2, NR4A1, FOXN2, ANKRD1, CITED2
	GO:0000122	Negative regulation of transcription from RNA polymerase II promoter	1.08E-06	MEF2C, ATF7IP, SATB2, FGF9, EDN1, KLF11, NR4A2, CBX4, NR4A3, GLI3, DLX2, OSR2, REL, IRF2BPL, TRPS1, JUN, ID4, SKIL, LCOR
	GO:0045444	Fat cell differentiation	2.81E-06	EGR2, ARID5B, GDF6, NR4A2, NR4A1, ID4, NR4A3, KLF4
	GO:0008584	Male gonad development	3.65E-06	INHBA, COL9A3, FGF9, ARID5B, BCL2, KITLG, LHX9, CITED2
	GO:0060021	Palate development	1.28E-05	MEF2C, INHBA, SATB2, OSR2, ARID5B, CSRNP1, COL2A1, GLI3
3h vs 12h	GO:0043066	Negative regulation of apoptotic process	1.82E-05	IER3, IL6, EGR3, ACTC1, LIMS2, NUAQ2, SOCS3, SMAD6, BTC, ASNS, GLI2, IGF1R, ATF5, SPRY2, STK40, PTK2B, ID1, CHST11, THOC6, AVEN, RARA, MYC
	GO:0030335	Positive regulation of cell migration	1.05E-04	FLT1, WNT5B, PDGFB, LYN, TGFBR1, FAM110C, EDN1, MYADM, CCL26, CCL11, IGF1R, NOTCH1, SEMA3E, CEMIP, PLAU
	GO:0035994	Response to muscle stretch	3.26E-04	FOS, SLC8A1, NFKBIA, NFKB1, ANKRD1
	GO:0035556	Intracellular signal transduction	3.68E-04	SGK1, SOCS2, LYN, NUAQ2, SPSB1, TGFBR1, PRKAG2, PRKCI, CXCL8, PRKCG, RPS6KA5, HUNK, RASSF5, STAC, TIAM2, CDC42BPA, GUCY1A2, SH2B3, RGS7, NIM1K, PAG1, TEC, RASA2
	GO:0006954	Inflammatory response	5.46E-04	TNFRSF21, TLR10, IL6, CCL2, ELF3, LYN, IL18, CSF1, TLR1, ACKR1, ANXA1, CXCL8, NFKB1, BDKRB1, CCL26, CCL11, TNFRSF1B, CCL20, RELT
12h vs 96h	GO:0001666	Response to hypoxia	2.17E-06	CAV1, CRYAB, EGLN3, TGFBR3, SMAD3, EGLN1, DDIT4, VEGFC, HSP90B1, LONP1, MYOCD, VEGFA, LOXL2, ALKBH5
	GO:0006096	Glycolytic process	2.57E-05	GPI, TPI1, ALDOC, PGAM1, HK1, PGK1, GAPDH, ENO1
	GO:0007155	Cell adhesion	1.22E-04	B4GALT1, ATP1B1, ACHE, ICAM5, PCDH10, ITGA1, ACKR3, ITGA3, TINAGL1, ADGRG1, ITGAX, ITGA8, SULF1, CNTN2, GP1BA, CNTN4, THBS2
	GO:0002931	Response to ischemia	1.33E-04	HYOU1, CAV1, UCHL1, CAMK2A, FAIM2, CIB1
	GO:0007507	Heart development	1.47E-04	PDGFB, FOXJ1, SOD2, GATA2, ACVR2B, ADAP2, ECE1, ADM, OSR1, JMJD6, GYS1, EPOR, LOX, BCOR

<sup>a</sup>Modified Fisher Exact P-Value was used to evaluate enrichment of genes in GO annotation terms.

Table 3.6. Ten most highly correlated hub genes with time course in interzone cell cultures during entire experimental period

Gene	ME <sup>a</sup> group	Gene significance	P-value	MM <sup>b</sup> to Turquoise	P-value
FOLR2	Turquoise	0.988	2.19E-24	0.954	3.80E-16
ITGA10	Turquoise	0.983	5.42E-22	0.942	7.88E-15
PRELP	Turquoise	0.981	1.45E-21	0.967	4.23E-18
LOXL4	Turquoise	0.981	1.50E-21	0.931	8.56E-14
KCNN4	Turquoise	0.977	1.95E-20	0.955	2.87E-16
FMOD	Turquoise	0.976	4.26E-20	0.944	5.52E-15
COL24A1	Turquoise	0.975	6.34E-20	0.972	3.31E-19
PODNL1	Turquoise	0.971	4.93E-19	0.955	2.26E-16
GABRE	Turquoise	0.971	7.01E-19	0.960	5.97E-17
AEBP1	Turquoise	0.969	1.32E-18	0.958	8.74E-17

<sup>a</sup>ME, Module eigengene

<sup>b</sup>MM, Module membership

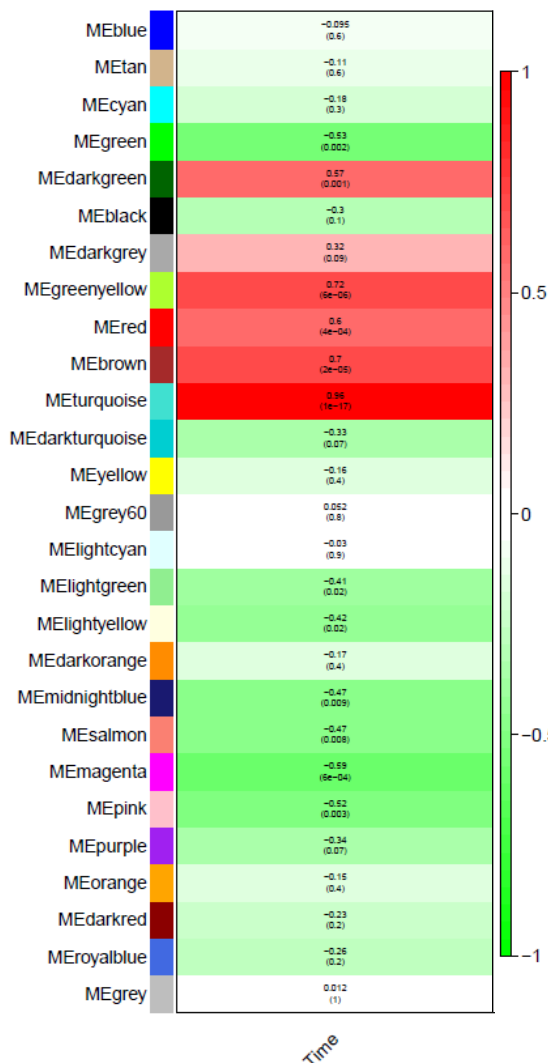


Figure 3.3. Module eigengene (ME) groups with module-trait relationship scores in interzone cell cultures over time. P-values are shown in the parentheses. The higher relationship score represents the greater correlation between a ME group and the time sequence.

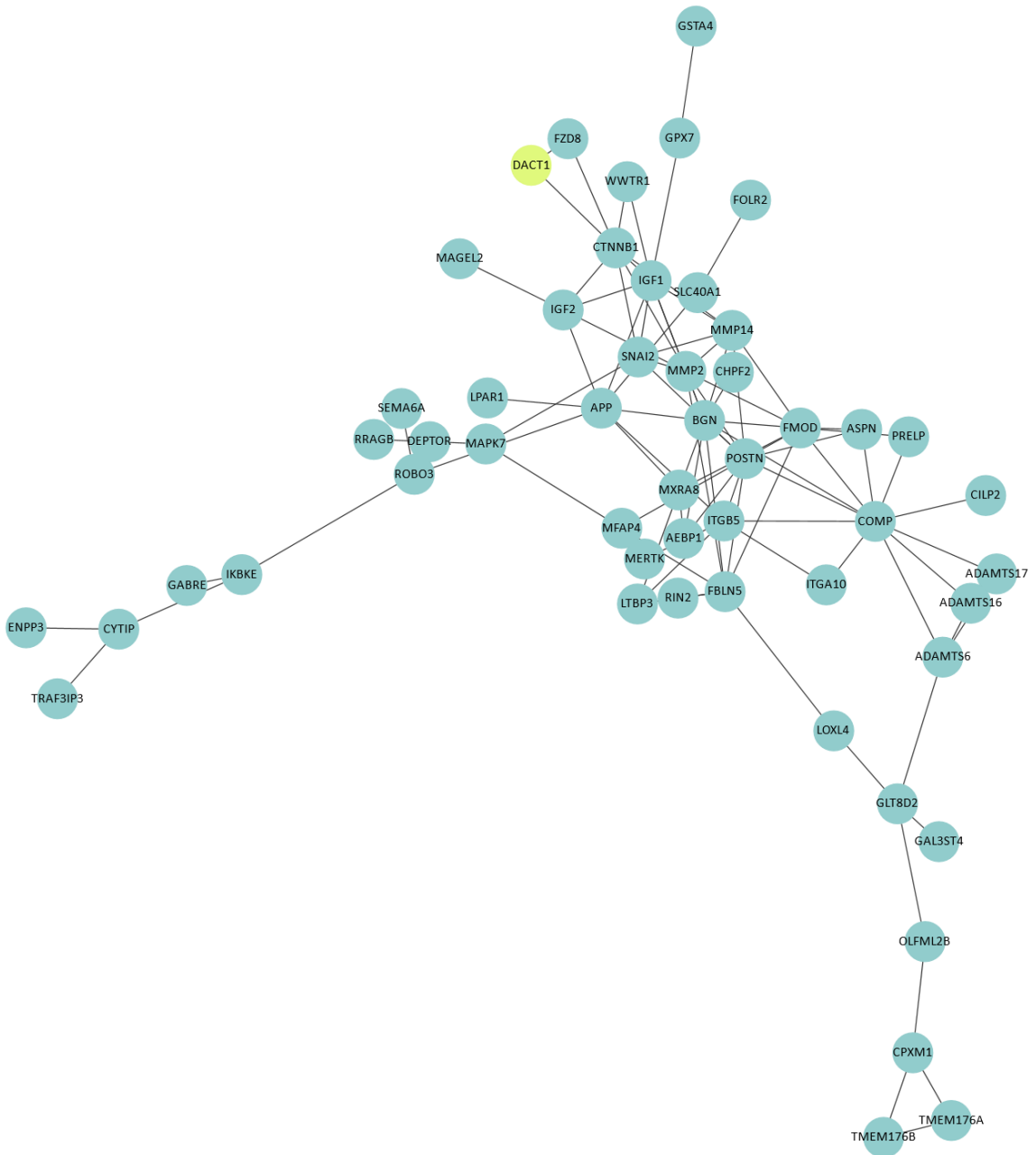


Figure 3.4. Hub gene co-expression network in interzone cell cultures during 96 hours of *in vitro* chondrogenesis. Colored by module eigengene group.



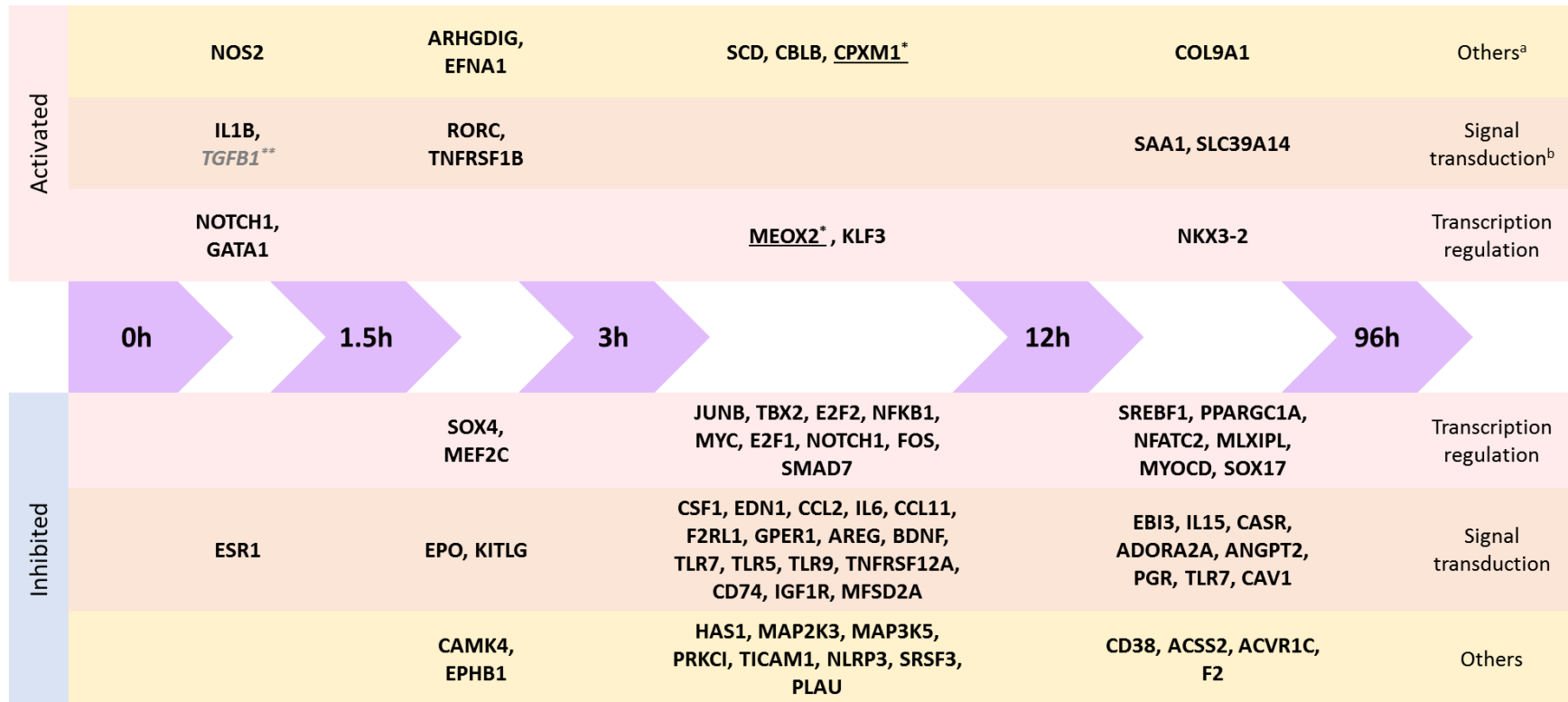


Figure 3.5. Upstream regulators predicted by Ingenuity Pathway Analysis (IPA) between subsequent time points in interzone cell cultures. The activation status thresholds were activation z-score > |2| and P-value of overlap < 0.05. Only genes that have log<sub>2</sub> fold change (FC) ≥ |1| in the dataset were presented in the figure. Positive z-scores represent activated upstream regulators, and negative values indicate inhibition. When a predicted activation status (activation or inhibition) matched the regulation pattern analyzed from the dataset (positive log<sub>2</sub>FC=upregulation; negative log<sub>2</sub>FC=downregulation), the prediction was accepted. <sup>a</sup>Others include enzymes, kinases, and others. <sup>b</sup>Signal transduction includes growth factors, ligands, receptors, and cytokines.

\*CPMX1 and MEOX2 are also hub genes in interzone cell cultures whose expression patterns are highly correlated with the time course.

\*\*In the data, log<sub>2</sub>FC of TGFB1 from 0h to 1.5h was 0.94.

## 2) Anlagen cell analyses

### 2-a) Expression patterns of DEGs during the experimental period

Over the entire experimental period, 3,784 DEGs were observed in Anlagen cell cultures: 2,242 genes were upregulated, 2,252 genes were downregulated, and 710 genes were overlapped between the lists of upregulated and downregulated DEGs. The transcriptomic profiles were not discriminative from the baseline until 3h (Figure 3.6.a). Of note, one biological replicate (Horse #6) at 0h and 1.5h had gene expression patterns more closely clustered to those of its own at 3h than those of the other biological replicates at 0h and 1.5h. Except for this case, the overall gene expression profiles were clustered by time point. A clear distinction of the transcriptomic characteristics was observed by 12h, and the initial gene expression patterns became almost opposite by 96h. Similar to interzone cell cultures, more common DEGs were recorded between chronologically closer time point comparisons, and fewer common DEGs were shared between time point comparisons with greater time gaps (Figure 3.6.b, and 6.c).

### 2-b) Gene enrichment analysis

The five most overrepresented biological processes at each time sequence analyzed from the upregulated and downregulated DEGs in Anlagen cell cultures are listed in Table 3.7 and 3.8, respectively. Only 3 biological processes were found to be over-represented from the downregulated DEGs at 1.5h based on 0h. During the first 1.5 hours, transcription regulatory events were most significantly upregulated. Regulatory processes in signaling cascades were observed during the entire experimental period. Finally, biological processes related to ECM metabolism were overrepresented in the last time comparison (between 12h and 96h).

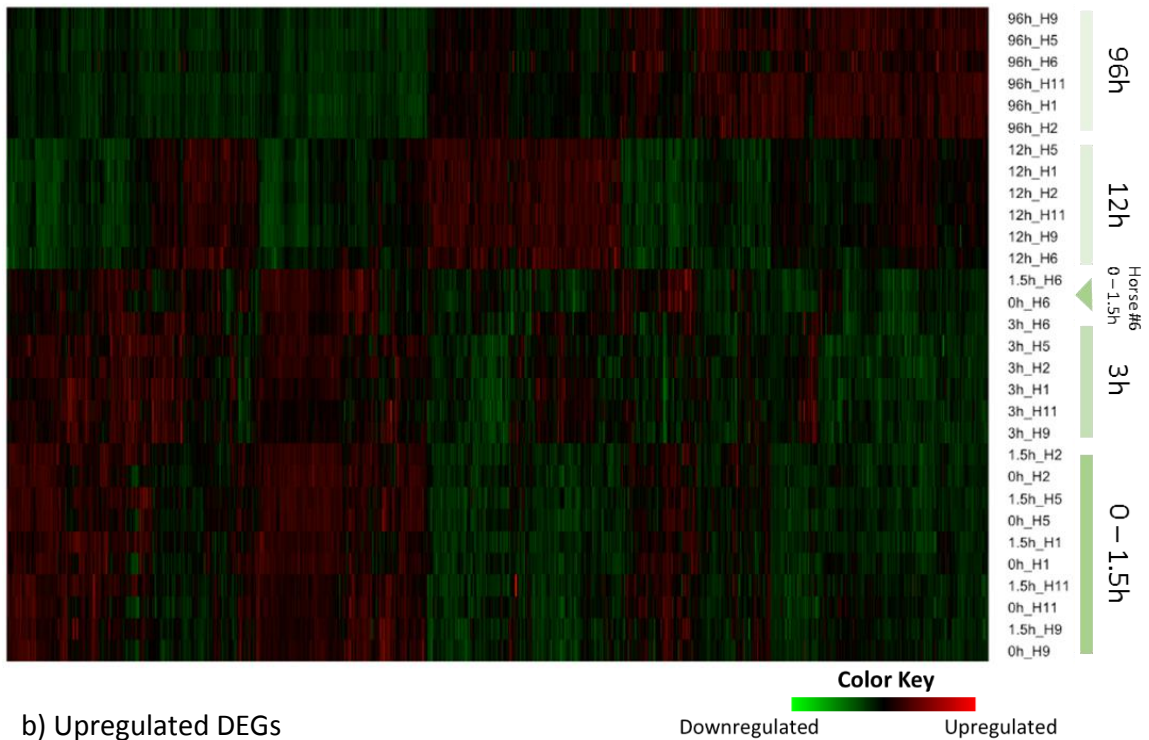
## 2-c) Hub genes and co-expression network

The most significant hub gene correlated with the time sequence in analgen cell cultures was SEMA6A (GS=0.99; P-value=6.4E-25), with the top ten listed in Table 3.9. The ME identification categorized these ten genes into the Brown group, and this group had the highest module–trait relationship (0.96; P-value=2E-17) with the time course in analgen cell cultures (Figure 3.7). The co-expression network of the hub genes that were greatly correlated with the Brown group is shown in Figure 3.8. The gene that showed the highest MM with the Brown group was GXYLT2 (MM=0.99; P-value=1.6E-24), however, interactions of this gene with the other members in the Brown group are unknown. FMOD had the second highest MM with the Brown group (MM=0.99; P-value=3.2E-23) and centered in the co-expression network, interacting with 9 members in the Brown group.

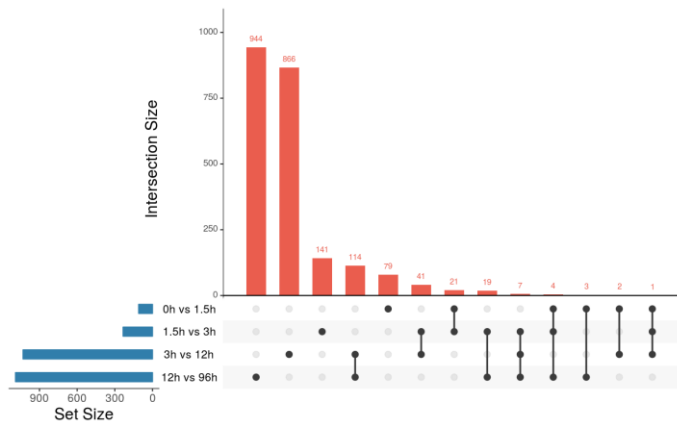
## 2-d) Upstream regulator prediction

Similar to interzone cell cultures, greater upstream regulators were predicted towards the later time sequences in both activation and inhibition status (Figure 3.9). Also, TGFB1 was predicted to be an activated upstream regulator during the first 1.5 hours (z-score=4.19; P-value=9.4E-36). However, during the last time sequence (12h – 96h), inhibition of TGFB1 was predicted (z-score=-3.45; P-value=5.7E-52).

a) Heatmap



b) Upregulated DEGs



c) Downregulated DEGs

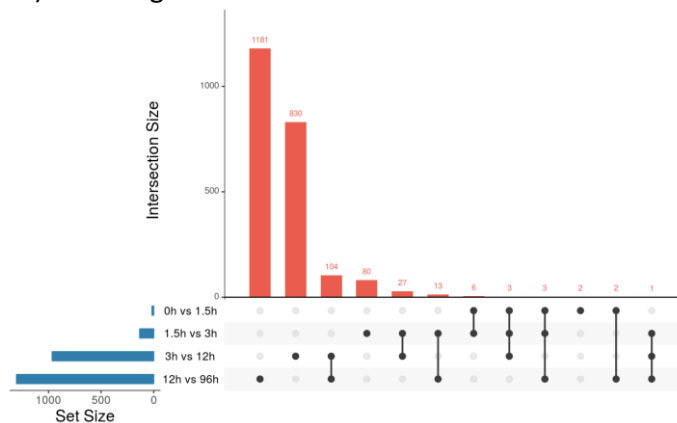


Figure 3.6. Differentially expressed genes (DEGs; n=3,784) over time within anlagen cell cultures. a) DEGs plotted in a heatmap. On the right side, time points (#h) and biological replicates (H#) are labeled. b) UpSet figure of upregulated DEGs (n=2,242) between time points. The set size graph represents the number of DEGs in each time point comparison. c) UpSet figure of downregulated DEGs (n=2,252) between time points. The set size graphs of the UpSet figures represent the number of DEGs in each time point comparison. Note that each time comparison has a different set size.

Table 3.7. Top five biological process gene ontology (GO) terms of differentially upregulated genes in anlagen cell cultures during time courses

Time course	GO ID	Term	P-value <sup>a</sup>	Differentially expressed genes in the dataset
0h vs 1.5h	GO:0000122	Negative regulation of transcription from RNA polymerase II promoter	2.21E-06	DLX2, PLK3, IRF2BPL, NRARP, SMAD7, EFNA1, NR4A2, LMCD1, PER1, FOXC2, ID3, SOX9, JUNB, HIC1
	GO:0045892	Negative regulation of transcription, DNA-templated	9.95E-06	CRY2, CEBPB, PDGFB, ELF3, CCDC85B, KLF10, IRF1, ID3, SOX9, SOX8
	GO:0045944	Positive regulation of transcription from RNA polymerase II promoter	1.48E-05	CRTC1, NFKBIA, NR4A1, DLL1, FADD, SOX9, SOX8, JUNB, IL11, LIF, IRF2BPL, VEGFA, ZC3H12A, PER1, NFATC1
	GO:1900745	Positive regulation of p38mapk cascade	5.39E-05	GADD45G, ZC3H12A, GADD45B, GADD45A
	GO:0043153	Entrainment of circadian clock by photoperiod	1.17E-04	CRY2, PER1, BHLHE40, SIK1
1.5h vs 3h	GO:0071456	Cellular response to hypoxia	3.22E-04	PDGFB, HMOX1, VEGFA, FAM162A, AQP1, ANGPTL4
	GO:0048589	Developmental growth	0.002	ADM, CHST11, SMAD3, GLI2
	GO:0006355	Regulation of transcription, DNA-templated	0.002	ZNF566, PPARG, HR, ZNF175, ABCG1, ZNF2, RGS20, ZNF691, JMJD6, ZSCAN20, HEYL, ZNF404, KDM3A, LIMD1, ALX3, CSDC2, NFATC1
	GO:0030206	Chondroitin sulfate biosynthetic process	0.007	CHST11, CHST3, CHSY1
3h vs 12h	GO:0019221	Cytokine-mediated signaling pathway	5.70E-04	FLRT3, ASPN, IL22RA1, LRRTM1, LRRC4B, GP1BA, EPOR, DCN, LRRC15, GREM2, CISH, GHR
	GO:0046426	Negative regulation of JAK-STAT cascade	7.06E-04	FLRT3, ASPN, LRRTM1, LRRC4B, GP1BA, DCN, LRRC15, CISH
	GO:0006469	Negative regulation of protein kinase activity	0.003	FLRT3, ASPN, LRRTM1, LRRC4B, TRIB3, GP1BA, DCN, LRRC15, CISH, TRIB2
	GO:0035023	Regulation of Rho protein signal transduction	0.003	PLEKHG4, DLC1, OBSCN, PLEKHG1, ARHGEF26, ARHGEF25, TIAM1, PLEKHG5, ARHGEF9, EPS8L1, FGD4
	GO:0007507	Heart development	0.003	ZFP36L1, ADAP2, ADM, FOXJ1, TRPS1, COL3A1, ACAN, GYS1, NFATC4, EPOR, CACNA1C, CXADR, SOD2
12h vs 96h	GO:0030574	Collagen catabolic process	7.22E-06	CTSK, MMP9, MMP16, CTSS, MMP14, MMP13, MMP2
	GO:0032331	Negative regulation of chondrocyte differentiation	1.94E-05	CHADL, RARG, GDF5, NKX3-2, GLI2, SNAI2, SOX9, GREM1
	GO:0090090	Negative regulation of canonical Wnt signaling pathway	2.90E-05	NOG, BMP2, FOXO1, RGS19, LEF1, SOX9, GREM1, FRZB, SNAI2, GLI1, DKK2, AMER1, WNT4, DACT1, PRICKLE1, SOSTDC1
	GO:0030198	Extracellular matrix organization	8.37E-05	MPZL3, EGFL6, MMP9, OLFML2B, ADAMTSL4, POSTN, SOX9, CSGALNACT1, SMOC2, TNFRSF11B, FBLN1, COL27A1, KAZALD1
	GO:0001837	Epithelial to mesenchymal transition	2.21E-04	NOG, WNT4, HIF1A, LEF1, LOXL3, SOX9

<sup>a</sup>Modified Fisher Exact P-Value was used to evaluate enrichment of genes in GO annotation terms.

Table 3.8. Top five biological process gene ontology (GO) terms of differentially downregulated genes in anlagen cell cultures during time courses

Time course	GO ID	Term	P-value <sup>a</sup>	Differentially expressed genes in the dataset
0h vs 1.5h	GO:0035914	Skeletal muscle cell differentiation	6.65E-04	FOS, EGR2, CITED2
	GO:0070373	Negative regulation of ERK1 and ERK2 cascade	0.028	SPRY1, DUSP6
	GO:0007179	Transforming growth factor beta receptor signaling pathway	0.034	FOS, CITED2
1.5h vs 3h	GO:0035914	Skeletal muscle cell differentiation	9.84E-07	EGR1, FOS, ATF3, EGR2, BTG2, MYF5, ANKRD1
	GO:0008543	Fibroblast growth factor receptor signaling pathway	8.15E-05	FLRT3, FGF5, FLRT2, FGF7, FGF21
	GO:0006355	Regulation of transcription, DNA-templated	8.22E-04	HES1, EGR1, HHEX, EGR3, EGR2, MEOX2, ARID5B, MYF5, VGLL2, TLE1, NPAS4, MEIS1, LHX9
	GO:0032496	Response to lipopolysaccharide	0.008	CEBPB, PTGER4, DUSP10, TRIB1
	GO:0045444	Fat cell differentiation	0.009	EGR2, ARID5B, NR4A2, KLF4
3h vs 12h	GO:0006270	DNA replication initiation	2.80E-08	CCNE2, CCNE1, CDC45, MCM7, POLA2, MCM2, MCM3, MCM10, MCM4, MCM5, MCM6
	GO:0030335	Positive regulation of cell migration	4.36E-06	WNT5B, FLT1, PDGFB, LYN, FAM110C, TGFBR1, EDN1, F2RL1, FER, CCL26, IGF1R, NOTCH1, SEMA3E, SEMA3D, ADRA2A, CEMIP, HAS2, PDGFC, PDGFD, PLAU, F2R
	GO:0045740	Positive regulation of DNA replication	8.47E-06	IGF1R, DNA2, PDGFB, PCNA, FGF10, KITLG, PDGFC, AREG, GLI2
	GO:0006364	rRNA processing	7.45E-04	DIS3, EBNA1BP2, CCDC86, RRP1B, RRP9, BOP1, GTF2H5, MDN1, MRTO4
	GO:0032870	Cellular response to hormone stimulus	0.001	CGA, NPFFR2, ADRA2A, OXTR, JUNB, SLIT2, SLIT3
12h vs 96h	GO:0071456	Cellular response to hypoxia	1.13E-06	P4HB, EPAS1, PTGS2, PINK1, PMAIP1, NPEPPS, RORA, PRKCE, HYOU1, HMOX1, VEGFA, MDM2, FAM162A, NDRG1, ANGPT4
	GO:0000070	Mitotic sister chromatid segregation	1.47E-06	CEP57, PLK1, CENPA, SPAG5, NEK2, ZWINT, KIF18A, KIF18B, ESPL1, KNSTRN
	GO:0007059	Chromosome segregation	9.19E-06	CENPN, KIF11, NEK2, NEK10, BIRC5, KNSTRN, BRCA1, ESCO2, SPC25, HJURP, CDCA2, CENPW, SKA3, SKA1
	GO:0007018	Microtubule-based movement	2.77E-05	KIF14, KIF23, KIF22, KIF4A, KIF3A, KIF11, KIF15, KIF18A, KIF18B, CENPE, DNAH2, KIF3C, KIF2C, KIF1B, DYNC1H1, KIF20A
	GO:0007076	Mitotic chromosome condensation	4.92E-05	NCAPH, NCAPG, NUSAP1, CDCA5, NCAPD3, NCAPD2, SMC4

<sup>a</sup>Modified Fisher Exact P-Value was used to evaluate enrichment of genes in GO annotation terms.

Table 3.9. Ten most highly correlated genes with time course in anlagen cell culture during entire experimental period

Gene	ME <sup>a</sup> group	Gene significance	P-value	MM <sup>b</sup> to Brown	P-value
SEMA6A	Brown	0.989	6.36E-25	0.977	3.22E-20
ANKRD35	Brown	0.981	1.33E-21	0.952	5.83E-16
PODNL1	Brown	0.979	8.50E-21	0.932	7.30E-14
AEBP1	Brown	0.978	1.44E-20	0.950	1.20E-15
ENSECAG0000022553	Brown	0.978	1.45E-20	0.962	2.22E-17
PLEKHA4	Brown	0.972	4.80E-19	0.935	4.03E-14
COMP	Brown	0.971	5.05E-19	0.962	2.77E-17
MAPK7	Brown	0.970	1.15E-18	0.948	2.07E-15
CHST6	Brown	0.968	2.15E-18	0.980	3.82E-21
LTBP3	Brown	0.968	2.39E-18	0.951	8.04E-16

<sup>a</sup>ME, Module eigengene

<sup>b</sup>MM, Module membership

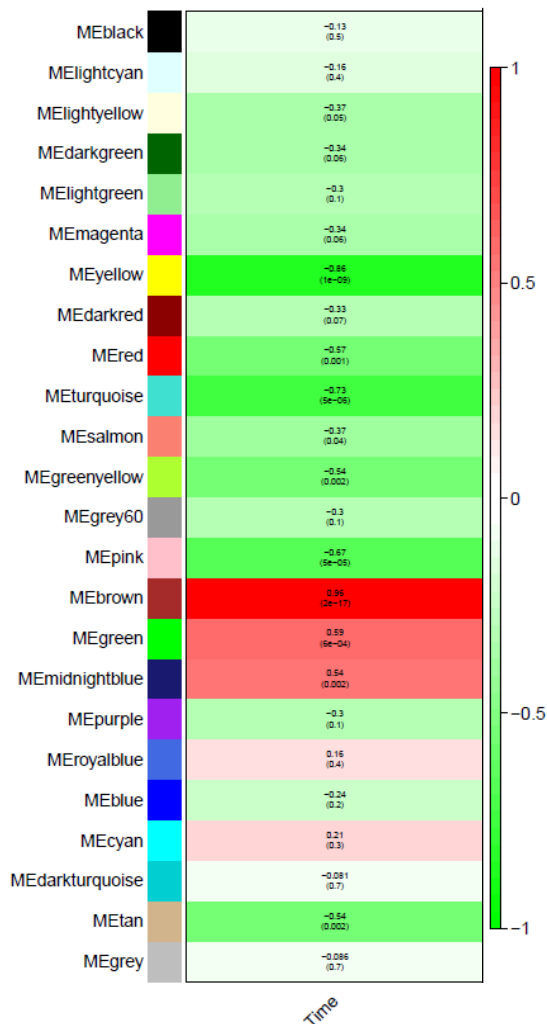


Figure 3.7. Module eigengene (ME) groups with module-trait relationship scores in anlagen cell cultures over time. P-values are shown in parentheses. The higher relationship score represents the greater correlation between a ME group and the time sequence.

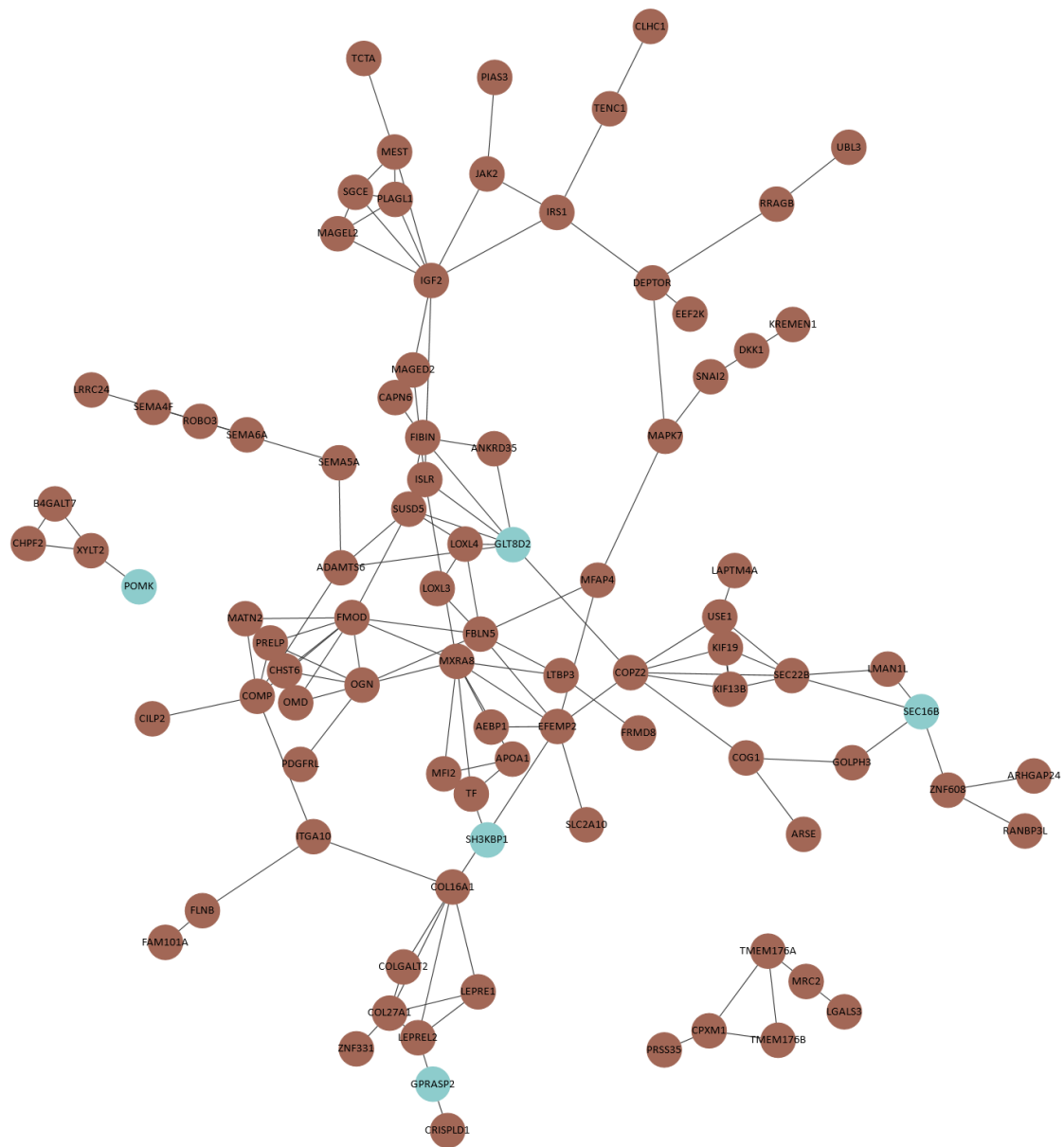


Figure 3.8. Hub gene co-expression network in anlagen cell cultures during 96 hours of *in vitro* chondrogenesis. Colored by module eigengene group.



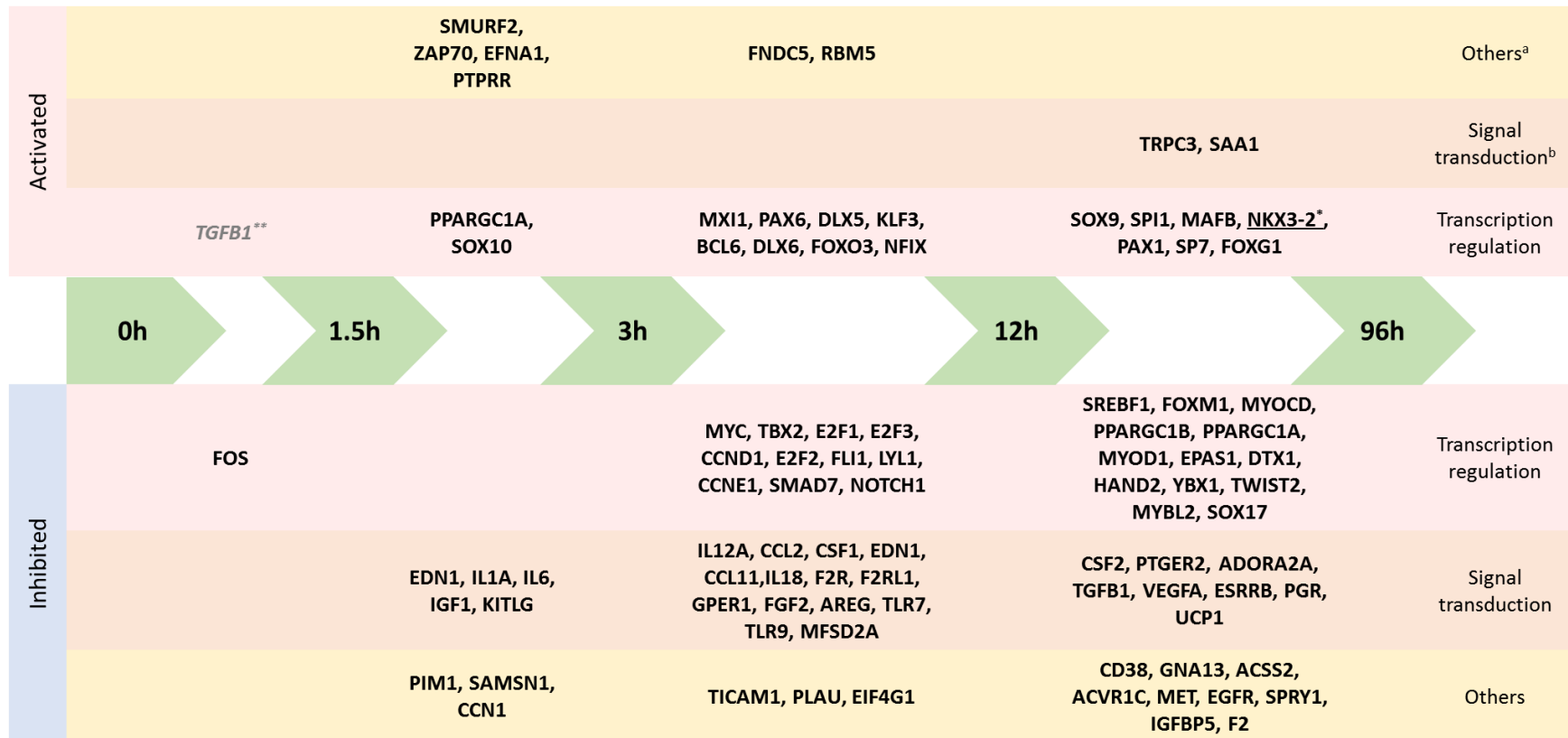


Figure 3.9. Upstream regulators predicted by Ingenuity Pathway Analysis (IPA) between subsequent time points in anlagen cell cultures. The activation status thresholds were activation z-score > |2| and P-value of overlap < 0.05. Only genes that have log<sub>2</sub> fold change (FC) ≥ |1| in the dataset were presented in the figure. Positive z-scores represent activated upstream regulators, and negative values indicate inhibition. When a predicted activation status (activation or inhibition) matched the regulation pattern analyzed from the dataset (positive log<sub>2</sub>FC=upregulation; negative log<sub>2</sub>FC=downregulation), the prediction was accepted. <sup>a</sup>Others include enzymes, kinases, and others. <sup>b</sup>Signal transduction includes growth factors, ligands, receptors, and cytokines.

\*NKX3-2 is also a hub gene in anlagen cell cultures whose expression pattern is highly correlated with the time course.

\*\*In the data, log<sub>2</sub>FC of TGFB1 from 0h to 1.5h was 0.93.

### 3) Fibroblast analyses

#### 3-a) Expression patterns of DEGs during the experimental period

In the fibroblast negative control group 3,212 DEGs were measured over time: 2,026 genes were upregulated, 1,861 genes were downregulated, and 675 genes switched their regulation patterns during the 96 hours. Similar to the other cell cultures, gene expression patterns were less distinctive between the baseline and the first collection point (1.5h), and gradually changed until 3h after the initiation of *in vitro* chondrogenesis (Figure 3.10.a). Notable transcriptomic changes were accumulated towards the later time points. Fibroblast cultures also had a greater number of common DEGs between closer time point comparisons, and fewer DEGs were overlapped between distant time point comparisons (Figure 3.10.b and c).

#### 3-b) Gene enrichment analysis

With an approach similar to the other cell types, overrepresented biological processes in each time sequence were analyzed from upregulated and downregulated DEGs in fibroblast cultures, and the five most significant GO terms are listed in Table 3.10 and 3.11, respectively. Due to the small size of the downregulated DEG list (n=9) in the comparison between 0h and 1.5h, no downregulated biological processes were identified in this time sequence. Overall, similar biological processes were overrepresented in fibroblast cultures to the chondrogenic cell cultures: transcription regulatory events were observed at earlier time points, and processes related to signaling transduction happened during the whole experimental period. Also, the profiles of downregulated DEGs represented skeletal muscle cell and fat cell differentiation processes between 1.5h and 12h, and these terms were also observed in interzone and anlagen cell cultures between 1.5h and 3h (Table 3.5 and 3.9).

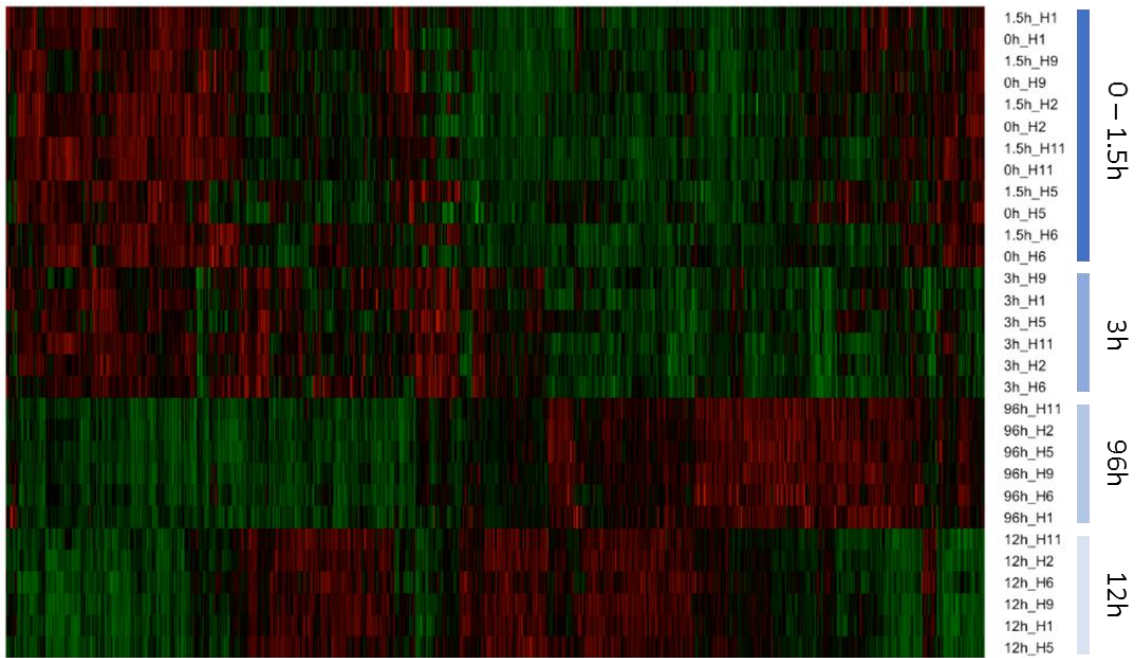
### 3-c) Hub genes and co-expression network

The hub gene that was most correlated with the time sequence in fibroblasts was NEB (GS=0.99; P-value=4.6E-26), and the ten most correlated genes with the time course are shown in Table 3.12. The ME identification categorized nine genes out of these ten genes into the Turquoise group, which had the highest module-trait relationship (0.98; P-value=1E-21) with the time sequence (Figure 3.11). The other, ITGA10, was identified as a member of the Red group with the second highest module-trait relationship (0.76; P-value=1E-06). The co-expression network of hub genes that were closely correlated with the Turquoise group is shown in Figure 3.12. The gene showed the highest MM with the Turquoise group was HTRA1 (MM=0.99; P-value=1.9E-23), interacting with two hub genes, VCAM and BGN in the co-expression network.

### 3-d) Upstream regulator prediction

Predicted upstream regulators in fibroblast cultures during the experimental period are shown in Figure 3.13. In common with the chondrogenic cell cultures, activation of TGFB1 was predicted (z-score=4.04; P-value=4.7E-28) during the first 1.5 hours. Activation of KLF3 between 3h and 12h and SAA1 between 12h and 96h was observed as a common prediction in all cell types. On the other hand, sixteen genes were predicted to be commonly inhibited in the chondrogenic cell lines and the negative control at a point during the experiment: KITLG, E2F2, F2RL1, TICAM1, EDN1, GPER1, IL6, AREG, TLR9, PLAU, SREBF1, CD38, PPARGC1A, ACVR1C, ADORA2A.

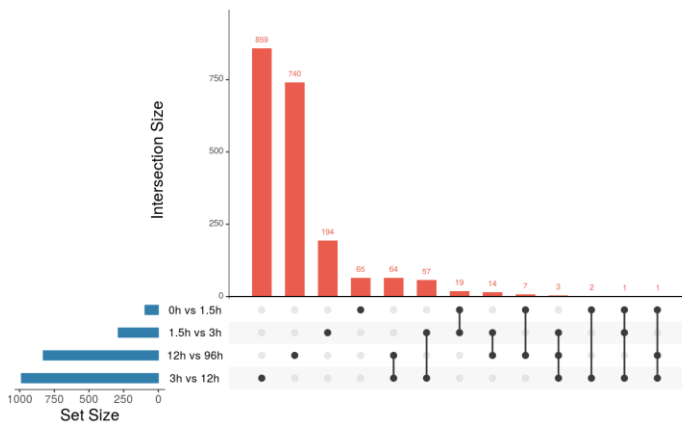
a) Heatmap



Color Key



b) Upregulated DEGs



c) Downregulated DEGs

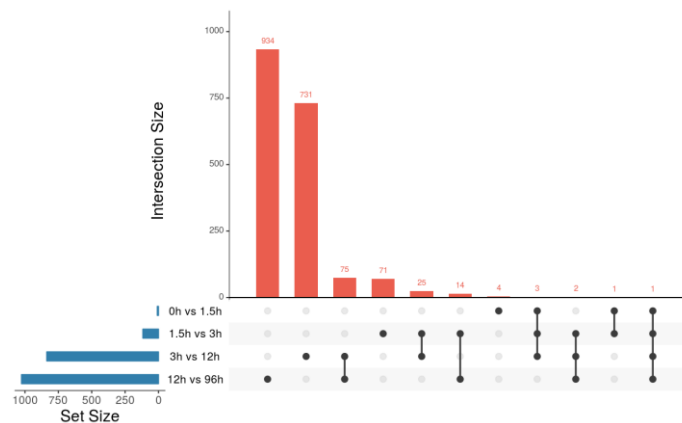


Figure 3.10. Differentially expressed genes (DEGs;  $n=3,212$ ) over time within fibroblast cultures. a) DEGs plotted in a heatmap. On the right side, time points (#h) and biological replicates (H#) are labeled. b) UpSet figure of upregulated DEGs ( $n=2,026$ ) between time points. The set size graph represents the number of DEGs in each time point comparison. c) UpSet figure of downregulated DEGs ( $n=1,861$ ) between time points. The set size graphs of the UpSet figures represent the number of DEGs in each time point comparison. Note that each time comparison has a different set size.

Table 3.10. Top five biological process gene ontology (GO) terms of differentially upregulated genes in fibroblast cultures during time courses

Time course	GO ID	Term	P-value <sup>a</sup>	Differentially expressed genes in the dataset
0h vs 1.5h	GO:0030335	Positive regulation of cell migration	7.69E-04	IRS2, PDGFB, FAM110C, ADRA2A, HAS2, SNAI1
	GO:0045892	Negative regulation of transcription, DNA-templated	0.001	CRY2, CEBPB, PDGFB, CCDC85B, KLF10, RELB, IRF1
	GO:1900745	Positive regulation of p38mapk cascade	0.002	GADD45G, ZC3H12A, GADD45B
	GO:0032922	Circadian regulation of gene expression	0.002	CRY2, RELB, PER1, BHLHE40
	GO:0006357	Regulation of transcription from RNA polymerase II promoter	0.002	MAFF, ETS2, RELB, RFX2, RUNX3, ELMSAN1, ETV3
1.5h vs 3h	GO:0006954	Inflammatory response	1.99E-04	IRAK2, TNFRSF21, S1PR3, TLR10, TNFRSF1B, HRH1, PTGIR, TNFRSF11A, CYP26B1, RELB, NFKB1, AFAP1L2
	GO:0048167	Regulation of synaptic plasticity	0.001	HRH1, HRH2, BAIAP2, ITPKA
	GO:0045892	Negative regulation of transcription, DNA-templated	0.002	ZNF263, WNT4, CRY2, PDGFB, RELB, HR, LIMD1, RFX3, CBFA2T3, SOX8, PPARGC1B
	GO:0006355	Regulation of transcription, DNA-templated	0.003	ZNF566, SOX13, HR, ZFP3, ZNF175, ZNF2, AHRR, RGS20, NOTCH1, ZNF235, ZSCAN20, ZNF404, HEYL, LHX4, LIMD1, RFX3, ZNF436
	GO:0001666	Response to hypoxia	0.006	VEGFA, EGLN3, SMAD3, LIMD1, CBFA2T3, DDIT4
3h vs 12h	GO:0042391	Regulation of membrane potential	8.96E-04	KCNMA1, HCN2, NEDD4, GRIK3, ATP1A4, POPDC3, SLC26A10, KCNH3, SLC26A11, CHRNE
	GO:0007155	Cell adhesion	0.002	HAPLN2, HAPLN4, ACHE, ICAM5, ITGAE, PCDH10, COL15A1, FERMT1, FES, PRKCE, COL5A1, NCAM1, ITGA9, COMP, TGFBI, LMLN, HRC
	GO:0010107	Potassium ion import	0.005	KCNJ4, ATP1B2, ATP1A3, ATP1A4, KCNJ12, KCNJ2
	GO:0060412	Ventricular septum morphogenesis	0.005	SOX4, EGLN1, WNT11, CITED2, PITX2
	GO:0071526	Semaphorin-plexin signaling pathway	0.006	SEMA5A, SEMA6C, SEMA4G, SEMA6D, SEMA4B, SEMA4A
12h vs 96h	GO:0030335	Positive regulation of cell migration	1.56E-06	BMP2, WNT5B, F2RL1, SPHK1, LRRC15, MMP14, SNAI2, AQP1, CCL26, SEMA5A, SEMA6D, SEMA3G, SEMA3E, PDGFRA, CEMIP, PDGFC, PDGFD, FGF1, F2R
	GO:0030574	Collagen catabolic process	2.96E-06	CTSK, MMP9, CTSS, MMP14, MMP13, MMP2, MMP11
	GO:0014068	Positive regulation of phosphatidylinositol 3-kinase signaling	4.86E-04	PTPN6, WNT16, F2RL1, PDGFC, ANGPT1, DCN, PDGFD, NRG1, F2R
	GO:0070374	Positive regulation of ERK1 and ERK2 cascade	5.39E-04	NOX4, BMP2, C5AR2, CCL2, F2RL1, FGF10, CD74, CCL26, SPRY2, PDGFRA, ANGPT1, PDGFC, PDGFD, PLA2G5, F2R
	GO:0006954	Inflammatory response	7.71E-04	C5AR2, RARRES2, BMP2, CCL2, IL2RA, TSPAN2, TNFRSF25, TLR1, PTGS1, SPHK1, CXCL8, PTGFR, CCL26, S1PR3, AGTR2, TNFRSF11B, CCR7, BMPR1B

<sup>a</sup>Modified Fisher Exact P-Value was used to evaluate enrichment of genes in GO annotation terms.

Table 3.11. Top five biological process gene ontology (GO) terms of differentially downregulated genes in fibroblast cultures during time courses

Time course	GO ID	Term	P-value <sup>a</sup>	Differentially expressed genes in the dataset
0h vs 1.5h	<i>No biological processes identified due to a small size of the DEG list (n=9) in this comparison</i>			
1.5h vs 3h	GO:0035914	Skeletal muscle cell differentiation	1.12E-05	FOS, ATF3, EGR2, NR4A1, ANKRD1, CITED2
	GO:0045444	Fat cell differentiation	4.64E-05	EGR2, GDF6, ARID5B, NR4A2, NR4A1, KLF4
	GO:0008543	Fibroblast growth factor receptor signaling pathway	0.001	FLRT3, FLRT2, FGF7, FGF21
	GO:0051965	Positive regulation of synapse assembly	0.005	FLRT3, FLRT2, BDNF, LRRTM1
	GO:0007264	Small GTPase mediated signal transduction	0.015	RND3, RASL11A, RND1, RAB30, RGL1, ARL4A
3h vs 12h	GO:0035914	Skeletal muscle cell differentiation	8.76E-05	EGR1, HLF, FOS, NOTCH1, ATF3, BTG2, HIVEP3, NR4A1, ANKRD1, SOX8
	GO:0009791	Post-embryonic development	1.52E-04	ATF5, SGPL1, KMT2A, BCL2, TGFB1, CSRNP1, CHST11, NR4A2, SLC18A2, HEG1, ASL, ITPR1
	GO:0048008	Platelet-derived growth factor receptor signaling pathway	2.93E-04	SGPL1, NRP1, PDGFB, BCAR1, CSRNP1, PDGFC, PTPN11
	GO:0045740	Positive regulation of DNA replication	2.93E-04	IGF1R, DNA2, PDGFB, FGF10, KITLG, PDGFC, AREG
	GO:0030335	Positive regulation of cell migration	3.05E-04	WNT5B, PDGFB, FAM110C, TGFB1, F2RL1, EDN1, CCL26, IGF1R, CORO1A, NOTCH1, DAB2, SEMA3E, CEMIP, HAS2, PDGFC, PLAU
12h vs 96h	GO:0071456	Cellular response to hypoxia	3.19E-07	SLC8A3, P4HB, EPAS1, PTGS2, PINK1, NPEPPS, RORA, PRKCE, HYOU1, HMOX1, VEGFA, NDRG1, FAM162A, ANGPT4
	GO:0001666	Response to hypoxia	2.01E-05	CAV1, CRYAA, EGLN3, SMAD3, EGLN1, DDIT4, VEGFC, HSP90B1, LONP1, MYOCD, VEGFA, P2RX2, RYR2, LOXL2, ALKBH5
	GO:0006096	Glycolytic process	2.82E-05	GPI, TPI1, ALDOC, PGAM1, ENO3, HK1, PGK1, GAPDH, ENO1
	GO:0000226	Microtubule cytoskeleton organization	4.09E-04	PRKCZ, EML1, CRMP1, MAP1A, CNTN2, NEFH, MAP6, GAPDH, NEFL, MAP7D2, TACC1
	GO:0018401	Peptidyl-proline hydroxylation to 4-hydroxy-L-proline	6.48E-04	P4HB, P4HA1, EGLN3, EGLN1

<sup>a</sup>Modified Fisher Exact P-Value was used to evaluate enrichment of genes in GO annotation terms.

Table 3.12. Ten most highly correlated genes with time course in anlagen cell culture during entire experimental period

Gene	ME <sup>a</sup> group	Gene significance	P-value	MM <sup>b</sup> to Turquoise	P-value
NEB	Turquoise	0.991	4.60E-26	0.983	4.07E-22
OLFML2B	Turquoise	0.989	1.26E-24	0.963	1.70E-17
GSTA4	Turquoise	0.984	1.24E-22	0.955	2.77E-16
NAAA	Turquoise	0.982	8.12E-22	0.969	1.30E-18
OMD	Turquoise	0.982	9.56E-22	0.973	2.38E-19
HTRA1	Turquoise	0.982	1.06E-21	0.986	1.87E-23
SLC29A1	Turquoise	0.979	6.44E-21	0.960	5.07E-17
IL1R1	Turquoise	0.979	8.39E-21	0.948	1.76E-15
ITGA10	Red	0.978	1.10E-20	0.944	5.48E-15
AMDHD2	Turquoise	0.978	1.38E-20	0.983	4.38E-22

<sup>a</sup>ME, Module eigengene

<sup>b</sup>MM, Module membership

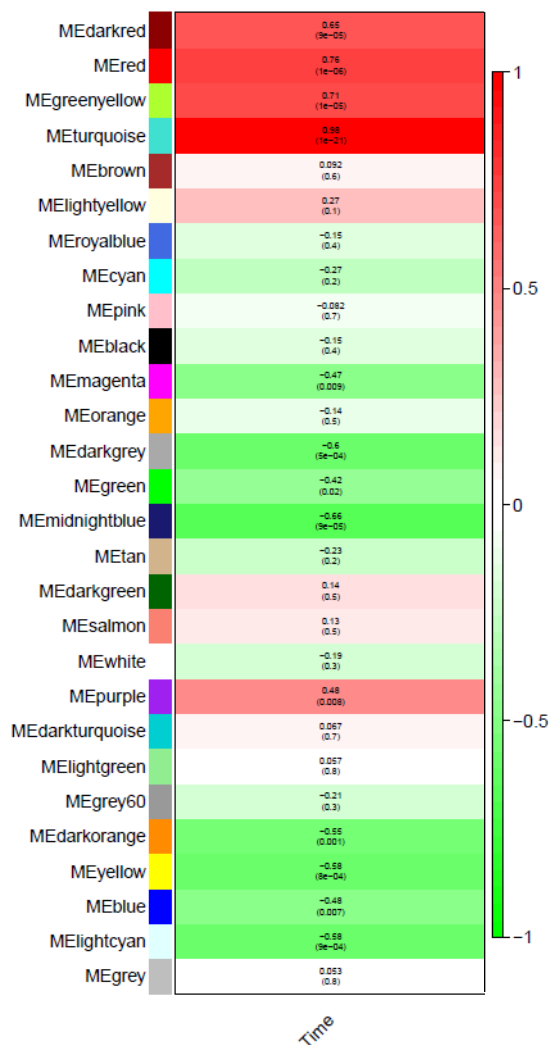


Figure 3.11. Module eigengene (ME) groups with module-trait relationship scores in fibroblast cultures over time. P-values are shown in parentheses. The higher relationship score represents the greater correlation between a ME group and the time sequence.

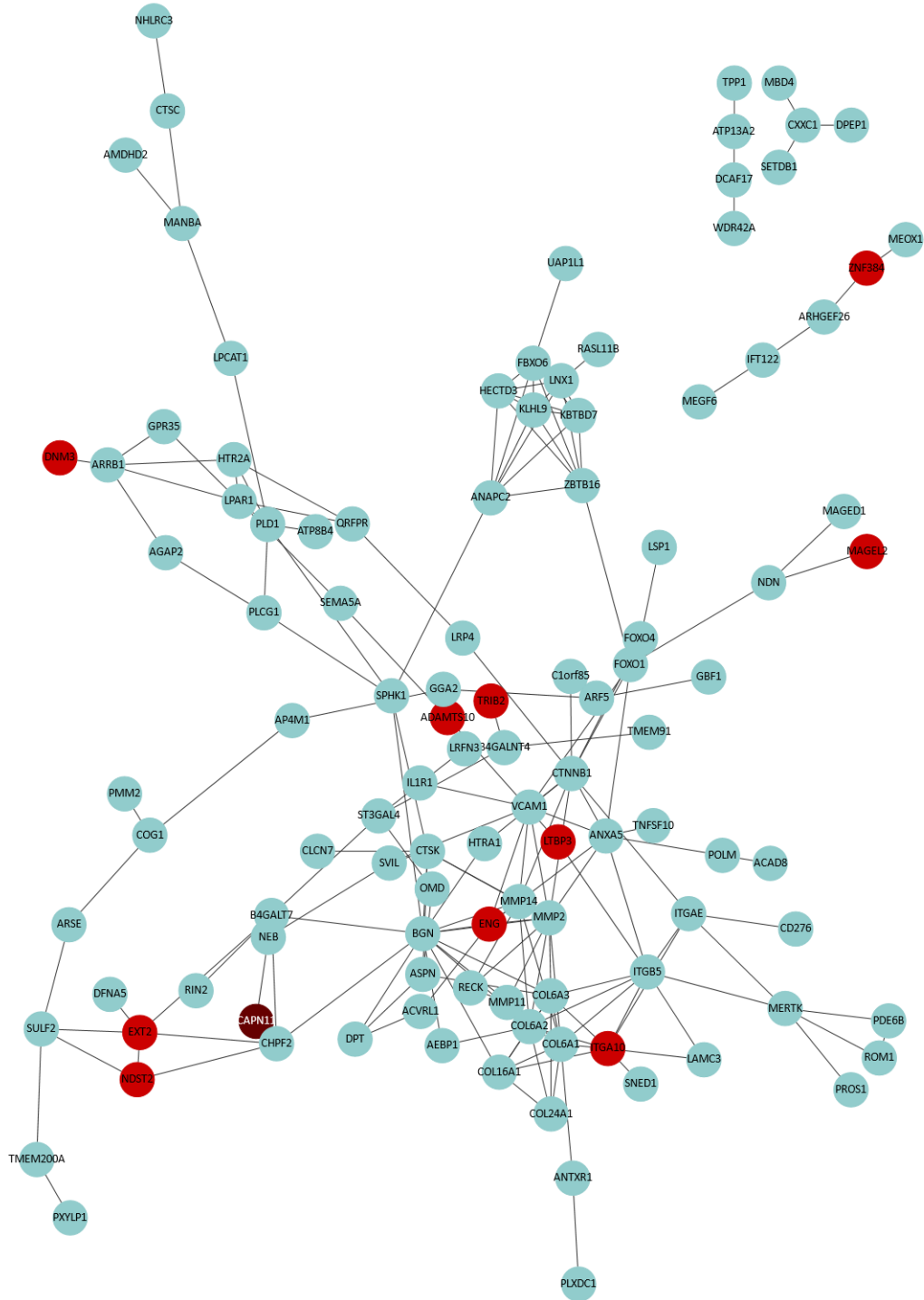


Figure 3.12. Hub gene co-expression network in fibroblast cultures during 96 hours of *in vitro* chondrogenesis. Colored by module eigengene group.



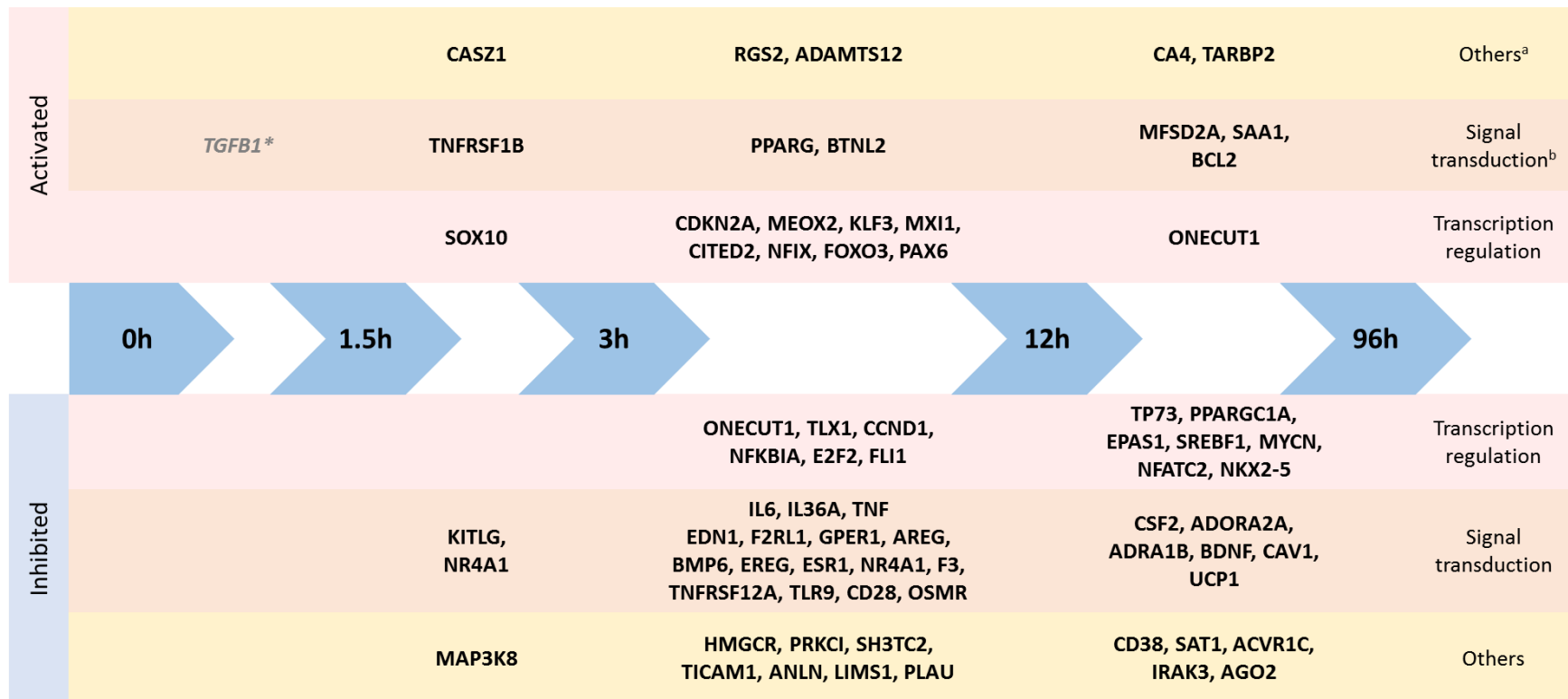


Figure 3.13. Upstream regulators predicted by Ingenuity Pathway Analysis (IPA) between subsequent time points in fibroblast cultures. The activation status thresholds were activation z-score > |2| and P-value of overlap < 0.05. Only genes that have log<sub>2</sub> fold change (FC) ≥ |1| in the dataset were presented in the figure. Positive z-scores represent activated upstream regulators, and negative values indicate inhibition. When a predicted activation status (activation or inhibition) matched the regulation pattern analyzed from the dataset (positive log<sub>2</sub>FC=upregulation; negative log<sub>2</sub>FC=downregulation), the prediction was accepted. <sup>a</sup>Others include enzymes, kinases, and others. <sup>b</sup>Signal transduction includes growth factors, ligands, receptors, and cytokines.

\*In the data, log<sub>2</sub>FC of TGFB1 from 0h to 1.5h was 0.38.

### ***Common chondrogenic characteristics of interzone and anlagen cell cultures***

Differentially expressed genes in the interzone and anlagen experimental groups relative to the negative fibroblasts were evaluated at each time point (interzone cell cultures vs. fibroblast cultures and anlagen cell cultures vs. fibroblast cultures). The number of upregulated and downregulated DEGs are listed in Table 3.13. Then, DEGs that were commonly observed in both chondrogenic cell cultures compared to the negative control at each time point are identified (Table 3.13) and processed through gene enrichment analysis using the DAVID Functional Annotation Tool (v6.8; Huang et al., 2009; Table 3.14).

In this section, the DEGs were further classified; when a DEG's expression was measured in a condition (either chondrogenic cell lines or negative control) but not at all in the other condition, this gene was considered as a candidate switch being turned on or off depending on a condition. The numbers of common chondrogenic molecular switches at each time point are present in Table 3.15.

Table 3.13. Number of differentially expressed genes (DEGs) in interzone and anlagen cell cultures compared to the negative control at each time point

Upregulated			
Time point	Interzone cell cultures	Common DEGs in chondrogenic cell lines	Anlagen cell cultures
0h	450	[274]	769
1.5h	436	[239]	779
3h	434	[242]	708
12h	470	[274]	629
96h	645	[366]	617
Downregulated			
Time point	Interzone cell cultures	Common DEGs in chondrogenic cell lines	Anlagen cell cultures
0h	611	[447]	850
1.5h	662	[471]	864
3h	664	[472]	868
12h	895	[590]	921
96h	793	[550]	877

The numbers in the square brackets ([ ]) represent the number of commonly upregulated or downregulated DEGs observed in both interzone and anlagen cell cultures at each time point. These numbers are included in the numbers of DEGs in each chondrogenic cell line in the same row.

Table 3.14. Number of overrepresented biological process gene ontology (GO) terms identified in the chondrogenic cell cultures compared to fibroblast cultures

	0h	1.5h	3h	12h	96h	Total unique GO terms
Upregulated	26	33	28	58	70	128
Downregulated	43	33	59	48	50	133

Table 3.15. Number of differentially expressed genes (DEGs) and candidate molecular switches in the chondrogenic cell cultures compared to fibroblast cultures

Upregulated						
	0h	1.5h	3h	12h	96h	
Total DEGs	247	239	242	274	366	
Molecular switches	[1]	[1]	[2]	[1]	[3]	
Downregulated						
	0h	1.5h	3h	12h	96h	
Total DEGs	447	471	472	590	550	
Molecular switches	[9]	[7]	[3]	[7]	[10]	

The numbers in the square brackets ([ ]) represent the number of genes whose expression was exclusively turned on or off in the chondrogenic cell cultures. These numbers are included in the numbers of DEGs in the cells above them.

Candidate molecular switches exclusively turned on in chondrogenic cell lines

Compared to fibroblast cultures, there were 4 common genes that were only expressed in interzone and anlagen cell cultures at the same time points (Table 3.16). At the base line, there were no common genes exclusively expressed in the chondrogenic cell cultures. After inducing chondrogenesis, GSC expression was turned on in both chondrogenic cell lines by 1.5h and until the last collection point (96h). Expression of a novel gene, ENSECAG00000040027, was turned on by 3h. Exclusive expression of another novel gene, ENSECAG00000034476, and HOXD13 was observed at 96h in the chondrogenic cell cultures.

Table 3.16. Genes that were exclusively expressed in both interzone and anlagen cell cultures at the same time points while not expressed in fibroblast cultures

0h	1.5h	3h	12h	96h
<i>None</i>	GSC	ENSECAG 00000040027	GSC	ENSECAG 00000034476
		GSC		GSC
				HOXD13

Candidate molecular switches exclusively turned off in chondrogenic cell lines

While expressed in fibroblast cultures, 19 genes were not expressed in both chondrogenic cell cultures at the same time points (Table 3.17). These genes include four of the Homeobox-containing transcription factor genes in cluster B (HOXB4, HOXB5, HOXB8, and HOXB9). More candidate molecular switches were turned off in the chondrogenic cell cultures towards the baseline and the last time point (96h).

Table 3.17. Genes that were not expressed exclusively in both interzone and anlagen cell cultures at the same time points while expressed in fibroblast cultures

0h	1.5h	3h	12h	96h
EN1	ENSECAG 00000031562	EN1	EN1	ENSECAG 00000031562
ENSECAG 00000031562	HOXB4	HOXB4	ENSECAG 00000030268	ENSECAG 00000034623
HOXB5	HOXB5	HOXB9	HOXB5	ENSECAG 00000039182
HOXB8	NTM		HOXB9	ENSECAG 00000043423
HOXB9	RBFOX1		SERPINB10	HOXB5
NTM	SIM1		SIM1	HOXB9
OLFM3	ZIC4		ZIC4	MMP26
RBFOX1				TLX3
ZIC4				TNFSF4
				ZIC4

### Upregulated DEGs in chondrogenic cell lines

The profiles of commonly upregulated chondrogenic DEGs, including the switches, resulted in 128 significant biological processes across the entire experimental period by gene enrichment analysis (Table 3.14). A greater number of upregulated biological processes were identified at later time points. Among the five collection time points, three GO terms were consistently observed as significantly overrepresented biological processes in interzone and analgen cell cultures compared to the negative control (Table 3.18). These processes are 1) anterior/posterior pattern specification, 2) embryonic skeletal system morphogenesis, and 3) positive regulation of chondrocyte differentiation.

### Downregulated DEGs in chondrogenic cell lines

The commonly downregulated DEGs as well as the turned-off molecular switches in the chondrogenic cell cultures presented 133 significant biological processes throughout the whole experimental period (Table 3.14). The profiles of downregulated DEGs in the chondrogenic cell lines relative to fibroblast cultures continuously overrepresented four biological processes at the all collection points (Table 3.19). These processes are 1) anterior/posterior pattern specification, 2) immune response, 3) negative regulation of cell proliferation, and 4) regulation of cell migration.

Table 3.18. Upregulated overrepresented biological processes in the chondrogenic cell lines compared to fibroblast cultures at all time points

GO:0009952 <sup>a</sup> , Anterior/posterior pattern specification			
	Time point	Genes in the process	P-value <sup>b</sup>
	0h	HOXD9, HOXA5, HOXA11, HOXA6, HOXA7, HOXA10, HOXD13, HOXD10	1.67E-05
	1.5h	HOXA11, HOXA6, HOXA7, HOXA10, HOXD10	4.99E-03
	3h	HOXD9, HOXA5, HOXA11, HOXA6, HOXA7, HOXA10, HOXD10	1.69E-03
	12h	HOXA11, HOXA6, HOXA7, HOXA10, HOXD13, HOXD10	1.13E-04
	96h	RARG, HOXA5, HOXA11, HOXA10, HOXA9, GLI3, HOXD10	2.87E-03
GO:0048704, Embryonic skeletal system morphogenesis			
	Time point	Genes in the process	P-Value
	0h	HOXD9, HOXA5, HOXA6, HOXA7, HOXD10	1.84E-04
	1.5h	GSC, HOXA6, HOXA7, HOXD10	1.81E-03
	3h	HOXD9, GSC, HOXA5, HOXA6, HOXA7, HOXD10	2.98E-03
	12h	GSC, HOXA6, HOXA7, HOXD10	1.19E-05
	96h	GSC, HOXA5, GLI3, HOXD10	2.97E-02
GO:0032332, Positive regulation of chondrocyte differentiation			
	Time point	Genes in the process	P-Value
	0h	HOXA11, SOX5, SOX6	1.20E-02
	1.5h	HOXA11, SOX5, SOX6, SOX9	5.22E-04
	3h	HOXA11, SOX5, SOX6, SOX9	2.76E-05
	12h	HOXA11, GDF5, SOX5, SOX6, SOX9	5.14E-04
	96h	HOXA11, SOX5, SOX6, SOX9, GLI3	7.55E-05

<sup>a</sup>GO:#, gene ontology ID

<sup>b</sup>Modified Fisher Exact P-Value was used to evaluate enrichment of genes in GO annotation terms.

Table 3.19. Downregulated overrepresented biological processes in the chondrogenic cell lines compared to fibroblast cultures at all time points

GO:0009952 <sup>a</sup> , Anterior/posterior pattern specification			
	Time point	Genes in the process	P-value <sup>b</sup>
	0h	HOXB3, HOXC8, HOXC9, HOXB2, HOXC4, HOXB5, HOXB6, EMX2, HOXB9	4.95E-04
	1.5h	HOXB3, HOXC6, HOXB4, HOXC8, HOXC9, HOXC4, HOXB5, HOXB6, EMX2	7.33E-04
	3h	HOXB4, HOXC8, HOXC9, HOXC4, EMX2, HOXB9	2.29E-05
	12h	HOXB3, HOXB4, HOXC8, HOXC9, HOXB2, HOXC4, HOXD3, HOXB5, HOXB6, EMX2, HOXB9, ZBTB16	4.60E-02
	96h	HOXC8, HOXB2, HOXC9, HOXC4, HOXD3, EMX2, HOXB9, ZBTB16	2.17E-03
GO:0006955, Immune response			
	Time point	Genes in the process	P-Value
	0h	C7, IL18, TGFBR3, OAS1, COLEC12, FAS, OAS2, TINAGL1, BMP6, B2M	2.07E-02
	1.5h	C7, HRH2, TGFBR3, OAS1, COLEC12, FAS, OAS2, TINAGL1, TNFAIP3, BMP6, B2M, TNFSF8	4.01E-03
	3h	TNFRSF21, C7, HRH2, IRF8, TGFBR3, OAS1, COLEC12, FAS, TNFAIP3, BMP6, B2M	3.74E-02
	12h	TNFRSF21, C7, IL18, IRF8, OAS3, OAS1, TNFRSF14, COLEC12, NGFR, TNFAIP3, BMP6	1.13E-02
	96h	CCL11, C7, TNFSF10, TNFSF4, HRH2, IL18, TNFRSF8, TGFBR3, COLEC12, FAS, BMP6, TNFSF8, B2M	1.14E-02
GO:0008285, Negative regulation of cell proliferation			
	Time point	Genes in the process	P-Value
	0h	CEBPA, ATF5, CDKN1A, PODN, PTGES, PTH1R, TFAP2B, PTPRU, SLC9A3R1, SKAP2, SLIT3, DPT	7.96E-03
	1.5h	ATF5, CDKN1A, PODN, PTGES, BCL11B, PTH1R, TFAP2B, PTPRU, SLC9A3R1, SKAP2, SLIT3, TP53INP1, DPT	4.58E-03
	3h	CEBPA, PODN, PTH1R, KLF11, FGF10, PTPRU, SLC9A3R1, SKAP2, SLIT3, ATF5, CDKN1A, PTGES, BCL11B, DPT, TP53INP1	8.20E-03
	12h	CEBPA, PODN, PTH1R, FGF10, ZBTB16, PTPRU, PROX1, CDH5, SLIT3, RERG, ATF5, PTK2B, BCL11B, KLF4	5.18E-04
	96h	NOX4, B4GALT1, PODN, ADARB1, ZBTB16, CBFA2T3, SKAP2, SLIT3, SPRY2, SPRY1, PTGES, PTK2B, BCL11B, SFRP4, ASPH, WNT9A	6.22E-04
GO:0030334, Regulation of cell migration			
	Time point	Genes in the process	P-Value
	0h	LAMA2, LAMA1, PLXNC1, PLXNA4, PLXNA2, AMOT	4.35E-03
	1.5h	LAMA2, LAMA1, PLXNC1, PLXNA4, PLXNA2, AMOT	5.59E-03
	3h	LAMA2, LAMA1, PLXNC1, PLXNA4, PLXNA2, AMOT	1.24E-02
	12h	LAMA2, PLXNC1, PLXNA4, LAMA3, PLXNA2, DOCK10	5.59E-03
	96h	LAMA2, PLXNC1, PLXNA4, AMOT, NTN1	4.18E-02

<sup>a</sup>GO:#, gene ontology ID

<sup>b</sup>Modified Fisher Exact P-Value was used to evaluate enrichment of genes in GO annotation terms.



### ***Distinct traits between interzone and anlagen cell cultures***

Differential gene expression was evaluated at each time point between interzone and anlagen cell cultures (Figure 3.14). More DEGs were observed at the earlier time points (772 – 710 genes between 0h and 3h), and less DGEs were counted towards the later time points (543 – 551 genes between 12h and 96h; Table 3.20). Then, the DEGs were further categorized if they were expressed exclusively in one type of cell culture while not expressed at all in the other (Table 3.21). The ten most upregulated DEGs in interzone cell cultures and analgen cell cultures compared to each other at each time point are listed in Table 3.22 and 3.23, respectively. Gene enrichment analysis was conducted to identify overrepresented biological processes that were upregulated in each cell type (Figure 3.15 and Table 3.24). Finally, upstream regulators between the two chondrogenic cell cultures at each time point were predicted by Ingenuity Pathway Analysis (IPA; Qiagen; Figure 3.16).

Table 3.20. Number of differentially expressed genes (DEGs) between interzone and analgen cell cultures at each time point

	0h	1.5h	3h	12h	96h
Total DEGs between interzone and anlagen cell cultures	772	770	710	543	551
Upregulated in interzone cell cultures	416	406	356	316	275
Exclusively expressed in interzone cell cultures	[1]	[1]	[2]	[2]	[1]
Upregulated in anlagen cell cultures	356	364	354	227	276
Exclusively expressed in anlagen cell cultures	[2]	[0]	[1]	[2]	[2]

The numbers of exclusively expressed genes in the square brackets ([ ]) are also included in the numbers of upregulated or downregulated DEGs in the cell above them.

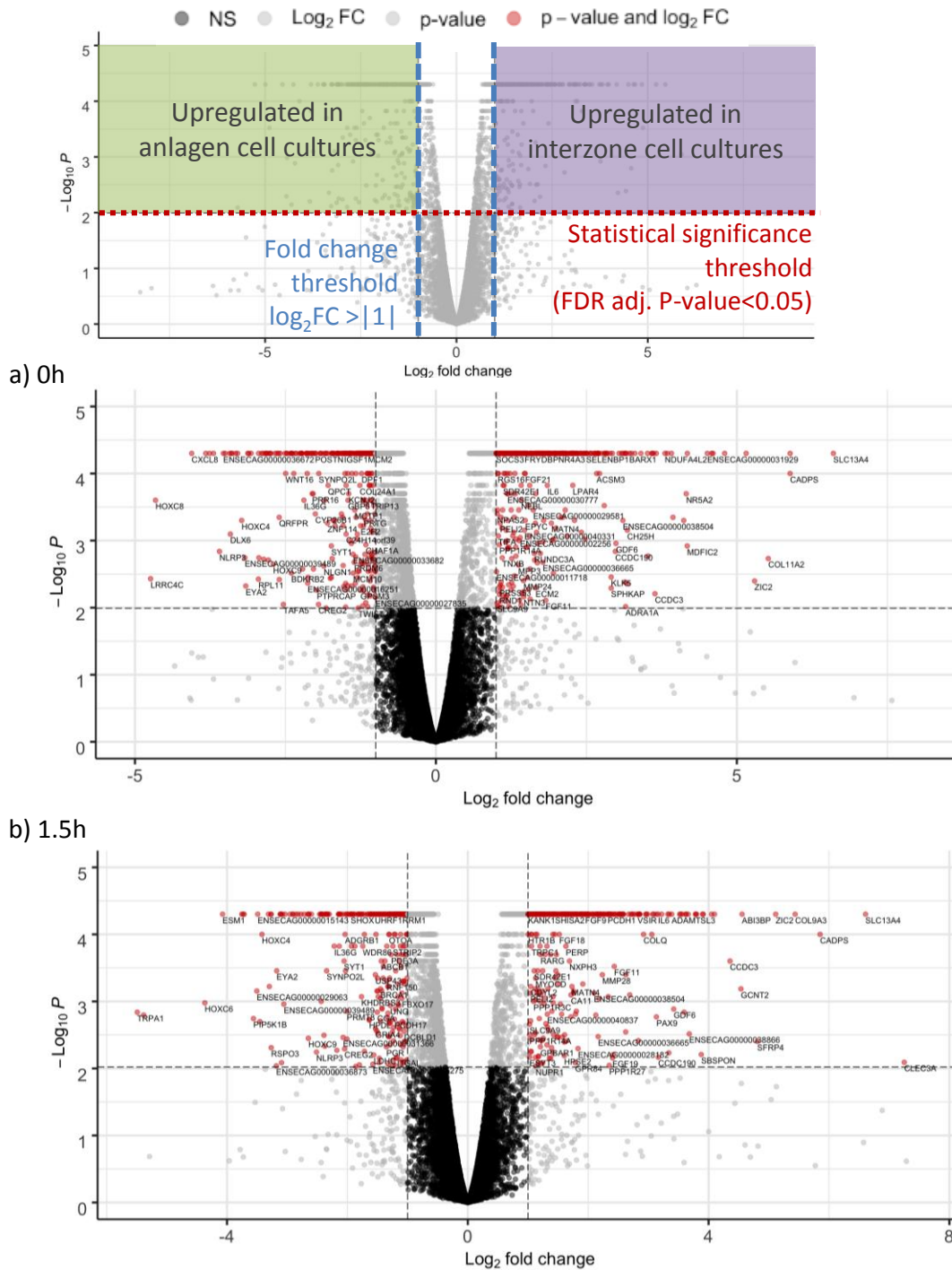


Figure 3.14. Differentially expressed genes (DEGs) between interzone and anlagen cell cultures at each time point plotted in volcano plots. Differential expression was determined using fold change (FC) and statistical significance thresholds,  $\log_2 FC > |1|$  and false discovery rate (FDR) adjusted P-value  $< 0.05$ , respectively. Note that  $-\log_{10} P$ -values defining an FDR  $< 0.05$  threshold in each graph are different. The DEGs are shown in red dots. Exclusively expressed genes in a cell type with infinite FC are not displayed in the volcano plots (continued on the next page).



Candidate molecular switches exclusively expressed in one cell type

<Exclusively expressed genes in interzone cells>

From the baseline until 3h, ENSECAG00000037611 was only expressed in interzone (FDR adjusted P-value=0.001; Table 3.21). At 3 and 12h, LY6G6C and C7H11orf52 were turned on (FDR adjusted P-value=0.001), respectively while there was no expression in anlagen cell cultures. Then, expression of ZIC3 was observed at 12h as well as at the last collection point (FDR adjusted P-value=0.001).

<Exclusively expressed genes in anlagen cells>

Before inducing chondrogenesis, HOXB2 (FDR adjusted P-value=0.001) and H4C9 (FDR adjusted P-value=0.003) were only expressed in analgen cell cultures (Table 3.21). Between 1.5h and 3h, HOXB2 expression was also detected in interzone cell cultures, but it was turned off again by 12h and at 96h. Expression of ENSECAG00000038392 (FDR adjusted P-value<0.01), HOXB3 (FDR adjusted P-value=0.001), and PLAC8B (FDR adjusted P-value=0.001) was exclusively measured in analgen cell cultures at 3h, 12h, and 9h, respectively.

Table 3.21. Exclusively expressed genes in either interzone or analgen cell cultures

	0h	1.5h	3h	12h	96h
Interzone cell cultures	ENSECAG00000037611	ENSECAG00000037611	ENSECAG00000037611	C7H11orf52	ZIC3
			LY6G6C	ZIC3	
Anlagen cell cultures	H4C9		ENSECAG00000038392	HOXB2	HOXB2
	HOXB2			HOXB3	PLAC8B

The expression levels in the fragments per kilobase of transcript per million fragments mapped read (FPKM) format, and false discovery rate adjusted P-values are shown in Table 3.22 and 3.23.

### Differentially expressed genes between interzone and anlagen cell cultures

The ten most upregulated genes in interzone and anlagen cell cultures are listed in Table 3.22 and 3.23, respectively. Gene enrichment analysis was conducted with the total DEGs between the two chondrogenic cell cultures at each time point. Then, the ten most overrepresented biological processes from the profiles of the DEGs between interzone and anlagen cell cultures at each time point were plotted into a circle plot (Figure 3.15 and Table 3.24). From the baseline until 3h, upregulated DEGs in anlagen cell cultures dominated in greater numbers of the significant overrepresented biological processes. However, at the last time point, only one process among the ten most significant processes was dominated by upregulated DEGs in anlagen cell cultures; this biological process is endochondral ossification.

Table 3.22. Ten most upregulated differentially expressed genes in interzone cell cultures compared to anlagen cell cultures at each time point

Time point	Genes	Expression levels, FPKM <sup>a</sup>		log <sub>2</sub> FC <sup>b</sup>	adj. P-value <sup>c</sup>
		Interzone	Anlagen		
0h	ENSECAG00000037611	2.94	0.00	inf <sup>d</sup>	0.001
	SLC13A4	81.98	0.84	6.60	0.001
	COL9A3	5.40	0.09	5.88	0.001
	CADPS	14.00	0.24	5.88	0.001
	COL11A2	2.62	0.06	5.52	0.013
	ZIC2	8.69	0.22	5.29	0.024
	HAPLN1	5.99	0.17	5.15	0.001
	COL2A1	17.18	0.62	4.80	0.001
	CLEC3A	5.96	0.25	4.60	0.001
	ABI3BP	87.94	3.67	4.58	0.001
1.5h	ENSECAG00000037611	2.59	0.00	inf	0.002
	CLEC3A	0.59	0.00	7.25	0.044
	SLC13A4	68.66	0.70	6.61	0.001
	CADPS	14.81	0.26	5.85	0.001
	COL9A3	2.65	0.06	5.44	0.001
	ZIC2	7.00	0.20	5.11	0.001
	SFRP4	0.57	0.02	4.81	0.025
	ABI3BP	88.27	3.75	4.56	0.001
	GCNT2	2.26	0.10	4.54	0.006
	CCDC3	3.43	0.17	4.36	0.003
3h	ENSECAG00000037611	2.28	0.00	inf	0.004
	LY6G6C	1.23	0.00	inf	0.001
	SLC13A4	50.92	0.91	5.81	0.001
	RPL11	11.54	0.21	5.75	0.047
	ZIC2	8.22	0.16	5.66	0.001
	CADPS	16.98	0.34	5.63	0.001
	IL6	17.13	0.55	4.97	0.001
	ABI3BP	88.35	4.00	4.46	0.001
	TFAP2A	2.38	0.15	4.03	0.001
	ENSECAG00000038866	7.61	0.53	3.84	0.017
12h	ZIC3	0.64	0.00	inf	0.001
	C7H11orf52	0.96	0.00	inf	0.001
	SLC13A4	127.92	2.59	5.63	0.001
	CADPS	13.25	0.36	5.22	0.001
	ZIC2	12.14	0.37	5.04	0.001
	RPL11	13.65	0.43	4.98	0.001
	ENSECAG00000031929	4.10	0.14	4.87	0.001

	PLA2G2A	6.80	0.25	4.79	0.047
	HOXA13	3.48	0.18	4.28	0.001
	CDH1	4.33	0.26	4.06	0.001
96h	ZIC3	0.96	0.00	inf	0.001
	ZIC2	10.23	0.17	5.88	0.016
	ADAMTS15	2.27	0.05	5.46	0.001
	THBS4	19.09	0.51	5.22	0.001
	ENSECAG00000033567	6.91	0.24	4.88	0.006
	ENSECAG00000031929	5.83	0.20	4.85	0.001
	HOXA13	6.68	0.26	4.67	0.001
	ENSECAG00000034282	4.61	0.19	4.63	0.001
	DRA	2.45	0.10	4.57	0.026
	ABI3BP	30.03	1.33	4.49	0.001

<sup>a</sup>FPKM, fragments per kilobase of transcript per million fragments mapped read

<sup>b</sup>FC, fold change

<sup>c</sup>False discovery rate adjusted P-value

<sup>d</sup>inf; Gene expression was not detected in anlagen cell cultures, therefore resulted in infinite fold change.

Table 3.23. Ten most upregulated differentially expressed genes in anlagen cell cultures relative to interzone cell cultures at each time point

Time point	Genes	Expression levels, FPKM <sup>a</sup>		log <sub>2</sub> FC <sup>b</sup>	adj. P-value <sup>c</sup>
		Interzone	Anlagen		
0h	H4C9	0.00	1.08	inf <sup>d</sup>	0.003
	HOXB2	0.00	0.67	inf	0.001
	LRRC4C	0.04	1.04	-4.740	0.023
	HOXC8	0.10	2.52	-4.658	0.003
	CXCL8	8.43	140.48	-4.058	0.001
	IGFBP3	2.82	39.98	-3.824	0.001
	ESM1	0.31	4.25	-3.760	0.001
	MEIS1	0.68	8.83	-3.695	0.001
	NLRP3	0.07	0.83	-3.596	0.011
	ENSECAG00000036672	0.66	7.67	-3.532	0.001
1.5h	TRPA1	0.02	0.88	-5.491	0.012
	HOXC8	0.05	2.25	-5.385	0.012
	HOXC6	0.51	10.54	-4.369	0.009
	ESM1	0.19	3.16	-4.074	0.001
	IGFBP3	2.93	39.57	-3.756	0.001
	CXCL8	7.68	102.08	-3.733	0.001
	JAM2	0.49	6.49	-3.716	0.001
	PIP5K1B	0.14	1.60	-3.561	0.014
	ENSECAG00000029063	0.14	1.60	-3.505	0.006
	ENSECAG00000015143	1.43	16.07	-3.490	0.001
3h	ENSECAG00000038392	0.00	1.34	inf	0.009
	HOXC6	0.59	12.87	-4.443	0.003
	HOXC8	0.16	2.62	-4.042	0.004
	JAM2	0.47	6.53	-3.780	0.001
	PDGFD	1.08	13.10	-3.595	0.001
	IGFBP3	2.75	32.14	-3.546	0.001
	TRPA1	0.05	0.59	-3.539	0.001
	HOXC4	0.12	1.40	-3.529	0.001
	MEIS1	0.41	4.66	-3.503	0.001
	CXCL8	2.81	29.53	-3.392	0.001
12h	HOXB2	0.00	0.96	inf	0.001
	HOXB3	0.00	0.75	inf	0.001
	HOXC6	0.76	25.41	-5.058	0.001
	HOXC8	0.15	3.44	-4.525	0.001
	SMPD3	0.04	0.80	-4.253	0.023
	HOXC4	0.15	2.26	-3.897	0.001
	ENSECAG00000039489	0.06	0.90	-3.891	0.017



	NRN1	20.85	246.23	-3.562	0.001
	IRX6	0.05	0.55	-3.551	0.013
	EYA2	0.72	8.14	-3.489	0.001
96h	PLAC8B	0.00	1.17	inf	0.001
	HOXB2	0.00	1.46	inf	0.001
	KIF19	0.37	14.41	-5.269	0.001
	HOXC6	0.51	16.09	-4.992	0.001
	HOXC4	0.08	2.46	-4.887	0.006
	SMPD3	0.59	13.81	-4.548	0.001
	IFIT3	0.27	5.74	-4.428	0.018
	HOXC8	0.12	2.40	-4.261	0.001
	FAM20A	0.03	0.60	-4.222	0.022
	DLX6	0.07	1.24	-4.104	0.002

<sup>a</sup>FPKM, fragments per kilobase of transcript per million mapped reads

<sup>b</sup>FC, fold change

<sup>c</sup>False discovery rate adjusted P-value

<sup>d</sup>inf; Gene expression was not detected in anlagen cell cultures, therefore resulted in infinite fold change.

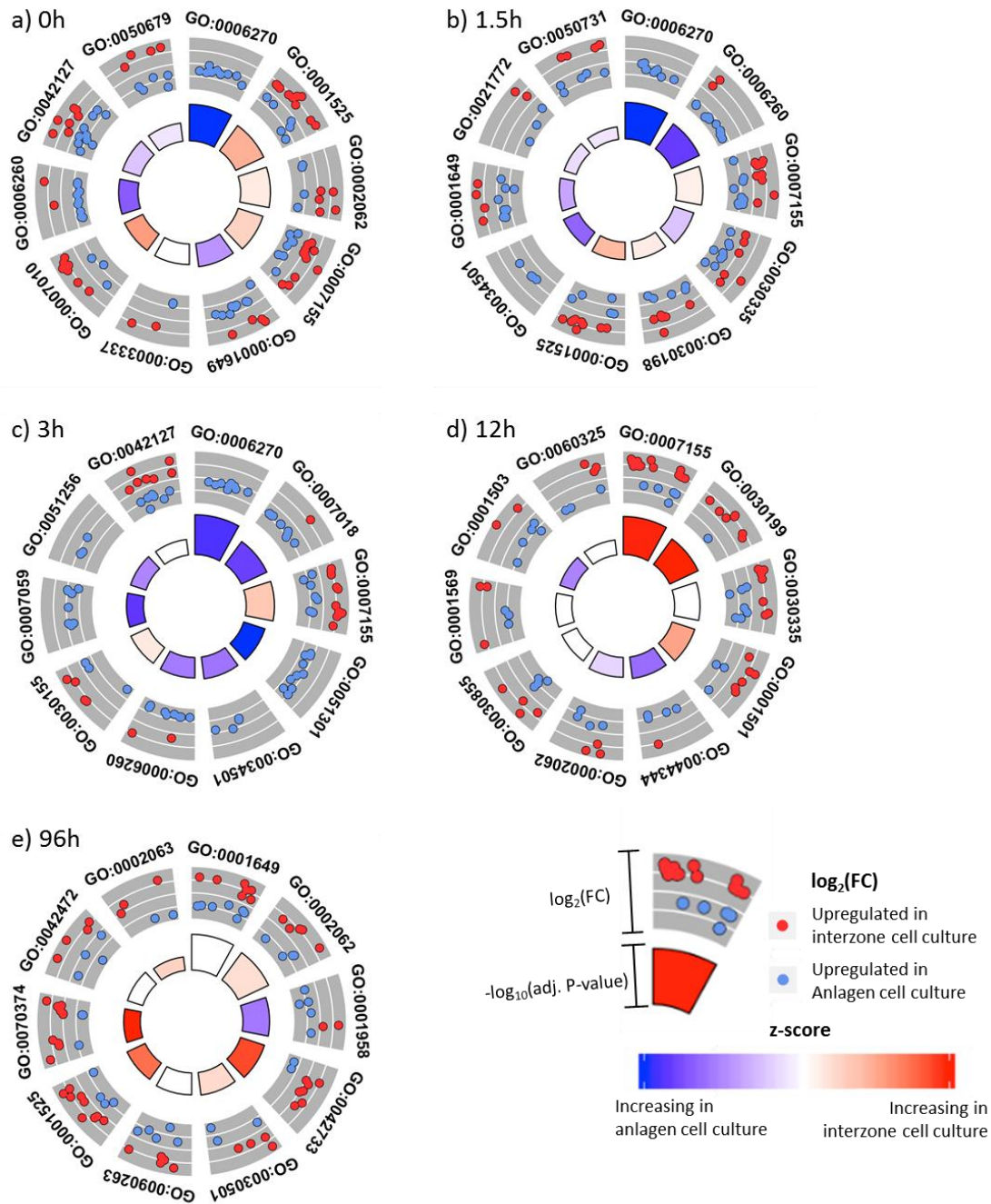


Figure 3.15. Gene ontology (GO) enrichment analysis on differentially expressed genes (DEGs) between interzone and anlagen cell cultures at each time point, plotted by GOCircle plot. The GO IDs and corresponding terms are listed in Table 3.24. The inner ring is a bar chart colored by z-score, and the height of the bars represents the significance of the GO term determined by  $-\log_{10}(\text{adjusted P-value})$ . The outer ring shows dotted plots of the  $\log_2$  fold change (FC) for the DEGs in each term. Upregulated DEGs in interzone cell cultures are shown in red dots while upregulated DEGs in anlagen cell cultures are shown in blue dots. Exclusively expressed genes in a cell type were excluded because of infinite FC. a) 0h, n=769 DEGs; b) 1.5h, n=769 DEGs; c) 3h, n=707 DEGs; d) 12h, n=539 DEGs; e) 96h, n=548 DEGs.

Table 3.24. Gene ontology (GO) IDs and corresponding terms present in Figure 3.15

Gene ontology ID	Term	0h	1.5h	3h	12h	96h
GO:0001501	Skeletal system development				12h	
GO:0001503	Ossification				12h	
GO:0001525	Angiogenesis	0h	1.5h			96h
GO:0001569	Patterning of blood vessels				12h	
GO:0001649	Osteoblast differentiation	0h	1.5h			96h
GO:0001958	Endochondral ossification					96h
GO:0002062	Chondrocyte differentiation	0h			12h	96h
GO:0002063	Chondrocyte development					96h
GO:0003337	Mesenchymal to epithelial transition involved in metanephros morphogenesis	0h				
GO:0006260	DNA replication	0h	1.5h	3h		
GO:0006270	DNA replication initiation	0h	1.5h	3h		
GO:0007010	Cytoskeleton organization	0h				
GO:0007018	Microtubule-based movement			3h		
GO:0007059	Chromosome segregation			3h		
GO:0007155	Cell adhesion	0h	1.5h	3h	12h	
GO:0021772	Olfactory bulb development		1.5h			
GO:0030155	Regulation of cell adhesion			3h		
GO:0030198	Extracellular matrix organization		1.5h			
GO:0030199	Collagen fibril organization				12h	
GO:0030335	Positive regulation of cell migration		1.5h		12h	
GO:0030501	Positive regulation of bone mineralization					96h
GO:0030855	Epithelial cell differentiation				12h	
GO:0034501	Protein localization to kinetochore		1.5h	3h		
GO:0042127	Regulation of cell proliferation	0h		3h		
GO:0042472	Inner ear morphogenesis					96h
GO:0042733	Embryonic digit morphogenesis					96h
GO:0044344	Cellular response to fibroblast growth factor stimulus				12h	
GO:0050679	Positive regulation of epithelial cell proliferation	0h				
GO:0050731	Positive regulation of peptidyl-tyrosine phosphorylation		1.5h			
GO:0051256	Mitotic spindle midzone assembly			3h		
GO:0051301	Cell division			3h		
GO:0060325	Face morphogenesis				12h	
GO:0070374	Positive regulation of ERK1 and ERK2 cascade					96h
GO:0090263	Positive regulation of canonical Wnt signaling pathway					96h

The cells are colored by the z-score scale shown in Figure 3.15. Red colors represent upregulation in interzone cells, and blue colors represent upregulation in anlagen cells.

### Predicted upstream regulators

Upstream regulators were estimated by analyzing differences of gene expression levels between the chondrogenic cell cultures at each time point using the IPA software (Figure 3.16). Twelve upstream regulators were predicted to be more activated in interzone cell cultures while twenty-three upstream regulators were predicted to be more activated in anlagen cell cultures over the experimental period. The upstream analysis projected NUPR1 as an activated regulator in interzone cell cultures and E2F1 in anlagen cell cultures at 0h, 1.5h, and 3h. In anlagen cell cultures, RUNX2 was expected as an upstream regulator at the baseline, 1.5h, and 96h.

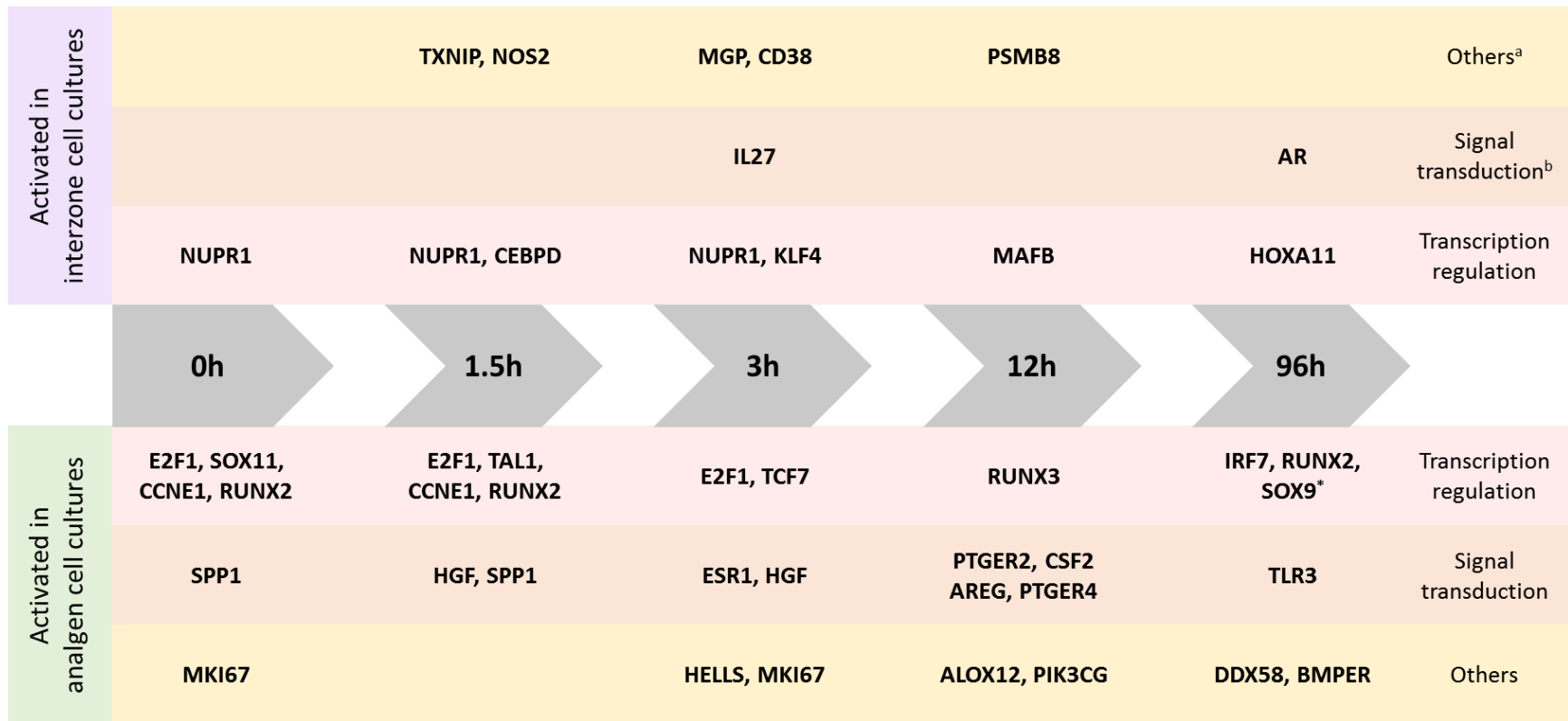


Figure 3.16. Prediction of upstream regulators that differentially regulate the chondrogenic pathways between interzone and analgen cultures by Ingenuity Pathway Analysis (IPA). The activation status thresholds were activation z-score > |2| and P-value of overlap < 0.05. Only genes that have log<sub>2</sub> fold change (FC) ≥ |1| in the dataset were presented in the figure. Positive z-scores represent activated upstream regulators in interzone cell cultures, and negative values indicate activation in analgen cell cultures. When a predicted activation in a cell type matched the regulation pattern analyzed from the dataset (positive log<sub>2</sub>FC=upregulation in interzone cell cultures; negative log<sub>2</sub>FC=upregulation in analgen cell cultures), the prediction was accepted. <sup>a</sup>Others include enzymes, kinases, and others. <sup>b</sup>Signal transduction includes growth factors, ligands, receptors, and cytokines.

\*In the data, log<sub>2</sub>FC of SOX9 at 96h was -0.96.

## Discussion

The present study tested the hypothesis that differentially expressed regulatory genes during the first 1.5 hours will initiate chondrogenic divergence between interzone and anlagen cell cultures, and further distinctions will accumulate as time passes in culture, represented by the activation of different pathways that leads to different ECM profiles. Accepting this hypothesis, the transcriptomic analyses in the current chapter demonstrated the kinetics of gene expression changes over 96 hours in the experimental and control cell cultures, characterized common molecular properties of the two chondrogenic cell lines compared to the negative control fibroblasts, and identified cell type-specific signatures in interzone and anlagen cell cultures.

### ***Kinetics of gene expression patterns during 96 hours of in vitro chondrogenesis***

#### TGF- $\beta$ 1 induced chondrogenesis

It should be mentioned that the upstream regulator prediction analysis projected TGF $\beta$ 1, which encodes the experimental stimulus (TGF- $\beta$ 1), as a common activated upstream regulator only between the baseline and the first collection point (1.5h) in all three cell types (Figure 3.5, 3.9, and 3.13). This result confirms that each one responded to the experimental chondrogenic stimulation induced by the TGF- $\beta$ 1 treatment during the first 1.5 hours. Since TGF- $\beta$ 1 is auto-regulated and this analysis is based on computational prediction, an important question arises: would different chondrogenic factors also activate TGF- $\beta$ 1 pathways during *in vitro* chondrogenesis? A mechanistic study with other chondrogenic inductive stimuli, such as IGF1, or blocking TGF- $\beta$ 1 signal is required to answer the question.

In addition, the hub gene analysis identified LTBP3 as a common hub gene in all three cell cultures (Figure 3.4, 3.8, and 3.12), and its expression was differentially upregulated at 96h compared to 12h in each cell type ( $\log_2FC=1.13$  in interzone cell cultures;  $\log_2FC=1.19$  in anlagen cell cultures;  $\log_2FC=1.45$  in fibroblast cultures; FDR adjusted P-value<0.001). This gene encodes an ECM component, and one of its important functions is to regulate the activity of the TGF- $\beta$  family (Robertson et al., 2015). By binding to TGF- $\beta$  propeptides, LTBP3 maintains their latent state, being stored in extracellular space. In the experimental setting, all three cell cultures were continuously exposed to the TGF- $\beta$ 1 treatment, and thus, the result presenting LTBP3 as a common hub gene may indicate that TGF- $\beta$  signaling mediation commonly occurred in all three cell cultures and was highly correlated to the time course.

#### Overall expression patterns of differentially expressed genes

All three cell cultures, including the negative control, shared similar expression pattern changes of their own DEGs responding to the chondrogenic stimulation during the 96-hour experimental period. Fewer DEGs were detected between 0h and 1.5h (106 – 170 DEGs), and greater DEGs were detected between 12h and 96h (1,433 – 2,395 DEGs; Table 3.3). In most cases, except the upregulated DEGs at 96h from 12h in interzone cell and fibroblast cultures, the numbers of DEGs between two consecutive time points within a cell type kept increasing throughout the time sequence. This result may imply exponential responses to an upstream change in the downstream; a transcription factor can regulate expression of multiple genes, a ligand can interact with different receptors, various signaling cascades can be turned on or off from the same upstream signal, and all of these processes orchestrate complex gene expression. Thus, the transcriptomic changes might accumulate over time, being represented

by the increased numbers of DEGs towards the delayed time points.

It is also interesting to note that most of the DEGs upregulated at the earlier time points became downregulated at the later time points, and *vice versa*. Between the baseline and the first collection point (1.5h), the DEG expression patterns in all cell types were more correlated with biological replicates than time points (Figure 3.2.a, 3.6.a, and 3.10.a). After 3 hours being grown in the chondrogenic medium, gene expression profiles were distinctively clustered by time point. The chondrogenic stimulation gradually shifted the regulation patterns (upregulated and downregulated) of DEGs to the opposite towards the end of the experimental period. When comparing the baseline gene expression to that of 96h, a very clear trend of upregulated genes and downregulated genes switching the patterns was observed in all three cell cultures.

#### Overrepresented biological processes in the time course

While the DEG expression pattern changes during the 96-hour *in vitro* chondrogenesis were similar among the cell types, overrepresented biological processes in each cell culture were evaluated to further characterize common and cell type-specific responses to the chondrogenic stimulation (Table 3.4, 3.5, 3.7, 3.8, 3.10, and 3.11). From the baseline to the first collection point (1.5h) after inducing chondrogenesis, the upregulated DEG profiles in all cell types were highly related to transcription regulatory events. Between 1.5h and 3h, skeletal muscle cell and fat cell differentiation processes were significantly downregulated in all cell types under the chondrogenic stimulation. Since mesenchymal stem cells derived from embryonic mesoderm have multipotency to differentiate into myocytes, adipocytes,



osteocytes, as well as chondrocytes, the downregulation of skeletal muscle cell and fat cell differentiation processes may confirm that all cell lines were receiving chondrogenic stimulation, and this stimulation initiated the downregulation of the other cell lineage differentiation pathways. Towards the last collection point (96h), it is noticeable that biological processes related to ECM metabolic events and involving ECM genes were overrepresented in all cell cultures. Altogether, the gene ontology analysis demonstrated the model of molecular event changes that was derived from Chapter 2 in which altered transcriptional events would lead to subsequent changes in signaling cascades and eventually in ECM biology.

#### ***Chondrogenic characteristics observed in both interzone and analgen cell cultures***

Common chondrogenic characteristics of the two skeletal cell lines were evaluated by analyzing the overlapping DEGs from the comparisons of each chondrogenic cell type to the negative control at each time point. These common chondrogenic DEGs also include the candidate molecular switches that were exclusively turned on or off in both chondrogenic cell lines, but not in the negative control fibroblasts (Table 3.16 and 3.17). Along with the DEGs, several interesting common hub genes correlated with the time course and projected upstream regulators are proposed in this section.

#### **Presumptive chondrogenic molecular switches**

GSC was exclusively expressed in chondrogenic cell lines but not in fibroblast cultures at all time points after inducing chondrogenesis (1.5 – 96h). This gene is a marker of mesoderm during gastrulation (Blum et al., 1992) and mesenchymal cell lineages in fetal limb buds (Gaunt et al., 1993). A gain-of-function experiment conducted in a chick embryo suggested that GSC

participated in limb formation and patterning through regulating expression of HOXA and HOXD clusters (Heanue et al., 1997). In the dataset, HOXD13 expression was also exclusively observed in the chondrogenic cell cultures by the last time point (96h). This gene is expressed in distal regions of limb buds and regulates digit patterning (Nelson et al., 1996). The exclusive expression of GSC and one member of the posterior HOX genes in cluster D might suggest that the chondrogenic cell cultures retained some regulatory processes observed during the fetal limb formation.

On the other hand, there were greater numbers of genes that were not expressed in the chondrogenic cell lines while exclusively expressed in fibroblasts. These genes include several central and posterior Homeobox transcription factor genes in cluster B (HOXB4, HOXB5, HOXB8, and HOXB9). In addition, EN1 was recorded as a candidate molecular switch turned off in the skeletal cell lines at four time points including the baseline (0, 3, and 12h).

Both GSC and EN1 are members of Homeobox transcription factor gene families, acting as a transcription repressor, and they were expressed in the opposite way between the chondrogenic cell lines and the negative control. Thus, a gain-of-function or loss-of-function study might elaborate if these genes are involved in the cellular fate decision whether to commit chondrogenic differentiation. Also, further research on the regulatory processes of these two genes and their relationship with other HOX gene families as well as other genes involved in development might bring important information to understand the regulation of chondrogenic pathways during the limb development.

### Overrepresented biological processes in chondrogenic cell lines

Gene enrichment analysis identified 128 upregulated and 133 downregulated overrepresented biological processes in the chondrogenic cell lines based on the negative control across the entire experimental period (Table 3.14). It is notable that a greater number of significant biological processes were observed towards the later time points (from 69 processes at 0h to 120 processes at 96h). This result might suggest that under the chondrogenic inductive stimulation, further differences between the skeletal cell lines and the non-skeletal fibroblast cultures were developed over time.

Among the biological processes that were upregulated in the chondrogenic cell lines, three GO terms were repeatedly observed at all collection points (Table 3.18). Noticeably, several central and posterior Homeobox transcription factor genes in clusters A and D were consistently upregulated in the chondrogenic cell cultures based on the negative control and categorized into these three common biological processes identified at all time points. In the meantime, the commonly downregulated DEGs in the chondrogenic cell cultures showed four common biological processes that were overrepresented across the whole experimental period (Table 3.19).

Before inducing chondrogenesis and also at all post-chondrogenic induction time points, the positive regulation of chondrocyte differentiation process was identified as one of the overrepresented biological processes from the common upregulated DEG profile in interzone and analgen cell cultures compared to the negative control. This process includes the SOX gene family (SOX5, SOX6, and SOX9). SOX6 negatively regulated RUNX2—an osteogenic

transcription factor—and activated S100A1, which inhibits chondrocyte terminal differentiation (Saito et al., 2007), and therefore prevented chondrocytes from entering pre-hypertrophic stages (Smits et al., 2004). At the earlier stages of embryonic development, SOX9 is expressed in mesenchymal cells and a mesenchymal condensation—the cartilaginous tissue prior to hypertrophic differentiation (Decker et al., 2014). When SOX9 was knocked out, it resulted in apoptosis (Akiyama et al., 2002) as well as decreased mesenchymal condensation and chondrocyte formation (Bi et al., 1999). Also, SOX9 is required for SOX5 and SOX6 expression (Akiyama et al., 2002) although SOX5 and SOX6 are not essential for chondrogenic differentiation unlike SOX9 (Liu and Lefebvre, 2015). Thus, expression profiles of these genes in the dataset might demonstrate close relationships between SOX genes and chondrocyte differentiation and also confirm chondrogenic differentiation in interzone and analgen cell cultures.

Of note, both upregulated and downregulated DEGs in the chondrogenic cell lines resulted in anterior/posterior pattern specification as an overrepresented biological process. In this process defined by the upregulated DEG profile, HOX genes in clusters A and D were involved. On the other hand, the downregulated DEGs in this biological event were HOX genes in clusters B and C as well as EMX2, a Hox-related gene. It is noteworthy that some central and posterior HOX genes in cluster B were exclusively turned off in the chondrogenic cell lines while they were expressed in the negative control (Table 3.17).

Homeobox genes encoding transcription factors are highly conserved between species and known to regulate embryonic development and morphogenesis. While the HOX families

are involved in axial skeleton patterning, the posterior HOX paralogs (HOX9-13) are also involved in the development of appendicular skeleton, the limb skeletal elements (Pineault and Wellik, 2014). The posterior HOX genes in clusters A and D are expressed in forelimb buds, and the posterior HOX genes in cluster C are expressed in hindlimb buds along the proximodistal axis. From the dataset, the central and posterior HOX gene clusters A and D were upregulated in chondrogenic cell lines while the anterior and central HOX gene clusters B and C were either switched off or downregulated. The contrasting HOX gene expression patterns between the chondrogenic cell lines and the negative control may confirm the origin of the cells. In addition, GSC was turned on in both interzone and analgen cell cultures at all time points after inducing chondrogenesis, and this gene is involved in limb formation and patterning through regulating expression of HOX clusters A and D, but GSC did not affect expression of HOX cluster C (Heanue et al., 1997). Thus, further investigation in regulatory events involving the HOX gene families might elaborate the current knowledge of chondrogenic differentiation pathways during limb development.

#### Common chondrogenic hub genes correlated with the time sequence

The chondrogenic cell lines shared 53 hub genes highly correlated with the time sequence, and IGF2, SNAI2, and COMP are examples of them (Figures 3.4 and 3.8). IGF1 is a hub gene in only interzone cell cultures and interacts with another hub genes highly related to IGF2 and SNAI2 (Figure 3.4). Both IGF isoforms can bind both IGF receptors (IGF1R and IGF2R) and activate downstream, but the downstream pathways of these receptors have different effects. IGF1R have been reported to induce hypertrophy in skeletal myocyte (Musarò et al., 1999) and chondrocyte (Rokutanda et al., 2009, Sun and Beier, 2014). On the other hand, IGF2R lacks

kinase domains, and thus, it is known to result in no downstream signaling events, only attenuating the IGF1R pathway by sequestering IGF ligands. However, both IGF receptors were not differentially expressed in any comparisons from the dataset. Further research on the IGF ligands and receptors would be beneficial to better understand the roles of the IGF signaling pathways in chondrogenic differentiation.

Although SNAI2 was a common hub gene highly correlated with the time course in the chondrogenic cell lines, SNAI1—another Snail transcription factor involved in epithelial-mesenchymal transition—was not identified as a hub gene in either of cell cultures. However, SNAI1 was differentially upregulated at 1.5h compared to the baseline in both interzone and analgen cell cultures ( $\log_2FC=3.27$  and  $2.77$  in interzone and analgen cell cultures, respectively; FDR adjusted P-value=0.003). SNAI2 and SNAI1 compensate the loss of each other and regulate chondrogenic differentiation during limb formation (Chen and Gridley, 2013). Also, TGF- $\beta$  sustained upregulation of both SNAI2 and SNAI1 in human corneal epithelial cell cultures (Aomatsu et al., 2011). Therefore, the profiles of Snail transcription factors in the chondrogenic cell lines may suggest that SNAI1 might be responsible for the initial responses to the TGF- $\beta$ 1 treatment, and SNAI2 expression might interact with other molecular responses to the chondrogenic stimulation along the experimental period.

COMP, a classic cartilaginous biomarker, was a hub gene in the chondrogenic cell lines, but not in fibroblast cultures, and the expression levels were increased over time. Also, its expression was significantly higher in the skeletal cell cultures at all post-chondrogenic induction time points compared to the negative control (1.5 – 96h;  $\log_2FC$  ranged from 2.00 –

4.85; FDR adjusted P-value<0.01). Thus, the expression profiles of COMP evaluated in the study may confirm that interzone and analgen cell cultures indeed underwent chondrogenic differentiation, accumulating cartilaginous ECM.

#### Common chondrogenic upstream regulators

While no common chondrogenic upstream regulators were predicted until 3h, inhibition of MYC, TBX2, NOTCH1, SMAD7 (involved in transcription regulation), CCL2, CCL11, CSF1, TLP7, and MFSD2A (involved in signaling transduction) was projected in the chondrogenic cell lines between 3h and 12h. One of the inhibited upstream regulators, SMAD7, represses canonical TGF- $\beta$  pathways regulated by Smad2/3 and Smad1/5/8 (Wang et al., 2014). The SMAD2/3 involved pathway induces cartilaginous phenotypes. On the other hand, the SMAD1/5/8 downstream pathway exhibits hypertrophic marker genes, such as COLX and MMP13.

Between the two last collection points (12h and 96h), NKX3-2 activation was predicted in both chondrogenic cell cultures. This gene was also a hub gene in anlagen cell cultures, but it was not significantly correlated with the time course in interzone cell cultures. In the literature, NKX3-2 was not expressed in cells at the surface of developing bones, including cells that would differentiate into articular chondrocytes; on the other hand, its expression was detected in proliferative chondrocytes at pre-hypertrophic stages (Church et al., 2005). When GO enrichment analysis was conducted on the upregulated DEGs in interzone cell cultures at 96h compared to 12h, endochondral ossification and negative regulation of chondrocyte differentiation were significantly overrepresented (Table 3.4). NKX3-2 was one of the players in the overrepresented process of “negative regulation of chondrocyte differentiation.”

However, the expression levels of NKX3-2 were significantly lower in interzone cell cultures compared to anlagen cell cultures at all collection points ( $\log_2FC$  ranged from  $|2.12 - 3.62|$ ; FDR adjusted  $P$ -value $<0.01$ ). Based on the literature, the findings from the current study may suggest that interzone cell cultures might obtain pre-hypertrophic characteristics, at least at a minimal level, towards the last time point while anlagen cell cultures constantly presented greater hypertrophic characteristics over time.

### ***Differential transcriptomic signatures between interzone and anlagen cells***

The transcriptomic signatures in interzone and anlagen cell cultures demonstrated the chondrogenic divergence directed towards articular and hypertrophic developmental pathways, respectively. The candidate molecular switches that were identified between these chondrogenic cell lines may be responsible for the decisions whether to commit one of the different chondrogenic programs. Furthermore, predicted upstream regulators within a cell type over the time course or between the cell types at each time point provided further evidence that interzone and anlagen cells were indeed diverged into their corresponding developmental pathways.

#### **Exclusively expressed genes in interzone cell cultures**

Among four exclusively expressed genes in interzone cell cultures over time, ENSECAG00000037611 (from 0h – 3h) is a novel gene, and C7H11orf52 (at 12h) encodes an uncharacterized protein (Table 3.21). The exclusive expression of ENSECAG00000037611 was continuous during the first 3 hours after inducing chondrogenesis, therefore this gene might be an interesting gene for a further investigation to define its role in chondrogenic inductive



mechanisms, especially in the articular chondrogenic pathway.

Expression of LY6G6C was turned on in interzone cell culture at 3h. This gene is a member of a cluster of leukocyte antigen-6 (LY6) and encodes proteins containing Ly6/uPAR (LU) domains in integral membrane receptors. The LU domains have various biological functions in eukaryotes, and one of their important roles is being the extracellular ligand-binding domain in the TGF- $\beta$  receptor family, including TGF- $\beta$ R1, BMPR1A, TGF- $\beta$ R2, and BMPR2 (Leth et al., 2019). Not only TGF- $\beta$ 1 was the experimental stimulus, but also is a key upstream regulator of chondrogenic differentiation pathways. Interacting with different TGF- $\beta$  receptors, this cytokine can activate both canonical and non-canonical pathways and regulate both articular and hypertrophic chondrogenic differentiation (Wang et al., 2014). Therefore, localizing this specific LU domain encoded by LY6G6C in different TGF- $\beta$  receptors may better describe its role in interzone cell-specific chondrogenic pathways.

#### Exclusively expressed genes in anlagen cell cultures

Members of the anterior HOX gene cluster B were exclusively expressed in anlagen cell cultures at the baseline (HOXB2) and at delayed time points (HOXB2 and HOXB3 at 12h and 96h; Table 3.21). Another gene exclusively expressed in anlagen cells is PLAC8B, which was turned on by 96h. In addition to HOXB2 that has been reported to regulate self-renewal processes (Phinney et al., 2005), PLAC8B is assigned to a GO term of “chromatin binding.” Thus, the results may suggest that regulating DNA replication and proliferation might be important processes in anlagen cell cultures under the chondrogenic stimulation.

## Differentially regulated biological processes and gene expression between chondrogenic cell lines

During the first 3 hours after inducing chondrogenesis, the DEG profiles between the chondrogenic cell lines represented a higher number of biological processes that were significantly upregulated in anlagen cell cultures (Figure 3.15 and Table 3.24). Similar to what the exclusive gene expression in anlagen cell cultures may suggest, DNA replication was upregulated in anlagen cell cultures between 0h and 3h compared to interzone cell cultures. At the baseline, regulation of cell proliferation and positive regulation of epithelial cell proliferation were overrepresented by the upregulated DEG profile in anlagen cells compared to interzone cells. Also, chromosome segregation and cell division were upregulated in anlagen at 3h. Altogether, more proliferative events might occur in anlagen cell cultures at the earlier stages after inducing *in vitro* chondrogenesis. The present findings are aligned with the biology; proliferating anlagen chondrocytes enter the hypertrophic final differentiation in a developing limb bud.

In contrast, towards the later time points, the DEG profiles between interzone and anlagen cell cultures represented greater biological processes upregulated in interzone cell cultures. Cell adhesion and collagen fibril organization were upregulated in interzone cell cultures at 12h. Positive regulation of ERK1/2 cascade, embryonic digit morphogenesis, chondrocyte development, and chondrogenic differentiation processes involved greater DEGs upregulated in interzone cell cultures at 96h. At the same time, DEGs upregulated in anlagen cells dominated in a GO term—ossification—at 12h and in endochondral ossification at 96h. The results may confirm that anlagen cell cultures entered the chondrogenic pathways that would

lead towards the bone forming processes.

In Chapter 2, FGF1 was upregulated in anlagen cultures from 6 – 48h compared to interzone cultures. Along with this previous finding, from this RNA-seq dataset, differential expression of FGF1 between the two chondrogenic cell lines was detected at 12h ( $\log_2FC=|3.43|$ ; FDR adjusted P-value=0.001) and at 96h ( $\log_2FC=|1.89|$ ; FDR adjusted P-value=0.001), and the gene was upregulated in anlagen cells. When human mesenchyme stromal cells were co-cultured with chondrocytes from osteoarthritic cartilage, FGF1 expression was increased in mesenchymal cells and promoted proliferation of osteoarthritic chondrocytes (Wu et al., 2013). When FGF1 activity was inhibited, proliferation of chondrocytes was also downregulated. In addition, FGF1 expression was detected from the proliferative and hypertrophic zones of fetal growth plate but not from the resting zone (Krejci et al., 2007). Thus, based on the previous studies, the present finding may suggest that anlagen cell cultures might undergo proliferation by 12h after inducing chondrogenesis and might become further hypertrophic over time.

Several HOX genes in clusters B and C were either exclusively expressed (HOXB2 and HOXB3; Table 3.21) or significantly upregulated (HOXC4, HOXC6, HOXC8, and HOXC9; Table 3.23) in anlagen cell cultures compared to interzone cell cultures. On the other hand, some posterior HOX genes in clusters A and D (HOXA11, HOXA13, HOXD11, and HOXD13) were upregulated in interzone cell cultures. It is worth mentioning that other HOX genes in cluster B around the central position were not expressed at all in both chondrogenic cell lines (HOXB4, HOXB5, HOXB8, and HOXB9; Table 3.17) while some anterior HOX genes (HOXB2 and HOXB3)

were only expressed in anlagen cell cultures but not in interzone cells. In addition, differential expression of HOX genes in cluster C around the central region (HOXC4, HOXC6, HOXC8, and HOXC9) were differentially upregulated in anlagen cultures compared to interzone cell cultures. It is known that the HOX genes' function is more similar to that of their paralogs in different chromosomes relative to that of other HOX genes that are closely located in the same chromosome. Therefore, the roles of HOXC4-9 in anlagen cell cultures might be related to non-chondrogenic pathways. Further research is required to better describe the mechanisms of HOX genes regulating limb development.

Also, one example of an interesting DEG between the two fetal skeletal cell cultures is TGFBR3. While multiple TGF- $\beta$  ligands can interact with multiple TGF- $\beta$  receptors, the downstream events of TGF- $\beta$  are mediated by different combinations of receptors, resulting in different chondrogenic outcomes. Among different TGF- $\beta$  receptor genes, TGFBR3 was differentially upregulated in interzone cell cultures at 12h ( $\log_2FC=1.01$ ; FDR adjusted P-value=0.001). TGFBR3 interacts with ALK5 and transduces the noncanonical SMAD-independent pathway (Iwata et al., 2012), and this receptor provides stable ligands to the receptor combination of TGFBR2 and ALK5, which facilitates the SMAD2/3 mediated canonical TGF- $\beta$  pathway (Shi and Massagué, 2003). SMAD2/3 are phosphorylated in this signaling cascade and become active transcription regulators, upregulating COL2A1 and ACAN expression. In addition, phosphorylated SAMD3 inhibits SMAD1/5/8, which induces hypertrophic differentiation, resulting in articular cartilage phenotypes (Chen et al., 2012, Wang et al., 2014). Thus, upregulation of TGFBR3 in interzone cell cultures may indicate that interzone cells at 12h might be directed to articular chondrogenic pathways. However, SMAD3

expression was differentially downregulated in interzone cell cultures at 12h ( $\log_2FC = |1.07|$ ; FDR adjusted P-value=0.005), and the other TGF- $\beta$  receptor genes or SMAD genes were not differentially expressed between the two chondrogenic cell cultures at any collection time points. To better understand TGF- $\beta$  signaling pathways involved in chondrogenesis may require further research on phosphorylation and protein level evaluation.

#### Upstream regulator prediction

<Upstream regulators in interzone cell cultures between time points>

Between the baseline and 1.5h, NOTCH1 and GATA1 involved pathways were predicted to be activated upstream of interzone cell cultures (Figure 3.5). Both NOTCH1 and GATA1 downregulate chondrogenesis. When NOTCH signaling was constitutively activated, formation of chondrogenic nodules was significantly reduced in mice fetal limb bud mesenchymal cells (Tian et al., 2015). GATA1 was reported to be induced by BMPs and correlated with BMPs' anti-chondrogenic activity, reducing expression of collagen type 2 (Karamboulas et al., 2010). Predicted activation of NOTCH1 and GATA1 regulatory pathways in interzone cell cultures during the first 1.5 hours might represent temporarily paused chondrogenesis in interzone tissue during earlier stages of limb development.

The activation of MEOX2 was projected as an upstream event in interzone cell cultures between 3h and 12h. Also, this gene was evaluated as a hub gene correlated with the time course in interzone cell cultures. Its expression levels were significantly increased between 3h and 12h ( $\log_2FC=3.73$ ; FDR adjusted P-value=0.015) and between 12h and 96h ( $\log_2FC=2.40$ ; FDR adjusted P-value<0.001) in interzone cell cultures. MEOX2 regulates chondrogenesis in

fetal axial skeleton as an upstream of PAX genes (Mankoo et al., 2003). Therefore, the results may suggest that chondrogenesis might resume in interzone cell cultures, 3 hours after introducing the TGF- $\beta$ 1 stimulation, and MEOX2 might be an interesting gene for further investigation to better understand the articular chondrogenic differentiation pathway.

During the last time sequence (12h to 96h), COL9A1 was predicted as an activated upstream in interzone cell cultures. This gene encodes an ECM component of hyaline cartilage. Also, the null mutation of this gene resulted in increased ossification in mice femoral heads and also enhanced invasion of vessels (Heilig et al., 2018). Therefore, this predicted upstream between the two last time points in interzone cell cultures, treated with TGF $\beta$ -1, may indicate that the cells might be towards the articular chondrogenic pathway in this experimental setting.

<Upstream regulators in anlagen cell cultures between time points>

In anlagen cell cultures, SOX10 activation was projected as an upstream event from 1.5h to 3h (Figure 3.9). SOX10 was expressed in hypertrophic cutaneous scars (Febres-Aldana and Alexis, 2020) and in hypertrophic nerve trunks (Sham et al., 2001). Also, some genes involved in hypertrophic differentiation during endochondral ossification—DLX5 and DLX6—were predicted as upstream regulators between 3h and 12h in anlagen cell cultures. Both DLX5 and DLX6 were involved in hypertrophic differentiation, and DLX5 was able to compensate the absence of DLX6 expression (Zhu and Bendall, 2009). Between the two last time points (12h and 96h), SP7, a classic marker of bone, was an activated upstream regulator. By cooperating with RUNX2, this gene upregulated MMP13 expression and cartilage mineralization

(Nishimura et al., 2012). Thus, the results altogether may indicate that anlagen cell cultures might enter pathways towards hypertrophic differentiation at a relatively earlier time point (between 1.5h and 3h) and remain in the pathways leading to bone formation throughout the experimental period.

<Upstream regulators between interzone and anlagen cell cultures at each time point>

Including the baseline, NUPR1 was predicted to be an active regulator in interzone cell cultures compared to anlagen cell cultures during the first 3 hours (Figure 3.16). In the literature, it downregulated cell survival pathways, inducing chondrocyte apoptosis (Tan and Yammani, 2019). The present result agreed with a previous RNA-seq study conducted with murine tissues from mandibular condyle articular and mature zones (Zhou et al., 2018). The authors also analyzed the data using IPA, and NUPR1 was predicted as an upstream regulator distinguishing articular cartilage and the mature zone, which is located between articular cartilage and hypertrophic cartilage.

E2F1 was predicted as an active upstream regulator in anlagen cell cultures during the earlier time points (from the baseline until 3h). In a previous study, its overexpression prevented chondrocytes from hypertrophic differentiation and interrupted endochondral ossification (Scheijen et al., 2003). On the other hand, one of the activated upstream regulators in anlagen cell cultures at the baseline and 1.5h was SPP1. This gene encoding osteogenic glycoprotein attaches osteoclasts to bone matrix and is required for biomineralization (Peacock et al., 2011). While SOX9 binds to SPP1 and inhibits transactivation, at 96h, SOX9 was predicted to be an upstream gene, and at the same time, SPP1 was not

predicted to be an upstream regulator in anlagen cell cultures. RUNX2 (at 0h, 1.5h, and 96h) and RUNX3 (at 12h), well-established bone biomarkers, were two of the predicted active upstream regulators in anlagen cell cultures treated with the TGF- $\beta$ 1 treatment. Thus, the results may verify that anlagen cell cultures in the experimental setting might be heading towards the bone forming-hypertrophic chondrogenic pathway.

### ***Future direction***

This RNA-seq study identified several important DEGs, candidate molecular switches, and upstream regulators that were either common in the chondrogenic cell lines or specific to different chondrogenic pathways. Since the current data were generated by measuring mRNA levels, a new question arises: if the transcripts were translated to proteins. To be functional in biological processes, transcripts need to be translated, and the resulting proteins need to be in an active form. Therefore, a cross-validation at a protein level would further confirm the biological relevance of the new information obtained from the present study. Also, the functional annotations and regulatory mechanisms of transcriptomic signatures that were proposed in this chapter could be elucidated by mechanistic studies. In addition, with the emergent high-throughput sequencing technologies that have enabled evaluation of gene expression at a single cell level, profiling characteristics of subpopulations of each cell preparation would add beneficial information to better understand the biology behind these chondrogenic pathways.



## Conclusion

In summary, the transcriptomic data generated from equine fetal interzone and analgen cell pellet cultures during the 96-hour of *in vitro* chondrogenesis characterized kinetics of gene expression of each cell type. In agreement with the study conducted in Chapter 2, transcription regulatory responses were initiated by 1.5h after inducing chondrogenesis, alterations in signaling transduction were observed throughout the experimental period, and ECM-related signatures became more evident towards the latest time point (96h). The study provided evidence that the transcriptomic profile of interzone cell cultures might be directed to articular chondrogenic pathways, and that of analgen cell cultures might be towards hypertrophic pathways, leading to bone formation, while both cell cultures shared some common chondrogenic characteristics. Furthermore, this study proposed candidate molecular switches, important DEGs, and predicted upstream regulators between interzone and analgen cell cultures, and various Homeobox transcription factor genes are some of the examples. The transcriptomic signatures of interzone and analgen cell cultures defined by the current study further described the chondrogenic divergence of these two cell lines.

## **Chapter 4. A pilot single cell RNA-seq study: Evaluation of chondrogenic divergence between equine fetal interzone and anlagen cell cultures at the single cell level**

### **Introduction**

Experimental *in vitro* chondrogenic differentiation models frequently utilize TGF- $\beta$  containing induction media and measure classic cartilaginous biomarkers, such as cartilage-specific ECM genes or proteins after 14-21 days (McCarthy et al., 2012, Rakic et al., 2018, Adam et al., 2019). One such study from our laboratory compared equine fetal interzone and anlagen cell pellets and demonstrated a clear distinction of histological characteristics (Adam et al., 2019). While anlagen cell pellets showed relatively homogenous histology in sagittal sections, interzone cell pellets usually exhibited a heterogenous zonal structure.

To investigate these differences further, a preliminary histology analysis was conducted on equine interzone and anlagen cell pellets (500,000 cells/pellet) that were grown in the TGF- $\beta$ 1 chondrogenic induction medium and collected at three different time points, 1d (24h), 2d (48h), and 21d (Figure 4.1). In agreement with the previous study (Adam et al., 2019), day-21 anlagen cell pellets displayed homogenous proteoglycan staining across the entire section as well as relatively stable cellular morphology and arrangement. In contrast, and also consistent with the previous findings, interzone cell pellets after 21 days presented the distinctive zonal configuration. In the periphery region, ECM was rich in proteoglycan, and cells were densely arranged parallel to the surface. Across the intermediate zone, proteoglycan staining and cell density became decreased towards the core. The histology in the core region was relatively variable among technical replicates; some exhibited lower cell

density and latticed ECM while others showed evidence of cellular necrosis. The cellular morphology, orientation, and ECM proteoglycan staining observed in day-21 interzone cell pellets have some resemblance to the architecture of articular cartilage—almost suggesting an organoid structure.

In contrast to the 21d time point, anlagen pellets at 1d (24h) and 2d (48h) had significantly less proteoglycan staining (Figure 4.1). Differences between interzone and anlagen cell pellets were minimal in terms of cell size and histology at days 1 and 2.

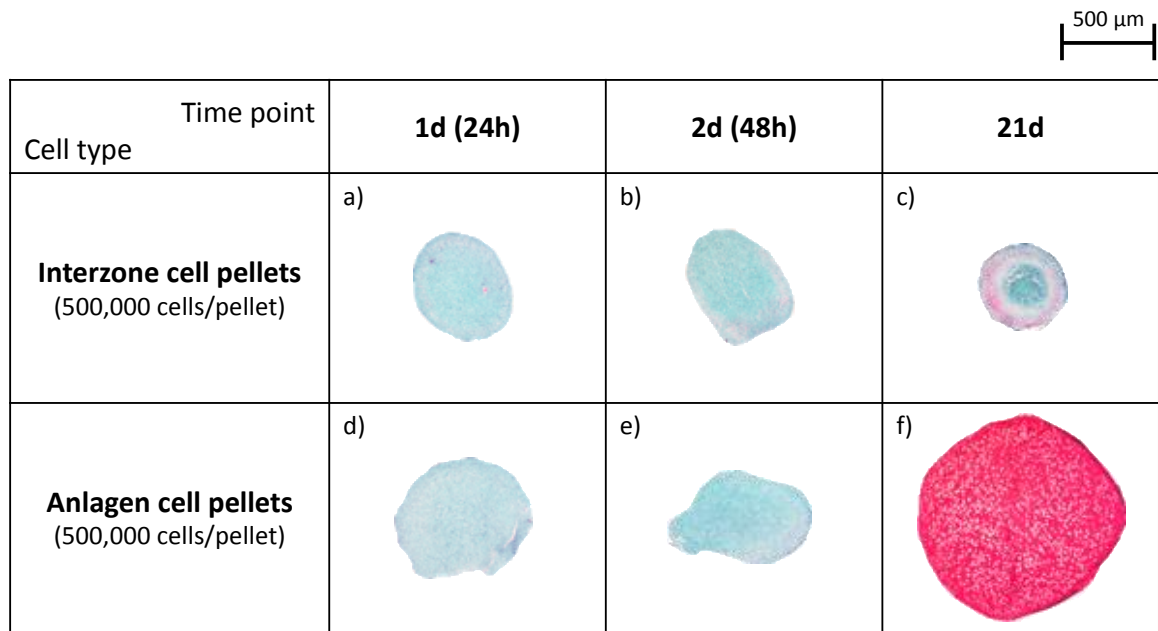


Figure 4.1 Cell pellet proteoglycan staining. Passage 5 equine fetal interzone and anlagen cell pellets were collected after 1, 2, and 21d of culture in chondrogenic induction medium and stained with Safranin-O and Fast Green. Proteoglycan in the extracellular matrix stains red, while regions devoid of proteoglycan stains blue-green.

The different levels of variation and intensity of proteoglycan staining between 21d interzone and anlagen cell pellets may indicate different cell subpopulations and cellular heterogeneity. With the complexity of gene expression profiles in kinetic data reported in the

previous chapters, in which changes induced by TGF- $\beta$ 1 were demonstrated as early as 1.5 hours after the initiation of *in vitro* chondrogenic induction, different cell subpopulations might be present by the 1d or 2d time points. Since differential expression of ECM effector genes was delayed, cell subtype differences may well be present even though histological staining characteristics are still similar.

In this pilot study, single cell RNA-seq analysis was conducted to further describe the differential chondrogenic pathways between interzone and anlagen cell cultures by profiling their cell subpopulations. The hypotheses tested in the study were that 1) interzone cell cultures will develop relatively heterogenous cell subpopulations while anlagen cell cultures will display less heterogenous subpopulations, and 2) these fetal skeletal cell lines will not only share common chondrogenic cell subpopulations but also show unique cell subpopulations that distinguish the two cell types.

## **Materials and Methods**

### ***Cell culture***

Primary interzone and anlagen cells collected from a single 45-day equine fetus (Adam et al., 2019) were used for the present study. Passage 5 cell pellets were established following the protocol described in Chapter 2 and cultured in the TGF- $\beta$ 1 chondrogenic induction medium. Samples were collected at two time points, 24 and 48 hours. These two time points were selected with consideration of gene expression profiles and for logistical considerations related to the technical ability to reestablish single cell suspensions from chondrogenic cell

pellet cultures. Regenerating cell suspensions was feasible until 48h without substantially compromising the cell number recovery rate or cell viability.

### ***Single cell suspension preparation***

The preparation of cell suspensions for the single cell cDNA library construction followed protocols recommended by the manufacturer (10X Genomics, Pleasanton, CA). Briefly, cell pellets, five per cell-type and time point, were washed with PBS to remove the chondrogenic medium. The five replicates were then transferred as a group into a cell strainer with a luer-lock flow controller (pluriStrainer® 5 µm; catalog No. 43-50005-13; pluriSelect Life Science, Germany) connected to a 50 ml conical tube. The pellets were incubated for 2.5 hours at 37°C in 1.5 ml of a collagenase cocktail, DMEM containing 20% (v/v) FBS, 0.2% (w/v) collagenase II (catalog No. LS004177; Worthington Biochem), 0.1% (w/v) collagenase type 4 (catalog No. LS004186; Worthington Biochem), and 2mM CaCl<sub>2</sub> (catalog No. 10043-52-4; Fisher Scientific). The collagenase medium was then drained through the filter by opening the luer-lock. Retained cells were washed twice with PBS, retrieved into a fresh 50 ml conical tube, and resuspended in DMEM.

The cell counting was conducted by both an automated cell counter (EVE™ Automated Cell Counter; catalog No. EVE-MC; NanoEnTek) and a manual hemocytometer with trypan blue staining. Cell viability was >95% and the concentration of cells in DMEM adjusted to 1,000 cells/µl, within the range recommended by 10X Genomics. The targeted number of cells for barcoded library construction was 5,000 cells for each sample. Per manufacturer recommendations, the starting cell number was set to 1.7 times greater per sample.

### ***Single cell cDNA library construction***

Barcoded gel beads are the key feature in 10X Genomics' single cell gene expression system. These beads comprised 1) an Illumina Read 1 primer site, 2) a 16 bp 10X barcode (specific to each cell), 3) a 10 bp unique molecular identifier (UMI; specific to each transcript), and 4) an oligo-dT sequence. To construct single cell cDNA libraries using a Chromium Single Cell 3' Library & Gel Bead v2 Kit (cat No. 120267; 10X Genomics) and following the company's protocols, the cell suspensions, barcoded gel beads, partitioning oil, and reverse-transcription reagents were channeled through a microfluidic chip using a Chromium controller (cat No. 110203; 10X Genomics). In this process, individual cells were captured with one barcoded gel bead and the RT reagents in the partitioning oil, forming a Gel bead-in-EMulsion (GEM). Within each GEM, RNA transcripts were released by cell lysis and then captured by oligo-dT sequences on the barcoded bead. A reverse transcription reaction was then conducted using a Veriti® 96-Well Thermal Cycler (cat No. 4375786; Thermo Fisher) for 45 min at 53°C and for another 5 min at 85°C. Finally, all of the GEM droplets were broken open to generate a pool of cell-specific barcoded cDNAs for each experimental group—interzone or anlagen cell type and either the 24h or 48h time point. The cDNAs in each sample were then amplified for 12 cycles in the Veriti® 96-Well Thermal Cycler, followed by an assessment of cDNA quality and quantity using the Agilent High Sensitivity DNA Kit (catalog No. 5067-4626; Agilent Technologies) reagents and a Bioanalyzer 2100 (Agilent Technologies) instrument.

After the amplification step, cDNA molecules were randomly fragmented. For fragments that contained the Read 1 primer sequence, Illumina P5 flow cell binding sequences, Illumina Read 2 primer sites, sample indexes, and Illumina P7 sequences were added using a Chromium

i7 Multiplex Kit (cat No. 120262; Table 4.1). Indexing PCR was conducted for 13 cycles in the Veriti® 96-Well Thermal Cycler. Final library concentration and quality were evaluated by a Qubit® 3.0 Fluorometer (cat No. Q33216; Life Technologies) with a Qubit® dsDNA HS Assay Kit (catalog No. Q32854; Life Technologies) and by a Bioanalyzer 2100, respectively (Table 4.2 and Figure 4.2).

Table 4.1. 10X Genomics' sample indexes

Sample ID	Index location	Index sequences (four oligonucleotides/sample)			
Interzone cells, 24h	SI-GA-A1	GGTTTACT	CTAAACGG	TCGGCGTC	AACCGTAA
Interzone cells, 48h	SI-GA-A2	TTTCATGA	ACGTCCCT	CGCATGTG	GAAGGAAC
Anlagen cells, 24h	SI-GA-A3	CAGTACTG	AGTAGTCT	GCAGTAGA	TTCCCGAC
Anlagen cells, 48h	SI-GA-A4	TATGATTC	CCCACAGT	ATGCTGAA	GGATGCCG

Chromium i7 Multiplex kit (cat No. 120262)

Table 4.2. cDNA library concentration and average size of fragments

Sample ID	Concentration <sup>a</sup> , ng/μl	Average fragment size <sup>b</sup> , bp
Interzone cells, 24h	47.6	498
Interzone cells, 48h	31.2	505
Anlagen cells, 24h	37.0	515
Anlagen cells, 48h	41.6	538

<sup>a</sup>Library concentration was measured by a Qubit® 3.0 Fluorometer (cat No. Q33216; Life Technologies) with a Qubit® dsDNA HS Assay Kit (catalog No. Q32854; Life Technologies).

<sup>b</sup>Average fragment size was evaluated by a Bioanalyzer 2100 (Agilent Technologies) using an Agilent High Sensitivity DNA Kit (catalog No. 5067-4626; Agilent Technologies).

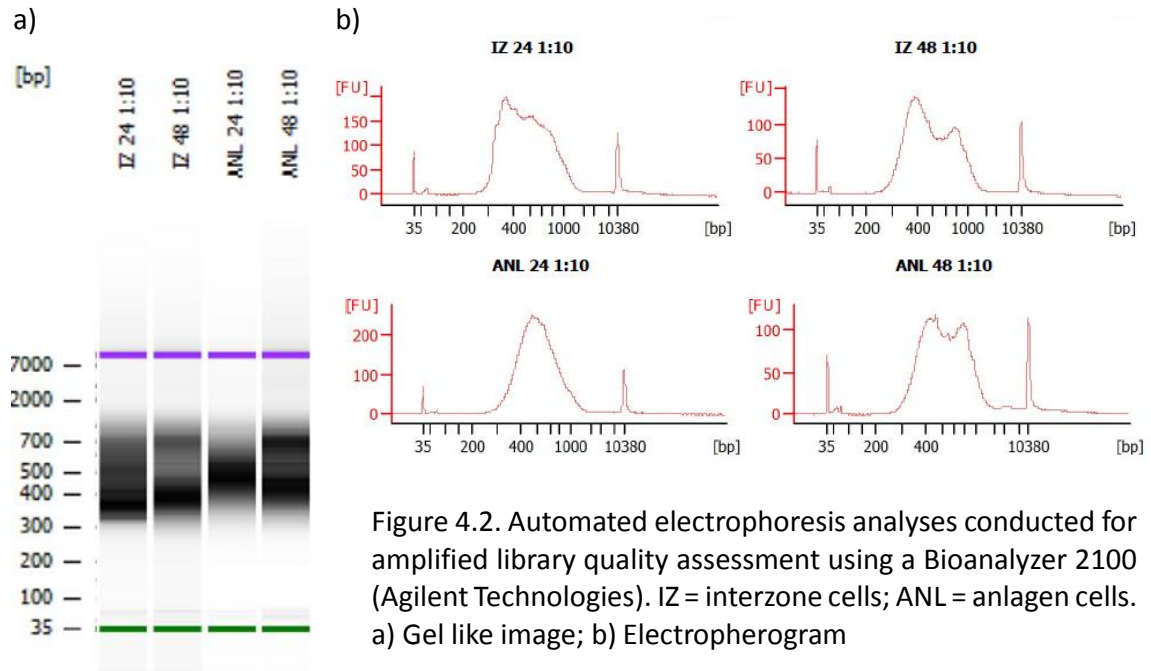


Figure 4.2. Automated electrophoresis analyses conducted for amplified library quality assessment using a Bioanalyzer 2100 (Agilent Technologies). IZ = interzone cells; ANL = anlagen cells. a) Gel like image; b) Electropherogram



### ***RNA-sequencing***

Aliquots of the four cDNA libraries (100 ng each) were pooled into a single tube totaling 400 ng in 20  $\mu$ l (20 ng/ $\mu$ l) and sent to the Roy J. Carver Biotechnology Center (University of Illinois at Urbana-Champaign, Urbana, IL, USA). The cDNA libraries were then sequenced (NovaSeq 6000, Illumina) on an S4 lane generating paired end reads, 2 $\times$ 150 nucleotides in length. Three read files were generated per sample: Read1 (reads for barcodes and UMIs), Read2 (reads for the transcripts), and Index1 (sequences of indexes) files. This is why the libraries in this system have to be sequenced in paired end reads. Also, the workflow does not include a fragmentation step for transcripts before being assigned to the UMIs; rather, poly-A tails of transcripts are captured by a gel bead in their full length. Therefore, levels of gene expression are expressed in UMI counts, and no further normalization for gene length is required.

### ***Data analysis pipeline***

Raw reads data were processed using the Cell Ranger software (v 3.0.2; 10X Genomics). In the layout for this software, STAR (v2.7) was used for aligning reads onto the latest horse reference genome (EquCab 3.0, GCA\_002863925.1; Kalbfleisch et al., 2018), and an ENSEMBL equine gene annotation file (v97) was used for expression quantification. Gene expression levels were expressed in UMI counts. Fold-change was calculated based on the ratio of the normalized mean UMI counts of a gene in each cell cluster relative to all of the other clusters. The thresholds defining DEGs were FDR adjusted P-value<0.05 and  $\log_2FC > |1|$ .

For non-linear dimensionality reduction analysis, uniform manifold approximation and projection (UMAP) was conducted using an R package, Seurat v 3.0 (Stuart et al., 2019). Then, the results were visualized by the Loupe Cell Browser (v 3.1.1; 10X Genomics). Expression of selected chondrocyte biomarkers (COL2A1, COMP, and ACAN), interzone cell biomarkers (GDF5, WNT9A, and ENPP2), and biomarkers of chondrocyte hypertrophic differentiation (COL10A1, MMP13, and DLX5) were visualized in UMAP graphs to evaluate subpopulations of cells with those characteristics. Gene ontology enrichment analysis was conducted by the Functional Annotation Tool from DAVID Bioinformatics Resources (v6.8; <https://david.ncifcrf.gov>; Huang et al., 2009) to describe cell clusters with a focus on biological processes.

## Results

While the targeted number of cells/sample was 5,000 cells, the recovered cell numbers ranged from 2,032 – 2,845 cells (Table 4.3). Although the cell number recovery rates were low (47 – 54%), other parameters indicated the abundance and high quality of the sequencing dataset. Ratios of GEMs containing 100% cell-associated UMIs based on the total cell number were greater than 72.5% in all samples, suggesting that most cells formed proper GEMs with 100% cell-originated UMIs. The percentages of fraction reads in cells, which indicates how many of the reads were integrated with barcodes that were associated with cells, ranged from 91.4 – 93.0%. This result confirmed that the vast majority of reads were associated with cells. Also, the mean reads/cell numbers that ranged from 219,117 – 332,136 were more than sufficient for a gene expression profiling study, suggested in the manufacturer's user guide, where the recommended minimum mean reads/cell number was 50,000. The average mapping efficiency from the four samples was 90.1%, and only reads that were mapped confidently to the genome (84.6 – 88.2% of the total reads) were processed through the downstream analyses.

Table 4.3. Sequencing results

Cell type	Interzone cell cultures		Anlagen cell cultures	
Time point	24h	48h	24h	48h
Total cell number	2,750	2,633	2,032	2,845
Number of GEMs containing 100% cell-associated UMIs <sup>b</sup>	2,255	2,107	1,474	2,094
Ratio of GEMs with 100% cell-associated UMIs based on total cell number, %	82.0	80.0	72.5	73.6
Mean reads/cell	273,499	227,277	332,136	219,117
Fraction reads in cells, %	92.7	93.0	91.4	92.3
Mapping efficiency, %	91.8	90.5	88.3	89.9
Reads mapped confidently to genome, %	88.2	86.7	84.6	86.0

<sup>a</sup>GEM, Gel bead-In Emulsion

<sup>b</sup>UMIs, Unique molecular identifiers

#### ***Cell subpopulations within a sample***

The analysis of each sample yielded 7-8 cell clusters (Table 4.4). Differentially expressed genes (DEGs) were identified in each cluster compared to the other clusters within a sample (Table 4.5), and the top five DEGs with the greatest fold change in each sample are listed in Table 4.6 – 4.9. Although the profiles of each cluster were different, subpopulation diversity or homogeneity was not remarkably different between interzone and anlagen samples in terms of the number of identified cell clusters and the distribution of cell numbers across the clusters.

Table 4.4. Number of cells in each cluster within sample

Interzone cell cultures				Anlagen cell cultures			
24h		48h		24h		48h	
Total	2,750	Total	2,633	Total	2,032	Total	2,845
Cluster 1	581	Cluster 1	477	Cluster 1	447	Cluster 1	543
Cluster 2	467	Cluster 2	433	Cluster 2	414	Cluster 2	478
Cluster 3	378	Cluster 3	412	Cluster 3	332	Cluster 3	431
Cluster 4	352	Cluster 4	385	Cluster 4	266	Cluster 4	399
Cluster 5	295	Cluster 5	385	Cluster 5	255	Cluster 5	397
Cluster 6	278	Cluster 6	372	Cluster 6	166	Cluster 6	357
Cluster 7	247	Cluster 7	169	Cluster 7	152	Cluster 7	240
Cluster 8	152						

Table 4.5. Number of differentially expressed genes in each cluster within sample

	Interzone cells, 24h			Interzone cells, 48h	
Cluster #	Upregulation	Downregulation	Cluster #	Upregulation	Downregulation
Cluster 1	20	49	Cluster 1	32	70
Cluster 2	2	89	Cluster 2	74	44
Cluster 3	1	59	Cluster 3	1	38
Cluster 4	25	0	Cluster 4	53	0
Cluster 5	85	8	Cluster 5	4	132
Cluster 6	16	43	Cluster 6	149	48
Cluster 7	74	22	Cluster 7	0	0
Cluster 8	0	0			
	Anlagen cells, 24h			Anlagen cells, 48h	
Cluster #	Upregulation	Downregulation	Cluster #	Upregulation	Downregulation
Cluster 1	20	162	Cluster 1	67	27
Cluster 2	0	95	Cluster 2	57	64
Cluster 3	33	45	Cluster 3	509	0
Cluster 4	97	34	Cluster 4	13	36
Cluster 5	458	0	Cluster 5	1	34
Cluster 6	0	0	Cluster 6	0	27
Cluster 7	0	0	Cluster 7	0	0

Differential expression was determined at thresholds where statistical significance and fold change (FC) are  $P\text{-value} < 0.05$  and  $\log_2\text{FC} > |1|$  ( $>1$ , upregulation;  $<-1$ , downregulation), respectively.

Table 4.6. Top five upregulated and downregulated differentially expressed genes in each cluster of interzone cell pellet sample at 24h

		Cluster 1			Cluster 2			Cluster 3			Cluster 4			Cluster 5			Cluster 6			Cluster 7			Cluster 8		
		Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value
Upregulation	NEFH	2.27	***	ENSECAG0000033856	1.22	***	IGF2	1.08	*	RF00100	7.47	***	CENPF	4.74	***	SCRG1	2.36	***	SDF2L1	3.92	***	None			
	SLC2A1	1.64	***	CCL2	1.16	***				ENSECAG0000029964	6.80	***	TPX2	4.52	***	COL8A1	1.64	***	HSPH1	3.34	***				
	IRF7	1.60	***							RF00619	6.56	***	ASPM	4.48	***	SELENBP1	1.40	***	HSPA1A	2.74	***				
	ANKRD37	1.46	***							ENSECAG0000037588	6.54	***	ENSECAG0000036105	4.46	***	LYPLAL1	1.26	**	ENSECAG0000028889	2.61	***				
	TNFRSF6B	1.44	***							ENSECAG0000039811	6.29	***	HMMR	4.16	***	RASL11B	1.24	**	CRELD2	2.42	***				
Downregulation	TPX2	-2.98	***	HMMR	-2.91	***	CENPF	-3.93	***	None			SCRG1	-1.78	***	ENSECAG0000036105	-3.60	***	RF00100	-2.66	**	None			
	CENPF	-2.87	***	ENSECAG0000014536	-2.73	***	TPX2	-3.29	***				OGN	-1.59	***	TPX2	-3.13	***	ASPM	-2.09	**				
	RF00100	-2.64	***	TPX2	-2.58	***	ASPM	-3.28	***				SDF2L1	-1.50	***	SGO1	-2.94	***	NEFH	-2.01	**				
	UBE2C	-2.58	***	ANKRD37	-2.53	***	ENSECAG0000036105	-3.05	***				ASPN	-1.45	**	ASPM	-2.87	***	SCRG1	-2.00	**				
	ENSECAG0000036105	-2.47	***	SGO1	-2.44	***	CENPE	-3.01	***				ADIRF	-1.30	**	UBE2C	-2.73	***	ANKRD37	-1.94	***				

log<sub>2</sub>FC = log<sub>2</sub>(fold change)

\*\*\*, P<0.0001; \*\*,P<0.01; \*, P<0.05; P-values are adjusted for false discovery rate.

Five genes with the greatest fold change were listed. Positive log<sub>2</sub>FC values represent upregulation, and negative log<sub>2</sub>FC values represent downregulation.

Table 4.7. Top five upregulated and downregulated differentially expressed genes in each cluster of interzone cell pellet sample at 48h

	Cluster 1			Cluster 2			Cluster 3			Cluster 4			Cluster 5			Cluster 6			Cluster 7		
	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value
Upregulation	CCL2	2.05	***	HILPDA	2.62	***	NEFH	1.28	*	ENSECAG0000031325	11.22	***	FIBIN	1.37	***	SDF2L1	4.16	***	None		
	NEXN	1.77	***	ADM	2.57	***				RF00100	7.95	***	ENSECAG0000036264	1.32	**	TRIB3	3.20	***			
	GALNT1	1.66	***	ANKRD37	2.51	***				ENSECAG0000039811	7.55	***	SCRG1	1.24	**	ZFAND2A	3.11	***			
	ACTA2	1.58	***	MSMO1	2.46	***				ENSECAG0000037588	7.41	***	PRELP	1.08	**	HERPUD1	2.86	***			
	SNX10	1.57	***	HMGB2	2.35	***				RAD51AP2	7.25	***				HSPA5	2.55	***			
Downregulation	TRIB3	-2.86	***	SDF2L1	-2.54	***	RF00619	-4.10	**	None			NUSAP1	-4.44	***	RF00100	-2.78	***	None		
	HILPDA	-2.75	***	FIBIN	-2.53	***	RF00100	-3.58	***				CENPF	-3.81	***	FBLN5	-2.49	***			
	ENSECAG0000029190	-2.74	*	ENSECAG0000037588	-2.47	*	ENSECAG0000037588	-3.23	***				UBE2C	-3.65	***	PRELP	-2.48	***			
	ENSECAG0000039811	-2.50	*	CCL2	-2.36	***	SDF2L1	-3.18	***				CDCA3	-3.27	***	SCRG1	-2.37	***			
	IRF7	-2.47	***	ENSECAG0000022459	-2.09	***	CENPF	-2.77	***				TANC2	-3.15	***	COMP	-2.34	***			

log<sub>2</sub>FC = log<sub>2</sub>(fold change)

\*\*\*, P<0.0001; \*\*, P<0.01; \*, P<0.05; P-values are adjusted for false discovery rate.

Five genes with the greatest fold change were listed. Positive log<sub>2</sub>FC values represent upregulation, and negative log<sub>2</sub>FC values represent downregulation.

Table 4.8. Top five upregulated and downregulated differentially expressed genes in each cluster of anlagen cell pellet sample at 24h

		Cluster 1			Cluster 2			Cluster 3			Cluster 4			Cluster 5			Cluster 6			Cluster 7		
		Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value
Upregulation	CCL2		2.11	***	None			SLC2A1	1.65	***	TPX2	4.2	***	RAD51AP2	9.55	***	None			None		
	ENSECAG0000033856		1.8	***				ENSECAG0000014086	1.54	***	SGO1	4.14	***	RF00100	9.28	***						
	GALNT1		1.59	***				GJB2	1.47	***	ENSECAG0000036105	3.95	***	MAFB	9.11	***						
	BLOC1S1		1.47	***				CALCB	1.36	**	UBE2C	3.85	***	ENSECAG0000039113	9.03	***						
	DSE		1.38	***				P4HA1	1.32	***	CENPF	3.83	***	ENSECAG0000039811	8.58	***						
		Cluster 1			Cluster 2			Cluster 3			Cluster 4			Cluster 5			Cluster 6			Cluster 7		
		Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value
Downregulation	ENSECAG0000037127		-4.57	***	TPX2	-3.45	***	MAFB	-3.87	*	RF00015	-4.94	***	None			None			None		
	RF00100		-3.91	***	NUSAP1	-2.86	***	RAD51AP2	-3.34	**	RF00100	-3.04	***									
	RAD51AP2		-3.10	***	RF00100	-2.96	***	CENPF	-3.29	***	MAFB	-2.89	*									
	ENSECAG0000029190		-2.88	**	UBE2C	-2.65	***	TPX2	-3.05	***	RAD51AP2	-2.62	*									
	MAFB		-2.87	**	CENPE	-3.22	***	CDC20	-3.02	***	ENSECAG0000029964	-2.54	***									

log<sub>2</sub>FC = log<sub>2</sub>(fold change)

\*\*\*, P<0.0001; \*\*, P<0.01; \*, P<0.05; P-values are adjusted for false discovery rate.

Five genes with the greatest fold change were listed. Positive log<sub>2</sub>FC values represent upregulation, and negative log<sub>2</sub>FC values represent downregulation.



Table 4.9. Top five upregulated and downregulated differentially expressed genes in each cluster of anlagen cell pellet sample at 48h

	Cluster 1			Cluster 2			Cluster 3			Cluster 4			Cluster 5			Cluster 6			Cluster 7		
	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value
Upregulation	TRIB3	1.76	***	CXCL8	3.16	***	RF00100	8.63	***	GJB2	2.10	***	IGF2	1.22	*	None			None		
	UBE2C	1.72	***	CRELD2	2.68	***	ARC	8.27	***	HILPDA	1.98	***									
	ARL6IP1	1.71	***	EMC9	2.15	***	HIST1H2AJ	8.16	***	IRF7	1.88	***									
	P4HA1	1.64	***	SDF2L1	2.04	***	ENSECAG00000039811	8.01	***	TANC2	1.73	***									
	SRPRB	1.62	***	MRPS16	1.86	***	MAFB	7.95	***	ANKRD37	1.72	***									

	Cluster 1			Cluster 2			Cluster 3			Cluster 4			Cluster 5			Cluster 6			Cluster 7		
	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value
Downregulation	RF00100	-3.19	***	RAD51AP2	-3.37	**	None			CXCL8	-3.90	**	CENPF	-3.67	***	ENSECAG00000039811	-3.47	***	None		
	MAFB	-2.71	**	RF00100	-3.17	***				CENPF	-3.85	***	HIST1H1A	-3.56	*	HIST1H2AJ	-3.16	*			
	ENSECAG00000029964	-2.67	***	ENSECAG00000037127	-3.12	**				HIST1H1B	-3.80	*	UBE2C	-3.46	***	RF00100	-2.73	***			
	ARC	-2.39	*	HIST1H2BB	-2.91	**				ANKRD1	-3.52	**	NUSAP1	-3.27	***	ANKRD1	-2.73	**			
	RAD51AP2	-2.31	*	ENSECAG00000037588	-2.80	***				RF00100	-3.23	***	RF00100	-3.23	***	ENSECAG00000037127	-2.68	*			

log<sub>2</sub>FC = log<sub>2</sub>(fold change)

\*\*\*, P<0.0001; \*\*,P<0.01; \*, P<0.05; P-values are adjusted for false discovery rate.

Five genes with the greatest fold change were listed. Positive log<sub>2</sub>FC values represent upregulation, and negative log<sub>2</sub>FC values represent downregulation.

### Biomarker expression

Steady state mRNA levels of established gene biomarkers in three different categories (chondrocyte, interzone, and hypertrophic chondrocyte) was visualized in UMAP graphs plotted for each sample (Figure 4.3, 4.4, and 4.5, respectively). Each dot in the graphs represents each cell and is colored by the maximum expression level ( $\text{Log}_2$  max feature UMI count) of the three biomarker genes in each category.

#### <Chondrocyte biomarkers>

In all samples, chondrocyte biomarker genes (COL2A1, COMP, and ACAN) were highly expressed in most of the cell subpopulations, although a couple of cell clusters had fewer cells expressing these transcripts (Figure 4.3). Expression of these ECM genes were higher at 48h compared to 24h in both cell types, interzone and anlagen.

#### <Interzone biomarkers>

Steady state levels of transcripts from interzone biomarker gene loci (GDF5, WNT9A, and ENPP2) were higher in the interzone cell cultures compared to anlagen samples (Figure 4.4). The expression of WNT9A was minimally detected in four clusters within the interzone cell samples at 24h and 48h (individual data not shown), while no cells from anlagen cell cultures at either time point expressed WNT9A. The clusters of cells that expressed greater chondrocyte biomarkers also showed greater expression of interzone marker genes.

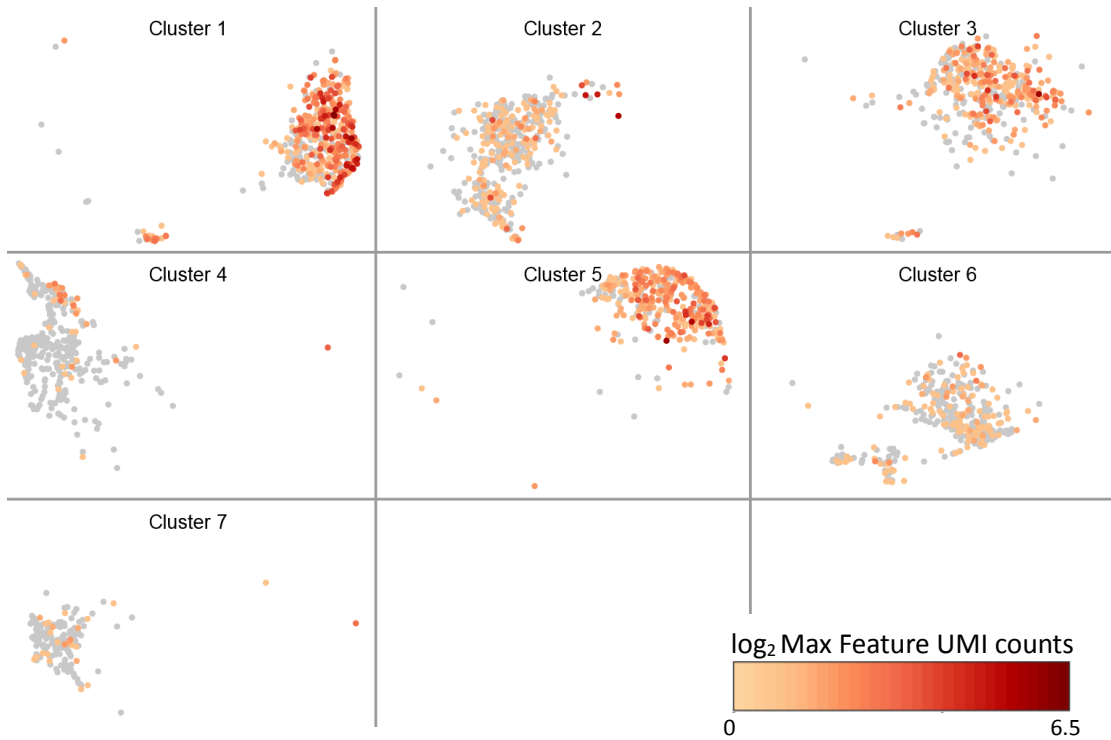
<Chondrocyte hypertrophy biomarkers>

While COL10A1 was not expressed in any sample, the two other chondrocyte hypertrophy biomarker genes (MMP13 and DLX5) were expressed at higher levels in anlagen cell cultures compared to interzone cell cultures over the experimental period (Figure 4.5). MMP13 expression in interzone cell cultures was not detected at 24h and minimal at 48h. In anlagen cell cultures, expression of MMP13 at both time points were modest compared to expression of DLX5 (individual data not shown).

4.3.a) Chondrocyte biomarker gene expression in interzone cell cultures at 24h



4.3.b) Chondrocyte biomarker gene expression in interzone cell cultures at 48h



4.3.c) Chondrocyte biomarker gene expression in anlagen cell cultures at 24h



4.3.d) Chondrocyte biomarker gene expression in anlagen cell cultures at 48h

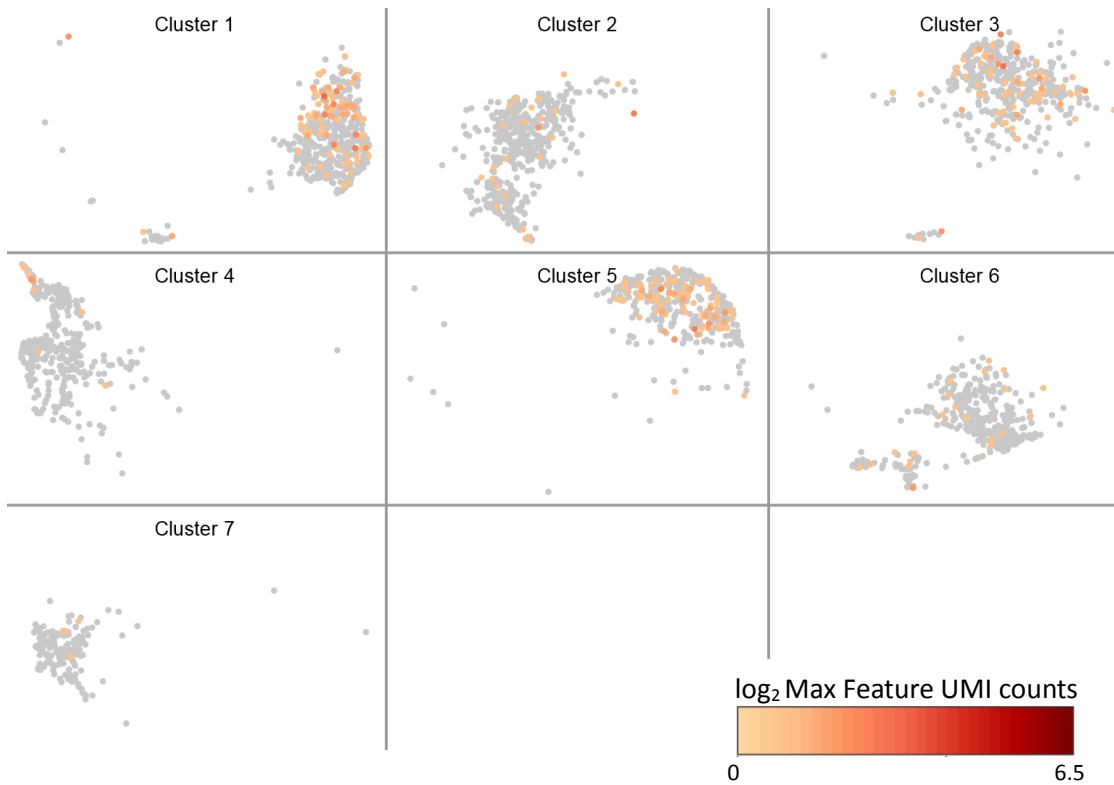


Figure 4.3. Chondrocyte biomarker gene (COL2A1, COMP, and ACAN) expression in each sample, plotted by uniform manifold approximation and projection (UMAP) analysis. a) Interzone cells at 24h; b) Interzone cells at 48h; c) Anlagen cells at 24h; d) Anlagen cells at 48h

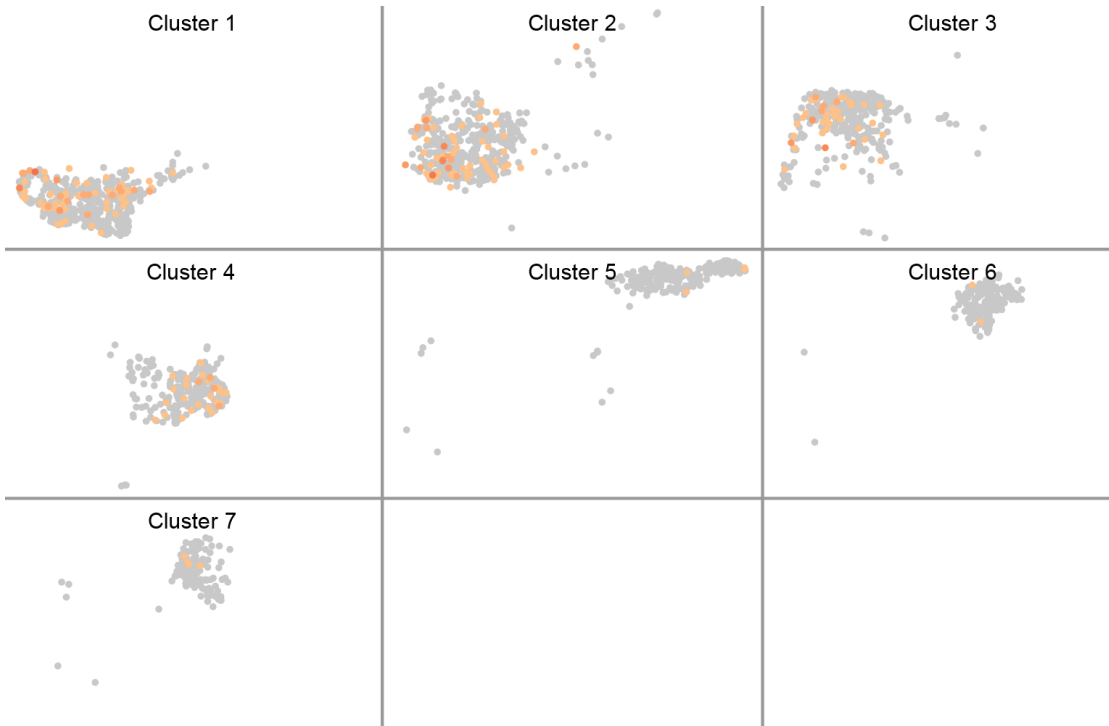
4.4.a) Interzone biomarker gene expression in interzone cell cultures at 24h



4.4.b) Interzone biomarker gene expression in interzone cell cultures at 48h



4.4.c) Interzone biomarker gene expression in anlagen cell cultures at 24h



4.4.d) Interzone biomarker gene expression in analgen cell cultures at 48h



Figure 4.4. Interzone biomarker gene (GDF5, WNT9A, and ENPP2) expression in each sample, plotted by uniform manifold approximation and projection (UMAP) analysis. a) Interzone cells at 24h; b) Interzone cells at 48h; c) Anlagen cells at 24h; d) Analgen cells at 48h

4.5.a) Chondrocyte hypertrophy biomarker gene expression in interzone cell cultures at 24h



4.5.b) Chondrocyte hypertrophy biomarker gene expression in interzone cell cultures at 48h





4.5.c) Hypertrophy marker gene expression in anlagen cell cultures at 24h



4.5.d) Chondrocyte hypertrophy biomarker gene expression in anlagen cell cultures at 48h

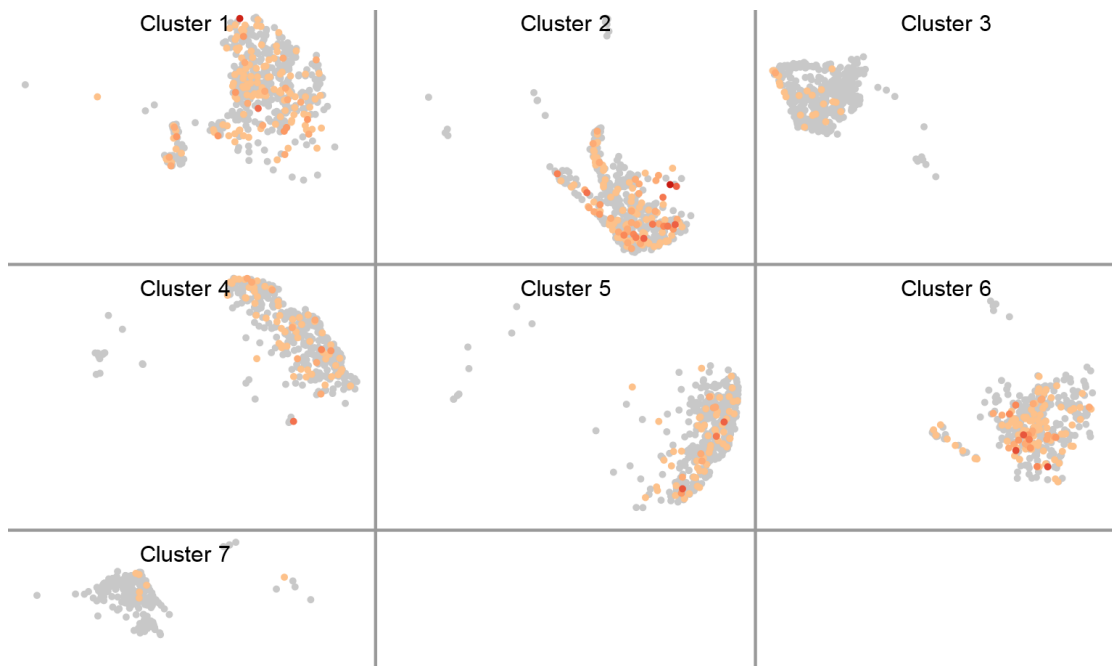


Figure 4.5. Chondrocyte hypertrophy biomarker gene (MMP13 and DLX5) expression in each sample, plotted by uniform manifold approximation and projection (UMAP) analysis. a) Interzone cells at 24h; b) Interzone cells at 48h; c) Anlagen cells at 24h; d) Anlagen cells at 48h

### ***Aggregation of the entire sample set***

To evaluate both common traits and differential signatures across the sample set, the single cell RNA-seq data from all samples were aggregated for the differential gene expression analysis and the dimensionality reduction analysis. By pooling the data from the four samples, the transcriptomic profiles of the entire sample set resulted in 13 cell clusters with measurements of 0 – 393 DEGs relative to the other clusters' gene expression profiles (Table 4.10). The five most differentially upregulated and downregulated genes in each cluster were tabulated (Table 4.11).

The gene expression patterns were visualized into UMAP graphs. Each dot in the graphs, representing each cell, is colored by cluster (Figure 4.6.a) or by sample (Figure 4.6.b). The UMAP analysis indicated that Cluster 8, 9, and 13 are characterized by commonly regulated genes among the samples (Figure 4.6.c and d). On the other hand, Cluster 6 and 12 were highly represented by interzone cells, while Cluster 2 and 4 were dominantly occupied by anlagen cells.

Table 4.10. Number of differentially expressed genes in each cluster of the aggregate of all four samples

	Upregulation	Downregulation
Cluster 1	5	137
Cluster 2	28	135
Cluster 3	28	32
Cluster 4	37	147
Cluster 5	18	41
Cluster 6	15	86
Cluster 7	121	18
Cluster 8	0	0
Cluster 9	311	0
Cluster 10	0	0
Cluster 11	60	4
Cluster 12	1	3
Cluster 13	393	0

Table 4.11. Five most upregulated and downregulated differentially expressed genes in each cluster of the aggregate of all four samples

	Cluster 1			Cluster 2			Cluster 3			Cluster 4			Cluster 5			Cluster 6			Cluster 7		
	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value
Upregulation	COL8A1	1.74	***	GJB2	2.18	***	DEFB1	2.01	***	ENSECAG0000028889	2.00	***	ANKRD37	1.59	***	IGF2	1.61	***	TPX2	5.02	***
	NEXN	1.41	***	NRN1	2.13	***	SCD	1.98	***	CXCL8	1.98	***	RASL11A	1.58	***	ENSECAG0000030236	1.56	***	CENPF	4.93	***
	RASL11B	1.17	***	ZBTB16	1.90	***	HILPDA	1.94	***	CCL2	1.58	***	GJB2	1.58	***	ASPN	1.51	***	ENSECAG0000036105	4.85	***
	FABP3	1.13	***	SNED1	1.82	**	HK2	1.94	***	ITGA10	1.54	***	HILPDA	1.51	***	RASL11B	1.33	**	SGO1	4.72	***
	ENSECAG0000012818	1.04	**	MRPS6	1.79	***	ADM	1.81	***	SLC20A1	1.48	***	SLC2A1	1.50	***	ADIRF	1.32	**	ENSECAG0000014536	4.70	***

	Cluster 1			Cluster 2			Cluster 3			Cluster 4			Cluster 5			Cluster 6			Cluster 7		
	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value
Downregulation	RF00100	-3.84	***	DEFB1	-5.62	***	ENSECAG0000039905	-3.47	**	DEFB1	-5.99	***	RAD51AP2	-4.61	**	ANKRD1	-4.80	***	RF00100	-2.91	***
	GATM	-3.69	*	CENPE	-4.07	***	RF00015	-3.19	**	ENSECAG0000033567	-4.38	***	MAFB	-3.71	**	HIST1H2BB	-4.58	**	MAFB	-2.70	*
	RF00030	-3.57	***	CENPF	-3.86	***	RF00100	-2.96	***	ENSECAG0000036746	-4.02	***	RF00100	-2.82	***	RF00100	-3.73	***	RF00015	-2.30	**
	RF00619	-2.94	***	ENSECAG0000036105	-3.59	***	KNL1	-2.81	**	CALCB	-3.43	***	CENPF	-2.77	***	TPX2	-3.59	***	OGN	-2.26	***
	RAD51AP2	-2.87	***	TPX2	-3.50	***	CENPF	-2.75	***	CDC20	-2.75	***	DEFB1	-2.76	***	CXCL8	-3.55	**	ASPN	-2.04	***

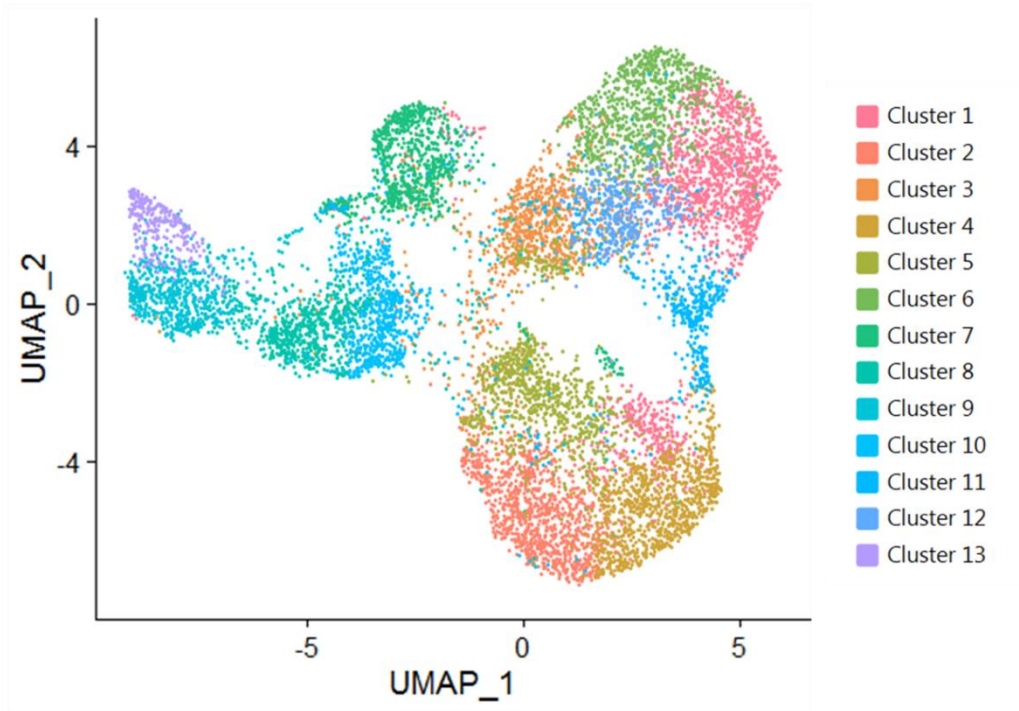
	Cluster 8			Cluster 9			Cluster 10			Cluster 11			Cluster 12			Cluster 13		
	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value
Upregulation	None			RF00003	6.57	***	None			SDF2L1	4.39	***	DEFB1	1.74	*	CITED4	10.96	***
				RF00100	6.13	***				ZFAND2A	2.99	***				MAFB	10.23	***
				ENSECAG0000034602	6.03	***				CRELD2	2.79	***				ARC	10.07	***
				RF00015	5.94	***				HERPUD1	2.78	***				HIST1H2AJ	10.04	***
				ENSECAG0000032769	5.84	***				TRIB3	2.70	***				GATM	9.86	***

	Cluster 8			Cluster 9			Cluster 10			Cluster 11			Cluster 12			Cluster 13		
	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value	Genes	log <sub>2</sub> FC	P-value
Downregulation	None			None			None			RF00100	-3.51	**	ENSECAG0000039811	-4.86	*	None		
										ASPM	-2.73	*	RF00619	-4.28	*			
										CENPF	-2.65	*	RF00100	-3.13	*			
										TPX2	-2.27	*						

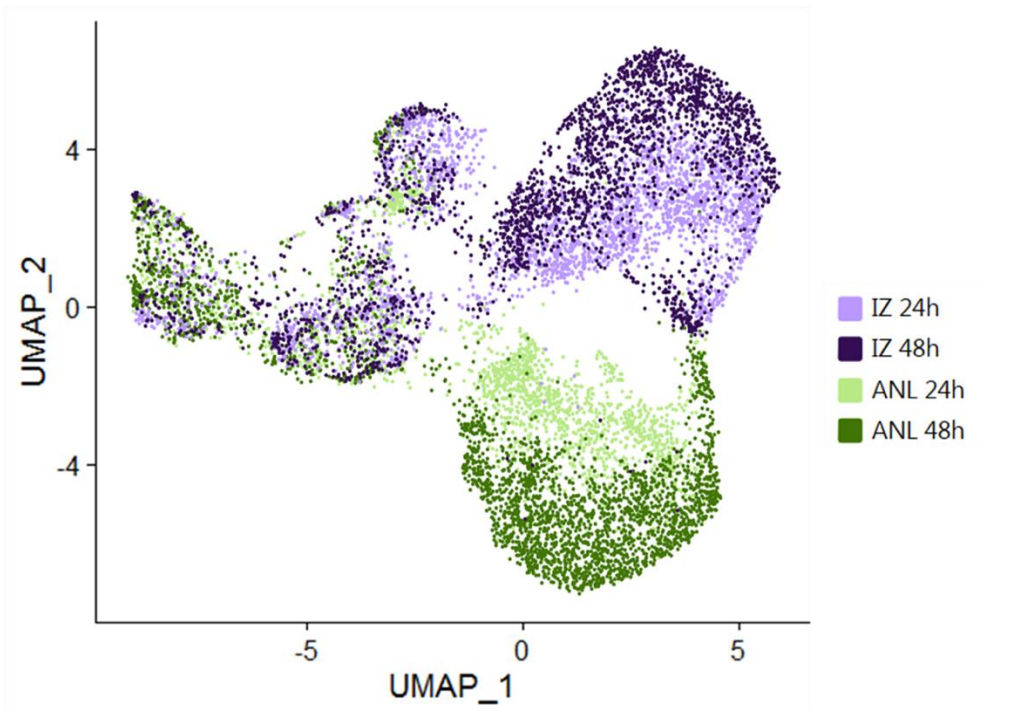
log<sub>2</sub>FC = log<sub>2</sub>(fold change); P-values are adjusted for false discovery rate; \*\*\*, P<0.0001; \*\*, P<0.01; \*, P<0.05

The five most upregulated DEGs and the five most downregulated DEGs in each cluster are listed. Positive log<sub>2</sub>FC values represent upregulation, and negative log<sub>2</sub>FC values represent downregulation.

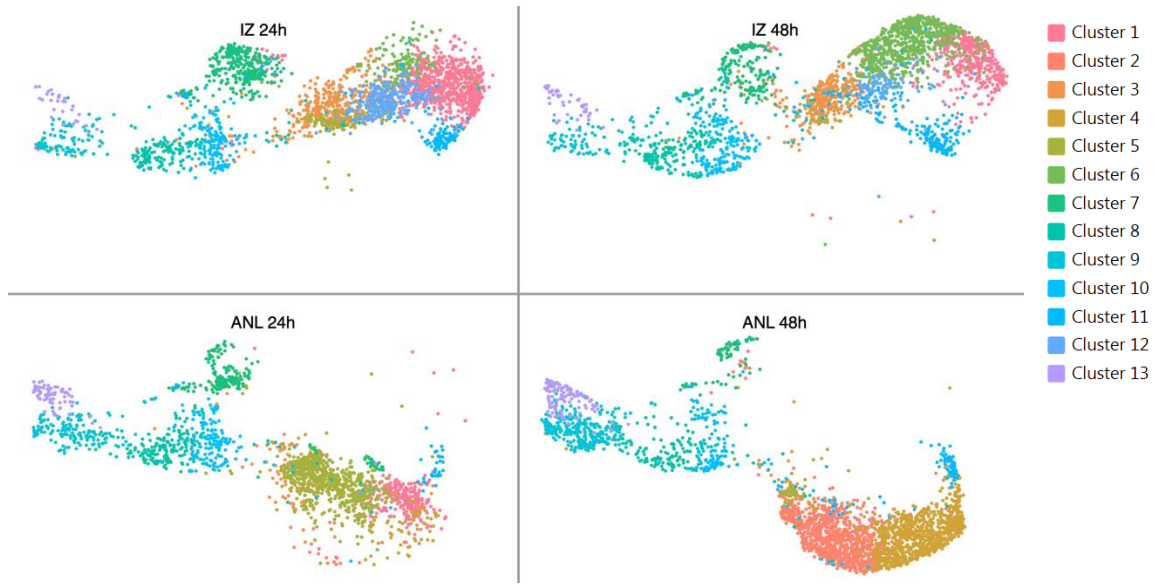
4.6.a) Colored by cluster



4.6.b) Colored by sample



4.6.c) Separated view by sample, colored by cluster



4.6.d) Separated view by cluster, colored by sample

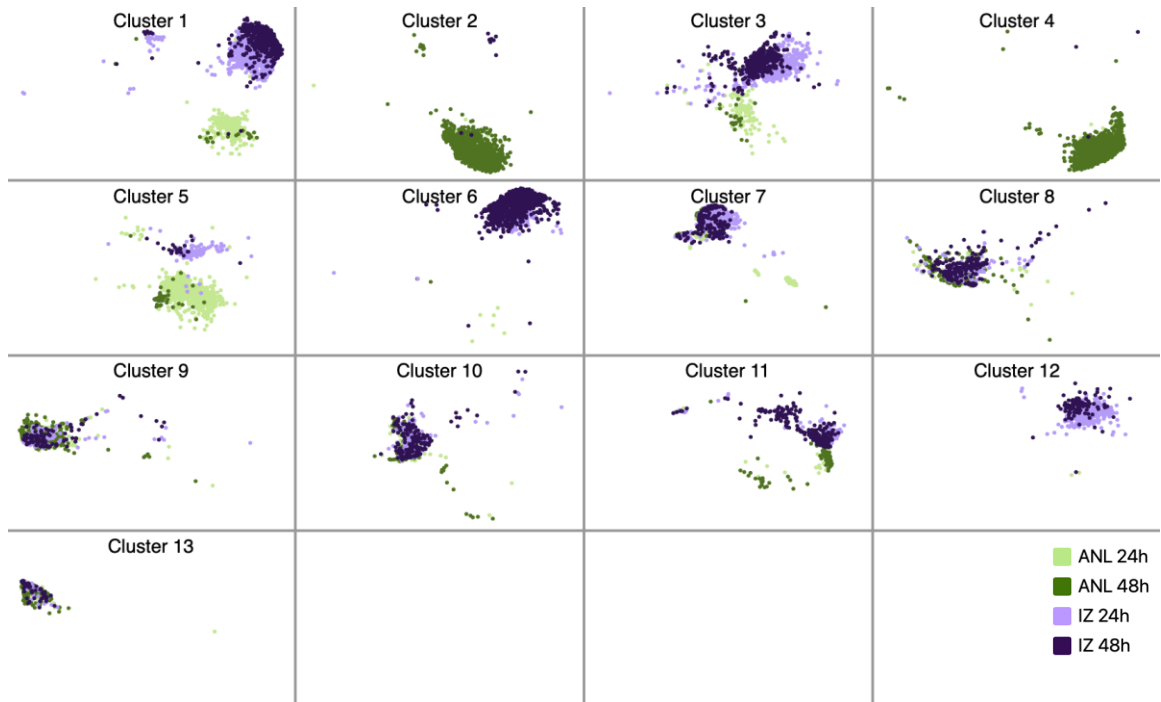


Figure 4.6. Non-linear dimensionality reduction analysis (uniform manifold approximation and projection; UMAP) conducted on the aggregate of all four samples. IZ = interzone cell cultures; ANL = anlagen cell cultures; a) Colored by cluster; b) Colored by sample; c) Separated view by sample, colored by cluster; d) Separated view by cluster, colored by sample

### Common gene expression patterns among the interzone and anagen samples

Among the DEGs from the cell clusters, in which all four samples overlapped, a gene ontology (GO) enrichment analysis was conducted on only upregulated DEGs in Cluster 9 and 13 (Table 4.10). The ten most overrepresented biological processes of Cluster 9 and 13 are listed in Table 4.12 and Table 4.13, respectively. In both clusters, cell migration was one of the most significant biological processes. The upregulated DEG profile in Cluster 9 showed positive regulation of chondrocyte differentiation as an overrepresented biological process, and the upregulated DEG profile in Cluster 13 showed positive regulation of epithelial to mesenchymal transition, by which mesenchymal progenitor cells for skeletal elements arise within the presumptive sites of limbs, as an overrepresented biological process.

In these three cell clusters—Cluster 8, 9, and 13, the relative expression levels of chondrocyte biomarkers were lower compared to the other clusters, including the cell type-specific subpopulations (Figure 4.7.a). Simultaneously, the expression of either interzone markers (Figure 4.7.b) or hypertrophic markers (Figure 4.7.c) was also relatively lower in these cell subpopulations.

Table 4.12. Overrepresented biological processes from upregulated differentially expressed gene profile in Cluster 9 of the aggregate of all four samples

GO <sup>a</sup> ID	Term	Genes	P-Value <sup>b</sup>
GO:0016477	Cell migration	ARC, PLCG1, PTK7, CDC42BPA, TNK2, LAMC1, EPHB3, NFATC2, MMP14, USP24	1.55E-05
GO:0034446	Substrate adhesion-dependent cell spreading	MICALL2, LAMC1, LAMB1, EPHB3, MERTK, FN1	8.64E-05
GO:0051726	Regulation of cell cycle	CCNE2, TARDBP, JUND, TSC2, PUM1, GADD45B	0.002
GO:0030217	T cell differentiation	EGR1, DLL4, SOX4, RUNX2	0.002
GO:0035987	Endodermal cell differentiation	COL6A1, LAMB1, MMP14, FN1	0.004
GO:0043517	Positive regulation of DNA damage response, signal transduction by p53 class mediator	SPRED2, ANKRD1, ATR	0.004
GO:0008380	RNA splicing	SRSF2, PRPF4B, TARDBP, PTBP1, MBNL1	0.009
GO:0032332	Positive regulation of chondrocyte differentiation	SOX5, LOXL2, RUNX2	0.009
GO:0031175	Neuron projection development	MICALL2, APP, BTG2, CAMSAP2, LAMB1	0.012
GO:0045600	Positive regulation of fat cell differentiation	ZFP36, SH3PXD2B, CEBPB, ID2	0.012

<sup>a</sup>GO, gene ontology

<sup>b</sup>Modified Fisher Exact P-Value was used to evaluate enrichment of genes in GO annotation terms, and all entities listed in the table have P-value<0.05.



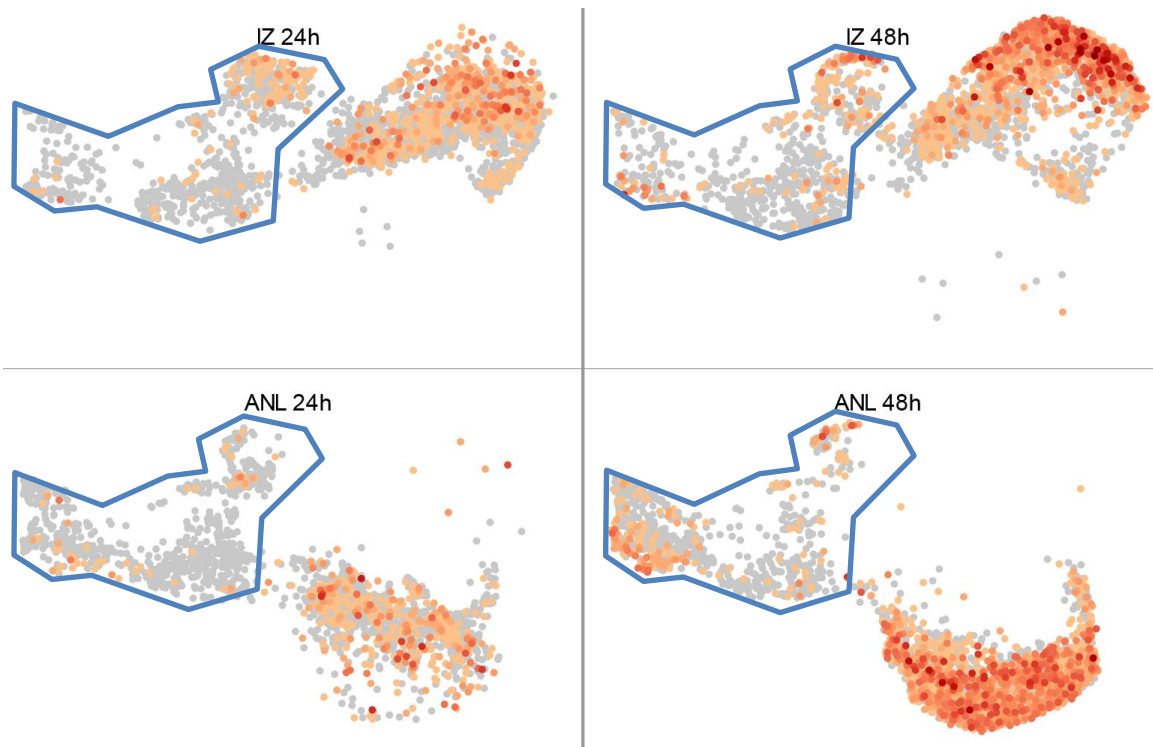
Table 4.13. Overrepresented biological processes from upregulated differentially expressed gene profile in Cluster 13 of the aggregate of all four samples

GO <sup>a</sup> ID	Term	Genes	P-Value <sup>b</sup>
GO:0098532	Histone H3-K27 trimethylation	HIST1H1E, HIST1H1D, HIST1H1C, EPHB3	4.90E-05
GO:0006334	Nucleosome assembly	TSPYL1, H1FO, HIST1H2BB, HIST1H1E, HIST1H1D, HIST1H1C, HIST1H1B, HIST1H1A	1.47E-04
GO:0016477	Cell migration	ARC, PLCG1, ARF4, PTK7, CSPG4, MMP14, NFATC2, SNAI1, USP24	4.52E-04
GO:0048538	Thymus development	MAPK1, MAFB, BCL2L11, SLC46A2, CTNNB1, CITED2	5.16E-04
GO:0000122	Negative regulation of transcription from RNA polymerase II promoter	ZFP36, ERF, FZD8, HIST1H1E, HIST1H1D, HIST1H1C, HIST1H1B, SMAD3, CTNNB1, OSR1, HEXIM1, ID1, SQSTM1, DLL4, ID4, RBM15, VLDLR	0.001
GO:0043066	Negative regulation of apoptotic process	ATF5, CDKN1B, PLK2, NUA2, OSR1, ID1, ARF4, PIM3, MYC, SLC40A1, CITED2, ANGPTL4	0.001
GO:0034333	Adherens junction assembly	ZNF703, CTNNB1, VCL	0.006
GO:0016584	Nucleosome positioning	HIST1H1E, HIST1H1D, HIST1H1C	0.008
GO:0033077	T cell differentiation in thymus	ZFP36L2, FZD8, MAFB, CTNNB1	0.008
GO:0010718	Positive regulation of epithelial to mesenchymal transition	ZNF703, SMAD3, EPHB3, SNAI1	0.012

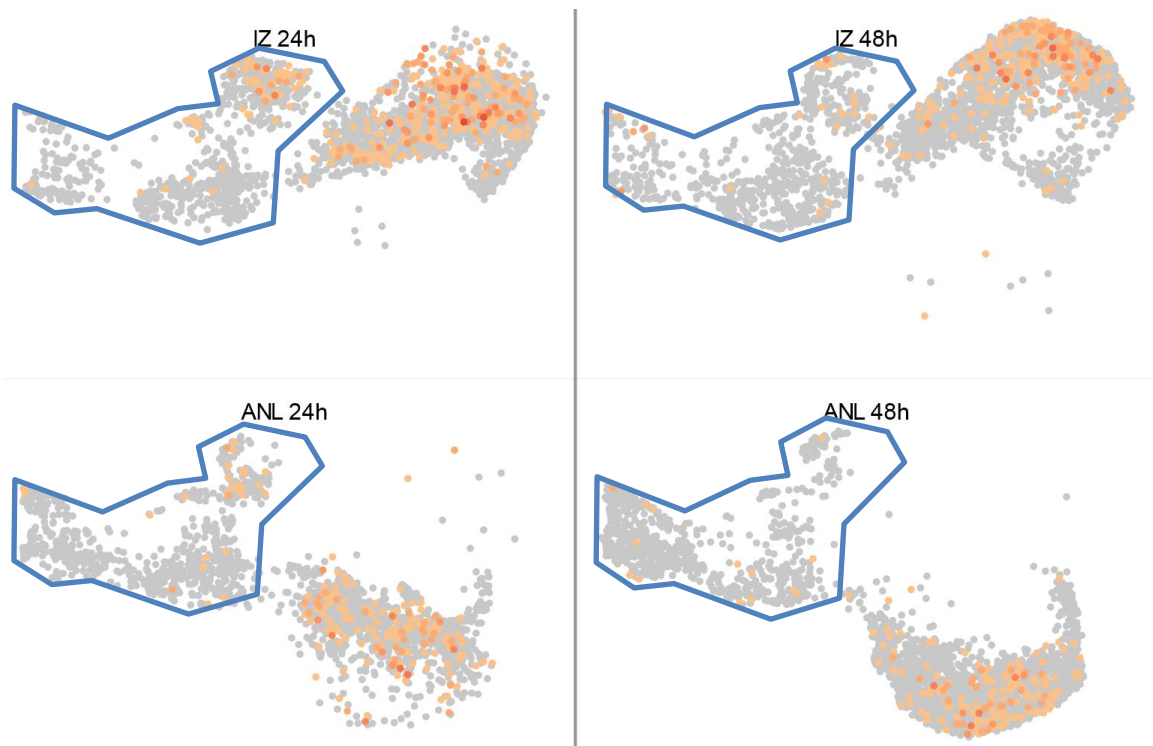
<sup>a</sup>GO, gene ontology

<sup>b</sup>Modified Fisher Exact P-Value was used to evaluate enrichment of genes in GO annotation terms, and all entities listed in the table have P-value<0.05.

4.7.a) Chondrocyte biomarker gene expression in the aggregate of samples



4.7.b) Interzone biomarker gene expression in the aggregate of samples



4.7.c) Chondrocyte hypertrophy biomarker gene expression in the aggregate of samples

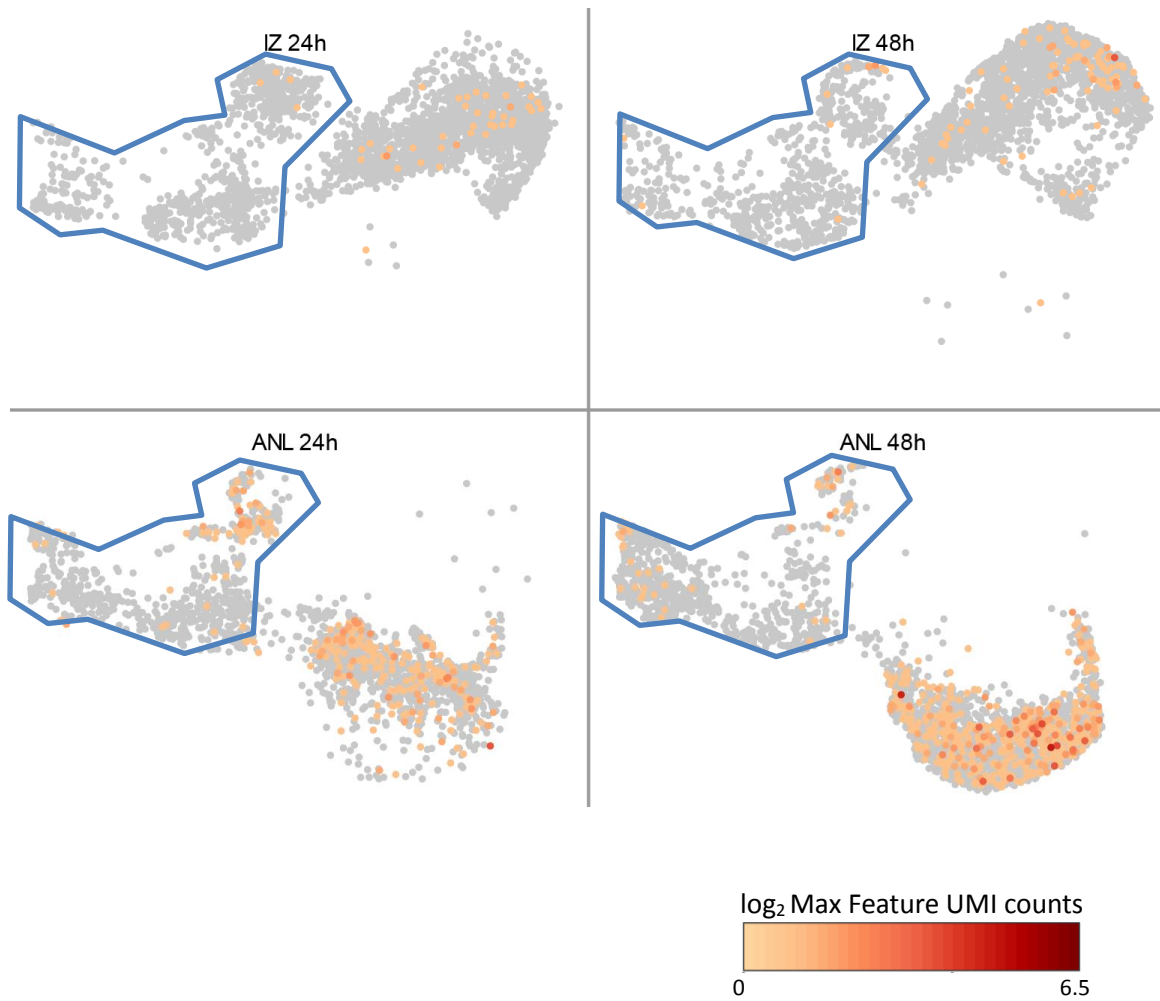


Figure 4.7. Biomarker gene expression in the aggregate of samples, plotted by uniform manifold approximation and projection (UMAP) analysis, with a separated view by sample. The areas demarcated by the blue line indicate Cluster 8, 9, and 13, in which all four samples overlapped. IZ = interzone cell cultures; ANL = anlagen cell cultures; a) Chondrocyte biomarker genes (COL2A1, COMP, and ACAN); b) Interzone biomarker genes (GDF5, WNT9A, and ENPP2); c) Chondrocyte hypertrophy biomarker genes (MMP13 and DLX5)

### Characteristics of interzone cells

No cluster had a clear distinction between interzone cell cultures at 24h and 48h (Figure 4.6.d). Cluster 5 and 12 had relatively greater interzone cells at 24h compared to 48h, and Cluster 6, 10, and 11 had relatively greater interzone cells at 48h compared to 24h. Among these clusters, only Cluster 6 and 12 dominantly consisted of interzone cells with a minimal number of analgen cells. Interzone cells at 48h were dominant in Cluster 6, and relatively similar proportions of interzone cells at both time points were observed in Cluster 12.

Gene ontology enrichment analysis was conducted with the DEGs from Cluster 6 and 12 to evaluate interzone cell-specific characteristics. Out of the ten most significant biological processes represented by the downregulated DEG profile in Cluster 6 (Table 4.14), eight GO terms are related to mitosis: 1) cell division, 2) mitotic nuclear division, 3) mitotic chromosome condensation, 4) mitotic cytokinesis, 5) microtubule-based movement, 6) mitotic sister chromatid segregation, 7) mitotic spindle organization, and 8) mitotic metaphase plate congression. Cluster 6 had only 15 upregulated DEGs, and Cluster 12 had one upregulated DEG and three downregulated DEGs; thus, no overrepresented biological processes were identified due to the small size of the DEG lists.

The majority of interzone biomarker (GDF5, WNT9A, and ENPP2) expressing cells were located within the cell type-specific areas in both interzone and analgen cell cultures at both time points (Figure 4.7.b).

Table 4.14. Top ten overrepresented biological processes from downregulated differentially expressed gene profile in Cluster 6 of the aggregate of all four samples

GO <sup>a</sup> ID	Term	Genes	P-Value <sup>b</sup>
GO:0051301	Cell division	SPC24, CCNB1, SPC25, CCNB2, TPX2, CKS2, UBE2C, ASPM, REEP4, CDCA3	1.76E-08
GO:0007067	Mitotic nuclear division	SPC24, SPC25, CCNB2, PLK1, NUF2, CENPW, ASPM, REEP4	3.71E-07
GO:0007076	Mitotic chromosome condensation	NCAPG, NUSAP1, NCAPD3, SMC4	1.75E-05
GO:0000281	Mitotic cytokinesis	CKAP2, PLK1, NUSAP1, STMN1	1.55E-04
GO:0042026	Protein refolding	HSPA6, HSPA1A, HSPA8	0.001
GO:0007018	Microtubule-based movement	KIF11, KIF18A, KIF20B, CENPE	0.002
GO:0000070	Mitotic sister chromatid segregation	CDCA8, PLK1, KIF18A	0.003
GO:0007052	Mitotic spindle organization	CCNB1, SPC25, STMN1	0.005
GO:0007080	Mitotic metaphase plate congression	CCNB1, CDCA8, KIF18A	0.008
GO:0034976	Response to endoplasmic reticulum stress	HYOU1, PDIA6, CXCL8	0.018

<sup>a</sup>GO, gene ontology

<sup>b</sup>Modified Fisher Exact P-Value was used to evaluate enrichment of genes in GO annotation terms, and all entities listed in the table have P-value<0.05.

### Characteristics of analgen cells

Cluster 2 and 4 were mostly composed of analgen cells at 48h, with these clusters containing much fewer analgen cells at 24h or interzone cells at either time point (Figure 4.6.d). On the other hand, analgen cells at 24h were dominant in Cluster 5 although several interzone cells were also observed in this cluster.

Gene enrichment analyses were conducted with DEGs from Cluster 2 and 4. Seven and one significantly overrepresented biological processes were identified from the upregulated DEG profiles of Cluster 2 and 4, respectively (Table 4.15 and 4.16). The ten most downregulated biological processes in Cluster 2 and 4 are also shown in the tables. In both cell clusters, regulatory processes involved in mitotic events were significantly downregulated: upregulation of positive regulation of cell cycle arrest (Cluster 2) and downregulation of mitotic sister chromatid segregation, microtubule-based movement, microtubule bundle formation, mitotic cytokinesis (Cluster 2), mitotic nuclear division (Cluster 4), cell division, and mitotic chromosome condensation (Cluster 2 and 4).

Similar to the interzone biomarker gene expression (Figure 4.7.b), most of hypertrophic biomarkers (COL10A1, MMP13, and DLX5) were observed in the cell type-specific areas in both types of cell cultures at both time points (Figure 4.7.c).

Table 4.15. Top ten overrepresented biological processes from differentially expressed gene profile in Cluster 2 of the aggregate of all four samples

GO <sup>a</sup> ID	Term	Genes	P-Value <sup>b</sup>
(Upregulated <sup>c</sup> )			
GO:0060428	Lung epithelium development	HMGA2, ERFF1	0.009
GO:0043405	Regulation of MAP kinase activity	TRIB3, TRIB1	0.011
GO:0045892	Negative regulation of transcription, DNA-templated	ATF5, TRIB3, HMGA2, TGFB1	0.012
GO:2000679	Positive regulation of transcription regulatory region DNA binding	HMGA2, TGFB1	0.022
GO:0008016	Regulation of heart contraction	S100A1	0.026
GO:0001837	Epithelial to mesenchymal transition	HMGA2, TGFB1	0.035
GO:0071158	Positive regulation of cell cycle arrest	HMGA2, TGFB1	0.041
(Downregulated)			
GO:0007076	Mitotic chromosome condensation	NCAPG, NUSAP1, NCAPD3, SMC4	3.83E-05
GO:0085020	Protein K6-linked ubiquitination	UBE2S, UBE2T, BARD1	7.13E-04
GO:1904668	Positive regulation of ubiquitin protein ligase activity	PLK1, CDC20, UBE2S	9.93E-04
GO:0031145	Anaphase-promoting complex-dependent catabolic process	CDC20, UBE2C, UBE2S	0.001
GO:0000070	Mitotic sister chromatid segregation	CDCA8, PLK1, KIF18A	0.005
GO:0007018	Microtubule-based movement	KIF11, KIF18A, KIF20B, CENPE	0.006
GO:0001578	Microtubule bundle formation	PRC1, PLK1, MAP1B	0.008
GO:0051301	Cell division	CDCA8, CDC20, ASPM, CDCA3	0.011
GO:0000281	Mitotic cytokinesis	CKAP2, PLK1, NUSAP1	0.012
GO:0035519	Protein K29-linked ubiquitination	UBE2S, UBE2T	0.014

<sup>a</sup>GO, gene ontology

<sup>b</sup>Modified Fisher Exact P-Value was used to evaluate enrichment of genes in GO annotation terms, and all entities listed in the table have P-value<0.05.

<sup>c</sup>Upregulated differentially expressed genes resulted in only seven significant biological processes.

Table 4.16. Top ten overrepresented biological processes from differentially expressed gene profile in Cluster 4 of the aggregate of all four samples

GO <sup>a</sup> ID	Term	Genes	P-Value <sup>b</sup>
(Upregulated <sup>c</sup> )			
GO:0019221	Cytokine-mediated signaling pathway	CCL2, PODNL1, IFNAR1	0.009
(Downregulated)			
GO:0051301	Cell division	SPC24, CCNB1, SPC25, CCNB2, GNAI3, TPX2, AURKA, UBE2C, ASPM, CDCA3	4.34E-06
GO:2000378	Negative regulation of reactive oxygen species metabolic process	HK2, BNIP3, PINK1, VDAC1	2.37E-04
GO:0007067	Mitotic nuclear division	SPC24, SPC25, CCNB2, PLK1, NUF2, GEM, ASPM	2.40E-04
GO:0001666	Response to hypoxia	CAV1, PLOD1, CRYAB, EGLN3, ALKBH5, DDIT4	3.49E-04
GO:0045766	Positive regulation of angiogenesis	GATA2, ADM, SFRP2, SERPINE1, RHOB, HSPB1	7.05E-04
GO:0030199	Collagen fibril organization	SFRP2, P4HA1, LOX, SERPINH1	0.002
GO:0010629	Negative regulation of gene expression	HMGB2, STC2, CRYAB, PINK1, UPK3B	0.003
GO:0045944	Positive regulation of transcription from RNA polymerase II promoter	PID1, CKAP2, NAMPT, HMGB2, HES1, HDAC5, CDH13, SFRP2, IRF7, SIX1, PSIP1, KDM3A, TOP2A	0.003
GO:0050829	Defense response to Gram-negative bacterium	HMGB2, ADM, SERPINE1, DEFB1	0.004
GO:0007076	Mitotic chromosome condensation	NCAPG, NUSAP1, SMC4	0.004

<sup>a</sup>GO, gene ontology

<sup>b</sup>Modified Fisher Exact P-Value was used to evaluate enrichment of genes in GO annotation terms, and all entities listed in the table have P-value<0.05.

<sup>c</sup>Upregulated differentially expressed genes resulted in only one significant biological process.



## Discussion

The first single cell RNA-seq study was published in 2009 (Tang et al., 2009), but this technology is still emerging and is relatively novel in the field of transcriptome research. In this pilot study, while asking scientific questions, this advanced technique was also evaluated for use in our laboratory with the experimental settings and typical specimens that the lab has been employing.

While the targeted cell number was 5,000 cells/sample, the numbers of recovered cells ranged from 2,032 – 2,845 (recovery rates of 47 – 54%). These lower recovery rates might be due to the characteristics of the sample type. The specimens studied in this experiment were chondrogenic cell pellets, which produced and accumulated ECM over time. To prepare single cell suspensions, the cell pellets had to be disaggregated by enzymatic digestion as well as filtration through a porous membrane. Even though cell pellets underwent those steps, remaining debris might affect the GEM formation.

Despite the lower cell number recovery rates, greater than 91.4% of the total reads were associated with cells in all samples. In addition, the total transcripts count/cell in all samples surpassed 200,000 reads/cell, which was well above the recommended minimum number of reads/cell (50,000 reads/cell) from the manufacturer's protocol. Thus, the sequencing generated abundant and high-quality data from the sample set (Table 4.3).

### ***Subpopulations of cells in each sample***

The data do not support the first hypothesis tested in the study that interzone cell cultures will develop relatively heterogenous cell clusters while anlagen cell cultures will result in relatively homogenous cell clusters. In terms of subpopulation homogeneity, there was no significant evidence that either cell lines at either time points had greater heterogeneity (Table 4.4). Also, chondrocyte biomarker gene expression (COL2A1, COMP, and ACAN) was observed from the majority of cell clusters in interzone and analgen cell cultures at 24h and 48h (Figure 4.3) with a few clusters showing lower levels. In agreement with findings in the previous chapters, the expression of these cartilaginous ECM genes was greater at 48h compared to 24h in both cell cultures, confirming that the cell lines were progressing towards a chondrogenic phenotype.

The interzone biomarker genes (GDF5, WNT9A, and ENPP2) selected for visualization are regulatory genes, unlike the ECM effector genes chosen for chondrocytes. Interestingly, though, the localization was generally concordant (Figures 4.3 and 4.4). This co-localization may suggest retention of interzone biomarkers in the subset of cells destined to form articular cartilage or downregulation in cell subsets moving towards non-cartilaginous joint tissues. Levels were lower in anlagen cell cultures, especially, WNT9A which was not detectable. The overall intensity of the interzone marker gene expression was not remarkably different between the time points. Since *in vivo* GDF5 and WNT9A expression drops significantly postnatally in mouse limb joints (Koyama et al., 2008), the equine primary cells used in the current study are consistent with an earlier stage of differentiation.

The chondrocyte hypertrophic biomarkers (MMP13, and DLX5) were present at higher levels in the anlagen samples (Figure 4.5), consistent with this tissue progressing to osteogenesis. Overall, however, COL10A1 and MMP13 were less informative. COL10A1 was not detectable and the MMP13 expression limited. In contrast, steady state mRNA levels of DLX5, a transcription factor gene, were relatively high in anlagen cell cultures and increased over time. These results may indicate that the anlagen cell cultures were at earlier stages of hypertrophic differentiation. Similar to the interzone biomarker gene expression, DLX5 was also expressed from the cell clusters that expressed the cartilage ECM genes. This finding is consistent with both the interzone- and hypertrophic biomarkers being closely linked to chondrogenic differentiation in the two skeletal cell lines.

#### ***Shared features and unique properties between interzone and anlagen cells***

The second hypothesis tested in the current experiments was that interzone and anlagen cells will share common chondrogenic characteristics, but also show cell type-specific properties. Supporting this hypothesis, the data clearly demonstrate both similarities and differences.

#### **Common gene expression patterns between interzone and anlagen cells**

Analyzing gene expression data from the four-sample aggregate, 13 cell clusters were identified, and all four samples overlapped in Cluster 8, 9, and 13 (Figure 4.6.c and d). Interestingly, the only common overrepresented GO term in these cell subpopulations was “cell migration” (Table 4.12 and 4.13). This result is consistent with data reported in Chapter 2 and may reflect the establishment of cell pellets. Cells were lifted from monolayers and

lightly centrifuged to aid in pellet formation, which might increase the expression of genes that regulate cell migration, regardless of cell types.

Cluster 9 is interesting because it shows an overrepresentation of positive chondrocyte differentiation regulators as a significant biological process, while expression of chondrocyte ECM biomarkers was relatively low. Perhaps the cells in Cluster 9 were regulating chondrogenesis in other cells, or alternatively were just earlier in the differentiation process. Among the significant biological processes analyzed from the upregulated DEG profile in Cluster 13, positive regulation of epithelial to mesenchymal transition was observed and is consistent with cellular processes undoubtedly also occurring during limb bud formation (Gros and Tabin, 2014).

#### Differences between interzone and anlagen cells

When plotting the gene expression in the aggregate of all samples, the chondrocyte biomarker gene expression (COL2A1, COMP, and ACAN) was higher in cell type-specific areas: Cluster 2 and 4 for anlagen samples, and Cluster 6 and 12 for interzone samples (Figure 4.7.a). Also, expression of interzone (GDF5, WNT9A, and ENPP2) and hypertrophic (MMP13 and DLX5) biomarkers had the same pattern especially at the 48h time point. When comparing the profiles of the most overrepresented biological processes from the interzone cell-specific cluster (Cluster 6) and the anlagen cell-specific clusters (Cluster 2 and 4), mitosis-related biological events were consistently downregulated. In addition, positive regulation of cell cycle arrest was upregulated in Cluster 2. More samples will be needed to investigate these relationships further.

There are several examples of genes that may diverge chondrogenic pathways between interzone and anlagen cell cultures. Among the DEGs in the anlagen cell-specific clusters within the aggregate of all samples (Cluster 2 and 4, which were dominated by anlagen cells at 48h), DEFB1 was the most downregulated gene ( $\log_2FC=-5.62$  in Cluster 2;  $\log_2FC=-5.99$  in Cluster 4;  $P\text{-value}<0.0001$ ; Table 4.11). At the same time, DEFB1 was the only upregulated DEG ( $\log_2FC=1.74$ ;  $P\text{-value}=0.019$ ; Table 4.11) in Cluster 12, in which interzone cells at both time points were dominant. This gene encodes a TLR ligand and activates downstream by interacting with a receptor, TLR4. Its signal activates RAC1 in MAPK pathways and HSP27 in that order (Melas et al., 2014). HSP27, a chaperone involved in proper protein folding, was significantly expressed in normal articular cartilage when compared to osteoarthritic cartilage (Lambrecht et al., 2010). Along with DEFB1, HSP27 was differentially downregulated in Cluster 4 ( $\log_2FC=-1.22$ ;  $P\text{-value}=0.006$ ). Altogether, these results may suggest that DEFB1 might participate in interzone-specific signaling pathways leading to articular cartilage formation, and these pathways might be significantly downregulated in anlagen cells.

On the other hand, CXCL8 was a key upregulated gene ( $\log_2FC=1.98$ ;  $P\text{-value}=6.96E-06$ ) in Cluster 4 within the aggregate of all samples, but this gene was differentially downregulated ( $\log_2FC=-3.55$ ;  $P\text{-value}=<0.001$ ) in Cluster 6 (Table 4.11). CXCL8 was also an upregulated DEG in anlagen cell cultures compared to interzone cell cultures before and after inducing *in vitro* chondrogenesis from the study reported in Chapter 3 (Table 3.23). In a previous study, although the expression of CXCL8 was detected in normal articular chondrocyte cultures, when its protein was excessively added (10 ng/ml) to the cultures, it increased MMP13 expression and calcification (Merz et al., 2003). In agreement with the literature as well as the

result from the previous chapter, the data may indicate that anlagen cell cultures showed greater levels of hypertrophic potential compared to interzone cell cultures.

### ***Future direction***

By profiling transcriptomic signatures at the single cell level, the present study showed that both interzone and anlagen cell cultures at 24h and 48h consisted of several cell subpopulations with different gene expression patterns. Although this study characterized cell clusters within a sample and within the aggregate of the sample set, spatial arrangement of these cell subpopulations and potential paracrine interactions between them are yet to be defined. These undefined properties could be further described when a spatial single cell RNA-seq is conducted on cell pellet sections. As shown in the histology images generated from the preliminary staining experiment (Figure 4.1), different ECM compositions and cellular arrangement were observed in a zonal structure on the day-21 interzone cell pellet section. In addition, the current dataset revealed that several genes were differentially upregulated in some cell clusters while differentially downregulated in other cell clusters within the same sample. Thus, the evaluation of spatial gene expression would provide a better understanding of the spatial arrangement of cell subpopulations and the molecular interaction between them in interzone and anlagen cell pellet cultures during chondrogenic differentiation.

Obviously, the current pilot experiments would greatly benefit from additional biological replicates. The observations are interesting, but clearly require more primary data to elucidate fully.

## Conclusion

This pilot transcriptomic study conducted at the single cell level proposes the single cell RNA-seq technology as a highly informative new approach for studying gene expression in chondrogenic cell pellet cultures. With the abundant and quality sequencing data generated, this study demonstrated subpopulations of cells in both interzone and anagen cell cultures grown in a chondrogenic induction medium containing TGF- $\beta$ 1 for 24 and 48 hours. Although interzone and analgen cell cultures at both time points did not display different degrees of heterogeneity in cell subpopulations, the sample set showed common traits as well as cell type-specific differences. It is interesting to note that the cell type-specific clusters exhibited greater cartilaginous ECM marker expression as well as either interzone-like characteristics or hypertrophic potentials compared to the clusters that were overlapped by all four samples. On the other hand, the gene expression pattern in the overlapping clusters presented positive regulation of chondrocyte differentiation as its significant biological process. Therefore, the data may suggest that upstream signal transduction and cartilaginous ECM production might be organized by different cell subpopulations during chondrogenesis. In conclusion, the present study added further evidence that interzone and analgen cell cultures progress down different chondrogenic pathways.

## Chapter 5. Overall summary and future directions

In this dissertation project, two studies and a pilot experiment were conducted to investigate chondrogenic divergence of equine fetal interzone and anlagen cells in tissue culture using various high throughput gene expression analytic technologies. Tissue culture protocols were held constant in all three studies. Interzone and anlagen cells collected from gestational day 45 equine fetuses were grown in 3-dimensional cell pellet cultures, and the *in vitro* chondrogenesis was induced by TGF- $\beta$ 1, which is a well-established chondrogenic factor. Kinetics of gene expression changes were evaluated in each chapter over different experimental periods using different methods.

In the first study, 93 targeted genes were selected for characterizing expression kinetics over 336 hours in Chapter 2. These genes are either involved in biological processes related to fetal skeletal development or differentially expressed between equine interzone and anlagen tissue lineages. By profiling their expression under the chondrogenic at ten different time points (0, 1.5, 3, 6, 12, 24, 48, 96, 168, and 336h), this study yielded important new information on changes in steady state mRNA levels between the cell lines, and importantly also the timing of those in changes in response to the *in vitro* chondrogenic induction.

Both shared and cell type-specific differences before and after inducing chondrogenesis were observed. Before inducing chondrogenesis, 73 out of 87 targeted genes were not differentially expressed between interzone and anlagen cells, while 14 genes already had different gene expression levels at the baseline. Among 73 genes that were not differentially expressed at 0h, 47 genes responded differently to the TGF- $\beta$ 1 treatment over time.



The first responses to the chondrogenic induction stimulus occurred within the first 1.5 hours. Interestingly, these initial gene loci all had functional annotation related to either transcription regulation or signal transduction. In fact, all transcription regulatory genes that were selected for the panel changed their expression patterns within the first 24 hours in the two chondrogenic cell lines, while the alterations in signaling events were relatively evenly distributed across all time points during the 336-hour experimental period. On the other hand, effector genes maintaining ECM composition had more delayed responses. No ECM effector genes changed their basal mRNA level at the first collection point (1.5h), with the earliest effects in this functional annotation group observed at 3h. Taken together, the results accepted the hypothesis that divergent chondrogenic pathways between interzone and anlagen cell cultures will become evident within the first 24 hours after inducing *in vitro* chondrogenesis.

From the findings in Chapter 2, five time points, including the baseline, were selected (0, 1.5, 3, 12, and 96h) for full mRNA transcriptome assessment based on the following reasons. By 1.5h, regulatory responses to the chondrogenic stimulation were initiated. At 3h, a peak of initial gene expression responses was recorded. A second major peak of first responses in transcription regulatory genes occurred at 12h. Finally, most of effector genes regulating ECM profiles (85.4%) responded by 96h. The method used to evaluate the transcriptome was RNA-sequencing.

Data from the Chapter 3 support that chondrogenic differentiation in interzone cell cultures are in the direction of an articular chondrocyte phenotype, while analgen cell cultures

are directed more towards a cartilage that will progress through hypertrophic differentiation. At the same time, a number of the gene expression changes in response to induced chondrogenesis are shared between these two cell lines. This study identified interesting DEGs and candidate molecular regulators that may diverge chondrogenic pathways between the two skeletal cell lines. In agreement with the Chapter 2 data, the Chapter 3 results support accepting the hypothesis that differential responses of regulatory genes start changing expression patterns within the first 1.5 hours, and more distinctive transcriptomic signatures between interzone and anlagen cells will accumulate as chondrogenesis progresses.

In Chapter 4, a pilot single cell RNA-seq experiment was conducted to characterize cell subpopulations in interzone and anlagen cell cultures at 24h and 48h by testing two hypotheses. The first hypothesis was that interzone cell cultures will develop relatively heterogeneous cell subpopulations while anlagen cell cultures will retain more cellular homogeneity. However, the data indicated that interzone and anlagen cell cultures at both time points resulted in 7 – 8 cell subset clusters, and the distribution of cell numbers in the clusters was not notably different among the sample set, rejecting this hypothesis.

The second hypothesis tested in this pilot experiment was that interzone and anlagen cell cultures share common chondrogenic characteristics and also show unique traits within a cell type. By plotting the expression of well-established cartilage biomarkers, the chondrogenic potential of interzone and anlagen cells was confirmed. Visualizing the interzone biomarker expression and the hypertrophy biomarker expression demonstrated that interzone cells retained interzone-like characteristics while anlagen cells retained hypertrophic potential in

culture. The expression of genes representing common chondrogenic potential and cell type-specific traits became greater at 48h compared to 24h. It is interesting to note that groups of cells that showed either interzone-like or hypertrophic characteristics had greater expression of cartilaginous ECM biomarkers compared to cell subsets that were overlapped, suggesting the same cell subsets that achieve unique features also accumulate functional outcomes of chondrogenic differentiation generally. In sum, the second hypothesis was accepted.

Overall, data in this dissertation research were generated with a balanced sample set composed of paired cell types (interzone cells, anlagen cells, and fibroblasts) between the biological replicates. The studies conducted provide new information on divergent chondrogenic pathways between fetal interzone and anlagen cells at the molecular level from the perspective of time course kinetics and have derived a molecular regulatory model (Figure 5.1). Early regulation of the different chondrogenic pathways may involve activation of NUPR1 in interzone cells and RUNX2 in anlagen cells within the first 3 hours after inducing chondrogenesis (Figure 3.16). In interzone cell cultures, chondrogenic differentiation might temporarily pause immediately after treating the chondrogenic inductive medium containing TGF $\beta$ -1, represented by activated NOTCH1 and GATA1 signaling (between 0h and 1.5h; Figure 3.5). Then, interzone cells may resume chondrogenesis with the activated MEIS2 regulation between 3h and 12h. At later stages, the expression of COL9A1, which composes the ECM of hyaline cartilage, may indicate that interzone cells are in the chondrogenic pathways forming articular cartilage. On the other hand, anlagen cells appear to be directed to hypertrophic differentiation pathways at earlier stages (between 1.5h and 12h) with activated SOX10, DLX5,

and DLX6 signals, and subsequently, the activation of SP7 may induce ECM mineralization at later stage (Figure 3.9), consistent with the biology of cartilaginous anlagen of limbs.

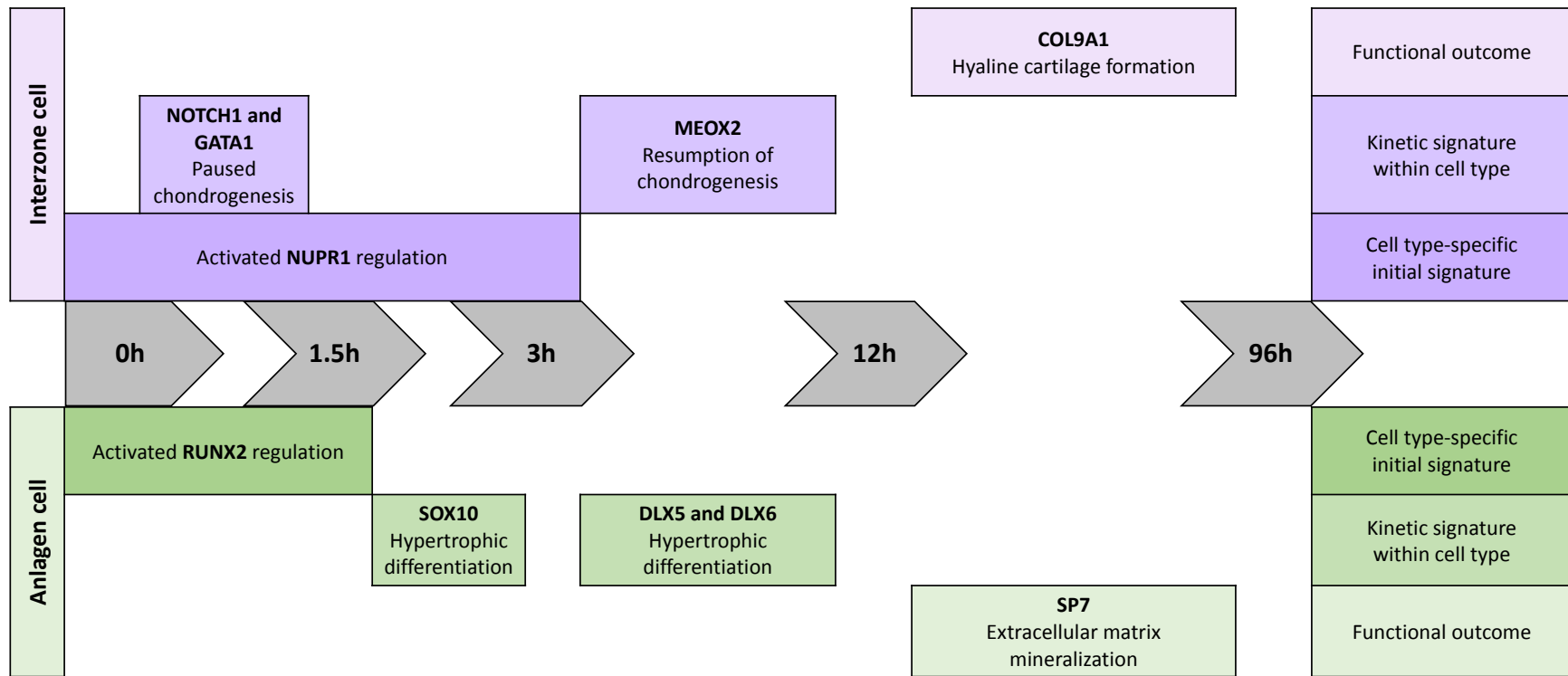


Figure 5.1. Molecular regulatory model in interzone and anlagen cell cultures during *in vitro* chondrogenesis. Within the first 3 hours, there is evidence for the activation of cell type-specific regulators that initiate divergence of the two different chondrogenic pathways (Figure 3.16). The interzone cells in culture appear to pause chondrogenesis immediately after treatment with the chondrogenic stimuli, but then resume at 3h (Figure 3.5). Anlagen cells, in contrast, progress towards a hypertrophic differentiation pathway (Figure 3.9). The distinguishing extracellular matrix profiles between interzone and anlagen cells became apparent at the later time points.

New knowledge obtained from the present research lays a foundation for mechanistic experiments and perhaps future translational clinical studies. The mechanistic experiments will be required to validate the model of molecular regulation (Figure 5.1) and elucidate the roles of candidate regulator genes that may dictate the cellular fates of these two fetal skeletal cell lines; a loss-of-function or gain-of-function study would add value to the present findings.

Concurrent to the proposed model, genes loci that are of higher biological interest were identified: DEGs, including candidate molecular switches between the chondrogenic cell lines, hub genes correlated with the time course, predicted upstream regulators, etc. Mechanistic studies would elaborate their biological relevance to the divergent chondrogenic pathways of interzone and anlagen cells. For example, some genes that were analyzed as potential molecular switches between the two cell lines were novel genes without known functional annotations at the current moment. Hypotheses to study functional roles in these differential chondrogenic pathways could be tested by conducting loss-of-function or gain-of-function experiments. Also, molecular interactions or hierarchy in signaling cascades of these interesting genes could be evaluated, adding more useful information. Although weighted gene co-expression network analysis (WGCNA) demonstrated interactions between some key hub genes in co-expression networks, the function of other hub genes have not well characterized in scientific literature and therefore could not be fully assessed in co-expression networks despite their greater relationship with other hub genes. In the same way, computational predictions of upstream regulators based on the differential expression data require mechanistic validation experiments to understand the relevant biology further.

In the present studies, a majority of primary data centered on measurements of steady state mRNA levels. Individual transcripts, however, still need to be translated, transported, and in some cases modified to be functional in biological processes. Although the expression of RNA and protein is correlated with each other via the central dogma, relationships between mRNA, protein, and biological activity are not always concordant (Liu et al., 2018). Steady state levels of mRNA are a function of both de novo synthesis and transcript degradation, and translation requires additional time which can add another source for discrepancy between the mRNA and protein levels. Not only do transcription and translation occur in different compartments in cells, but intron excision prior to translation also takes at least 5 – 10 minutes (Singh and Padgett, 2009) and can be regulated. Furthermore, proteins can exist in various forms with activation status regulated through post-translational modifications. Therefore, future cross-validation of the current RNA-based observations with protein expression and biological function would be valuable in efforts to more fully elucidate the biological relationships.

In addition, spatial expression evaluation at the transcript and/or protein levels would enrich the current understanding. A new method generating substantial interest in many biomedical areas is spatial single cell RNA-seq analyses to elucidate cell subpopulation arrangements *in situ*. In the current application, this would demonstrate morphological relationships of cell subgroups (clusters) within the interzone and anlagen cell pellets, with the clear potential to tie this back the developmental tissues themselves in fetal limbs.

In conclusion, this dissertation research confirmed and elaborated differential signatures of fetal interzone and anlagen cells. By suggesting important candidate molecular regulators that may direct these two skeletal cell lines towards articular cartilage development or hypertrophic chondrogenesis, the present studies propose future research directions, with hopes to improve the clinical approaches for supporting articular cartilage regeneration.



## References

- ADAM, E. N., JANES, J., LOWNEY, R., LAMBERT, J., THAMPI, P., STROMBERG, A. & MACLEOD, J. N. 2019. Chondrogenic differentiation potential of adult and fetal equine cell types. *Veterinary Surgery*, 48, 375-387. doi:10.1111/vsu.13183
- AKIYAMA, H., CHABOISSIER, M.-C., MARTIN, J. F., SCHEDL, A. & DE CROMBRUGGHE, B. 2002. The transcription factor Sox9 has essential roles in successive steps of the chondrocyte differentiation pathway and is required for expression of Sox5 and Sox6. *Genes & development*, 16, 2813-2828. doi:10.1101/gad.1017802
- ALFORD, J. W. & COLE, B. J. 2005. Cartilage restoration, part 1: basic science, historical perspective, patient evaluation, and treatment options. *The American Journal of Sports Medicine*, 33, 295-306. doi:10.1177/0363546504273510
- AOMATSU, K., ARAO, T., SUGIOKA, K., MATSUMOTO, K., TAMURA, D., KUDO, K., KANEDA, H., TANAKA, K., FUJITA, Y., SHIMOMURA, Y. & NISHIO, K. 2011. TGF- $\beta$  induces sustained upregulation of SNAI1 and SNAI2 through Smad and Non-Smad pathways in a human corneal epithelial cell line. *Investigative Ophthalmology & Visual Science*, 52, 2437-2443. doi:10.1167/iovs.10-5635
- BERIS, A. E., LYKISSAS, M. G., PAPAGEORGIOU, C. D. & GEORGOULIS, A. D. 2005. Advances in articular cartilage repair. *Injury*, 36, S14-S23. doi:10.1016/j.injury.2005.10.007
- BI, W., DENG, J. M., ZHANG, Z., BEHRINGER, R. R. & DE CROMBRUGGHE, B. 1999. Sox9 is required for cartilage formation. *Nature genetics*, 22, 85-89. doi:10.1038/8792
- BLIGHE, K., RANA, S. & LEWIS, M. 2019. EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling. R package version 1.2.0. doi: 10.18129/B9.bioc.EnhancedVolcano
- BLUM, M., GAUNT, S. J., CHO, K. W. Y., STEINBEISSER, H., BLUMBERG, B., BITTNER, D. & DE ROBERTIS, E. M. 1992. Gastrulation in the mouse: The role of the homeobox gene gooseoid. *Cell*, 69, 1097-1106. doi:10.1016/0092-8674(92)90632-M
- BOND, S. R., LAU, A., PENUELA, S., SAMPAIO, A. V., UNDERHILL, T. M., LAIRD, D. W. & NAUS, C. C. 2011. Pannexin 3 is a novel target for Runx2, expressed by osteoblasts and mature growth plate chondrocytes. *Journal of Bone and Mineral Research*, 26, 2911-2922. doi: 10.1002/jbmr.509
- CALDWELL, K. L. & WANG, J. 2015. Cell-based articular cartilage repair: the link between development and regeneration. *Osteoarthritis and Cartilage*, 23, 351-362. doi:10.1016/j.joca.2014.11.004
- CAMERON, D. A. 1963. The fine structure of bone and calcified cartilage. A critical review of the contribution of electron microscopy to the understanding of osteogenesis. *Clinical orthopaedics and related research*, 26, 199-228.
- CHEN, C. G., THUILLIER, D., CHIN, E. N. & ALLISTON, T. 2012. Chondrocyte-intrinsic Smad3 represses Runx2-inducible matrix metalloproteinase 13 expression to maintain articular cartilage and prevent osteoarthritis. *Arthritis & Rheumatism*, 64, 3278-3289. doi:10.1002/art.34566
- CHEN, Y. & GRIDLEY, T. 2013. Compensatory regulation of the Snai1 and Snai2 genes during chondrogenesis. *Journal of Bone and Mineral Research*, 28, 1412-1421. doi:10.1002/jbmr.1871
- CHURCH, V., YAMAGUCHI, K., TSANG, P., AKITA, K., LOGAN, C. & FRANCIS-WEST, P. 2005.

- Expression and function of Bapx1 during chick limb development. *Anatomy and embryology*, 209, 461-469. doi:10.1007/s00429-005-0464-z
- COSDEN-DECKER, R. S., BICKETT, M. M., LATTERMANN, C. & MACLEOD, J. N. 2012. Structural and functional analysis of intra-articular interzone tissue in axolotl salamanders. *Osteoarthritis and Cartilage*, 20, 1347-1356. doi:10.1016/j.joca.2012.07.002
- COSDEN, R. S., LATTERMANN, C., ROMINE, S., GAO, J., VOSS, S. R. & MACLEOD, J. N. 2011. Intrinsic repair of full-thickness articular cartilage defects in the axolotl salamander. *Osteoarthritis and Cartilage*, 19, 200-205. doi:10.1016/j.joca.2010.11.005
- DECKER, R. S., KOYAMA, E. & PACIFICI, M. 2014. Genesis and morphogenesis of limb synovial joints and articular cartilage. *Matrix Biology*, 39, 5-10. doi:10.1016/j.matbio.2014.08.006
- DOBACZEWSKI, M., CHEN, W. & FRANGOGIANNIS, N. G. 2011. Transforming growth factor (TGF)- $\beta$ 2; signaling in cardiac remodeling. *Journal of Molecular and Cellular Cardiology*, 51, 600-606. doi:10.1016/j.yjmcc.2010.10.033
- ESER, P., DEMEL, C., MAIER, K. C., SCHWALB, B., PIRKL, N., MARTIN, D. E., CRAMER, P. & TRESCH, A. 2014. Periodic mRNA synthesis and degradation co-operate during cell cycle gene expression. *Molecular Systems Biology*, 10, 717. doi:10.1002/msb.134886
- FEBRES-ALDANA, C. A. & ALEXIS, J. 2020. Normal expression of SRY-related HMG-BOX gene 10 (SOX-10) in recent and old cutaneous scars is a potential mimicker of desmoplastic malignant melanoma. *Applied Immunohistochemistry & Molecular Morphology*, 28, 197-204. doi:10.1097/PAI.0000000000000729
- FISHER, M., ACKLEY, T., RICHARD, K., OEI, B. & DEALY, C. N. 2019. Osteoarthritis at the cellular level: mechanisms, clinical perspectives, and insights from development. In: NARAYAN, R. (ed.) *Encyclopedia of Biomedical Engineering*. Oxford: Elsevier.
- FLUIDIGM CORPORATION, F. 2018. Real-Time PCR Analysis, PN 68000088 N1 User Guide. South San Francisco, CA, USA.
- FRANCO, D. L., MAINEZ, J., VEGA, S., SANCHO, P., MURILLO, M. M., DE FRUTOS, C. A., DEL CASTILLO, G., LÓPEZ-BLAU, C., FABREGAT, I. & NIETO, M. A. 2010. Snail1 suppresses TGF- $\beta$ -induced apoptosis and is sufficient to trigger EMT in hepatocytes. *Journal of Cell Science*, 123, 3467-3477. doi:10.1242/jcs.068692
- GAUNT, S. J., BLUM, M. & DE ROBERTIS, E. M. 1993. Expression of the mouse gooseoid gene during mid-embryogenesis may mark mesenchymal cell lineages in the developing head, limbs and body wall. *Development*, 117, 769-778.
- GOLDSMITH, A. M., BENTLEY, J. K., ZHOU, L., JIA, Y., BITAR, K. N., FINGAR, D. C. & HERSHENSON, M. B. 2006. Transforming growth factor-beta induces airway smooth muscle hypertrophy. *American Journal of Respiratory Cell and Molecular Biology*, 34, 247-54. doi:10.1165/rcmb.2005-0166OC
- GOLUB, E. E. & BOESZE-BATTAGLIA, K. 2007. The role of alkaline phosphatase in mineralization. *Current Opinion in Orthopaedics*, 18, 444-448. doi:10.1097/BCO.0b013e3282630851
- GROS, J. & TABIN, C. J. 2014. Vertebrate limb bud formation is initiated by localized epithelial-to-mesenchymal transition. *Science*, 343, 1253-1256. doi:10.1126/science.1248228
- HAM, A. W. 1952. Some histophysiological problems peculiar to calcified tissues. *Journal of Bone and Joint Surgery*, 24 a, 701-28.
- HASHIMOTO, Y., ISHIDA, J., YAMAMOTO, R., FUJIWARA, K., ASADA, S., KASUYA, Y., MOCHIZUKI, N. & FUKAMIZU, A. 2005. G protein-coupled APJ receptor signaling induces focal

- adhesion formation and cell motility. *International journal of molecular medicine*, 16, 787-792.
- HEANUE, T. A., JOHNSON, R. L., IZPISUA-BELMONTE, J.-C., STERN, C. D., DE ROBERTIS, E. M. & TABIN, C. J. 1997. Goosecoid misexpression alters the morphology and Hox gene expression of the developing chick limb bud. *Mechanisms of Development*, 69, 31-37. doi:[https://doi.org/10.1016/S0925-4773\(97\)00149-4](https://doi.org/10.1016/S0925-4773(97)00149-4)
- HEILIG, J., ZAUCKE, F. & NIEHOFF, A. 2018. The role of collagen IX in the ossification of the murine femoral head. *Osteoarthritis and Cartilage*, 26, S98. doi:10.1016/j.joca.2018.02.211
- HESTAND, M. S., KALBFLEISCH, T. S., COLEMAN, S. J., ZENG, Z., LIU, J., ORLANDO, L. & MACLEOD, J. N. 2015. Annotation of the protein coding regions of the equine genome. *PLOS ONE*, 10, e0124375. doi:10.1371/journal.pone.0124375
- HISSNAUER, T. N., BARANOWSKY, A., PESTKA, J. M., STREICHERT, T., WIEGANDT, K., GOEPFERT, C., BEIL, F. T., ALBERS, J., SCHULZE, J., UEBLACKER, P., PETERSEN, J. P., SCHINKE, T., MEENEN, N. M., PÖRTNER, R. & AMLING, M. 2010. Identification of molecular markers for articular cartilage. *Osteoarthritis and Cartilage*, 18, 1630-1638. doi:10.1016/j.joca.2010.10.002
- HODGKINSON, C. P., NAIDOO, V., PATTI, K. G., GOMEZ, J. A., SCHMECKPEPER, J., ZHANG, Z., DAVIS, B., PRATT, R. E., MIROTSOU, M. & DZAU, V. J. 2013. Abi3bp is a multifunctional autocrine/paracrine factor that regulates mesenchymal stem cell biology. *Stem cells*, 31, 1669-1682. doi:10.1002/stem.1416
- HUANG, D. W., SHERMAN, B. T. & LEMPICKI, R. A. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, 4, 44-57. doi:10.1038/nprot.2008.211
- HYDE, G., DOVER, S., ASZODI, A., WALLIS, G. A. & BOOT-HANDFORD, R. P. 2007. Lineage tracing using matrilin-1 gene expression reveals that articular chondrocytes exist as the joint interzone forms. *Developmental biology*, 304, 825-833. doi:10.1016/j.ydbio.2007.01.026
- IWATA, J. I., HACIA, J. G., SUZUKI, A., SANCHEZ-LARA, P. A., URATA, M. & CHAI, Y. 2012. Modulation of noncanonical TGF- $\beta$  signaling prevents cleft palate in Tgfr2 mutant mice. *The Journal of Clinical Investigation*, 122, 873-885. doi:10.1172/JCI61498
- KALBFLEISCH, T. S., RICE, E. S., DEPRIEST, M. S., WALENZ, B. P., HESTAND, M. S., VERMEESCH, J. R., O'CONNELL, B. L., FIDDES, I. T., VERSHININA, A. O., SAREMI, N. F., PETERSEN, J. L., FINNO, C. J., BELLONE, R. R., MCCUE, M. E., BROOKS, S. A., BAILEY, E., ORLANDO, L., GREEN, R. E., MILLER, D. C., ANTCZAK, D. F. & MACLEOD, J. N. 2018. Improved reference genome for the domestic horse increases assembly contiguity and composition. *Communications Biology*, 1, 197. doi:10.1038/s42003-018-0199-z
- KARAMBOULAS, K., DRANSE, H. J. & UNDERHILL, T. M. 2010. Regulation of BMP-dependent chondrogenesis in early limb mesenchyme by TGF $\beta$  signals. *Journal of Cell Science*, 123, 2068-2076. doi:10.1242/jcs.062901
- KARSENTY, G. & WAGNER, E. F. 2002. Reaching a genetic and molecular understanding of skeletal development. *Developmental cell*, 2, 389-406. doi:10.1016/s1534-5807(02)00157-0
- KAWASAKI, Y., KUGIMIYA, F., CHIKUDA, H., KAMEKURA, S., IKEDA, T., KAWAMURA, N., SAITO, T., SHINODA, Y., HIGASHIKAWA, A., YANO, F., OGASAWARA, T., OGATA, N., HOSHI, K.,

- HOFMANN, F., WOODGETT, J. R., NAKAMURA, K., CHUNG, U.-I. & KAWAGUCHI, H. 2008. Phosphorylation of GSK-3 $\beta$  by cGMP-dependent protein kinase II promotes hypertrophic differentiation of murine chondrocytes. *The Journal of Clinical Investigation*, 118, 2506-2515. doi:10.1172/JCI35243
- KOYAMA, E., SHIBUKAWA, Y., NAGAYAMA, M., SUGITO, H., YOUNG, B., YUASA, T., OKABE, T., OCHIAI, T., KAMIYA, N., ROUNTREE, R. B., KINGSLEY, D. M., IWAMOTO, M., ENOMOTO-IWAMOTO, M. & PACIFICI, M. 2008. A distinct cohort of progenitor cells participates in synovial joint and articular cartilage formation during mouse limb skeletogenesis. *Developmental Biology*, 316, 62-73. doi:https://doi.org/10.1016/j.ydbio.2008.01.012
- KOYAMA, N., MIURA, M., NAKAO, K., KONDO, E., FUJII, T., TAURA, D., KANAMOTO, N., SONE, M., YASODA, A., ARAI, H., BESSHO, K. & NAKAO, K. 2013. Human induced pluripotent stem cells differentiated into chondrogenic lineage via generation of mesenchymal progenitor cells. *Stem Cells Dev*, 22, 102-113. doi:10.1089/scd.2012.0127
- KREJCI, P., KRAKOW, D., MEKIKIAN, P. B. & WILCOX, W. R. 2007. Fibroblast growth factors 1, 2, 17, and 19 are the predominant FGF ligands expressed in human fetal growth plate cartilage. *Pediatric Research*, 61, 267-272. doi:10.1203/pdr.0b013e318030d157
- LAMBRECHT, S., ALMQVIST, F., VERDONK, P., VERBRUGGEN, G., DEFORCE, D. & ELEWAUT, D. 2010. The small heat-shock protein HSP27 shows decreased expression in OA-chondrocytes and mediates IL-6 secretion in human articular chondrocytes. *Osteoarthritis and Cartilage*, 18, S107. doi:10.1016/S1063-4584(10)60258-0
- LANE, J. M. & WEISS, C. 1975. Review of articular cartilage collagen research. *Arthritis & Rheumatism*, 18, 553-562. doi:10.1002/art.1780180605
- LEONARD, C. M., FULD, H. M., FRENZ, D. A., DOWNIE, S. A., MASSAGUE, J. & NEWMAN, S. A. 1991. Role of transforming growth factor- $\beta$  in chondrogenic pattern formation in the embryonic limb: Stimulation of mesenchymal condensation and fibronectin gene expression by exogenous TGF- $\beta$  and evidence for endogenous TGF- $\beta$ -like activity. *Developmental Biology*, 145, 99-109. doi:https://doi.org/10.1016/0012-1606(91)90216-P
- LETH, J. M., LETH-ESPENSEN, K. Z., KRISTENSEN, K. K., KUMARI, A., LUND WINTHER, A.-M., YOUNG, S. G. & PLOUG, M. 2019. Evolution and Medical Significance of LU Domain-Containing Proteins. *International journal of molecular sciences*, 20, 2760. doi:10.3390/ijms20112760
- LIU, C. F. & LEFEBVRE, V. 2015. The transcription factors SOX9 and SOX5/SOX6 cooperate genome-wide through super-enhancers to drive chondrogenesis. *Nucleic acids research*, 43, 8183-8203. doi:10.1093/nar/gkv688
- LIU, Y., BEYER, A. & AEBERSOLD, R. 2016. On the dependency of cellular protein levels on mRNA abundance. *Cell*, 165, 535-550. doi:10.1016/j.cell.2016.03.014
- LONG, F. & ORNITZ, D. M. 2013. Development of the endochondral skeleton. *Cold Spring Harbor Perspectives in Biology*, 5, a008334. doi:10.1101/cshperspect.a008334
- LUO, G., DUCY, P., MCKEE, M. D., PINERO, G. J., LOYER, E., BEHRINGER, R. R. & KARSENTY, G. 1997. Spontaneous calcification of arteries and cartilage in mice lacking matrix GLA protein. *Nature*, 386, 78-81. doi:10.1038/386078a0
- MANKOO, B. S., SKUNTZ, S., HARRIGAN, I., GRIGORIEVA, E., CANDIA, A., WRIGHT, C. V. E., ARNHEITER, H. & PACHNIS, V. 2003. The concerted action of Meox homeobox genes is required upstream of genetic pathways essential for the formation, patterning and

- differentiation of somites. *Development*, 130, 4655-4664. doi:10.1242/dev.00687
- MCCARTHY, H. E., BARA, J. J., BRAKSPEAR, K., SINGHRAO, S. K. & ARCHER, C. W. 2012. The comparison of equine articular cartilage progenitor cells and bone marrow-derived stromal cells as potential cell sources for cartilage repair in the horse. *The Veterinary Journal*, 192, 345-351. doi:10.1016/j.tvjl.2011.08.036
- MELAS, I. N., CHAIRAKAKI, A. D., CHATZOPOULOU, E. I., MESSINIS, D. E., KATOPODI, T., PLIAKA, V., SAMARA, S., MITSOS, A., DAILIANA, Z., KOLLIA, P. & ALEXOPOULOS, L. G. 2014. Modeling of signaling pathways in chondrocytes based on phosphoproteomic and cytokine release data. *Osteoarthritis and Cartilage*, 22, 509-518. doi:10.1016/j.joca.2014.01.001
- MERZ, D., LIU, R., JOHNSON, K. & TERKELTAUB, R. 2003. IL-8/CXCL8 and Growth-related oncogene  $\alpha$ /CXCL1 induce chondrocyte hypertrophic differentiation. *The Journal of Immunology*, 171, 4406. doi:10.4049/jimmunol.171.8.4406
- MEVEL, R., DRAPER, J. E., LIE-A-LING, M., KOUSKOFF, V. & LACAUD, G. 2019. RUNX transcription factors: orchestrators of development. *Development*, 146, dev148296. doi:10.1242/dev.148296
- MOHANTY-HEJMADI, P., DUTTA, S. K. & MAHAPATRA, P. 1992. Limbs generated at site of tail amputation in marbled balloon frog after vitamin A treatment. *Nature*, 355, 352-353. doi:10.1038/355352a0
- MOSKALEWSKI, S., HYC, A., JANKOWSKA-STEIFER, E. & OSIECKA-IWAN, A. 2013. Formation of synovial joints and articular cartilage. *Folia Morphologica*, 72, 181-7. doi:10.5603/fm.2013.0031
- MUSARÒ, A., MCCULLAGH, K. J., NAYA, F. J., OLSON, E. N. & ROSENTHAL, N. 1999. IGF-1 induces skeletal myocyte hypertrophy through calcineurin in association with GATA-2 and NF-ATc1. *Nature*, 400, 581-585. doi:10.1038/23060
- NEJADNIK, H., DIECKE, S., LENKOV, O. D., CHAPELIN, F., DONIG, J., TONG, X., DERUGIN, N., CHAN, R. C. F., GAUR, A., YANG, F., WU, J. C. & DALDRUP-LINK, H. E. 2015. Improved approach for chondrogenic differentiation of human induced pluripotent stem cells. *Stem Cell Reviews and Reports*, 11, 242-253. doi:10.1007/s12015-014-9581-5
- NELSON, C. E., MORGAN, B. A., BURKE, A. C., LAUFER, E., DIMAMBRO, E., MURTAUGH, L. C., GONZALES, E., TESSAROLLO, L., PARADA, L. F. & TABIN, C. 1996. Analysis of Hox gene expression in the chick limb bud. *Development*, 122, 1449-1466.
- NG, L. J., WHEATLEY, S., MUSCAT, G. E. O., CONWAY-CAMPBELL, J., BOWLES, J., WRIGHT, E., BELL, D. M., TAM, P. P. L., CHEAH, K. S. E. & KOOPMAN, P. 1997. SOX9 binds DNA, activates transcription, and coexpresses with type II collagen during chondrogenesis in the mouse. *Developmental Biology*, 183, 108-121. doi:https://doi.org/10.1006/dbio.1996.8487
- NINOMIYA, K., MIYAMOTO, T., IMAI, J.-I., FUJITA, N., SUZUKI, T., IWASAKI, R., YAGI, M., WATANABE, S., TOYAMA, Y. & SUDA, T. 2007. Osteoclastic activity induces osteomodulin expression in osteoblasts. *Biochemical and Biophysical Research Communications*, 362, 460-466. doi:10.1016/j.bbrc.2007.07.193
- NISHIDA, T., KUBOTA, S., AOYAMA, E. & TAKIGAWA, M. 2013. Impaired glycolytic metabolism causes chondrocyte hypertrophy-like changes via promotion of phospho-Smad1/5/8 translocation into nucleus. *Osteoarthritis and Cartilage*, 21, 700-709. doi:https://doi.org/10.1016/j.joca.2013.01.013

- NISHIMURA, R., WAKABAYASHI, M., HATA, K., MATSUBARA, T., HONMA, S., WAKISAKA, S., KIYONARI, H., SHIHO, G., YAMAGUCHI, A., TSUMAKI, N., AKIYAMA, H. & YONEDA, T. 2012. Osterix regulates calcification and degradation of chondrogenic matrices through matrix metalloproteinase (MMP13) expression in association with transcription factor Runx2 during endochondral ossification. *Journal of Biological Chemistry*, 287, 33179-33190. doi:10.1074/jbc.M111.337063
- OGLE, M. E., OLINGY, C. E., AWOJODU, A. O., DAS, A., ORTIZ, R. A., CHEUNG, H. Y. & BOTCHWEY, E. A. 2017. Sphingosine-1-Phosphate Receptor-3 supports hematopoietic stem and progenitor cell residence within the bone marrow niche. *Stem cells*, 35, 1040-1052. doi:10.1002/stem.2556
- OH, S. K., SHIN, J. O., BAEK, J. I., LEE, J., BAE, J. W., ANKAMERDDY, H., KIM, M. J., HUH, T. L., RYOO, Z. Y., KIM, U. K., BOK, J. & LEE, K. Y. 2015. Pannexin 3 is required for normal progression of skeletal development in vertebrates. *The FASEB Journal*, 29, 4473-4484. doi:10.1096/fj.15-273722
- ONDRÉSIK M., OLIVEIRA J. M. & REIS R.L. 2017. Knee articular cartilage. In: Oliveira J., Reis R. (eds) *Regenerative Strategies for the Treatment of Knee Joint Disabilities. Studies in Mechanobiology, Tissue Engineering and Biomaterials*, vol 21. Springer International Publishing, Cham, Switzerland. doi:10.1007/978-3-319-44785-8
- PACIFICI, M., KOYAMA, E. & IWAMOTO, M. 2005. Mechanisms of synovial joint and articular cartilage formation: recent advances, but many lingering mysteries. *Birth Defects Research Part C: Embryo Today: Reviews*, 75, 237-248. doi:10.1002/bdrc.20050
- PEACOCK, J. D., HUK, D. J., EDIRIWEERA, H. N. & LINCOLN, J. 2011. Sox9 transcriptionally represses Spp1 to prevent matrix mineralization in maturing heart valves and chondrocytes. *PloS one*, 6. doi:10.1371/journal.pone.0026769
- PFEIFER, A., ASZÓDI, A., SEIDLER, U., RUTH, P., HOFMANN, F. & FÄSSLER, R. 1996. Intestinal secretory defects and dwarfism in mice lacking cGMP-dependent protein kinase II. *Science*, 274, 2082-6. doi:10.1126/science.274.5295.2082
- PHINNEY, D. G., GRAY, A. J., HILL, K. & PANDEY, A. 2005. Murine mesenchymal and embryonic stem cells express a similar Hox gene profile. *Biochemical and Biophysical Research Communications*, 338, 1759-1765. doi:10.1016/j.bbrc.2005.10.140
- PINEAULT, K. M. & WELLIK, D. M. 2014. Hox Genes and Limb Musculoskeletal Development. *Current Osteoporosis Reports*, 12, 420-427. doi:10.1007/s11914-014-0241-0
- PITSILLIDES, A. A. & ASHHURST, D. E. 2008. A critical evaluation of specific aspects of joint development. *Developmental Dynamics*, 237, 2284-2294. doi:10.1002/dvdy.21654
- RAKIC, R., BOURDON, B., DEMOOR, M., MADDENS, S., SAULNIER, N. & GALÉRA, P. 2018. Differences in the intrinsic chondrogenic potential of equine umbilical cord matrix and cord blood mesenchymal stromal/stem cells for cartilage regeneration. *Scientific Reports*, 8, 13799. doi:10.1038/s41598-018-28164-9
- RANGANATHAN, P., AGRAWAL, A., BHUSHAN, R., CHAVALMANE, A. K., KALATHUR, R. K. R., TAKAHASHI, T. & KONDAIAH, P. 2007. Expression profiling of genes regulated by TGF-beta: Differential regulation in normal and tumour cells. *BMC Genomics*, 8, 98. doi:10.1186/1471-2164-8-98
- ROBERTSON, I. B., HORIGUCHI, M., ZILBERBERG, L., DABOVIC, B., HADJIOLOVA, K. & RIFKIN, D. B. 2015. Latent TGF-β-binding proteins. *Matrix Biology*, 47, 44-53. doi:10.1016/j.matbio.2015.05.005

- ROKUTANDA, S., FUJITA, T., KANATANI, N., YOSHIDA, C. A., KOMORI, H., LIU, W., MIZUNO, A. & KOMORI, T. 2009. Akt regulates skeletal development through GSK3, mTOR, and FoxOs. *Developmental biology*, 328, 78-93. doi:10.1016/j.ydbio.2009.01.009
- RUDINI, N., FELICI, A., GIAMPIETRO, C., LAMPUGNANI, M., CORADA, M., SWIRSDING, K., GARRÈ, M., LIEBNER, S., LETARTE, M., TEN DIJKE, P. & DEJANA, E. 2008. VE-cadherin is a critical endothelial regulator of TGF-beta signalling. *The EMBO journal*, 27, 993-1004. doi:10.1038/emboj.2008.46
- SAITO, T., IKEDA, T., NAKAMURA, K., CHUNG, U. I. & KAWAGUCHI, H. 2007. S100A1 and S100B, transcriptional targets of SOX trio, inhibit terminal differentiation of chondrocytes. *EMBO reports*, 8, 504-509. doi:10.1038/sj.embor.7400934
- SCHEIJEN, B., BRONK, M., VAN DER MEER, T. & BERNARDS, R. 2003. Constitutive E2F1 overexpression delays endochondral bone formation by inhibiting chondrocyte differentiation. *Molecular and Cellular Biology*, 23, 3656-3668. doi:10.1128/mcb.23.10.3656-3668.2003
- SCOTT, I. C., MASRI, B., D'AMICO, L. A., JIN, S.-W., JUNGBLUT, B., WEHMAN, A. M., BAIER, H., AUDIGIER, Y. & STAINIER, D. Y. R. 2007. The G protein-coupled receptor AGTRL1B regulates early development of myocardial progenitors. *Developmental Cell*, 12, 403-413. doi:10.1016/j.devcel.2007.01.012
- SEKINE, K., OHUCHI, H., FUJIWARA, M., YAMASAKI, M., YOSHIZAWA, T., SATO, T., YAGISHITA, N., MATSUI, D., KOGA, Y., ITOH, N. & KATO, S. 1999. Fgf10 is essential for limb and lung formation. *Nature Genetics*, 21, 138-141. doi:10.1038/5096
- SETTLE, S. H., ROUNTREE, R. B., SINHA, A., THACKER, A., HIGGINS, K. & KINGSLEY, D. M. 2003. Multiple joint and skeletal patterning defects caused by single and double mutations in the mouse Gdf6 and Gdf5 genes. *Developmental Biology*, 254, 116-130. doi:10.1016/s0012-1606(02)00022-2
- SHAM, M. H., LUI, V. C. H., FU, M., CHEN, B. & TAM, P. K. H. 2001. SOX10 is abnormally expressed in aganglionic bowel of Hirschsprung's disease infants. *Gut*, 49, 220-226. doi:10.1136/gut.49.2.220
- SHI, Y. & MASSAGUÉ, J. 2003. Mechanisms of TGF- $\beta$  signaling from cell membrane to the nucleus. *Cell*, 113, 685-700. doi:10.1016/S0092-8674(03)00432-X
- SHIMIZU, H., YOKOYAMA, S. & ASAHARA, H. 2007. Growth and differentiation of the developing limb bud from the perspective of chondrogenesis. *Development, Growth & Differentiation*, 49, 449-454. doi:10.1111/j.1440-169X.2007.00945.x
- SHWARTZ, Y., VIUKOV, S., KRIEF, S. & ZELZER, E. 2016. Joint development involves a continuous influx of Gdf5-positive cells. *Cell Reports*, 15, 2577-2587. doi:https://doi.org/10.1016/j.celrep.2016.05.055
- SINGH, J. & PADGETT, R. A. 2009. Rates of in situ transcription and splicing in large human genes. *Nature Structural & Molecular Biology*, 16, 1128-1133. doi:10.1038/nsmb.1666
- SMITS, P., DY, P., MITRA, S. & LEFEBVRE, V. 2004. Sox5 and Sox6 are needed to develop and maintain source, columnar, and hypertrophic chondrocytes in the cartilage growth plate. *The Journal of cell biology*, 164, 747-758. doi:10.1083/jcb.200312045
- SOPHIA FOX, A. J., BEDI, A., & RODEO, S. A. 2009. The basic science of articular cartilage: structure, composition, and function. *Sports Health*, 1, 461-468. doi:10.1177/1941738109350438
- STUART, T., BUTLER, A., HOFFMAN, P., HAFEMEISTER, C., PAPALEXI, E., MAUCK, W. M., III, HAO,

- Y., STOECKIUS, M., SMIBERT, P. & SATIJA, R. 2019. Comprehensive integration of single-cell data. *Cell*, 177, 1888-1902.e21. doi:10.1016/j.cell.2019.05.031
- SUN, M. M. G. & BEIER, F. 2014. Chondrocyte hypertrophy in skeletal development, growth, and disease. *Birth Defects Research Part C: Embryo Today: Reviews*, 102, 74-82. doi:10.1002/bdrc.21062
- TAN, L. & YAMMANI, R. R. 2019. Nupr1 regulates palmitate-induced apoptosis in human articular chondrocytes. *Bioscience reports*, 39. doi:10.1042/BSR20181473
- TANG, F., BARBACIORU, C., WANG, Y., NORDMAN, E., LEE, C., XU, N., WANG, X., BODEAU, J., TUCH, B. B., SIDDIQUI, A., LAO, K. & SURANI, M. A. 2009. mRNA-Seq whole-transcriptome analysis of a single cell. *Nature Methods*, 6, 377-382. doi:10.1038/nmeth.1315
- TIAN, H., BIEHS, B., CHIU, C., SIEBEL, C. W., WU, Y., COSTA, M., DE SAUVAGE, FREDERIC J. & KLEIN, OPHIR D. 2015. Opposing activities of Notch and Wnt signaling regulate intestinal stem cells and gut homeostasis. *Cell Reports*, 11, 33-42. doi:10.1016/j.celrep.2015.03.007
- TRAPNELL, C., ROBERTS, A., GOFF, L., PERTEA, G., KIM, D., KELLEY, D. R., PIMENTEL, H., SALZBERG, S. L., RINN, J. L. & PACHTER, L. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols*, 7, 562-578. doi:10.1038/nprot.2012.016
- USDA 2017. Equine 2015, Equine Management and Select Equine Health Conditions in the United States. *United States Department of Agriculture, Animal and Plant Health Inspection Service Veterinary Services, National Animal Health Monitoring System Report 3*.
- WALTER, W., SÁNCHEZ-CABO, F. & RICOTE, M. 2015. GOplot: an R package for visually combining expression data with functional analysis. *Bioinformatics*, 31, 2912-2914. doi:10.1093/bioinformatics/btv300
- WANG, W., RIGUEUR, D. & LYONS, K. M. 2014. TGF $\beta$  signaling in cartilage development and maintenance. *Birth Defects Research Part C: Embryo Today: Reviews*, 102, 37-51. doi:10.1002/bdrc.21058
- WARNES, G. R., BOLKER, B., BONEBAKKER, L., GENTLEMAN, R., LIAW, W. H. A., LUMLEY, T., MAECHLER, M., MAGNUSSON, A., MOELLER, S. & SCHWARTZ, M. 2015. gplots: Various R programming tools for plotting data.
- WATANABE, H., DE CAESTECKER, M. P. & YAMADA, Y. 2001. Transcriptional cross-talk between Smad, ERK1/2, and p38 mitogen-activated protein kinase pathways regulates transforming growth factor- $\beta$ -induced aggrecan gene expression in chondrogenic ATDC5 cells. *Journal of Biological Chemistry*, 276, 14466-14473. doi:10.1074/jbc.M005724200
- WU, L., LEIJTEN, J., VAN BLITTERSWIJK, C. A. & KAPERIEN, M. 2013. Fibroblast growth factor-1 is a mesenchymal stromal cell-secreted factor stimulating proliferation of osteoarthritic chondrocytes in co-culture. *Stem Cells and Development*, 22, 2356-67. doi:10.1089/scd.2013.0118
- YAMAMOTO, K. R. & ALBERTS, B. M. 1976. Steroid receptors: elements for modulation of eukaryotic transcription. *Annual Review of Biochemistry*, 45, 721-46. doi:10.1146/annurev.bi.45.070176.003445
- YAMAZAKI, S., NAKANO, N., HONJO, A., HARA, M., MAEDA, K., NISHIYAMA, C., KITaura, J.,



- OHTSUKA, Y., OKUMURA, K., OGAWA, H. & SHIMIZU, T. 2015. The transcription factor Ehf is involved in TGF- $\beta$  induced suppression of Fc $\epsilon$ RI and c-Kit expression and Fc $\epsilon$ RI-mediated activation in mast cells. *The Journal of Immunology*, 195, 3427-3435. doi:10.4049/jimmunol.1402856
- YOKOTA, J., CHOSA, N., SAWADA, S., OKUBO, N., TAKAHASHI, N., HASEGAWA, T., KONDO, H. & ISHISAKI, A. 2014. PDGF-induced PI3K-mediated signaling enhances the TGF- $\beta$ -induced osteogenic differentiation of human mesenchymal stem cells in a TGF- $\beta$ -activated MEK-dependent manner. *International Journal of Molecular Medicine*, 33, 534-42. doi:10.3892/ijmm.2013.1606
- ZENG, X. X. I., WILM, T. P., SEPICH, D. S. & SOLNICA-KREZEL, L. 2007. Apelin and its receptor control heart field formation during zebrafish gastrulation. *Developmental Cell*, 12, 391-402. doi:10.1016/j.devcel.2007.01.011
- ZHOU, Y., CHEN, M., RICUPERO, C. L., HE, L., WU, J., CHEN, K., FRIEDMAN, R. A., GUARNIERI, P., WANG, Z., ZHOU, X. & MAO, J. J. 2018. Profiling of stem/progenitor cell regulatory genes of the synovial joint by genome-wide RNA-seq analysis. *BioMed Research International*, 2018, 9327487-9327487. doi:10.1155/2018/9327487
- ZHU, H. & BENDALL, A. J. 2009. Dlx5 is a cell autonomous regulator of chondrocyte hypertrophy in mice and functionally substitutes for Dlx6 during endochondral ossification. *PLOS ONE*, 4, e8097. doi:10.1371/journal.pone.0008097

**Vita**  
**ChanHee Mok**

---

**EDUCATION**

**University of Kentucky, 2015 – 2020** Lexington, KY  
Doctor of Philosophy in Veterinary Science  
- Research subject: Equine Orthopedics, Cell Biology  
**Mok, C. H.** 2020. Comparative chondrogenesis of interzone and anlagen cells in equine skeletal development. University of Kentucky. *Doctoral Dissertation*

**University of Kentucky, 2013 – 2015** Lexington, KY  
Master of Science in Animal Science  
- Research subject: Equine Nutrition  
**Mok, C. H.** 2015. Using the indicator amino acid oxidation technique to study threonine requirements in horses. University of Kentucky. *Master's Thesis*

**Konkuk University, 2007 – 2012** Seoul, South Korea  
Bachelor of Science in Animal Science & Environment  
- Research subject: Swine Nutrition  
**Mok, C. H.** 2011. A comparison of an oven-drying method and a freeze-drying method for measuring energy concentration in pig urine. Konkuk University. *Bachelor's Thesis*

**AWARDS and ASSISTANTSHIPS**

- Research Assistantship  
    Geoffrey C. Hughes Foundation, Fall 2015 – 2020
- Three-Minute Thesis (3-MT) competition for Graduate students, 1<sup>st</sup> place  
    College of Agriculture, Food and Environment, University of Kentucky, Spring 2019
- Three-Minute Thesis (3-MT) competition for PhD candidates, 2<sup>nd</sup> place  
    Department of Veterinary Science, University of Kentucky, Spring 2019
- Three-Minute Thesis (3-MT) competition for PhD candidates, 1<sup>st</sup> place  
    Department of Veterinary Science, University of Kentucky, Spring 2018
- Research Assistantship  
    Department of Animal and Food Sciences, University of Kentucky, Fall 2014 – Spring 2015
- Kentucky Opportunity Fellowship  
    University of Kentucky, Fall 2013 – Spring 2014
- Research Assistantship  
    Department of Animal Bioscience, Konkuk University, South Korea, 2011 – 2012
- Undergraduate Fellowship  
    Department of Animal Bioscience, Konkuk University, South Korea, 2007 – 2010

## **PUBLICATIONS**

**Mok, C. H.** and K. L. Urschel. 2020. Amino acid requirements in horses – A review. *Asian-Australas. J. Anim. Sci.* 33:679–695.

**Mok, C. H.**, C. L. Levesque, and K. L. Urschel. 2019. Precision nutrition for monogastric animals. Monogastric Feed Research Institute, Konkuk University, Seoul, South Korea. p81 – 94. ISBN: 979-11-965289-1-1.

**Mok, C. H.**, C. L. Levesque, and K. L. Urschel. 2018. Using the indicator amino acid oxidation technique to study threonine requirements in horses receiving a predominantly forage diet. *J. Anim. Physiol. Anim. Nutr.* 102:1366–1381.

**Mok, C. H.**, C. Kong, and B. G. Kim. 2015. Combination of phytase and  $\beta$ -mannanase supplementation on energy and nutrient digestibility in pig diets containing palm kernel expellers. *Anim. Feed. Sci. Technol.* 205:116–121.

**Mok, C. H.**, J. H. Lee, and B. G. Kim. 2013. Effects of exogenous phytase and  $\beta$ -mannanase on ileal and total tract digestibility of energy and nutrient in palm kernel expeller-containing diets fed to growing pigs. *Anim. Feed. Sci. Technol.* 186:209–213.

**Mok, C. H.**, S. Y. Shin, and B. G. Kim. 2013. Aflatoxin, deoxynivalenol, and zearalenone in swine diets: Predictions on growth performance. *Rev. Colomb. Cienc. Pecu.* 26:243–254.