# COMPARATIVE TRANSCRIPTOMICS REVEALS THE GENE REGULATORY NETWORK UNDERLYING THE DIFFERENTIATION AND EVOLUTION OF SKELETAL CELLS

A Thesis Submitted to the
College of Graduate and Postdoctoral Studies
In Partial Fulfillment of the Requirements
For the Degree of Doctor of Philosophy
In the Department of Anatomy, Physiology and Pharmacology
University of Saskatchewan
Saskatoon

By

PATSY GÓMEZ PICOS

## PERMISSION TO USE

In presenting this thesis/dissertation in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis/dissertation in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis/dissertation work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis/dissertation or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis/dissertation.

## DISCLAIMER

Reference in this thesis/dissertation to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement, recommendation, or favoring by the University of Saskatchewan. The views and opinions of the author expressed herein do not state or reflect those of the University of Saskatchewan and shall not be used for advertising or product endorsement purposes.

Requests for permission to copy or to make other uses of materials in this thesis/dissertation in whole or part should be addressed to:

Head of the Department of Anatomy, Physiology and Pharmacology
107 Wiggins Rd
University of Saskatchewan
Saskatoon, Saskatchewan S7N 5E5 Canada

OR

Dean
College of Graduate and Postdoctoral Studies
University of Saskatchewan
116 Thorvaldson Building, 110 Science Place
Saskatoon, Saskatchewan S7N 5C9 Canada

## ABSTRACT

Deciphering the gene regulatory network (GRN) underlying skeletal cells is key to understanding the differentiation and even origins of cartilage and bone, but current knowledge is limited to a few candidate genes. The vertebrate skeleton is mostly composed of three specific cell types: immature chondrocytes (IMM), mature (hypertrophic) chondrocytes (MAT), and osteoblasts (OST). Importantly, mature chondrocytes share the expression of many genes with both immature chondrocytes and osteoblasts. Currently little is known about mechanisms of gene regulation in mature chondrocytes, but previous studies suggest that overlapping actions between portions of the GRN active in IMM and OST direct the differentiation of MAT. While one GRN might give rise to distinct cell fates, studies analyzing the interaction between distinct portions of a GRN during cell differentiation are lacking. Immature chondrocytes, mature chondrocytes, and osteoblasts can have distinct embryological origins, mesoderm and neural crest, but the similarities in gene expression among skeletal cells of these two embryonic lineages remain controversial. Here, we aim to test the hypothesis that the skeletal cell GRN is generally conserved throughout the body and across vertebrate clades, but the molecular relationship between portions of this GRN active in IMM and OST has been modified during evolution (i.e. positive in earlier diverged vertebrates vs antagonistic in later diverged vertebrates). To test this hypothesis, laser- capture microdissection (LCM) coupled to RNA-seq was performed on skeletal cells (i.e. immature chondrocytes, mature chondrocytes, and osteoblasts) of mouse, chick, and gar. The hypothesis was partially supported by our results. First, transcriptomic analyses on a single species, the mouse embryo, suggested that one GRN drives the differentiation of IMM, MAT, and OST, and the overlapping actions of portions of this GRN active in IMM and OST regulate MAT differentiation. Second, to test whether this skeletal cell GRN is conserved throughout the body, limb (humerus, mesoderm-derived) and head cartilage (ceratobranchial, neural-crest derived) isolated from the chick embryo were compared and their transcriptomes showed a high degree of similarity. Third, regarding bone evolution, previous molecular studies suggest that gar OST express higher levels of chondrocyte genes (e.g. *Sox9*, *Col2a1*, and *Col10a1*) compared to mouse and chick OST, suggesting that the osteoblast might have evolved from a chondrocyte. To further test this hypothesis, several bioinformatic approaches including differential gene expression, model-based clustering, and gene ontology analyses were performed. Pairwise differential gene expression analyses revealed that gar OST indeed expressed

higher levels of some hallmark chondrogenic markers including *Col2a1*, *Sox6*, *Col10a1*, and *Acan* compared to mouse or chick OST. Moreover, model-based clustering analysis revealed one cluster that showed higher expression in the gar OST and included the hallmark mature chondrocyte gene *Col10a1*. Finally, with the goal of understanding how changes in GRNs underlie skeletal cell evolution, gene co-expression network (GCN) analysis was used to estimate skeletal cell GRNs. In mouse and chick, two portions of the GRN driving IMM and OST exhibit an antagonistic relationship, and they only interacted positively in MAT. In contrast, GCN analysis in gar skeletal cells showed that positive interactions between IMM and OST increased, supporting the hypothesis that the relationship between portions of the GRN active in IMM and OST has been modified over time. To provide insight into how portions of the GRN might regulate gene expression in MAT, model-based cluster analysis was performed, and specific categories of gene expression were identified. These categories of gene expression were conserved to some degree in skeletal cell transcriptomes of all vertebrate clades, but when gar skeletal cell transcriptomes were analyzed more clusters showed downregulation of gene expression levels in MAT compared to IMM and OST, while IMM and OST exhibited similar expression levels. This result in gar was particularly unexpected since results in mouse and chick showed that IMM and OST generally show opposite gene expression patterns, and MAT is mostly a combination of gene expression levels between both cell types. Adding more phylogenetic clades into these evolutionary comparisons will provide more insight into skeletal cell GRN regulation and structure. These results highlight the complexity of skeletal cell GRN organization and propose a novel unbiased approach through which to understand differentiation and evolutionary origins of cartilage and bone.

## ACKNOWLEDGEMENTS

I would like to thank my patient and supportive supervisor, Dr. Brian Eames, for his unconditional encouragement and mentorship throughout my PhD project. I am extremely grateful for his guidance and support in my personal and academic endeavors. Furthermore, I would like to thank my committee members Dr. Julia Boughner, Dr. Ian McQuillan, Dr. Troy Harkness, Dr. Patrick Krone, and Dr. William Kulyk whose assistance and valuable feedback was a milestone in completing my project. I would also like to express my gratitude to Dr. Paul Trainor who kindly agreed to participate as an external examiner of this thesis.

In addition to my committee members, I would also like to acknowledge my Eames lab mates, thanks for all the good moments we shared and your constant support. Special thanks to Katie Ovens for her unconditional help and patience, without her it would have been incredibly challenging to analyze the transcriptomic data. Thanks for teaching me so much about the bioinformatic aspect of my project and explaining in detail every analysis. I would also like to express my sincere thanks to Dr. Amir Ashique, who collected the mouse skeletal cell types and greatly helped me to get this project started by teaching me many important techniques that were crucial for this work. Also, thanks to Dr. Bogdan Popescu for the use of his laser capture equipment.

Many thanks to the Department of Anatomy, Physiology and Pharmacology, College of Graduate and Postdoctoral Studies, CONACYT Mexico, as well as supervisor's NSERC grant who provided funding for this project.

Finally, I would like to thank my family who have always motivated and supported me in any project I decide to pursue.

## DEDICATION

*To my parents and sister, who always support me and never give up on me; to my husband for his unconditional love and encouragement to pursue my goals; to my beloved daughter who inspires me every day to be the best person I can be.*

**TABLE OF CONTENTS**

## LIST OF TABLES

### CHAPTER 3. Portions of the gene regulatory network driving cartilage and bone formation interact via averaging and synergism during cartilage maturation



### CHAPTER 4: Comparative transcriptomics reveal that a conserved molecular program underlies mesoderm- and neural crest-derived chondrocytes



### CHAPTER 5: Comparative transcriptomics reveals the gene regulatory network driving osteoblast differentiation and identifies higher expression of cartilage genes in gar bone



### CHAPTER 6: Skeletal cell GRN structure and regulatory control of mature chondrocytes are highly conserved features between mouse and chick, while gar shows more variation

## LIST OF FIGURES

### CHAPTER 1. Introduction



### CHAPTER 2. On the evolutionary relationship between chondrocytes and osteoblasts



### CHAPTER 3. Portions of the gene regulatory network driving cartilage and bone formation interact via averaging and synergism during cartilage maturation

**CHAPTER 4: Comparative transcriptomics reveal that a conserved molecular program underlies mesoderm- and neural crest-derived chondrocytes**



**CHAPTER 5: Comparative transcriptomics reveals the gene regulatory network driving osteoblast differentiation and identifies higher expression of cartilage genes in gar bone**

**CHAPTER 6: Skeletal cell GRN structure and regulatory control of mature chondrocytes are highly conserved features between mouse and chick, while gar shows more variation**

**CHAPTER 7. Limitations and future directions**



**Appendix I. Histological characterization of frog skeletal tissues**



**Appendix II. Book chapter: On the evolution of skeletal cells before and after neural crest**

## LIST OF ABBREVIATIONS

| | |
|---|---|
| **A2P** | Ascorbate-2-phosphate |
| ***Acan*** | Aggrecan |
| **ALP** | Alkaline phosphatase protein |
| ***Alpl*** | Alkaline phosphatase |
| ***Angpt1*** | Angiopoietin 1 |
| **ARS** | Alizarin Red staining |
| **As** | Angulosplenial |
| **ATAC-Seq** | Assay for Transposase-Accessible Chromatin using sequencing |
| ***Atf3*** | Activating Transcription Factor 3 |
| **AVE** | average gene expression between IMM and OST |
| ***Bambi*** | BMP And Activin Membrane Bound Inhibitor |
| ***Bglap*** | Bone Gamma-Carboxyglutamate Protein |
| **BGP** | β-glycerophosphate |
| **Bh** | basihyal |
| ***Bmp5*** | Bone Morphogenetic Protein 5 |
| ***Bmp6*** | Bone Morphogenetic Protein 6 |
| ***Bmp7*** | Bone Morphogenetic Protein 7 |
| **c** | carpal |
| **Cb** | ceratobranchial |
| **cDNA** | complementary DNA |
| **Ch** | ceratohyal |
| **Chip-Seq** | Chromatin Immunoprecipitation Sequencing |
| **$CO_2$** | Carbon dioxide |
| ***Col10a1*** | Collagen Type X Alpha 1 Chain |
| ***Col11a1*** | Collagen Type XI Alpha 1 Chain |
| ***Col1a1*** | Collagen Type I Alpha 1 Chain |
| ***Col1a2*** | Collagen Type I Alpha 2 Chain |
| ***Col2a1*** | Collagen Type II Alpha 1 Chain |
| ***Col9a1*** | Collagen Type IX Alpha 1 Chain |
| ***Col9a2*** | Collagen Type IX Alpha 2 Chain |
| **COLX** | Collagen type X protein |
| ***Ctnnb1*** | Catenin beta-1 (β-catenin) |
| **d** | dentary |
| **DAVID** | Database for Annotation, Visualization and Integrated Discovery |
| ***Dcn*** | Decorin |
| **DIG** | Digoxigenin |
| ***Dlx2*** | Distal-Less Homeobox 2 |
| ***Dlx4*** | Distal-Less Homeobox 4 |

| | |
|---|---|
| ***Dlx5*** | Distal-Less Homeobox 5 |
| ***Dlx6*** | Distal-Less Homeobox 6 |
| **DMEM** | Dulbecco's Modified Eagle's Medium |
| **dpf** | days post-fertilization |
| **E14.5** | Embryonic day 14.5 |
| **Eb** | epibranchial |
| ***Ebf2*** | EBF Transcription Factor 2 |
| **ECM** | Extracellular matrix |
| **EDTA** | Ethylenediaminetetraacetic acid |
| ***Epyc*** | Epiphycan |
| ***Fbn2*** | Fibrillin 2 |
| ***Fn1*** | Fibronectin 1 |
| ***Foxa2*** | Forkhead Box protein A2 |
| ***Foxa3*** | Forkhead Box protein A3 |
| ***Foxo1*** | Forkhead Box O1 |
| ***Foxo3*** | Forkhead box O3 |
| **GCN** | Gene Co-expression Network |
| **GEO** | Gene Expression Omnibus |
| **GFP** | Green Fluorescent Protein |
| **GO** | Gene ontology |
| **GRN** | Gene regulatory network |
| **h** | humerus |
| **hCG** | Human chorionic gonadotropin |
| **HCL** | Hidrochloric acid |
| ***Hey1*** | Hes Related Family BHLH Transcription Factor With YRPW Motif 1 |
| **HH36** | Hamburger–Hamilton stage 36 |
| ***Hhip*** | Hedgehog Interacting Protein |
| ***Ibsp*** | Integrin Binding Sialoprotein |
| **IHC** | Immunohistochemistry |
| ***Ihh*** | Indian hedgehog |
| **IMM** | Immature chondrocyte |
| **IRX** | Iroquois gene family |
| ***Irx1*** | Iroquois Homeobox 1 |
| ***Irx3*** | Iroquois Homeobox 3 |
| ***Irx5*** | Iroquois Homeobox 5 |
| ***Irx6*** | Iroquois Homeobox 6 |
| **ISH** | *In situ* hybridization |
| **ITS** | Insulin-transferrin-sodium selenite |
| **KOH** | Potassium hydroxide |
| **LCM** | Laser capture microdissection |

| | |
|---|---|
| **m** | metacarpal |
| **MAT** | Mature chondrocyte |
| ***Matn1*** | Matrilin 1 |
| **MBS** | Modified Barth's saline |
| ***Mef2c*** | Myocyte Enhancer Factor 2C |
| **$Mgcl_2$** | Magnesium chloride |
| **Mgp** | Matrix Gla Protein |
| **MgSO4** | magnesium sulfate |
| **Mk** | Meckel's cartilage |
| **ml** | milliliter |
| **mM** | millimolar |
| ***Mmp13*** | Matrix Metallopeptidase 13 |
| ***Mmp2*** | Matrix Metallopeptidase 2 |
| ***Mmp9*** | Matrix Metallopeptidase 9 |
| **mRNA** | messenger RNA |
| ***Msx1*** | Msh homeobox 1 |
| ***Msx2*** | Msh homeobox 2 |
| **$N_2$** | Nitrogen |
| **NaCl** | Sodium Chloride |
| **NaHCO3** | Sodium Bicarbonate |
| **NC** | Neural crest |
| **NCBI** | National Center for Biotechnology Information |
| **NF** | Nieuwkoop and Faber |
| **NGS** | Normal goat serum |
| ***Nhej1*** | Non-Homologous End Joining Factor 1) |
| **NSS** | Normal sheep serum |
| **OCT** | Optimal cutting temperature compound |
| **OST** | Osteoblast |
| **p** | phalanges |
| ***Parm1*** | Prostate Androgen-Regulated Mucin-Like Protein 1 |
| **Pb** | Perichondral bone |
| **PBS** | Phosphate-buffered saline |
| **PC1** | Principal component 1 |
| **PC2** | Principal component 2 |
| **PCA** | Principal Component Analysis |
| ***Pde11a*** | phosphodiesterase 11A |
| **PFA** | paraformaldehyde |
| ***Postn*** | Periostin |
| ***Prom1*** | Prominin 1 |
| ***Pth1r*** | Parathyroid Hormone 1 Receptor |

| | |
|---|---|
| ***Pthlh*** | Parathyroid Hormone Like Hormone |
| **r** | radius |
| ***Rhox8*** | Reproductive homeobox 8 |
| **RNA** | Ribonucleic acid |
| **RNA-seq** | RNA sequencing |
| **rpm** | Revolutions per minute |
| **RT-PCR** | Reverse transcription polymerase chain reaction |
| ***Runx1*** | Runt-related transcription factor 1 |
| ***Runx2*** | Runt-related transcription factor 2 |
| ***Runx3*** | Runt-related transcription factor 3 |
| **s** | scapula |
| **SafO** | Safranin O |
| ***Satb1*** | SATB Homeobox 1 |
| ***Satb2*** | SATB Homeobox 2 |
| ***Smad1*** | SMAD Family Member 1 |
| ***Smad5*** | SMAD Family Member 5 |
| ***Sox10*** | SRY-Box Transcription Factor 10 |
| ***Sox4*** | SRY-Box Transcription Factor 4 |
| ***Sox5*** | SRY-Box Transcription Factor 5 |
| ***Sox6*** | SRY-Box Transcription Factor 6 |
| ***Sox8*** | SRY-Box Transcription Factor 8 |
| ***Sox9*** | SRY-Box Transcription Factor 9 |
| ***Sp7*** | Sp7 Transcription Factor |
| ***Sparc*** | secreted protein acidic and rich in cysteine (Osteonectin) |
| ***Spp1*** | Secreted Phosphoprotein 1 (Osteopontin) |
| ***Tcf7*** | Transcription Factor 7 |
| **TF** | Transcription factor |
| **Tri** | Trichrome |
| ***Ttll3*** | Tubulin Tyrosine Ligase Like 3 |
| **u** | ulna |
| **U/mL** | Units per milliliter |
| **Uh** | urohyal |
| ***Vdr*** | Vitamin D Receptor |
| ***Vegfa*** | Vascular endothelial growth factor A |
| ***Wnt11*** | Wnt Family Member 11 |
| ***Wnt5a*** | Wnt Family Member 5A |
| ***Wnt5b*** | Wnt Family Member 5B |
| ***Zbtb16*** | Zinc Finger And BTB Domain Containing 16 |
| **µm** | micrometers |

# CHAPTER 1

## Introduction

# CHAPTER 1. Introduction

The intent of the present chapter is to communicate the overall organization of the thesis in terms of chapter content and context to thesis, and to highlight my contributions in collaborative projects included in each chapter. This thesis is organized in the form of a manuscript-based document. It is composed of seven chapters and two appendices. Since each chapter was treated as a separate manuscript and sections of specific chapters include overlapping information with other chapters, in rare cases the same data and/or figures were included in different chapters to facilitate comparisons. Chapter 2 summarizes basic information on what is currently known about the development and evolution of skeletal cells in order to test the idea that skeletal cell types evolve, just as skeletal morphology does. The information included in this manuscript provides general background to understand Chapters 3, 4, 5, and 6. Chapter 2 has been published in *Frontiers in Genetics*: Gomez-Picos P and Eames BF. On the evolutionary relationship between chondrocytes and osteoblasts. *Front Genet.* 2015 Sep 23;6:297. doi: 10.3389/fgene.2015.00297. I performed an extensive literature review and co-authored with my supervisor Dr. Brian Eames who guided me through this process. In this review, we explain how most of the vertebrate skeleton is composed of three skeletal cell types: immature chondrocytes (IMM), mature chondrocytes (MAT), and osteoblasts (OST). The differentiation of these three skeletal cell types is driven by a gene regulatory network (GRN), which is defined as a set of genes that interact and underlie a given biological process (Davidson and Erwin 2006; Peter and Davidson 2011). In Chapter 2, we explore different scenarios on the evolution of the osteoblast GRN, a vertebrate evolutionary novelty. To do this, we present findings from traditional studies of the fossil record, embryology, and comparative anatomy, as well as modern molecular studies, which clearly show that immature cartilage evolved before mature cartilage or bone. We also propose that in order to reveal evolutionary origins of skeletal cells, particularly the osteoblast, GRNs in different clades need to be revealed and compared using comparative transcriptomics, which allows to quantitatively measure constraint and adaptation within skeletal cell GRNs. Evidence of constraint can be inferred when transcriptomes are conserved among different skeletal cell types or vertebrate clades, whereas adaptation can be seen when transcriptomes show more variation.

Following up on the proposal in the review paper of chapter 2, in chapter 3 we developed a novel experimental model using comparative transcriptomics to analyze *in vivo* GRN organization and interaction globally in skeletal cells of a single species: the mouse embryo. This chapter is based on manuscript that is currently in preparation: Gomez-Picos P, Ovens K, Ashique AM, Whalen M, Eskiw CH, McQuillan I, and Eames BF. Portions of the Gene Regulatory Network driving cartilage and bone formation interact via averaging and synergism during cartilage maturation. Specific aims for this chapter include: (1) a histological analysis of skeletal elements/cell types in the mouse embryo, (2) to capture IMM, MAT, and OST from mouse skeletal tissues, (3) to identify and compare transcriptomic profiles of IMM, MAT, and OST in mice (4) to reveal the GRN driving the differentiation of IMM, MAT, and OST and verify regulatory connections between hallmark candidate genes in these three skeletal cell types, (5) to validate the expression of classic and novel mature cartilage genes, (6) to analyze possible mechanisms of GRN regulation driving the differentiation of MAT, and (7) to functionally validate putative novel transcription factors in mature chondrocytes using ATDC5 cells. Although most of the bioinformatic data was generated in collaboration with our colleagues in the Department of Computer Science, MSc student Katie Ovens and her supervisor Dr. Ian McQuillan, I also contributed to these analyses by performing gene ontology in the gene lists generated. I did an extensive literature review for the introduction and discussion part which allowed me to write most of the manuscript. Furthermore, I applied the knowledge I gained in the 'Gene Regulatory Networks for Development', course that I took at the Marine Biological Laboratories (MBL) in Woods Hole Massachusetts in 2016. I was able to confirm many of the regulatory interactions using published work involving molecular genetic experiments using skeletal cells. Once these interactions were confirmed, I included them in the BioTapestry software and constructed a validated skeletal cell GRN. Laser-capture microdissection (LCM) on mouse skeletal tissues was performed by Dr. Amir Ashique, but I performed all the histological and molecular techniques, such as *in situ* hybridization, DNA digestions and purifications, cell culture experiments, transfections, and immunohistochemistry.

Using skeletal cells *in vivo* we focused on the idea that one GRN directs the differentiation of IMM, MAT, and OST, and different portions of it are active in each cell type (Fig. 1).

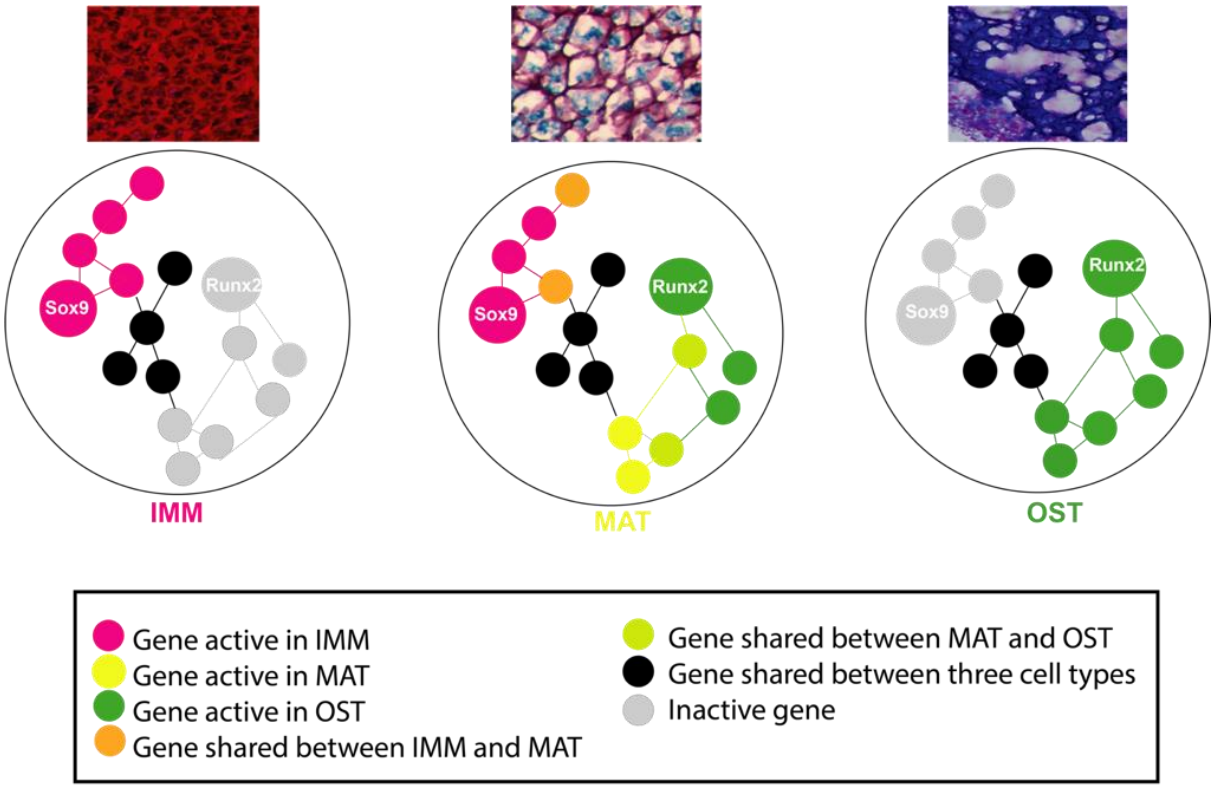



**Figure 1. A GRN drives the differentiation of skeletal cells and distinct portions are active in each cell type.** A *Sox9*-mediated GRN is active in IMM, a *Runx2*-mediated GRN is active in OST, and both programs are active in MAT.

Specifically, we hypothesize that two distinct portions of the GRN directing IMM and OST interact during the formation of MAT. To test this hypothesis, IMM, MAT, and OST were collected from mouse E14.5 skeletal tissues and then mRNA was extracted and processed for RNA-seq in order to characterize transcriptomic profiles of these skeletal cell types. Several findings supported the hypothesis. Gene co-expression network (GCN) analyses revealed that a portion of the GRN drives IMM, a portion of the GRN drives OST, and both transcriptional programs (i.e. portions of the GRN) operate in MAT. Model-based clustering provided insight into regulatory control of MAT. Some genes in MAT had expression levels similar to IMM or OST, suggesting that portions of the GRN active in IMM or OST are independently regulating the expression of distinct genes in MAT. Other genes had expression levels that represented an averaging or synergy between levels in the IMM and OST, suggesting that portions of the GRN directing IMM and OST are interacting in MAT. Besides providing insight into GRN structure and interaction, our approach identified novel genes in MAT, and tested the function of new MAT transcription factors in the mouse chondrogenic

cell line ATDC5.

IMM, MAT, and OST can all have two distinct embryonic origins, mesoderm and neural crest, depending on their location in the body, but the degree of similarity between these two embryonic lineages is still debatable. A great portion of the cranial skeleton such as the pharyngeal skeleton is of neural crest origin, whereas the axial skeleton including rib cage and vertebral column as well as the appendicular limb skeleton derives from the mesoderm. In chapter 4, we present a transcriptomic comparison between chick chondrocytes isolated from the limb (i.e. humerus) and head (i.e. ceratobranchial), which have two distinct embryological origins, mesoderm and neural crest, respectively. The same cells were used for subsequent evolutionary comparisons presented in chapter 6. To test the hypothesis that mesoderm- and neural crest-derived cartilage transcriptomes are conserved, we isolated immature and mature chondrocytes from the chick humerus and ceratobranchial using LCM, and their transcriptomes were compiled and compared. Specific objectives for this chapter include: (1) to make a histological characterization of chick skeletal tissues, (2) to capture IMM and MAT from the chick humerus and ceratobranchial using LCM, and (3) to obtain and compare chondrocyte (IMM and MAT) transcriptomes derived from the mesoderm and the neural crest. First, I performed a histological characterization of chick humerus and ceratobranchial in order to identify skeletal cells of interest (i.e. IMM and MAT). Then, I collected chick IMM and MAT from both the humerus and the ceratobranchial using LCM, extracted and amplified mRNA which was then submitted for RNA-seq. Again, most of the transcriptomic data was generated in collaboration with Katie Ovens, but I performed gene ontology analysis to identify biological processes and genes of interest that allowed me to derive insights into fundamental biological processes. In general, our data shows that there is a high degree of similarity between transcriptomes of mesoderm- and neural crest-derived chondrocytes, but a few differences in gene expression were also identified. The data presented in this chapter are still preliminary and additional analyses are needed for publication purposes.

Regarding OST evolution, in chapter 5, we made an evolutionary comparison between mouse, chick, and gar osteoblasts. Here, we aimed to test the hypothesis that gar osteoblasts express higher levels of 'chondrogenic' genes compared to tetrapod (i.e. mouse and chick) osteoblasts. A manuscript focusing on the osteoblast data is currently in preparation: Gomez-Picos P, Ovens K, Nguyen JKB, McQuillan I, and Eames BF. Comparative transcriptomics reveals the core gene regulatory network driving OST differentiation and identifies higher expression of cartilage genes

in gar bone. Specific aims for this chapter include: (1) to make a histological characterization of mouse, chick, and gar dentary bones in order to identify osteoblasts, (2) to isolate osteoblasts of chick and gar bones using LCM (mouse OST were the same ones used for analyses presented in chapter 3, LCM credit: Dr. Amir Ashique), (3) to reveal the conserved set of genes driving OST differentiation and validate regulatory interactions using published literature, and (4) to make an evolutionary comparison between fish and tetrapod bones using bioinformatic approaches. To fulfill these objectives, I performed histological analyses (i.e. whole mount and section histology) on mouse, chick, and gar bones in order to identify osteoblasts. Then, I isolated osteoblasts from the dentary bone, a homologous element in mouse, chick, and gar (the same cells were used for evolutionary comparisons in chapter 6). Bioinformatic data again was generated in collaboration with Katie Ovens, but using BioTapestry software coupled to gene ontology analyses, I constructed an osteoblast gene regulatory network by validating regulatory connections using molecular published work. In general, our analyses supported the hypothesis. Indeed, gar OST expressed higher levels of some hallmark chondrogenic genes including *Col2a1*, *Sox6*, *Ihh*, *Col10a1*, and *Col9a1* compared to mouse or chick. The results presented in this chapter are at a preliminary stage and more transgenic and molecular experiments are needed in order to be published. Adding a transitional water-to-land animal, such an anuran could help to shed light into these comparisons.

Following up the results in mouse skeletal cells discussed in chapter 3, in chapter 6, we explored the organization and regulation of skeletal cell GRNs in different vertebrate clades in order to provide insights into their evolution. Here, we aimed to test the hypothesis that skeletal cell GRN structure and regulatory control of gene expression in MAT are conserved across vertebrates. Specific objectives for Chapter 6 include: (1) a histological analysis of skeletal elements/tissues during the embryonic development of chick and gar, (2) identification and comparative transcriptomic profiles of IMM, MAT, and OST in mouse, chick, and gar, (3) a comparative GRN structural analysis between homologous skeletal cells in mouse, chick, and gar, (4) a determination of whether mechanisms of GRN regulation in MAT are conserved between the three species analyzed. To fulfill these objectives, transcriptomes of homologous skeletal cell types between mouse (*Mus musculus*), chick (*Gallus gallus*), and gar (*Lepisosteus oculatus*), which represent three of the main vertebrate clades, were compiled and compared (Fig. 2). For these evolutionary comparisons we analyzed osteoblasts isolated from the dentary, a neural crest-derived skeletal element, but depending on the species (i.e. mouse, chick, or gar) chondrocytes were isolated from

two distinct locations, the humerus and/or the ceratobranchial/ceratohyal, which are mesoderm- or neural crest-derived, respectively. Since the mouse does not have large pharyngeal cartilage that matures, we were not able to isolate a homologous element to the ceratobranchial/ceratohyal.

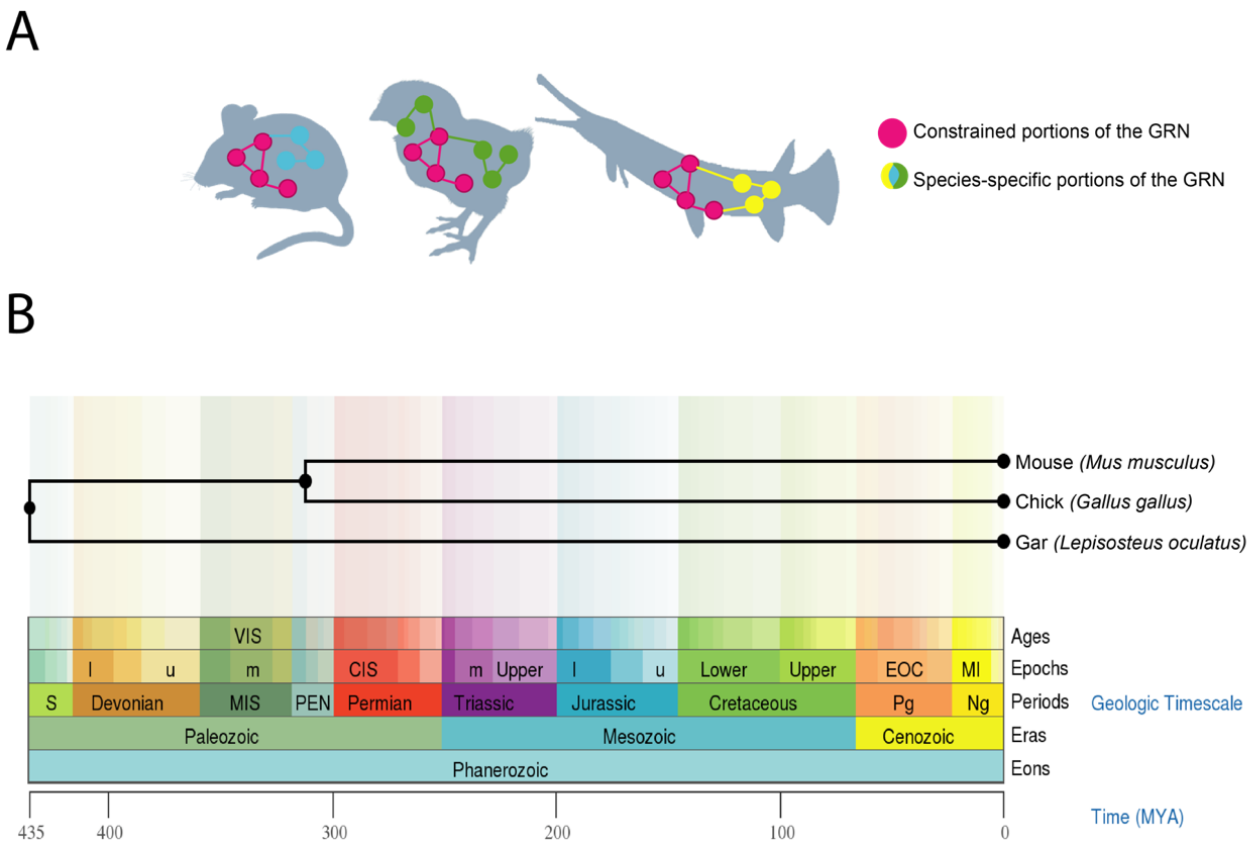



**Figure 2. Identifying conserved and species-specific portions of the skeletal cell GRN could provide insight into evolution.** (A) Mouse, chick, and gar skeletal cell transcriptomes, which represent three of the main phylogenetic clades, were analyzed and compared in this study. (B) Phylogenetic tree showing the time of divergence of different animal clades. While the most recent common ancestor between mouse and chick existed ~312 MYA, the most recent common ancestor between gar, mouse, and chick existed ~430 MYA. The tree was generated in Timetree (http://www.timetree.org/; Kumar et al. 2017). Abbreviations: S, Silurian; VIS, Visean; MIS, Mississipian; PEN, Pennsylvanian; Pg, Paleogene; N, Neogene; l, Lower; u, Upper; m, Middle; CIS, Cisuralian; EOC, Eocene; MI, Miocene.

I dissected and embedded skeletal elements in optimal cutting temperature compound (OCT), including the dentary, humerus, and ceratobranchial from chick and the dentary and ceratohyal for gar, and then collected frozen tissue sections which were used to do histological and molecular experiments. LCM was performed on frozen sections of skeletal tissues in order to collect IMM, MAT, and OST. Once cells were captured, I extracted and amplified RNA, and then submitted it for

RNA-seq. Mouse skeletal cells (IMM, MAT, and OST) were also used for analysis presented in chapter 3, chick chondrocytes (IMM and MAT) were also used for transcriptomic analyses in chapter 4, and mouse, chick and gar OST were also the same ones used for the evolutionary comparisons presented in chapter 5. Moreover, I performed *in situ* hybridizations experiments on conserved mature chondrocyte genes between mouse and chick. Furthermore, I was involved in analyzing the bioinformatic data generated by my colleague Katie Ovens, performed several gene ontology analyses and literature review in order to make a biological interpretation of the data. In general, our results partially supported our hypotheses. GRN structure and regulation is conserved to some extent in skeletal cells of all different species analyzed, but also important differences were revealed. Although analyses in chapter 4 showed that the transcriptomes of chick chondrocytes of distinct embryonic origins are highly conserved, differences in GRN structure were noted when chick neural crest vs mesoderm-derived tissues were compared. GRN structure was also conserved in gar when homologous neural crest-derived skeletal elements were analyzed, suggesting that embryonic origin might influence GRN topology.

In chapter 7, I summarize the findings of this thesis overall, limitations of experiments, and future directions for the present project. Finally, two appendices were included at the end of the thesis. The first appendix summarizes my contributions to a project focused on the development of frog skeletal tissues using *Xenopus tropicalis* as a model organism. During my PhD, I had the opportunity to train my first summer student, Yiwen Liu, who then spent one year at the lab working on her 401 project under my supervision. Together we started the *X. tropicalis* project. I taught her sample collection procedures, histological techniques (whole mount and section histology), cryosectioning and helped her setting up the crossing system. Then, I had the fortune to supervise Jason Nguyen during his Masters project, who greatly improved the crossing system that we had previously established. I taught him sample collection procedures, histological and molecular techniques, including DNA and RNA purification, RNA probe synthesis via *in vitro* transcription, and *in situ* hybridization. I also spent some time training him to do laser capture microdissection in *X.tropicalis* skeletal tissues. His samples have been already submitted for RNA-seq, and we are planning to include these data into our OST evolution paper, in order to make comparisons more robust. The second appendix includes a book chapter recently published: Eames BF, Gomez-Picos P, and Jandzik D (2020). *On the Evolution of Skeletal Cells before and after Neural Crest.* In Eames BF, Meulemans Medeiros D, and Adameyko I (Eds.), Evolving Neural Crest Cells (pp. 185-218).

Boca Raton: CRC Press, https://doi.org/10.1201/b22096. Using the fossil record, comparative, anatomy, and molecular genetics we analyze how skeletal tissues evolved relative to the neural crest, and we discuss different scenarios on the evolution of the GRN driving cartilage and bone. In the case of cartilage, we proposed two different scenarios: (a) cartilage evolved from acellular cartilage and (b) cartilage evolved from chordoid tissue. For bone, we proposed three different scenarios: (a) bone evolved from a deuterostome mineralization GRN, (b) bone evolved from a dentin/enamel GRN, and (c) bone evolved from a mature cartilage GRN. Specifically, I contributed to writing sections VI and VII which discuss these possible scenarios for the evolution of the GRNs driving cartilage and bone and put together figures 64 and 65 illustrating these different possibilities.

## CHAPTER 2:

**Review paper; On the evolutionary relationship between chondrocytes and osteoblasts (published in *Front Genet.* 2015 Sep 23;6:297. doi: 10.3389/fgene.2015.00297)**

Gomez-Picos P and Eames BF

# CHAPTER 2. On the evolutionary relationship between chondrocytes and osteoblasts

## 2.1. Abstract

Vertebrates are the only animals that produce bone, but the molecular genetic basis for this evolutionary novelty remains obscure. Here, we synthesize information from traditional evolutionary and modern molecular genetic studies in order to generate a working hypothesis on the evolution of the gene regulatory network (GRN) underlying bone formation. Since transcription factors are often core components of GRNs (i.e., kernels), we focus our analyses on Sox9 and Runx2. Our argument centers on three skeletal tissues that comprise the majority of the vertebrate skeleton: immature cartilage, mature cartilage, and bone. Immature cartilage is produced during early stages of cartilage differentiation and can persist into adulthood, whereas mature cartilage undergoes additional stages of differentiation, including hypertrophy and mineralization. Functionally, histologically, and embryologically, these three skeletal tissues are very similar, yet unique, suggesting that one might have evolved from another. Traditional studies of the fossil record, comparative anatomy and embryology demonstrate clearly that immature cartilage evolved before mature cartilage or bone. Modern molecular approaches show that the GRNs regulating differentiation of these three skeletal cell fates are similar, yet unique, just like the functional and histological features of the tissues themselves. Intriguingly, the Sox9 GRN driving cartilage formation appears to be dominant to the Runx2 GRN of bone. Emphasizing an embryological and evolutionary transcriptomic view, we hypothesize that the Runx2 GRN underlying bone formation was co-opted from mature cartilage. We discuss how modern molecular genetic experiments, such as comparative transcriptomics, can test this hypothesis directly, meanwhile permitting levels of constraint and adaptation to be evaluated quantitatively. Therefore, comparative transcriptomics may revolutionize understanding of not only the clade-specific evolution of skeletal cells, but also the generation of evolutionary novelties, providing a modern paradigm for the evolutionary process.

## 2.2. Introduction: cartilage and bone might share an evolutionary history

Most of evolutionary theory has focused on studies of morphological change (morphogenesis) among taxa, but the formation of tissue types (histogenesis) also can evolve in clade-specific manners. Therefore, we focus our attention on a relatively understudied subject of evolutionary research: the evolution of histogenesis. A classic problem in evolutionary theory is to explain novelties, or traits with no clear ancestral antecedent (Shubin 2002; Moczek 2008; Wagner and Lynch 2010). For example, vertebrates are the only animals that produce bone, but so far, the molecular genetic basis for this evolutionary novelty remains obscure. Here, we synthesize information from traditional evolutionary and modern molecular studies in order to generate a working hypothesis on the evolution of the genetic system underlying bone formation. Many studies argue that bone evolved from dentine (Kawasaki et al. 2004; Wagner and Aspenberg 2011). However using molecular genetic and embryological arguments that favor gradualism over saltationism (Gould 2002), we hypothesize that bone (and perhaps all mineralizing tissues, such as dentine) appeared during evolution by co-opting a gene regulatory network (GRN) that was under prior natural selection to mineralize cartilage. In order to present an argument for skeletal tissue development and evolution over the past 500 million years, we make some generalizations that may trouble some readers, of whom we ask their indulgence, hoping that such generalizations help to reveal broader trends during the evolution of skeletal tissues.

An introductory look at the similarities and differences among cartilage and bone suggests that the underlying GRNs may be related. Cartilage and bone are specialized connective tissues that provide form and structural support to the body, protect vital organs, and play a crucial role in locomotion through muscle attachments (Gray 2001). Despite these similarities, they also have distinct functions (Fig. 3). Cartilage typically offers a flexible structure to support soft tissues and also to serve as a load-bearing surface between bones. On the other hand, bone is a hard, rigid structure that protects vital organs and acts as a storage site for minerals, such as calcium and phosphorus (Smith and Hall 1990a; Volkmann and Baluska 2006). Also unlike cartilage, which has almost no capacity for regeneration, bone is a highly dynamic structure that undergoes constant remodeling, preserving bone strength and regulating calcium homeostasis (Datta et al. 2008). Perhaps related to regenerative capacity, these tissues differ in vascularity. Bone is highly vascularized, but cartilage typically is avascular. However important exceptions to cartilage

vascularization occur. Mature cartilage in tetrapods often is invaded by vasculature as it degrades, creating the marrow cavity (Johnson 1980; Roach 1997; Stricker et al. 2002; Ortega et al. 2004; Moriishi et al. 2005), and even immature cartilage is highly vascularized near articulating surfaces in some avian and mammalian species (Ytrehus et al. 2004; Blumer et al. 2005). When cartilage extracellular matrix (ECM) undergoes mineralization, its functions change. In some vertebrates, such as sharks, mineralized cartilage can serve as the major rigid structural support for the body, meanwhile providing a mineral reservoir (Daniel 1926; Kemp and Westrin 1979; Eames et al. 2007). In most extant vertebrates, however, mineralized cartilage mainly serves as a scaffold during endochondral ossification, outlined below.

During embryonic development, cartilage and bone formation share many features (Fig. 3). Both cartilage and bone are differentiated from common mesenchymal (osteochondral) progenitor cells (Fang and Hall 1997; Day et al. 2005; Hill et al. 2005). Both cartilage and bone initiate overt differentiation by aggregating mesenchymal cells into condensations, which can go on directly to secrete cartilage- or bone-specific matrix (Hall and Miyake 1995; Hall and Miyake 2000; Kronenberg 2003; Day et al. 2005). However a unique feature of bone formation is that, in addition to differentiating directly from an osteogenic condensation (intramembranous ossification), bone also forms on a pre-existing cartilage template (endochondral ossification). Endochondral ossification actually involves the formation of the three skeletal tissues that comprise the majority of the extant vertebrate skeleton: immature cartilage, mature cartilage and bone (Eames et al. 2003; Eames and Helms 2004; Eames et al. 2004). Some cartilage remains throughout development at the growth plates and throughout life at articular surfaces (we term this immature cartilage). Most of the cartilage produced during endochondral ossification, however, undergoes a series of changes, termed maturation (thus the terms immature vs. mature cartilage). In most vertebrates, cartilage maturation involves cell hypertrophy, matrix mineralization, cell death, and matrix degradation (Leboy et al. 1988; Hatori et al. 1995; Takeda et al. 2001; Miura et al. 2008). Although exceptions exist (Thorogood 1988; Hirasawa and Kuratani 2015), endochondral ossification typically gives rise to the bones of the endoskeleton, such as the chondrocranium or limb skeleton, whereas intramembranous ossification produces the exoskeleton, such as lateral plates in teleosts or the calvarium (Smith and Hall 1990).

Histologically, immature cartilage, mature cartilage, and bone are very similar, yet each also has some unique features (Fig. 3).

| | Similarities | Differences | |
|---|---|---|---|
| Functional | Connective tissues that enable: 1) Form and structural support 2) Movement 3) Protection of vital organs | Flexible structure; often serves as a load-bearing surface between bones | immature cartilage |
| | | Often serves a transient function; in chondrichthyans: rigid structure, mineral reservoir | mature cartilage |
| | | Rigid structure, mineral reservoir | bone |
| Embryonic | Share mesenchymal cell lineage (osteochondroprogenitor); initiate formation with mesenchymal condensation, then abundant ECM secretion | Persists through development; forms independently and during endochondral ossification | immature cartilage |
| | | Often cell death and matrix degradation; forms during endochondral ossification | mature cartilage |
| | | Matrix remodelling; forms during intramembranous and endochondral ossification | bone |
| Histological | Specialized cells embedded in abundant extracellular matrix (ECM) consisting of fibrous proteins and sulfated proteoglycans | Chondrocyte embedded in avascular, Col2/Aggrecan ECM | immature cartilage |
| | | Hypertrophic chondrocyte embedded in avascular, mineralized Col2/Col10/Aggrecan ECM | mature cartilage |
| | | Osteoblast/osteocyte embedded in vascular, Col1 ECM | bone |

**Figure 3. Similarities and differences among immature cartilage, mature cartilage, and bone suggest that these three skeletal tissues share an evolutionary story.**

All three skeletal tissues are composed of cells embedded in an ECM that is rich in collagens and proteoglycans (Hardingham 1981; Eames et al. 2003a; Eames and Helms 2004; Eames et al. 2004; Gentili and Cancedda 2009). Immature cartilage is formed by chondrocytes that deposit a network of loose collagen fibers and a rich substance of proteoglycans, whereas chondrocytes of mature cartilage alter the immature cartilage ECM by decreasing its proteoglycan sulfation and mineralizing it (Lohmander and Hjerpe 1975; Buckwalter et al. 1987; Bayliss et al. 1999). The

requirement of proteoglycan degradation for mature cartilage ECM mineralization is debated (Hirschman and Dziewiatkowski 1966; Granda and Posner 1971; Poole et al. 1982; Campo and Romano 1986). Bone is formed by osteoblasts that produce an ECM of tightly wound and highly cross-linked collagen fibers, and bone ECM has lower levels of proteoglycans than cartilage (Gentili and Cancedda 2009). As a result of these collagen and proteoglycan concentrations, these three skeletal tissues have overlapping and unique histological staining patterns. High concentrations of sulfated proteoglycans cause immature cartilage to stain with Alcian blue and Safranin O (by comparison, mature cartilage and bone bind these dyes with decreasing intensity, respectively). The tightly wound collagen fibers of bone stain with Direct red and Aniline blue (by comparison, loose collagen fibers of cartilage matrix bind these dyes with lower intensity; Villanueva et al. 1983; Hall 1986; Eames and Helms 2004; Eames et al. 2004; Eames et al. 2007). Alizarin red can stain mineralized tissues of mature cartilage and bone (Hogg 1982; Kirsch et al. 1997; Eames and Helms 2004; Eames et al. 2007).

Immature cartilage, mature cartilage, and bone have overlapping, but distinct, gene and protein expression profiles; Fig. 3). All these skeletal tissues express Collagen 11 and the proteoglycans Biglycan and Decorin (Li et al. 1998; Knudson and Knudson 2001; Rees et al. 2001; Roughley 2006). Immature cartilage expresses high levels of Collagens 2 and 9, as well as the proteoglycans Aggrecan, Fibromodulin, and Epiphycan, which distribute growth factors and provide swelling pressure due to water attraction (Yanagishita 1993; Lefebvre et al. 1997; Lefebvre and de Crombrugghe 1998; Watanabe et al. 1998; Liu et al. 2000). Mature cartilage has reduced expression of these same collagens and proteoglycans, while also expressing high levels of Collagen 10 (Orth et al. 1996; Eames et al. 2004; Talwar et al. 2006). In contrast to both types of cartilage, bone expresses high levels of Collagen 1(Yasui et al. 1984; Kream et al. 1995). Interestingly (and central to the argument of this review), both mature cartilage and bone share expression of genes not expressed in immature cartilage, including *Sp7* (formerly called Osterix), Matrix metallopeptidase 13 and Indian hedgehog (Vortkamp et al. 1996a; Inada et al. 1999; Neuhold et al. 2001; Zaragoza et al. 2006; Abzhanov et al. 2007a; Mak et al. 2008; Huycke et al. 2012; Nishimura et al. 2012; Weng and Su 2013). In fact, very few genes expressed in bone are not expressed in mature cartilage, and this list of genes decreases further when comparisons among mature cartilage and bone are carried out in actinopterygians (Eames et al. 2012). Multiple genes associated with matrix mineralization are expressed in both mature cartilage and bone, such as Alkaline phosphatase, liver/bone/kidney

(*Alpl*, formerly called Tissue-nonspecific alkaline phosphatase), Secreted phosphoprotein 1 (*Spp1*, formerly called Osteopontin or Bone sialoprotein), Secreted protein, acidic, cysteine-rich (*Sparc*, formerly called Osteonectin), and Bone gamma-carboxyglutamate protein (*Bglap*, formerly called Osteocalcin; Termine et al. 1981b; Pacifici et al. 1990; Chen et al. 1991; Bonucci et al. 1992; McKee et al. 1992; Mundlos et al. 1992; Nakase et al. 1994; Roach 1999; Sasaki et al. 2000).

Currently, the evolutionary relationship among skeletal tissues is unclear, but the similarities highlighted above suggest that immature cartilage, mature cartilage, and bone share an evolutionary history. From a molecular genetic perspective, these observations lead to the hypothesis that the GRNs governing the formation of these three skeletal tissues (in particular, the differentiation of three skeletal cell types) also share an evolutionary history. Indeed, the many varieties of skeletal tissues intermediate between cartilage and bone observed in extant and fossil vertebrates may owe their existence to this shared history (Benjamin 1990; Benjamin and Ralphs 1991; Benjamin et al. 1992; Mizoguchi et al. 1997; Hall 2005b; Witten et al. 2010). In this review, we explore this hypothesis using traditional evolutionary and modern molecular genetic studies. We are not focussing on the exact anatomical location of a tissue, given that once the GRN regulating formation of that skeletal tissue is established in the genome, any cell in the body can co-opt its expression. Traditional studies have provided insight into the evolutionary relationship among skeletal tissues, since they demonstrate that immature cartilage originated first during phylogeny (Mallatt and Chen 2003; Rychel et al. 2006). Interestingly, modern molecular genetic studies reveal that two GRNs dictate the formation of these three skeletal tissues (Bi et al. 1999; Inada et al. 1999; Eames et al. 2004; Hattori et al. 2010; Leung et al. 2011), and also that the GRN underlying cartilage formation is dominant to that of bone (Eames et al. 2004; Zhou et al. 2006). We expand upon this finding using an argument based on the relative parsimony of gradualism versus saltationism to hypothesize that bone evolved from a cartilage maturation program. In closing, we discuss how comparative transcriptomics will enhance dramatically our ability to test hypotheses on the evolution of the GRNs underlying cartilage and bone formation.

### 2.3. GRN underlying immature cartilage formation evolved first

Traditional studies, such as the fossil record, comparative anatomy, and embryology, demonstrate that the first skeletal tissue to evolve was immature cartilage (Fig. 4). The fossil record reveals a great diversity of mineralized tissues about 500 million years ago (Mya; Janvier 1996a; Donoghue and Sansom 2002; Donoghue et al. 2006), suggesting that GRNs of skeletal histogenesis were undergoing an adaptive radiation. So which skeletal tissue appeared first in the fossil record? This question is complicated by the facts that currently discovered fossils may represent a biased fraction of ancestral tissues, and that non-mineralized, lightly mineralized, or transiently mineralized tissues likely are not preserved well in the fossil record. Despite these limitations, however the oldest skeletal tissue in the fossil record is unmineralized cartilage in the chordate fossil *Haikouella* from 530 Mya (Fig. 4A; Mallatt and Chen 2003). Many specimens preserving soft tissues of this incredibly important fossil have been found, but they appear to be represented only in a small region of the Yunnan province in China (Chen et al. 1999), reflecting potential bias in the fossil record.

Bone and mature cartilage appeared much later than immature cartilage in the fossil record (Fig. 4A). Conodonts, a group of agnathans (jawless vertebrate fish), are the earliest (515 Mya) known fossils with a mineralized skeleton, characterized by pharyngeal tooth-like elements composed of tissues that were bone-like, enamel-like, and mineralized cartilage-like (Sansom et al. 1992). However subsequent analyses of conodont fossils refuted the conclusion that bone or mineralized cartilage was present in these primitive jawless fish, instead attributing the first appearance of bone in the fossil record to the exoskeleton of pteraspidormorphi (480 Mya), a group of armored agnathans (Janvier 1996a; Donoghue 1998; Donoghue et al. 2006).

Interestingly, some pteraspidomorph species (e.g., eriptychiids and arandaspids) and other, primitive fossil fish show traces of both mineralized cartilage and bone in their endoskeleton (Janvier 1996a; Zhang et al. 2009). Also, fossils of the ancestral vertebrate *Palaeospondylus gunni* (385 Mya) reveal an entire adult skeleton composed of hypertrophic, mineralized cartilage, while bone is completely absent (Johanson et al. 2010). Despite these findings, the current fossil record generally suggests that bone preceded mineralized cartilage (Smith and Hall 1990a; Janvier 1997; Donoghue et al. 2006), although the molecular genetic and embryological arguments of this review call into question the accuracy of this conclusion. What is clear from the fossil record is that unmineralized

cartilage was the first skeletal tissue to appear leading to the evolution of vertebrates (Northcutt and Gans 1983; Smith and Hall 1990).

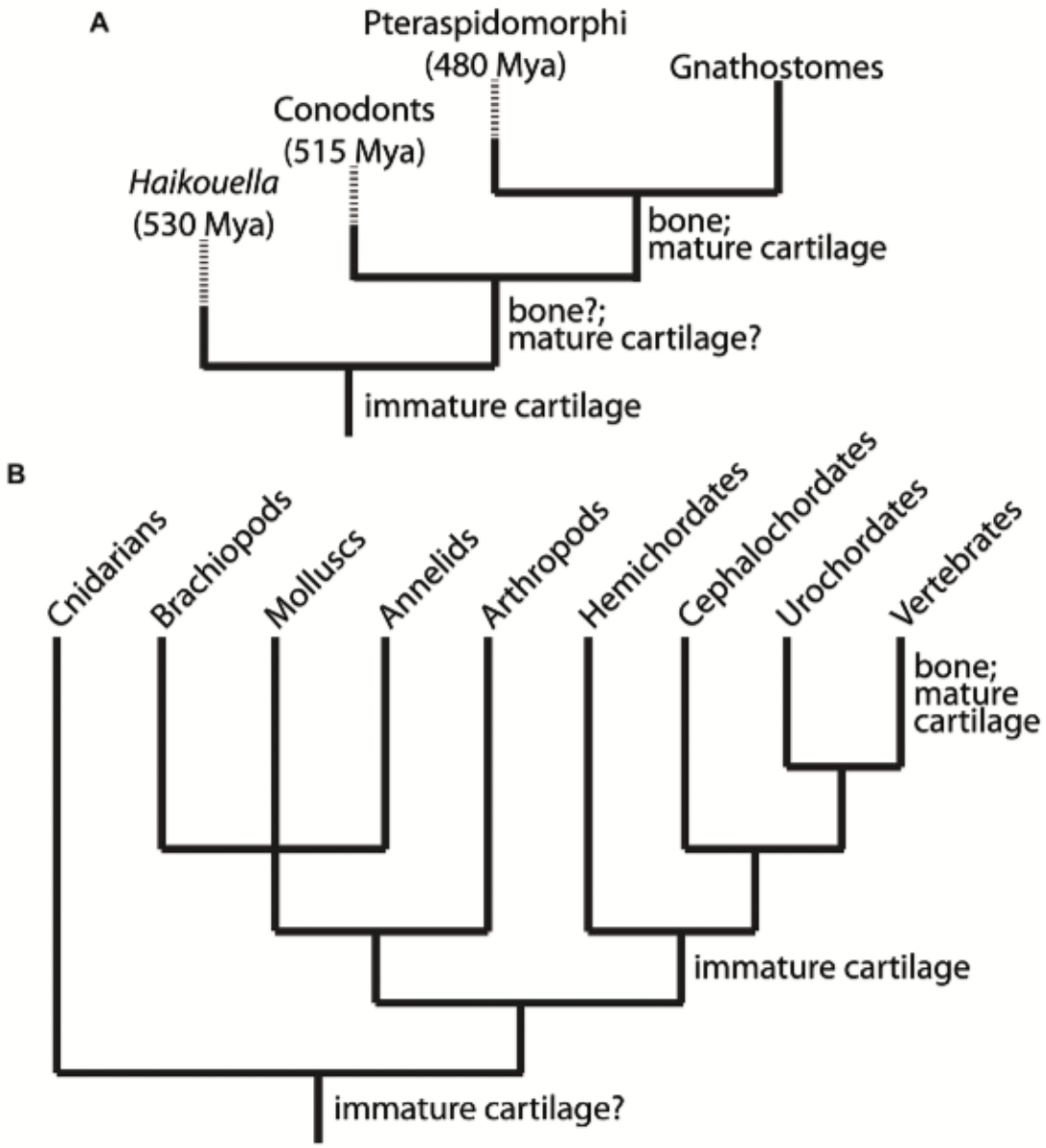



**Figure 4. Clues to the evolutionary relationship between the chondrocyte and osteoblast emerge from analyses of the fossil record and comparative anatomy.** (A) Appearance of immature cartilage, mature cartilage, and bone from available fossil record. These data indicate clearly that immature cartilage appeared first, then mature cartilage and bone. By extension, the chondrocyte preceded the osteoblast during evolution. (B) Extant taxa with at least one species containing cartilage or cartilage-like tissues, which are non-mineralized outside of vertebrates. These data suggest that a GRN driving differentiation of an immature chondrocyte evolved first, and then became established in the genome of chordates (along with the notochord, a cartilage-like tissue). Subsequently, this GRN was modified by another GRN that drove differentiation of a mature chondrocyte (and osteoblast) within vertebrates. Branch lengths in trees are arbitrary; dashed lines indicate extinct taxa.

Comparative anatomy also supports the notion that immature cartilage was the first skeletal tissue to evolve, because immature cartilage is distributed in a broader range of taxonomic lineages than mature cartilage or bone (Fig. 4B). Immature cartilage appears in both vertebrate and non-vertebrate species, whereas mature cartilage and bone are shared, derived traits of vertebrates only (Cole and Hall 2004b; Rychel et al. 2006). In a seminal study by Cole and Hall (2004), cartilage was demonstrated in a variety of taxonomically distinct invertebrates, such as polychaetes, arthropods, and molluscs. Reflecting the different evolutionary histories of immature and mature cartilage, cartilage in any invertebrate lineage, and also in extant agnathans, is unmineralized (Cole and Hall 2004b; Hall 2005b). The finding that lamprey cartilage can mineralize in vitro suggests that early agnathans may have possessed mineralized cartilage and these mineralization programs were repressed in cyclostomes (Langille and Hall 1993a).

The taxonomic distribution of cartilage suggests that the ancestor of vertebrates, cephalochordates, and hemichordates had an ability to make immature cartilage (Fig. 4B). In fact, the deuterostome ancestor was proposed to be a benthic worm with cartilaginous gill slits (Rychel et al. 2006). Homology between invertebrate and vertebrate cartilages is supported by biochemical and histological analyses, which demonstrate high amounts of fibrous proteins and mucopolysaccharides (Cole and Hall 2004b; Cole 2011). In fact, recent studies have shown that the cirri in amphioxus share many histological and molecular features with vertebrate immature cartilage (Kaneto and Wada 2011; Jandzik et al. 2015). However homology between deuterostome and protostome cartilage is still uncertain and must be confirmed by modern molecular analyses, including examination of gene expression patterns, GRN architectures, and GRN regulation. The ECM of hemichordate skeletal tissues may show features of both cartilage and bone (Cole and Hall 2004b), supporting the notion that these two tissues share an evolutionary history. Mineralized cartilage and bone, however, are only found in extant gnathostomes (Fig. 4). These comparative anatomy analyses suggest that immature cartilage evolved before mature cartilage and bone.

Final support for the idea that cartilage arose earlier in evolution than mature cartilage and bone comes from comparative embryology. While the Biogenetic Law of Ernst Haeckel definitely has its theoretical problems (Haeckel 1866), a general correlation (recapitulation) between the timing of events during ontogeny with events during phylogeny is undeniable. Indeed, many early evolutionary biologists assumed this to be true (Gould 2002). In this context, it is interesting to note that immature cartilage is the first skeletal tissue to undergo histogenesis during embryonic

development, while cartilage maturation and bone formation are later events. The relative timing of cartilage maturation to bone formation, on the other hand, appears to vary among vertebrate taxa (Mori-Akiyama et al. 2003; Eames et al. 2004; Moriishi et al. 2005). While such relationships between the timing of developmental events have been argued to reflect simply the increasing complexity of ontogeny during phylogeny (Wallace 1997), we believe that this issue, which has been debated for 100s of years, remains unresolved.

To sum up traditional studies of the fossil record, comparative anatomy, and embryology, the ability to make immature cartilage predates the ability to make mature cartilage or bone during evolution. Therefore, from a molecular genetic perspective, the GRN governing chondrocyte differentiation clearly appeared prior to that of the osteoblast. However traditional approaches are still unclear whether mature cartilage or bone appeared next during evolution. With hopes that modern molecular and embryological analyses can shed light into the evolutionary origins of the vertebrate skeleton, we next discuss how the GRNs underlying the formation of immature cartilage, mature cartilage, and bone are organized.

### 2.4. SOX9 GRN is dominant to the RUNX2 GRN

Skeletal histogenesis is governed by complex sets of genes, largely controlled by central transcription factors that are responsible for determining cell fate decisions (Eames et al. 2003; Kronenberg 2003; Karsenty et al. 2009). Molecular genetic experiments demonstrate that the transcription factors *Sox9* and *Runx2* are the "master regulatory genes" of skeletal histogenesis. *Sox9* and *Runx2* expression patterns during mesenchymal condensation predict whether osteochondroprogenitor cells differentiate into immature cartilage, mature cartilage, or bone (Eames and Helms 2004; Eames et al. 2004). Loss of *Sox9* function abrogated immature and mature cartilage formation, whereas *Runx2* loss of function blocked mature cartilage and bone formation (Hoshi et al. 1999; Inada et al. 1999; Kim et al. 1999; Enomoto et al. 2000). In gain-of-function experiments, *Sox9* mis-expression induced ectopic cartilage formation, whereas *Runx2* mis-expression induced ectopic mature cartilage and bone formation (Eames et al. 2004). These and other experiments show clearly that a SOX9 GRN regulates immature cartilage formation, a RUNX2 GRN drives bone formation, and a combination of SOX9 and RUNX2 GRNs produce mature cartilage (Fig. 5). We

emphasize the relevance of these transcription factors to the evolution of GRNs underlying skeletal histogenesis, since conserved, core components of GRNs (i.e., kernels) are often transcription factors (Levine and Davidson 2005; Davidson and Erwin 2006).

Expression studies of skeletal tissues in a range of organisms suggest an ancestral interaction between SOX and RUNX GRNs. *Runx2*, along with its related family members, *Runx1* and *3*, derive from gnathostome duplications of an ancestral *Runx*, while agnathan *Runx* genes may have undergone an independent duplication (Meulemans and Bronner-Fraser 2007; Hecht et al. 2008; Cattell et al. 2011; Kaneto and Wada 2011; Nah et al. 2014a). *Sox9*, along with its related family members, *Sox8* and *10*, derive from duplications to the ancestral *SoxE*, while agnathan *SoxE* genes may have undergone an independent duplication (Meulemans and Bronner-Fraser 2007; Ohtani et al. 2008a; Yu et al. 2008; Cattell et al. 2011; Uy et al. 2012; Jandzik et al. 2015). *Runx* and *SoxE* orthologs are expressed in cartilage of amphioxus, lamprey, and hagfish, suggesting that the gene ancestral to *Runx2* primitively functioned with the gene ancestral to *Sox9* in early cartilage formation (Hecht et al. 2008; Wada 2010; Kaneto and Wada 2011). Notably, these animals do not have bone, and they do not mineralize their skeletons. Interestingly, the amphioxus cirral skeleton shows features of both cartilage and bone, suggesting that this ancient skeleton might have diverged to form cellular cartilage and bone of vertebrates (Kaneto and Wada 2011). We argue that evaluating the interactions between SOX9 and RUNX2 GRNs leads to a novel hypothesis for the evolution of bone.

Many studies in mammals and chick demonstrate that the SOX9 GRN is at least partially dominant to the RUNX2 GRN. First, co-expression of *Sox9* and *Runx2* typically causes cartilage formation, not bone (Eames and Helms 2004; Eames et al. 2004). Second, ectopic expression of *Sox9* in *Runx2*-expressing cells of developing bone (achieved either normally during secondary cartilage formation or experimentally using *Sox9* mis-expression) diverts the cells to make cartilage, whereas ectopic *Runx2* expression in *Sox9*-expressing cells of developing cartilage does not divert them to make bone (Eames et al. 2004). Third, *Sox9* expression needs to be down-regulated in order for the full *Runx2*-dependent cartilage maturation program to be expressed (Akiyama et al. 2002; Eames et al. 2004). Fourth, *Sox9* over-expression can inhibit Runx2 expression (Eames et al. 2004). Finally, and most conclusively, *Sox9* directly binds to *Runx2*, inhibits its transcriptional activity, and increases ubiquitin-mediated degradation of *Runx2* (Zhou et al. 2006; Cheng and Genever 2010).

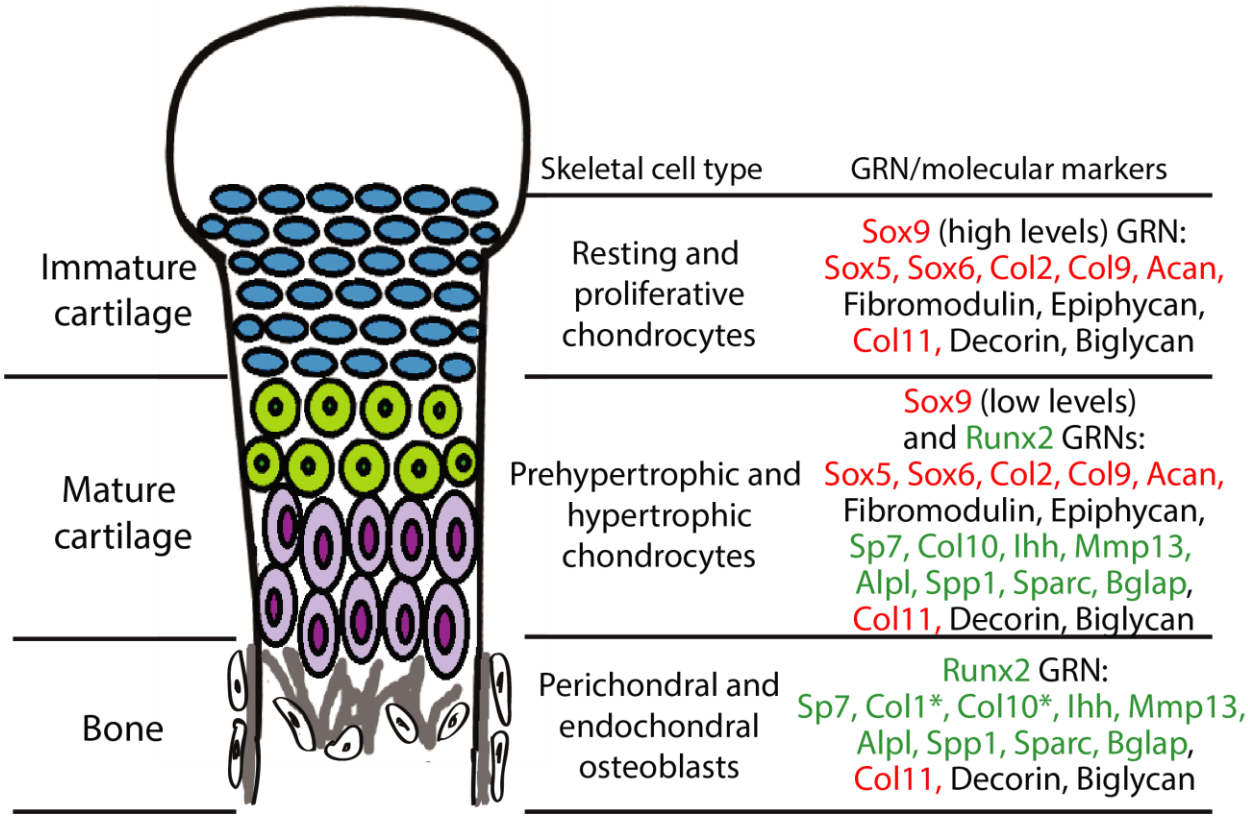



**Figure 5. During endochondral ossification, immature cartilage, mature cartilage, and bone differentiate under the control of SOX9 and RUNX2 GRNs.** Chondrocytes of immature cartilage, termed resting and proliferative chondrocytes during endochondral ossification, express high levels of genes in the SOX9 GRN. Genes known to be under direct transcriptional control of SOX9 or RUNX2 are highlighted in red or green text, respectively. Chondrocytes of mature cartilage, termed prehypertrophic and hypertrophic chondrocytes during endochondral ossification, express low levels of genes in the SOX9 GRN and also genes in the RUNX2 GRN. Osteoblasts in perichondral and endochondral bone during endochondral ossification express genes in the RUNX2 GRN. ∗*Col1* is one of the only genes expressed in osteoblasts that is not expressed in mature chondrocytes; *Col10* expression in osteoblasts is high only in some vertebrates. *Col11*, *Decorin*, and *Biglycan* are expressed in all three of these skeletal cell types. Similar gene expression patterns are seen in immature cartilage, mature cartilage, and bone developing in the articular surface (not shown).

Given evidence that the SOX9 GRN can dominate the RUNX2 GRN, the formation of mature cartilage during endochondral ossification, which requires both *Sox9* and *Runx2*, must be regulated exquisitely (Fig. 5). During early stages, both *Sox9* and *Runx2* are co-expressed in mesenchymal condensations (Akiyama et al. 2002; Eames and Helms 2004; Eames et al. 2004; Zhou et al. 2006), so *Sox9* must exert a dominant inhibitory effect over *Runx2* in order to produce immature cartilage. Later, *Sox9* is down-regulated and *Runx2* activity increases, triggering cartilage

maturation (Eames et al. 2004; Yoshida et al. 2004; Hattori et al. 2010). In fact, *Sox9* down-regulation is a crucial step for mature cartilage formation (Hattori et al. 2010). Despite this down-regulation, a role for *Sox9* in very late stages of cartilage maturation also has been revealed (Ikegami et al. 2011; Dy et al. 2012). One study even suggests that *Runx2* can inhibit *Sox9* activity (Cheng and Genever 2010), illustrating that complex feedback mechanisms are in place to achieve the appropriate relative levels of *Sox9* and *Runx2* activity. In summary, the preponderance of published literature on molecular genetics demonstrates that *Sox9* has dominant effects over *Runx2*, and we extend this conclusion to generate a new hypothesis on the evolution of bone.

### 2.5. Bone evolved from mature cartilage

Combining evidence from traditional and modern studies, we hypothesize that the GRN underlying bone formation evolved from a GRN underlying mature cartilage formation (Fig. 6). Functional, histological, embryological, and molecular similarities among immature cartilage, mature cartilage, and bone suggest that these tissues may share an evolutionary history (Fig. 3). The fossil record, comparative anatomy, and embryology demonstrate that immature cartilage evolved first (Fig. 4). When combined with molecular genetic data (Fig. 5), this means that the first evolved skeletal GRN was dominated by the gene ancestral to *Sox9*, driving immature cartilage formation. This GRN likely involved genes ancestral to *Runx2* in early phylogenetic (and ontogenetic) stages. In gnathostomes, a RUNX2 GRN drives formation of both mature cartilage and bone (Fig. 5), but how did this novel GRN evolve to produce these novel skeletal tissues?

We propose that immature cartilage provided a structural and molecular "buffer" for the gradual development of this novel, RUNX2 GRN. The structural buffering effect refers to the fact that immature cartilage already had a functional role as a skeletal tissue, allowing more freedom for the evolving RUNX2 GRN to develop new functions that simply modify a pre-existing skeletal tissue in a gradual, step-wise fashion. The molecular buffering effect refers to the partial dominance of the SOX9 GRN, which might have shielded to some extent the evolving RUNX2 GRN from natural selection. This concept recalls the principle of "weak linkage," which contributes to

evolvability by reducing the cost of generating variation (Kirschner and Gerhart 1998; Gerhart and Kirschner 2007).

We argue that these putative buffering effects provide a more parsimonious account for the gradual evolution of bone from mature cartilage than the alternative, which depends upon de novo establishment of bone in a more saltationist fashion (Fig. 6). If bone had evolved before mature cartilage, then the RUNX2 GRN would have been under much stronger natural selection than if it had been buffered by immature cartilage. Arguments that bone evolved from dentine suffer from the same limitations: how did dentine and its GRN appear? A new GRN appearing simultaneously with a completely new skeletal tissue, while possible, seems a less likely evolutionary scenario than the gradual establishment of the RUNX2 GRN during evolution of mature cartilage. Assembling a GRN driving bone formation de novo appears to depend upon saltationist genetic mechanisms, such as large-scale genomic changes or small genetic effects acting early in development. Regarding the latter possibility, chondrocytes and osteoblasts are known to share a relatively late embryonic progenitor (Day et al. 2005). Therefore, the former, "macromutational" saltationist mechanism, favored by Goldschmidt (Goldschmidt 1982), would have to have operated in the de novo appearance of the osteoblast. Even saltationists granted that gradualism is the more common evolutionary mechanism (Gould 2002). Therefore, based on the relative parsimony and abundance of gradualism versus saltationism, we favor a model in which the RUNX2 GRN evolved within immature cartilage to produce mature cartilage, and then a different mesenchymal (non-chondrogenic) cell population co-opted this GRN, producing the world's first example of bone formation (Fig. 6B).

The hypothesis that bone evolved from mature cartilage also is consistent with a variety of other observations on skeletal tissues (Fisher and Franz-Odendaal 2012). During evolution, the features of mature cartilage seen in various vertebrate taxa did not appear at the same time (Hall 1975; Smith and Hall 1990a). Hypertrophy and mineralization occurred first, followed by cartilage matrix degradation, replacement by fat and endochondral bone deposition, and finally, invasion by the vasculature (in tetrapods). These findings suggest that cartilage maturation is a highly evolvable process. Also, the progression from immature cartilage to mature cartilage to bone during evolution is mimicked during endochondral ossification. Recently, cell lineage analyses suggest that some cells that express immature cartilage genes go on to express mature cartilage genes, and finally they express bone genes, effectively transitioning from an immature chondrocyte to a mature chondrocyte

to an osteoblast (Hammond and Schulte-Merker 2009; Zhou et al. 2014; Park et al. 2015). Finally, gene expression patterns appear to overlap much more when comparing mature cartilage to bone in actinopterygians, such as teleosts, than in sarcopterygians, such as tetrapods (Eames et al. 2012). This may reflect differential retention of molecular signatures of the evolutionary history between mature cartilage and bone in earlier diverging versus later diverging vertebrates.

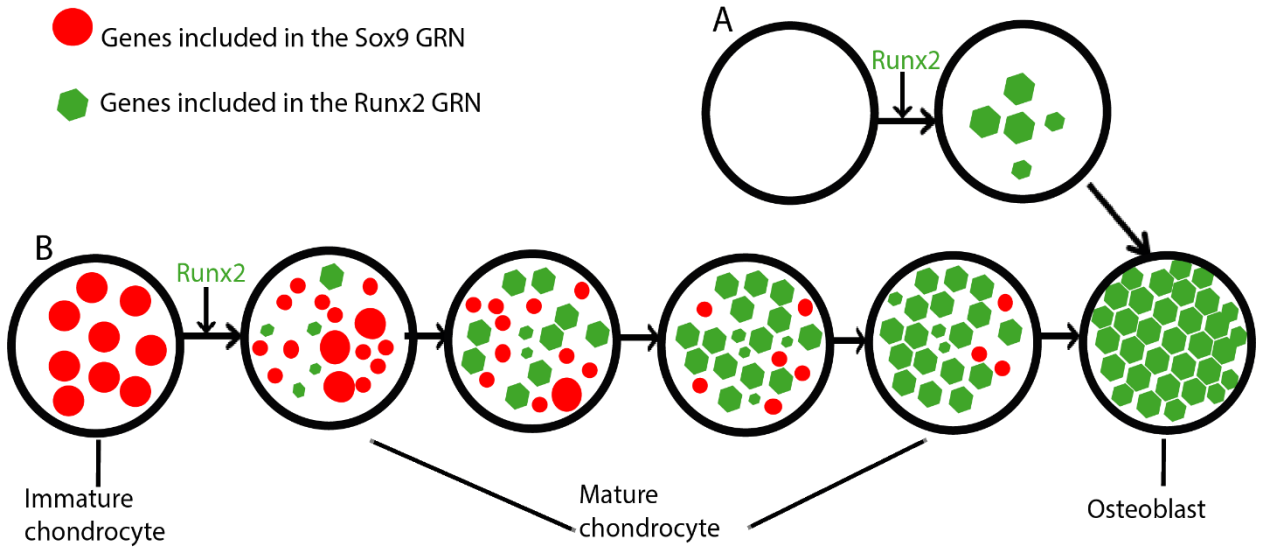



**Figure 6. Differing models for the appearance of the GRN driving osteoblast formation.** (A) In this scenario, the osteoblast (and the RUNX2 GRN that drives its formation) appeared de novo, independent of the chondrocyte. This model is consistent with saltational evolution, in which large-scale genomic changes may facilitate the evolution of novelty over short periods of geologic time. (B) In an alternative scenario, the osteoblast appeared after a series of step-wise additions to the mature chondrocyte (and thus the RUNX2 GRN that drives its formation). After establishment of the RUNX2 GRN in mature chondrocytes, the osteoblast appeared when another population of cells co-opted the RUNX2 GRN. This model is consistent with gradual evolution, in which a series of small changes over geologic time may facilitate the evolution of novelty. The size of the circles and polygons represent relative levels of up- or down-regulation of genes in the respective GRNs (see text for discussion of interactions between SOX9 and RUNX2 GRNs).

## 2.6. Comparative transcriptomics: a novel approach to solve Evo-Devo issues

Identification of homologous tissue types among different taxonomic lineages using histology and cell morphology has enabled evolutionary studies of histogenesis, but modern molecular techniques will dramatically expand this field. Traditionally, comparative anatomy established homologies at the levels of organs, tissues, and cells. Homology among cartilage-like tissues can be relatively clear for closely related species but can prove more difficult when comparing distant clades, where clade-specific differences can obscure homology. For example, histological features, such as cellularity of a tissue, may confuse homology designation; cartilage is cellular in vertebrates, but is acellular in hemichordates (Smith et al. 2003; Cole and Hall 2004b; Rychel et al. 2006). In addition, three types of agnathan cartilage have been distinguished by histology: hard cartilage, soft cartilage, and mucocartilage (Zhang and Cohn 2006; Zhang et al. 2006; Cattell et al. 2011). Which of these would be homologous to hyaline cartilage of gnathostomes, or are they all? Modern evolutionary thinking overlooks such superficial histological differences, emphasizing instead the importance of tracking changes to the underlying molecular genetic factors during trait evolution.

Evolutionary studies of skeletal cells will benefit from transcriptomic techniques, such as RNAseq, that enable characterization of their molecular fingerprints, which are the sets of genes expressed in a homogenous population of cells (Arendt 2003). Comparing the molecular fingerprint of distinct cell types has yielded insight into evolutionary relationships among remote animal clades (Arendt 2005; Arendt 2008; Eames et al. 2012). A few technologies can generate molecular fingerprints, but of these, RNAseq currently produces the most robust, unbiased results (Necsulea and Kaessmann 2014). Some advantages of RNA-seq include a higher dynamic range, allowing the detection of transcripts that are expressed at very high or low levels, and the ability to detect novel genes and alternative splice variants in samples from any animal (Wang et al., 2009). Important for evolutionary studies, then, RNAseq allows for an accurate comparison of molecular fingerprints in both closely and distantly related species (Necsulea and Kaessmann 2014; Pantalacci and Semon 2015).

Tracking gene expression patterns that underlie a homologous trait through phylogeny provides unparalleled insight into molecular mechanisms of evolution. In fact, comparative transcriptomics might reveal that two tissues are homologous (so-called “deep homology”; Shubin

et al. 2009), despite superficial histological or cellular differences. For example, the presence of immature cartilage in a variety of invertebrate taxa raises the possibility of a tissue with deep homology to cartilage present in the ancestor to all metazoans (Fig. 4B). Also, identifying invertebrate tissues that express "bone genes" may reveal deep homology of these cells to osteoblasts, potentially facilitating the de novo appearance of the Runx2 GRN underlying bone formation. Genes in the vertebrate Sparc family play a role in skeletal matrix mineralization in vitro (Termine et al. 1981; Pataquiva-Mateus et al. 2012). Although similar in vivo roles for Sparc genes have not been demonstrated clearly (Roach 1994; Gilmour et al. 1998; Rotllant et al. 2008), comparative genomics reveal a clear correlation between some Sparc genes and bone formation (Kawasaki and Weiss 2006; Martinek et al. 2007; Koehler et al. 2009; Bertrand et al. 2013; Venkatesh et al. 2014). Interestingly, Sparc genes are expressed in amphioxus, which do not have bone nor mineralize their tissues (Bertrand et al. 2013). If *Runx2* co-opted regulation of these genes during the de novo appearance of the osteoblast, then Sparc-expressing cells in amphioxus may have deep homology to osteoblasts.

Comparative transcriptomics can be used to evaluate quantitatively important features of GRN evolution, including constraint and adaptation. Although Gould recently revived the formalist pleas of Galton, Whitman, and others for constraint to have a positive role during evolution (Gould 2002), constraint commonly is considered a restriction or limitation on the evolutionary process (Arnold 1992). Evidence of constraint can be seen when transcriptomes are highly conserved among various tissues or clades, presumably due to genomic, developmental, or structural limitations. In addition to these constraints, a GRN under stabilizing selection would not vary much with respect to the genes expressed and their levels of expression, thus giving a transcriptomic signal of constraint. In fact, the architecture of GRN kernels, which usually consist of transcription factors and other regulatory genes, can remain highly conserved for a long period of time (Levine and Davidson 2005; Davidson and Erwin 2006). In contrast, adaptation commonly is considered positive for change during evolution (Gould 2002; Stayton 2008; Losos 2011). Evidence of adaptation can be seen when transcriptomes differ widely among various tissues or clades, presumably in response to tissue- or clade-specific selective pressures. A GRN under negative or positive selection would vary a lot in the genes expressed and their levels of expression.

Comparative transcriptomics has unraveled the complexity of several important developmental and evolutionary processes in both invertebrate (Levin et al. 2012; McKenzie et al.

2014) and vertebrate organisms (Chan et al. 2009; Brawand et al. 2011). A major challenge in evolutionary biology is to explain the appearance of novel traits and the GRNs underlying their formation. Two different models have been proposed, with only one currently receiving much experimental support. In the first model, a GRN driving a novel trait also evolved de novo (Fig. 6A). For example, orphan genes, or genes without clear family members, might be important drivers of evolutionary novelty. First described in the yeast genome (Dujon 1996), they occur also in many taxa, including rodents, primates, and humans (Heinen et al. 2009; Toll-Riera et al. 2009; Li et al. 2010). Orphan genes might have appeared de novo from non-coding sequences rather than from existing genes (Tautz and Domazet-Loso 2011). Subsequent interactions that these orphan genes establish among other genes would create a novel GRN with the capability of driving formation of a novel trait. This "de novo" model has received little experimental support in metazoans, but currently serves as the basis for the hypothesis that bone (or dentine, if dentine appeared before bone during evolution) evolved before mature cartilage (Fig. 6A). In molecular terms, the GRN driving formation of the osteoblast would have appeared de novo, presumably in a short evolutionary timeframe.

In the second model for appearance of evolutionary novelties, which is increasingly supported by the literature, a novel trait appears by co-opting a pre-existing GRN (Fig. 6B; Fisher and Franz-Odendaal 2012; Achim and Arendt 2014). For example, comparative genomic studies on muscle cells, immune cells, and neurons suggested that these cell types evolved by co-opting pre-existing genetic systems (Achim and Arendt 2014). In addition, the appearance of a novel embryonic cell lineage in vertebrates, the neural crest cell, has been argued to result from the co-option of pre-existing GRNs that were employed by cells in the neural tube, notochord, and pharynx in ancestral chordates (Baker and Bronner-Fraser 1997; Donoghue and Sansom 2002; Meulemans and Bronner-Fraser 2005; McCauley and Bronner-Fraser 2006; Zhang and Cohn 2006a; Meulemans and Bronner-Fraser 2007). In fact, the neural crest-derived vertebrate cartilaginous head skeleton might have arisen after neural crest cells co-opted an ancestral chordate GRN that was used for cartilage formation in other parts of the body (Jandzik et al. 2015). Here, we use the same argument to support our idea that the osteoblast appeared when a non-chondrogenic mesenchymal cell co-opted expression of the mature cartilage Runx2 GRN.

## 2.7. Comparative transcriptomics and skeletal tissue evolution

How extensive is our understanding of the GRNs driving cartilage and bone formation? As outlined above, Sox9 and Runx2 GRNs are critical in a variety of vertebrates, but is this the whole story? Few studies have analyzed the molecular fingerprint of the chondrocyte and osteoblast using unbiased transcriptomics, but such experiments may identify unknown GRN's driving formation of these cell types. The chondrocyte molecular fingerprint was estimated by compiling data from the literature and summarizing their interactions into a GRN (Cole 2011). Recently, transcriptomics on Sox9 and Runx2 loss-of-function skeletal cells in vitro have shed light on Sox9 and Runx2 GRNs that are relevant to chondrocyte and osteoblast differentiation (Oh et al. 2014; Wu et al. 2014). A promising future direction is to use transcriptomics to define these GRNs in vivo using Sox9 and Runx2 loss-of-function animals. Comparative transcriptomics between vertebrae and gill arch skeletal elements of a teleost demonstrated a high degree of overlap in gene expression between these two tissues (Vieira et al. 2013), but the presence of multiple cell types, including chondrocytes and osteoblasts, in both samples confounds attribution of these data to a particular cell type. Therefore, more specific techniques should be used to isolate a pure population of cells in vivo in order to accurately reveal and compare the molecular fingerprints of different skeletal cell types (Fig. 5).

Two related, fascinating questions remain for future research: how did the GRNs directing skeletal cell differentiation appear, and how did they evolve afterward? In this review, we argue that gradual establishment of the Runx2 GRN during evolution of the mature chondrocyte (subsequently co-opted by a non-chondrogenic mesenchymal cell to form bone) is more parsimonious than the de novo appearance of the Runx2 GRN in osteoblasts (Fig. 6). Given the latter possibility, however the tremendous gene expression similarities between mature cartilage and bone in tetrapods also may reflect co-option of the Runx2 GRN by the mature chondrocyte after it was established in the osteoblast. These possibilities predict divergent vs. convergent evolution, respectively, of the Runx2 GRN in mature chondrocytes after the appearance of the osteoblast. Therefore, we propose an examination of skeletal cell molecular fingerprints in a variety of vertebrates to resolve this issue. Our divergent model predicts that the overlap between mature chondrocyte and osteoblast molecular fingerprints will decrease in more recently diverged organisms (Fig. 7A). For example, molecular fingerprints of mature chondrocytes and osteoblasts would overlap more in earlier diverged lineages

of vertebrates, such as teleosts, than in later diverged lineages, such as amphibians or mammals. On the other hand, the convergent model predicts the opposite result (Fig. 7B).

But do skeletal cell molecular fingerprints evolve in clade-specific manners? A limited number of studies trying to answer this question suggest two competing ideas. On the one hand, molecular fingerprints of the chondrocyte and the osteoblast have been proposed to be highly constrained among various vertebrate clades (Fig. 8A; Fisher and Franz-Odendaal 2012; Vieira et al. 2013). On the other hand, gene expression comparisons between gar, zebrafish, chick, and mouse suggest that the chondrocyte molecular fingerprint is constrained among vertebrates, while the osteoblast molecular fingerprint varied, perhaps in response to clade-specific selective pressures (Fig. 8B; Eames et al. 2012). Interestingly, generalizing these results puts forward the hypothesis that earlier-evolved cell types, in this case chondrocytes, might be more constrained in their gene expression than cell types that appeared later, such as osteoblasts, perhaps due to stabilizing selection over geologic timescales. Comparative transcriptomics can quantitate constraint and adaptation, by measuring how transcript levels vary among samples from different taxonomic lineages.

In the future, comparative transcriptomics will elucidate the dynamics of skeletal cell type evolution, identifying lineage-specific changes in gene expression, providing quantitative measures of constraint and adaptation, and potentially establishing deep homology of skeletal cells with previously unappreciated cell types. Indeed, appropriate application of comparative transcriptomics has the potential to revolutionize understanding of the molecular mechanisms of trait evolution.

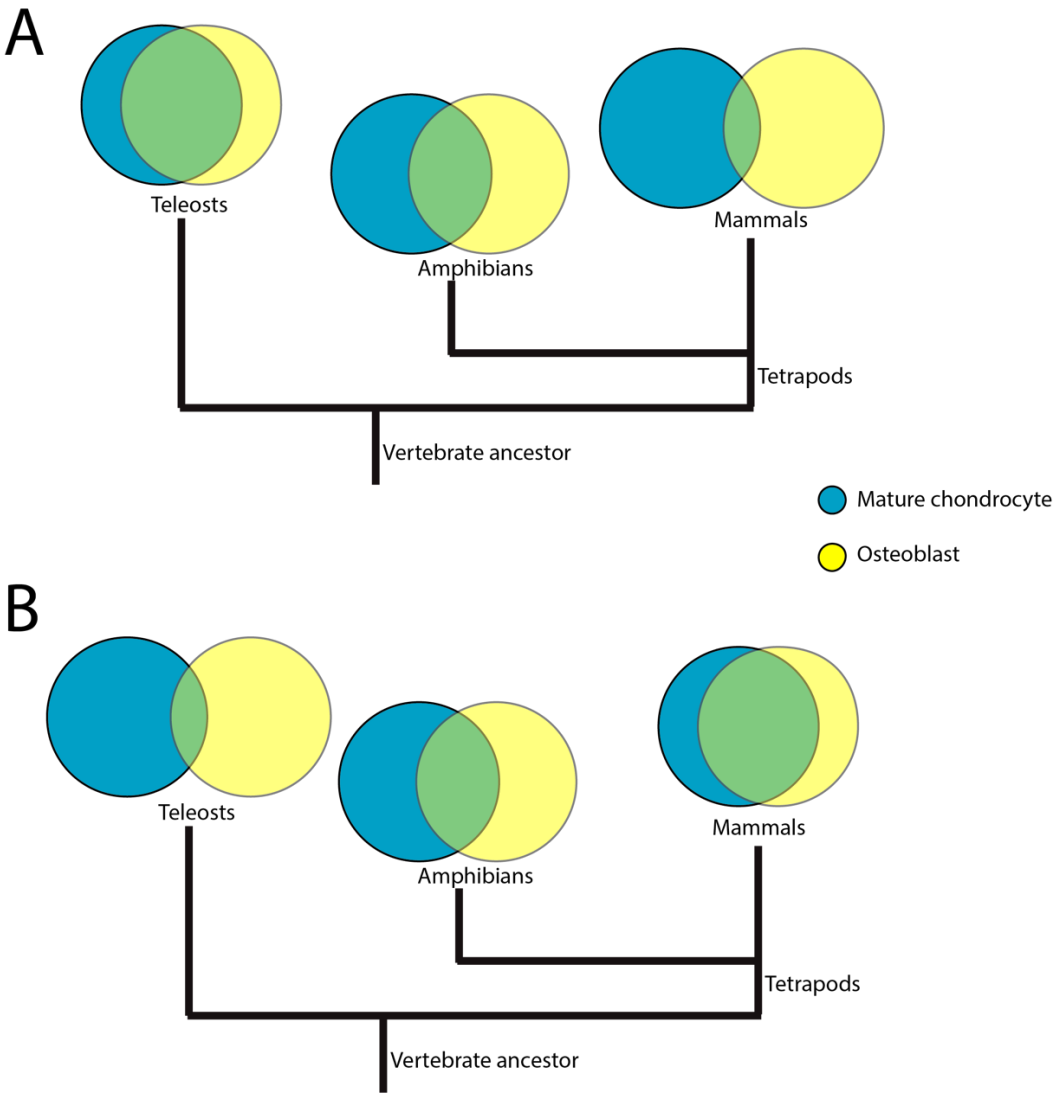



**Figure 7. Divergent vs. convergent evolution of the molecular fingerprints of mature chondrocytes and osteoblasts.** Venn Diagrams comparing putative molecular fingerprints between mature chondrocytes and osteoblasts in three distinct vertebrate clades may resolve among two hypotheses for the origins of the osteoblast. (A) Divergent model. Osteoblast evolved when a GRN was co-opted from mature chondrocytes. Differing selective pressures on ancestors of various lineages, followed by lineage-specific constraints, may have caused gradual divergence between the GRN of osteoblasts and mature chondrocytes during vertebrate evolution. If true, then the overlap between mature chondrocyte and osteoblast molecular fingerprints will be significantly higher in earlier diverged lineages, such as teleosts, than in later diverged lineages, such as mammals. (B) Convergent model. Osteoblast GRN evolved de novo. Similar selective pressures on osteoblasts and mature chondrocytes in ancestors of later diverging lineages may have caused convergence between the GRN of osteoblasts and mature chondrocytes during vertebrate evolution. If true, then the overlap between molecular fingerprints of mature chondrocytes and osteoblasts will be significantly lower in earlier diverged lineages. Branch lengths in trees are arbitrary; the overlap between molecular fingerprints is shown in green and, in the divergent model, may represent the ancestral GRN kernel of both mature chondrocyte and osteoblast.

## 2.8. Summary

Given the role that fossilized bones played in devising early evolutionary theory, skeletal tissue evolution has fascinated scientists for centuries. In particular, the appearance of bone as an evolutionary novelty demands explanation, which modern molecular and embryological techniques address in ways never imagined by studies of the fossil record alone. Here, we focus on the three main skeletal tissues present in vertebrates (immature cartilage, mature cartilage, and bone), and use findings from both traditional and modern studies to argue that bone evolved from mature cartilage. Standing in contrast to the available fossil record, which suggests that bone appeared prior to mature cartilage, this hypothesis posits that a GRN driving traits such as matrix mineralization in mature cartilage was co-opted by non-chondrogenic mesenchymal cells to produce bone. Alternatively, the GRN driving bone formation may have evolved first and subsequently was co-opted by mature cartilage, but we use an argument based on parsimony that this scenario would be more complicated to achieve. Comparing the molecular fingerprints of skeletal tissues in agnathans and sister chordate species with those in vertebrates might resolve among these possibilities. In addition to comparative transcriptomics revealing the origins of evolutionary novelties, tracking molecular fingerprints of skeletal cells in various vertebrate lineages can identify quantitative measures of constraint and adaptation within the GRNs that govern the formation of skeletal tissues. Therefore, we strongly believe that this novel approach may revolutionize understanding of the evolution of cartilage and bone and more generally provide a modern paradigm for molecular genetic changes during the evolutionary process.

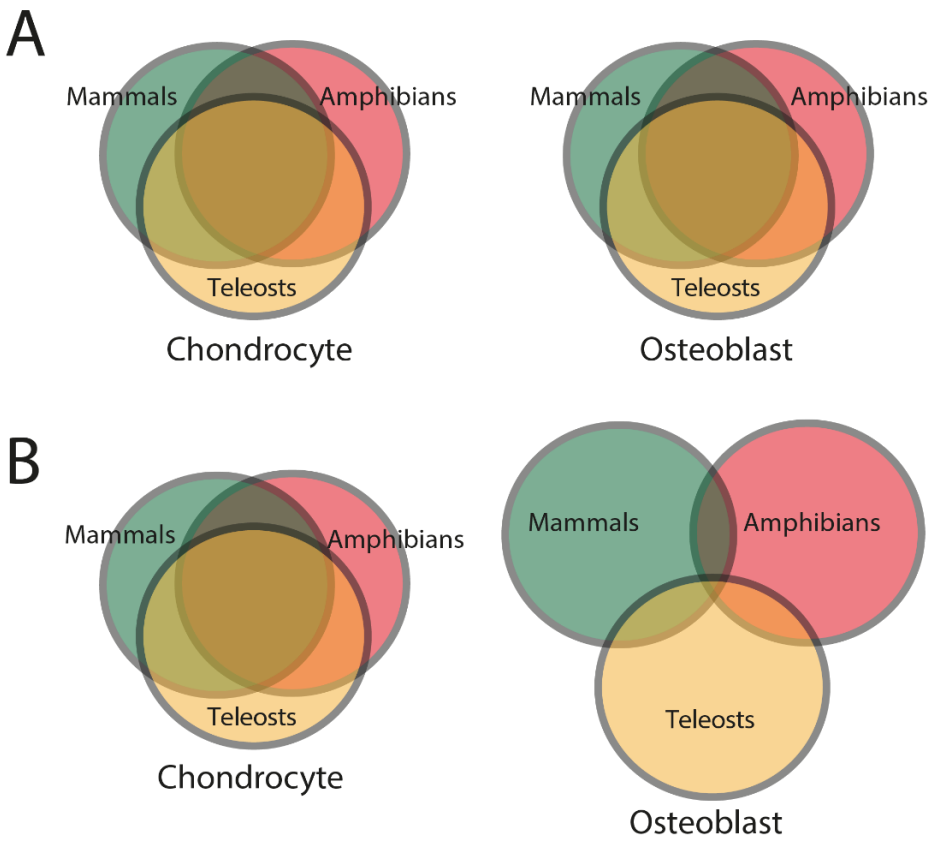



**Figure 8. Differing models for levels of constraint and adaptation among skeletal cells of different vertebrate lineages.** Venn diagrams comparing putative molecular fingerprints of chondrocytes and osteoblasts from three vertebrate clades. The ancestral chondrocyte and osteoblast GRN kernels are represented in the overlap of the circles. More overlap represents more constraint/less adaptation among clades. (A) The first scenario predicts that the molecular fingerprints of the chondrocyte and osteoblast (and thus the GRNs governing their formation) are constrained to equal extents among vertebrates (Fisher and Franz-Odendaal, 2012; Vieira et al., 2013). (B) The second scenario predicts that the chondrocyte molecular fingerprint is more constrained among vertebrate clades, while the osteoblast molecular fingerprint shows more signs of clade-specific adaptations (Eames et al., 2012). In general, this latter scenario posits that a cell type appearing later during animal phylogeny is more free to vary than a cell type appearing earlier, whose molecular fingerprint was fixed via stabilizing selection.

## CHAPTER 3:

**Manuscript 1; Portions of the gene regulatory network driving cartilage and bone formation interact via averaging and synergism during cartilage maturation (*in preparation*).**

Gomez-Picos P, Ovens K, Ashique AM, Whalen M, Eskiw CH, McQuillan I, and Eames BF

# CHAPTER 3. Portions of the gene regulatory network driving cartilage and bone formation interact via averaging and synergism during cartilage maturation

## 3.1. ABSTRACT

Transcriptional control of any biological process occurs through the action of gene regulatory networks (GRNs), which characterize specific regulatory states. While one GRN might give rise to distinct cell fates, studies analyzing the interaction between distinct portions of a GRN during cell differentiation are lacking. To address this, we tested *in vivo* the hypothesis that two distinct portions of the GRN, which are active during formation of two discrete skeletal cell types (IMM and OST), interact during differentiation of a third skeletal cell type (MAT). These three cell types were isolated from the mouse embryo using laser capture microdissection (LCM), and then RNA was extracted. Multiple analyses of corresponding RNA-seq data supported the hypothesis. Gene co-expression network analyses suggested that one portion of the GRN containing the chondrogenic transcription factor SOX9 characterizes immature chondrocytes, one portion of the GRN containing the osteogenic transcription factor RUNX2 characterizes osteoblasts, and both SOX9 and RUNX2 transcriptional programs operate in mature chondrocytes. Indeed, mature chondrocytes differentially expressed fewer genes than the other cell types, consistent with the idea that overlapping actions between the transcriptional programs active in immature chondrocyte and osteoblast dictate mature chondrocytes. Clustering analyses provided molecular insights into potential interactions between distinct transcriptional programs active in this skeletal cell GRN. Several genes in mature chondrocytes had expression levels that represented an averaging between levels in the immature chondrocyte and osteoblast. Expression levels of one gene cluster, containing the hallmark mature chondrocyte genes *collagen type 10a1 (Col10a1)* and *indian hedgehog (Ihh)*, suggested a synergistic interaction between portions of the GRN driving immature chondrocytes and osteoblasts. In addition to identifying novel genes expressed in mature chondrocytes and confirming the function of two new mature chondrocyte transcription factors, EBF2 and IRX6, these results

outline a novel *in vivo* experimental system through which to understand GRN organization and interaction.

## 3.2. INTRODUCTION

Every biological process is regulated in part by the action of gene regulatory networks (GRNs), each GRN being governed by one to many transcription factors that determine specific regulatory states (Davidson et al. 2003; Li et al. 2015). Interactions among different portions of one GRN might be important in determining specific phenotypes, but the mechanisms of GRN interaction have been largely unexplored. Here, we focus on skeletal cells *in vivo*, because previous studies suggested that distinct portions of the GRN driving skeletal cell differentiation might interact (Eames et al. 2004; Cheng and Genever 2010; Cole 2011; Kerkhofs et al. 2012), making it an ideal model to investigate this understudied issue.

Most of the vertebrate skeleton is composed of three main cell types: immature chondrocytes, mature chondrocytes (i.e., pre-hypertrophic and hypertrophic chondrocytes), and osteoblasts (Eames et al. 2003). During the skeletogenic process of endochondral ossification, immature and mature chondrocytes produce two distinct types of cartilage, whereas osteoblasts form bone (Eames and Helms 2004; Eames et al. 2004; Tamamura et al. 2005; Gentili and Cancedda 2009). Bone also can be formed by intramembranous ossification, where mesenchymal cells directly differentiate into osteoblasts without using a cartilage template (Smith and Hall 1990; Karsenty et al. 2009; Komori 2010).

Skeletal cells might be a good model to test GRN interaction since they are distinct, but they also share many features. Immature chondrocytes, mature chondrocytes, and osteoblasts derive from a common cell lineage (an osteochondroprogenitor) and share the expression of many genes (Fang and Hall 1997; Day et al. 2005; Vortkamp et al. 1996a; Li et al. 1998; Inada et al. 1999; Knudson and Knudson 2001; Rees et al. 2001; Eames and Helms 2004; Eames et al. 2004; Gentili and Cancedda 2009; Gomez-Picos and Eames 2015). Exemplifying this relatedness, cells transition during endochondral ossification from an immature to a mature chondrocyte, and then some mature chondrocytes even trans-differentiate into osteoblasts (Hammond and Schulte-Merker 2009; Yang et al. 2014; Zhou et al. 2014; Jing et al. 2015; Park et al. 2015). Immature chondrocytes and

osteoblasts express relatively few common genes (Hoffmann et al. 1996; Eames et al. 2003; Li et al. 2017). Also, mature chondrocytes are a distinct cell type that can remain throughout an animal's life, such as in articular cartilage, expressing a characteristic set of genes that are never expressed in immature chondrocytes or osteoblasts (Lefebvre et al. 1995; Sophia Fox et al. 2009; Decker et al. 2015; Wang et al. 2017). Indeed, many mature chondrocytes do not become osteoblasts during endochondral ossification (Yang et al. 2014; Zhou et al. 2014; Hinton et al. 2017; Aghajanian and Mohan 2018).

Transcription factors could explain the similarities in gene expression among immature chondrocytes, mature chondrocytes, and osteoblasts (Eames et al. 2003a; Kronenberg 2003; Eames et al. 2004; Karsenty et al. 2009). Two transcription factors, SOX9 and RUNX2, play roles consistent with being 'master regulators' of skeletal cell differentiation. *Sox9* is expressed in immature and mature chondrocytes, while *Runx2* is expressed in mature chondrocytes and osteoblasts (Eames and Helms 2004). Genetic experiments confirmed that SOX9 and RUNX2 drive cartilage and bone formation, respectively. Loss of *Sox9* results in the absence of chondrocytes (Bi et al. 1999), whereas loss of *Runx2* blocks osteoblast differentiation (Komori et al. 1997). Mature chondrocytes do not form without the function of either *Sox9* or *Runx2,* although the exact mechanisms through which they partner in this process are unknown (Yamashita et al. 2009; Cheng and Genever 2010; Gomez-Picos and Eames 2015). *Sox9* and *Runx2* mis-expression can toggle among these three skeletal fates, leading to the idea that a binary transcription factor code underlies skeletal cell specification (Eames et al. 2004). A SOX9 transcriptional program regulates immature chondrocytes, a RUNX2 transcriptional program drives osteoblasts, and both transcriptional programs are active in mature chondrocytes. We test this hypothesis *in vivo*, focusing on putative interactions between distinct transcriptional programs of the GRN active in mature chondrocytes.

Despite recent progress in identifying target genes for SOX9 and RUNX2 in skeletal cells (Oh et al. 2014; Wu et al. 2014; Ohba et al. 2015; Tarkkonen et al. 2017; Liu et al. 2018; Yamashita et al. 2019), the molecular mechanisms underlying the interaction between portions of the GRN mediated by SOX9 and RUNX2 during formation of mature chondrocytes are unknown. Only a few previous studies have analyzed GRN interactions in any system. For example, during sea urchin embryogenesis, transcription factors from a mesoderm GRN influence gene expression of an endoderm GRN (Peter and Davidson 2011). In cells of hybrid species of plants and rodents, specific categories of gene expression suggested that a portion of two GRNs might interact, while other

portions act independently (Brekke et al. 2016; Hu et al. 2016). Similar mechanisms of GRN interaction could be involved in mature chondrocytes.

To test the hypothesis, we developed a novel unbiased experimental model using comparative transcriptomics to analyze interaction globally between distinct transcriptional programs in the skeletal cell GRN. Specifically, laser capture microdissection (LCM) was used to isolate immature chondrocytes, mature chondrocytes, and osteoblasts from the mouse embryo *in vivo.* RNA from these cells underwent RNA-seq and transcriptomic analyses. The GRN directing the differentiation of immature chondrocytes, mature chondrocytes, and osteoblasts was estimated using gene co-expression network (GCN) analyses (Stuart et al. 2003; Mahanta et al. 2012; McCall 2013; Khosravi et al. 2015), and the relationships of *Sox9* and *Runx2* to portions of this GRN were evaluated. While a wealth of information was generated on genes associated with skeletal cell differentiation, our LCM-RNA-seq approach also can help to reveal GRNs in any cell type *in vivo*, meanwhile identifying genomic loci that are predicted to have discrete molecular mechanisms of GRN interaction.

## 3.3. RESULTS

### 3.3.1. While immature chondrocytes, mature chondrocytes, and osteoblasts shared the expression of many genes, immature chondrocytes and osteoblasts showed the least similarity

Immature and mature chondrocytes, the latter composed of a heterogeneous population of pre-hypertrophic and hypertrophic cells, were obtained from the mouse humerus at E14.5 (Fig. 9A,C-E). Bone is evident in the perichondrium surrounding mature cartilage of the E14.5 humerus, but vascular invasion into mature cartilage has not yet occurred (Maes et al. 2002), thus avoiding the introduction of contaminating cell types (Fig. 9A,C-E,G-J). Osteoblasts were isolated from the E14.5 dentary, an intramembranous bone in the jaw adjacent to Meckel's cartilage (Fig. 9B,F). This allowed capture of osteoblasts without chondrocyte contamination (Fig. 9F,K-L), a likely problem if harvesting perichondral osteoblasts. RNA-seq was then carried out on RNA isolated from

immature chondrocytes (IMM), mature chondrocytes (MAT), and osteoblasts (OST; n=3 for each cell type).

To evaluate similarities and differences among the IMM, MAT, and OST transcriptomes, a principal component analysis (PCA) was performed. PCA is a dimension reduction method and has been extensively used in gene expression analysis (McLachlan et al., 2004; Sharov et al., 2005) to emphasize variation and bring out strong patterns. This approach was used to test the similarities in gene expression between IMM, MAT, and OST. The variation in the samples was captured well with two components (71.1% variance explained by PC1 and PC2; Fig. 10A). PCA grouped samples according to cell types, suggesting that variation between cell types was greater than variation between replicates, in part validating the LCM technique. The IMM transcriptome was separated from other samples in PC1/PC2 with 95% confidence, while MAT and OST transcriptomes overlapped in PC1/PC2, suggesting overlapping and distinct aspects of gene expression among the three cell types (Fig. 10A). In order to explore further the relatedness among skeletal cell transcriptomes, correlation plots of $\log_2$ gene expression were generated (Fig. 10B-F). First, 2D plots showed that the overall gene expression profile of MAT was similarly high to both IMM ($R^2$ =0.80; Fig. 10B) and OST ($R^2$= 0.79; Fig. 10C), suggesting that portions of the GRN driving IMM and OST are also active in MAT. OST data were from the jaw, while IMM and MAT data were from the limb. By contrast, IMM and OST were much less similar ($R^2$=0.63; Fig. 10D), supporting the idea that distinct GRN portions operate in IMM and OST. These datasets were also compared to published transcriptomic data from grossly dissected articular cartilage and skeletal muscle from juvenile mice (Goodwin et al. 2014; Dudek et al. 2016). As expected, the gene expression profiles of IMM, MAT, and OST captured by LCM were more similar to grossly dissected cartilage than to skeletal muscle (Fig. 10E,F). Finally, to visualize the relationship among all three skeletal cell transcriptomes simultaneously, 3D plots were generated. The 3D plots showed a highly correlated relationship among IMM, MAT, and OST gene expression, resulting in a more restricted shape that did not take up as much area of the plot ($R^2$= 0.89; Fig. 10G) as when muscle was plotted with IMM and OST ($R^2$= 0.27; Fig. 10H).

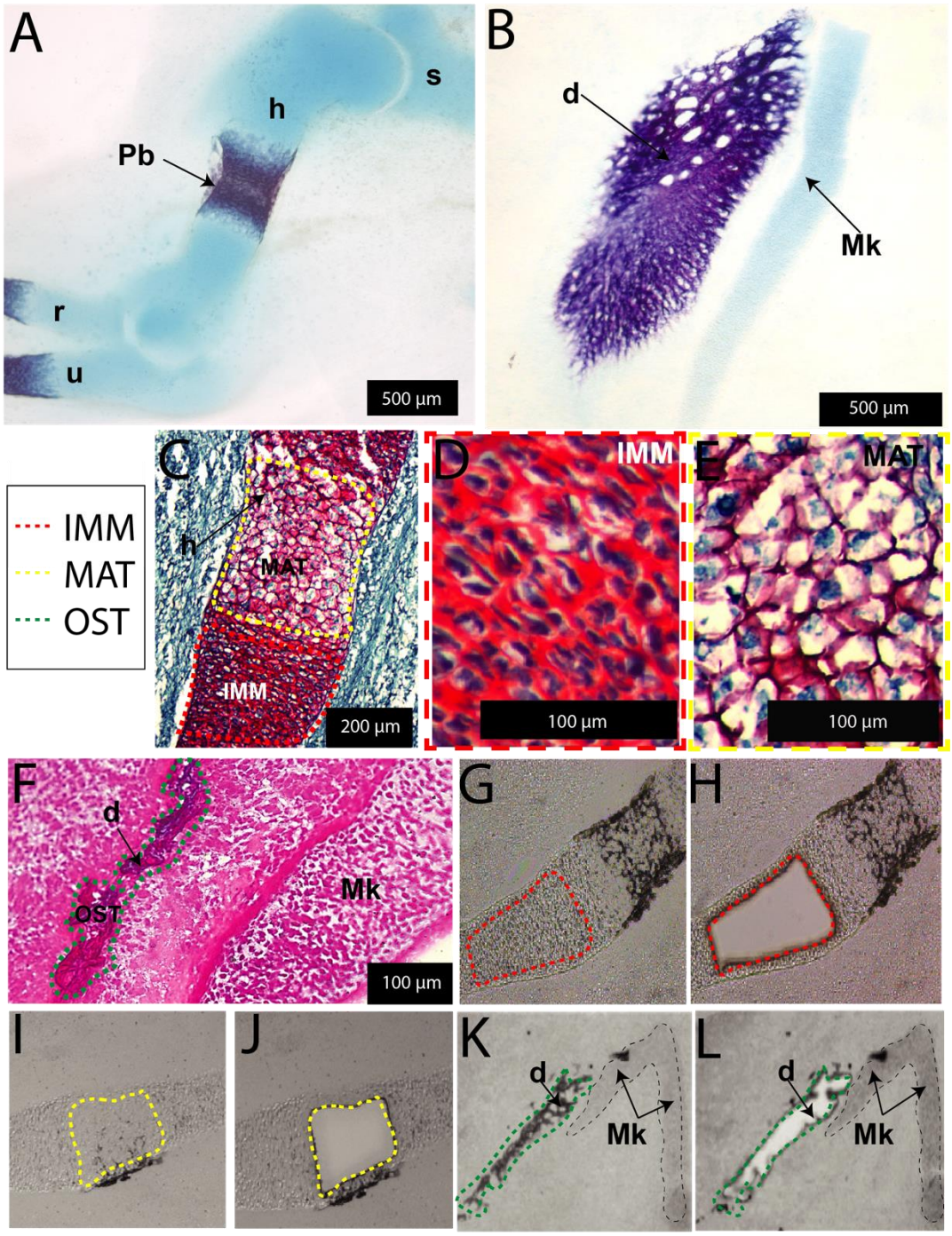



**Figure 9. Laser capture microdissection allowed the isolation of specific skeletal cell types from the mouse E14.5 humerus and mandible.** (A, B) Whole-mount Alcian blue and Alizarin red staining identified cartilage and bone. (C) Safranin O-stained section of E14.5 humerus highlighted the mature cartilage region (yellow dotted outline). High-magnification images of immature (D) and mature chondrocytes (E) illustrated cell hypertrophy in mature cartilage. (F) Trichrome-stained section of E14.5 mandible showed Aniline blue staining of bone matrix in the dentary (d). Unstained sections of E14.5 mouse humerus before (G) and after (H) laser capture of immature chondrocytes, and before (I) and after (J) laser capture of mature chondrocytes. Unstained sections of E14.5 mouse dentary before (I) and after (J) laser capture of osteoblasts. Abbreviations: IMM, immature chondrocytes; MAT, mature chondrocytes; OST, osteoblasts; Mk, Meckel's cartilage; d, dentary bone; Pb, perichondral bone; h, humerus; s, scapula; r, radius; u, ulna (LCM credit: Dr. Amir Ashique).

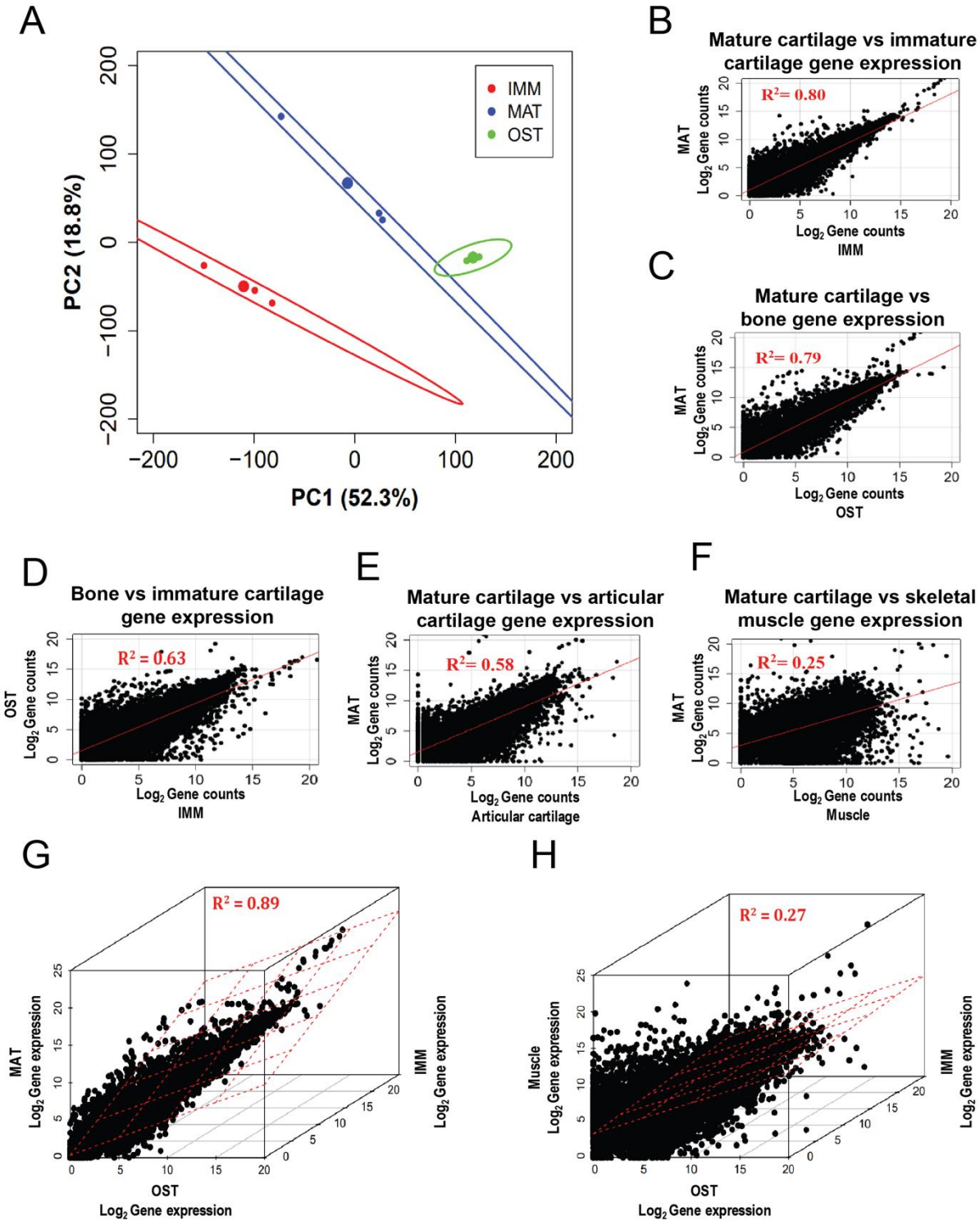



**Figure 10. PCA analysis emphasized similarities and differences in gene expression.** (A) OST transcriptomes showed the least variation along PC1/PC2 compared to the other cell types. Pairwise $\log_2$ gene expression plots (each dot is a gene) demonstrated that MAT was equally similar to both (B) IMM ($R^2$=0.80), and (C) OST ($R^2$=0.79). (D) IMM and OST were the most disparate datasets ($R^2$=0.63). Skeletal cells captured by LCM showed a more similar expression pattern to (E) grossly dissected cartilage, than to (F) skeletal muscle. 3D $\log_2$ gene expression plots confirmed similarities were greater among IMM, MAT, and OST (G) than in muscle (H), which was substituted for MAT (Credit: Katie Ovens).

### 3.3.2. Mature chondrocytes had fewer uniquely expressed genes than immature chondrocytes or osteoblasts

Gene expression level thresholds were established for each cell type based upon normalized count distributions, and comparative analyses identified genes that were unique to each skeletal cell transcriptome. The numbers of genes expressed in IMM, MAT, and OST were 11,485, 11,922, and 11,963, respectively. MAT shared expression of several genes with only one of the other cell types: 857 with OST, and 505 with IMM (Fig. 11A). Similar to the $\log_2$ gene expression plot (Fig. 10D), the least overlap between two cell types was between IMM and OST (228; Fig. 11A). "Unique" genes were defined to be expressed exclusively above threshold values in only one cell type. MAT had the fewest unique genes (321), compared to OST (639) and IMM (513). The presence of unique gene expression in MAT, however, argues that MAT is not simply a mixture of IMM GRN-dependent cells and OST GRN-dependent cells.

### 3.3.3. Novel transcripts expressed at very low levels in MAT demonstrated the unbiased sensitivity of LCM-RNA-seq

The unbiased nature of RNA-seq allows identification of genes that were not known previously to be involved in a biological process. The present datasets were cross referenced with published microarray data in order to identify unique genes that had not been associated with chondrocytes or osteoblasts; data not shown). Using RNA *in situ* hybridization, two previously uncharacterized unique MAT genes, phosphodiesterase 11A (*Pde11a*) and reproductive homeobox 8 (*Rhox8*), were confirmed to be expressed specifically in mature cartilage of the E14.5 humerus, along with the hallmark mature chondrocyte gene *Col10a1* (Fig. 11B-E). Overall, the average MAT gene had over 2000 transcript counts and the minimum counts for a gene considered expressed in this cell type was 24 counts. Given that *Pde11a* (150 counts) and *Rhox8* (73 counts) were expressed at very low levels, these results verified the sensitivity and accuracy of the RNA-seq dataset.

### 3.3.4. Mature chondrocytes had fewer differentially expressed genes than immature chondrocytes or osteoblasts

Similar to the unique gene analyses, the hypothesis predicts that MAT would have fewer differentially expressed genes (DEGs) than IMM or OST. A prerequisite for designating a gene as a DEG was that it be expressed above threshold in more than one cell type (otherwise, it was termed "unique"; see Fig. 11). Count thresholds were determined to be 25, 24, and 18 for IMM, MAT, and OST, respectively. Of those DEGs specific to one cell type (i.e., significantly upregulated or downregulated compared to both other cell types), MAT expressed far fewer (42) than IMM (270) and OST (328; Fig. 12A).

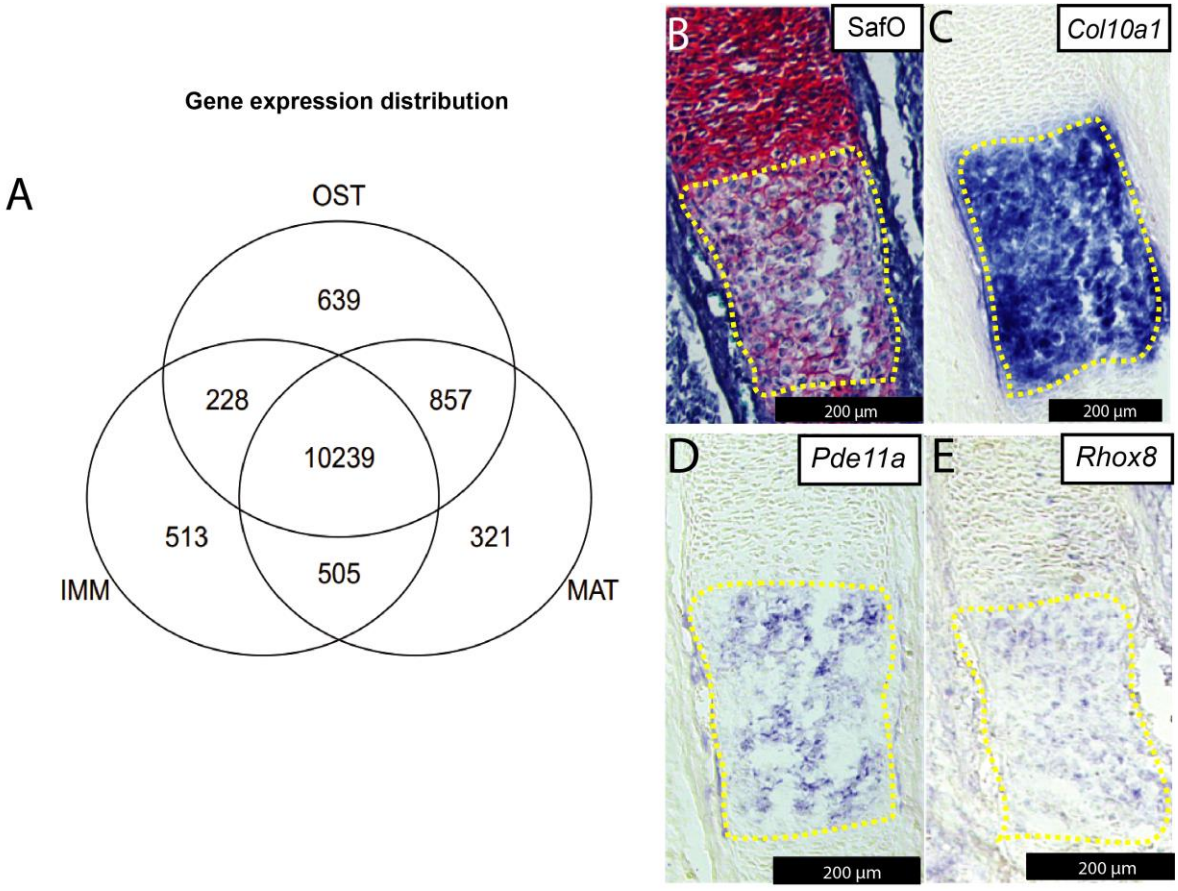



**Figure 11. Validation of novel 'unique' genes expressed in mature chondrocytes.** (A) Venn diagram of genes expressed above threshold demonstrated that MAT had the fewest unique genes **(321)**, compared to OST (639) and IMM (513). (B) Safranin O-stained section of E14.5 humerus highlighted the mature cartilage region (yellow dotted outline). RNA in situ hybridization validated (C) *Col10a1* as a positive control for MAT, and also (D) *Pde11a* and (E) *Rhox8* as novel unique MAT genes. Holes are artifacts and do not represent vascular invasion. Abbreviation: SafO=Safranin O (Venn diagram credit: Katie Ovens)

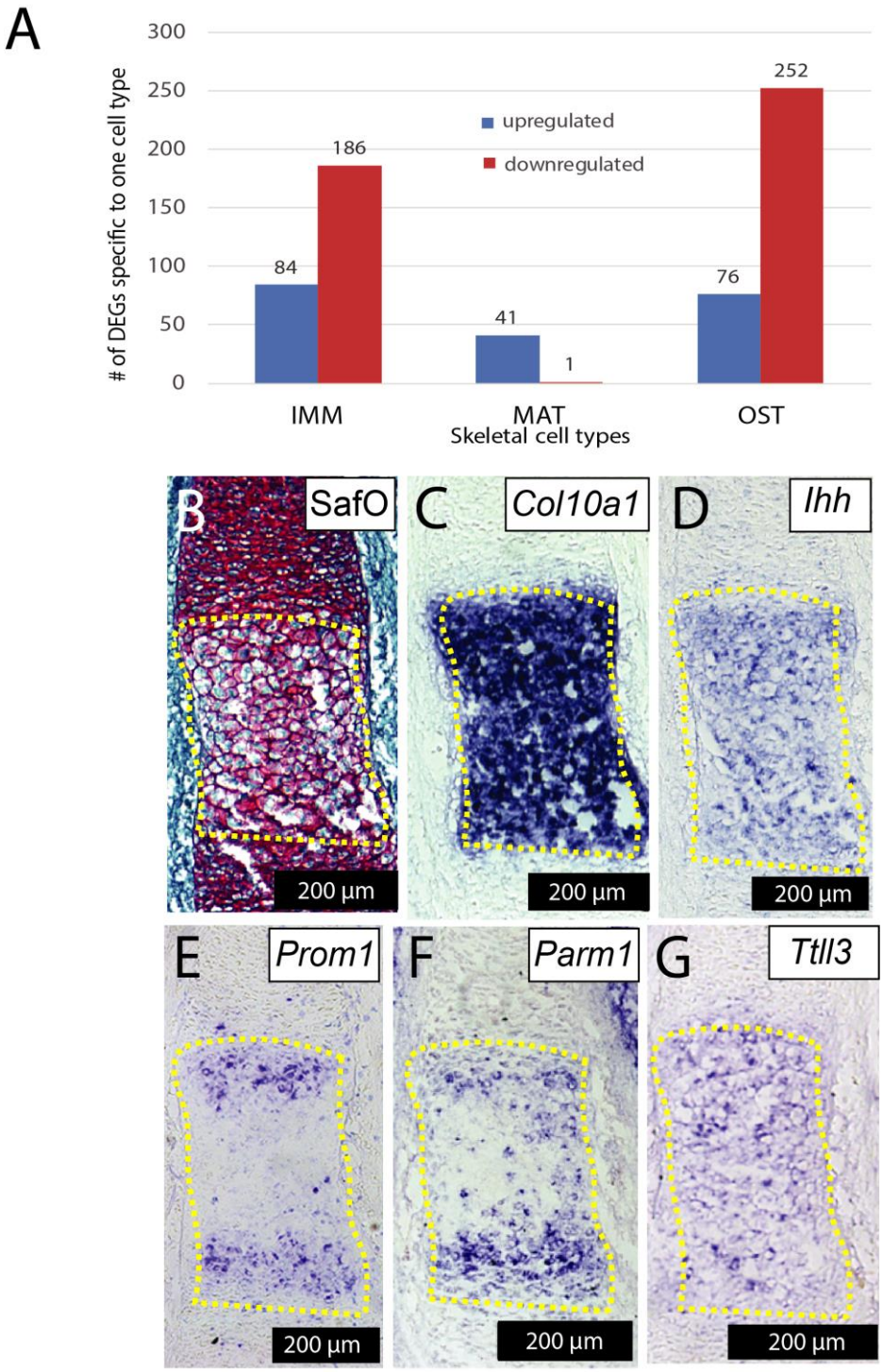



**Figure 12. Analyzing genes differentially expressed in one skeletal cell type compared to both other cell types demonstrated that MAT had the fewest.** (A) Of the 42 DEGs in MAT, only one was downregulated. (B) Safranin O-stained section of E14.5 humerus highlighted the mature cartilage region (yellow dotted outline). *In situ* hybridization of DEGs in MAT: confirmed (C) *Col10a1* and (D) *Ihh* as positive controls; and validated (E) *Prom1*, (F) *Parm1*, and (G) *Ttll3* as novel MAT genes. Holes are artifacts and do not represent vascular invasion. Abbreviation: SafO=Safranin O. Gene lists used for DEGs graphs were generated by Katie Ovens.

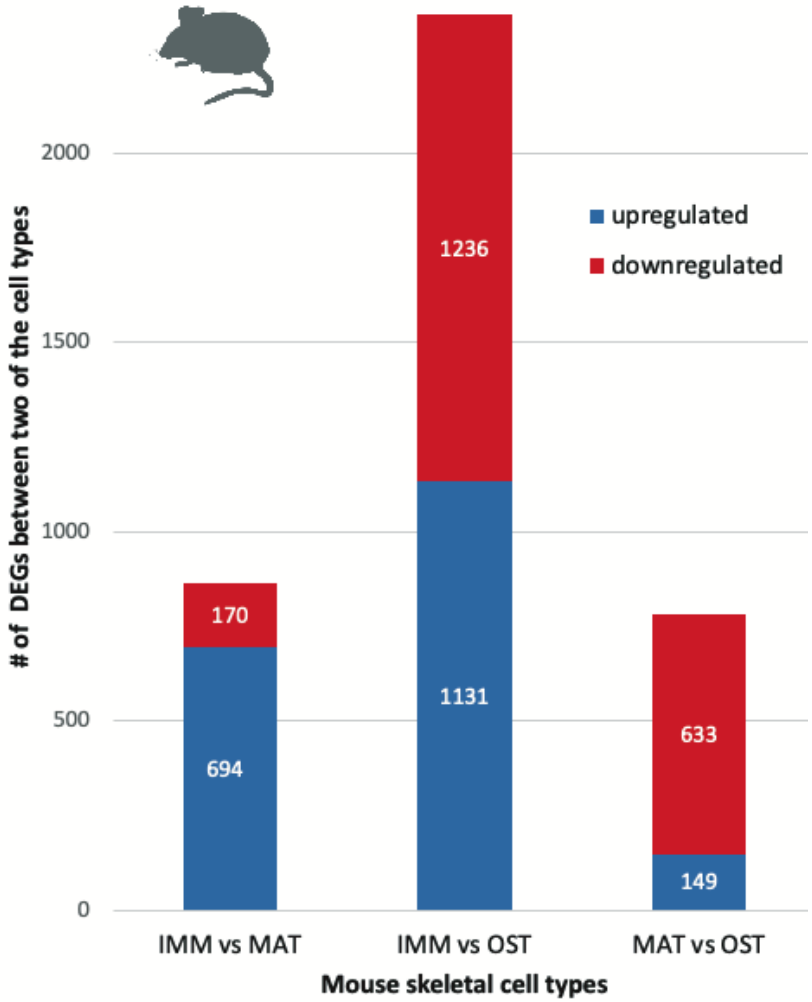



**Figure 13. Pairwise comparisons showed that many genes are differentially expressed in IMM compared to OST.** A gene was considered differentially expressed if it had an absolute log2 fold change greater than 2. A total of 864 genes are differentially expressed between IMM and MAT, 2367 genes are differentially expressed between IMM and OST, and 782 genes are differentially expressed between MAT and OST. This differential gene expression analysis also included genes in the overlap between the three cell types (see Fig. 11A). Gene lists used for DEGs graphs were generated by Katie Ovens.

Pairwise comparisons between two of the cell types showed that many genes are differentially expressed in IMM compared to OST, supporting the idea that distinct portions of the GRN are active in these two cell types (Fig 13). DEGs between IMM and OST suggest that These data also suggested that the MAT regulatory state is largely, but not exclusively, a mixture of gene expression between IMM and OST. Alternatively, the fact that mature chondrocytes had less DEGs could also be due to the heterogeneity of MAT population since mature chondrocytes at different stages were

captured. For instance, gene ontology analyses revealed that apoptosis was an enriched biological process in mature chondrocytes (Fig. 14), suggesting that some MAT at later stages of differentiation were already expressing genes required for this programmed cell death.

RNA *in situ* hybridization was used to confirm not only the upregulation of hallmark mature chondrocyte markers in our MAT RNA-seq dataset, but also to validate DEGs not previously associated with mature chondrocytes. Consistent with previous findings (Kielty et al. 1985; St-Jacques et al. 1999; Zheng et al. 2003), *Col10a1* and *Ihh* were differentially expressed in MAT, and their expression was confirmed in mature cartilage of the E14.5 humerus (Fig. 12B-D, Table 1). In addition, many MAT DEGs that had not been associated previously with skeletal development (Bruyninx et al. 1999; Park et al. 2013; Mak et al. 2014; Rocha et al. 2014) were identified, including the prostate androgen-regulated mucin-like protein 1 *(Parm1),* prominin 1 *(Prom1),* and the tubulin tyrosine ligase-like family, member 3 (*Ttll3;* Table 1). Expression of these genes was confirmed in mature cartilage of the E14.5 humerus (Fig. 12E-G), validating the reliability of the RNA-seq dataset. In addition, gene ontology analyses showed that genes enriched in MAT are involved in biological processes related to both cartilage and bone development (Fig. 14A-C).

**Table 1.** Differentially expressed genes in mature chondrocytes

| Upregulated genes in MAT | logFC (IMM) | logFC.2 (OST) | MinFC | Average counts |
|---|---|---|---|---|
| *Col10a1* | 11.94 | 10.28 | 10.28 | 28641.67 |
| *Ccdc80* | 5.82 | 5.68 | 5.68 | 2281.33 |
| *Ihh* | 8.42 | 5.64 | 5.64 | 2091.00 |
| *Prom1* | 4.24 | 4.58 | 4.24 | 518.33 |
| *Hhip* | 4.56 | 4.22 | 4.22 | 699.33 |
| *Stmn2* | 8.64 | 3.81 | 3.81 | 289.67 |
| *Fbln5* | 7.49 | 3.69 | 3.69 | 308.67 |
| *Rspo3* | 3.67 | 5.54 | 3.67 | 1047.00 |
| *Wnt5b* | 3.64 | 3.75 | 3.64 | 584.00 |
| *Dkk2* | 5.31 | 3.64 | 3.64 | 485.00 |
| *Gcnt2* | 5.12 | 3.63 | 3.63 | 288.33 |
| *Serinc5* | 4.31 | 3.44 | 3.44 | 3202.33 |
| *Nt5dc1* | 4.58 | 3.41 | 3.41 | 699.33 |
| *Pth1r* | 3.71 | 3.36 | 3.36 | 47625.33 |
| *Ttll3* | 3.10 | 4.06 | 3.10 | 933.33 |
| *Mbp* | 3.73 | 3.02 | 3.02 | 181.67 |
| *Slc35g1* | 3.00 | 4.10 | 3.00 | 243.00 |
| *Slc43a2* | 4.42 | 2.99 | 2.99 | 143.00 |
| *Rgs7bp* | 3.01 | 2.95 | 2.95 | 277.33 |
| *Eps8l2* | 2.77 | 8.71 | 2.77 | 542.33 |
| *Slc17a9* | 2.76 | 3.11 | 2.76 | 2465.67 |
| *Kirrel3* | 3.67 | 2.74 | 2.74 | 188.67 |
| *Rpl39l* | 2.71 | 2.86 | 2.71 | 173.00 |
| *Parm1* | 5.08 | 2.68 | 2.68 | 726.00 |
| *Dusp5* | 4.51 | 2.62 | 2.62 | 383.00 |
| *Arsi* | 2.59 | 9.11 | 2.59 | 1942.00 |
| *Nhej1* | 2.58 | 2.87 | 2.58 | 340.33 |
| *Arap2* | 3.51 | 2.54 | 2.54 | 254.33 |
| *Itga1* | 2.53 | 4.37 | 2.53 | 306.00 |
| *Cpa6* | 2.50 | 5.65 | 2.50 | 371.33 |
| *Tmie* | 2.37 | 3.54 | 2.37 | 227.00 |
| *Cttnbp2* | 2.36 | 3.01 | 2.36 | 260.00 |
| *Znhit6* | 2.33 | 2.53 | 2.33 | 1537.00 |
| *Nim1k* | 2.31 | 5.41 | 2.31 | 746.67 |
| *Abtb1* | 2.31 | 2.76 | 2.31 | 680.00 |
| *Rapgef3* | 2.42 | 2.25 | 2.25 | 202.33 |
| *Gm15712* | 2.21 | 4.34 | 2.21 | 170.00 |
| *Rbms3* | 2.35 | 2.14 | 2.14 | 788.67 |
| *Prkg2* | 2.12 | 3.53 | 2.12 | 1750.33 |
| *Comp* | 2.10 | 10.50 | 2.10 | 14353.00 |
| *Sidt2* | 2.18 | 2.02 | 2.02 | 1828.67 |

| Downregulated genes in MAT | Log FC (IMM) | log FC (OST) | MaxFC | Average counts |
|---|---|---|---|---|
| *Selenbp1* | -2.36 | -2.59 | -2.36 | 51.66 |

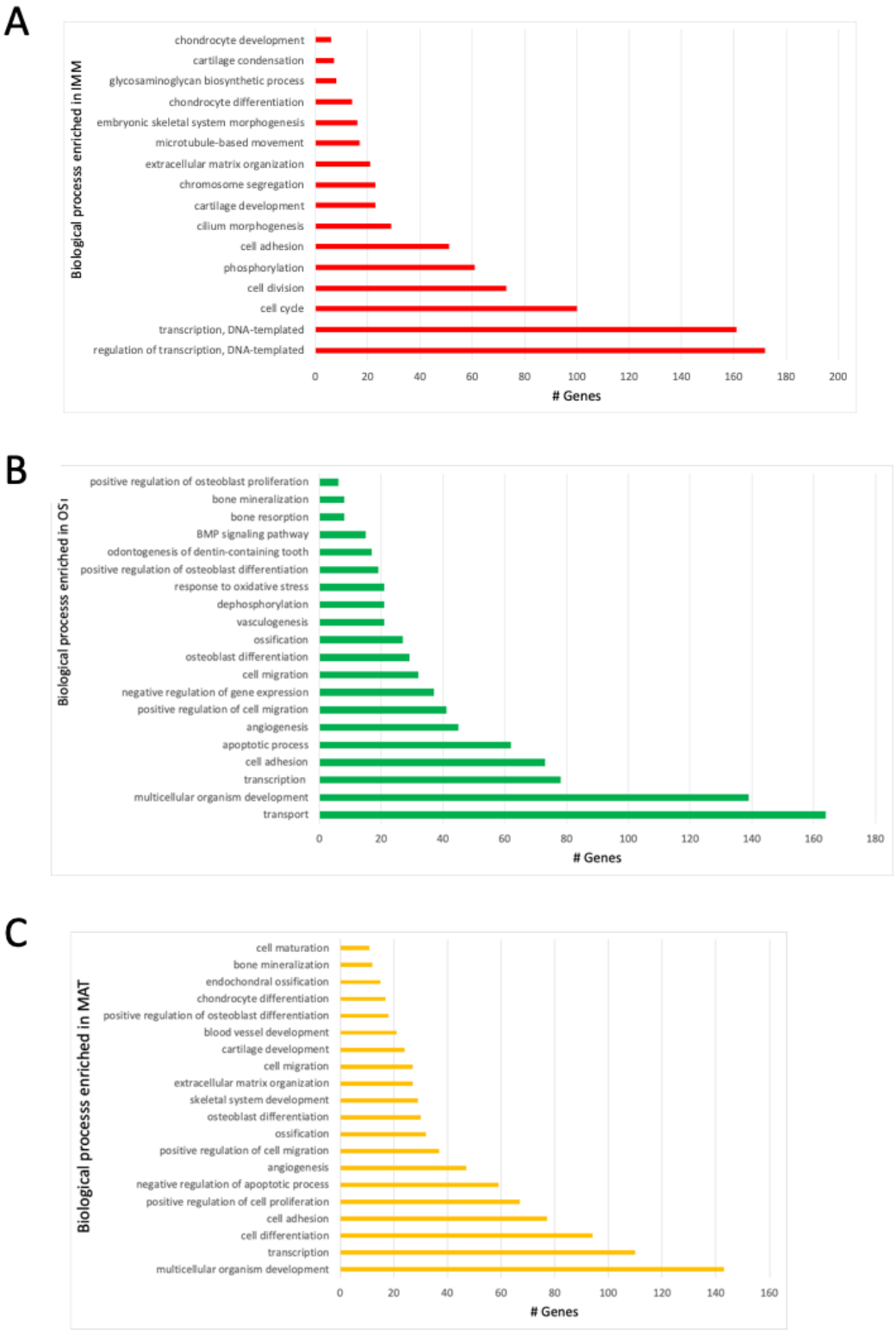



**Figure 14. Gene ontology analysis showed enriched biological processes in IMM, MAT, and OST.** (A) Enriched biological processes in IMM included chondrocyte differentiation and cartilage development. (B) Enriched biological processes in OST included bone mineralization and osteoblast differentiation. (C) Enriched terms in MAT included biological processes important during cartilage and bone development as well as, cell maturation, cell migration, and apoptosis.

### 3.3.5. Two independent portions of the GRN driving IMM and OST overlap partially in MAT

To determine whether portions of the GRN driving IMM and OST could possibly be interacting in MAT, the GRN driving the differentiation of these skeletal cells was estimated based upon co-expression analysis of all unique and differentially expressed (including pairwise comparisons) genes (DEGs). Gene co-expression networks (GCN) are graphs composed of edges that connect highly co-expressed genes (McCall 2013; Khosravi et al. 2015). Since GCNs can provide more information than clustering, enabling a geometric interpretation, and can be used to predict GRNs (Mahanta et al. 2012; McCall 2013), this approach was applied to the LCM-RNA-seq data obtained from all three cell types. A total of 3292 DEGs were included in this analysis. GCN analysis identified two large networks (i.e., inferred GRN portions) of positively correlated genes (i.e. genes positively interacting; Fig. 15A, red lines in GRN) that were negatively correlated with each other (i.e. genes negatively interacting; Fig. 15A, blue lines in GRN). MAT-enriched genes were in both networks, whereas genes enriched in IMM and OST were located in opposite networks (Fig. 15A, Fig. 16A). The yellow bridging portion of the GCN contained several hallmark mature chondrocyte genes, such as *Col10a1* and *Ihh*, as well as the novel MAT genes whose expression was validated, such as *Parm1*, *Prom1*, *Ttll3*, *Rhox8*, and *Pde11a* (Fig. 15A, Fig. 16A). Together, GCN analyses clearly demonstrated that portions of the GRN operating in IMM and OST, interact (yellow bridging portion) in MAT.

How well do *Sox9* and *Runx2* expression patterns support the idea that SOX9 directs one portion of the GRN and RUNX2 directs the other? GCN analysis placed *Sox9* and *Runx2* in separate portions of the GRN. *Sox9* was placed in the enriched IMM portion of the network and *Runx2* in the OST enriched portion of the network. Importantly, *Sox9* and *Runx2* were positively correlated with genes of their own network and negatively correlated with genes closer to the other network (Fig 15A,A',A''). Minimized to nearest neighbor nodes, *Sox9* had a higher degree of centrality (i.e., more correlated genes) compared to *Runx2* (135 vs. 18, respectively; Fig 15B,C, Fig. 16B). Many genes known to be associated with *Sox9* or *Runx2* were in the appropriate estimated networks. For example, several genes positively correlated with *Sox9* included classical cartilage genes (e.g., *Col2a1*, *Col9a1*, and *Acan*) that are under SOX9 transcriptional control (Lefebvre et al. 1997; Sekiya et al. 2000; Zhang et al. 2003). On the other hand, many genes related to bone formation were

negatively correlated with *Sox9* (e.g., *Dlx2*, *Dlx4*, and *Msx1*; Satokata and Maas 1994a; Panganiban and Rubenstein 2002; Li et al. 2008; Nassif et al. 2014; Wu et al. 2015). Genes with a known role during bone formation were positively correlated to *Runx2*, including *Bmp7, Bambi,* and *Angpt1* (Fig. 15C; Tou et al. 2003; Luu et al. 2007; Higashihori et al. 2008; Kishiya et al. 2008; Park et al. 2016; Colden et al. 2017). Importantly, *Sox9* was differentially expressed in IMM and MAT, whereas *Runx2* was differentially expressed in MAT and OST (Tables 2,3). Other transcription factors included in this GRN are *Sox5*, *Sox6*, and *Runx3* which are upregulated in IMM and MAT compared to OST, and *Sp7* which is upregulated in OST compared to IMM and MAT (Tables 2,3). When comparing gene expression levels of classic skeletal transcription factors, the relationship between IMM and OST is mostly antagonistic (Tables 2,3). Together, these data support the idea that a transcriptional program mediated by SOX9 drives cartilage formation, a transcriptional program mediated by RUNX2 drives bone formation, and both operate during cartilage maturation. Similarly, three distinct regulatory states were identified, with MAT as an intermediate state between the other two cell types, using published regulatory connections among these genes as input to the GRN modeling software BioTapestry (Fig. 17; Longabaugh et al. 2005). The topology of the GCN output was the same using the BioTapestry set of genes (Fig. 18).

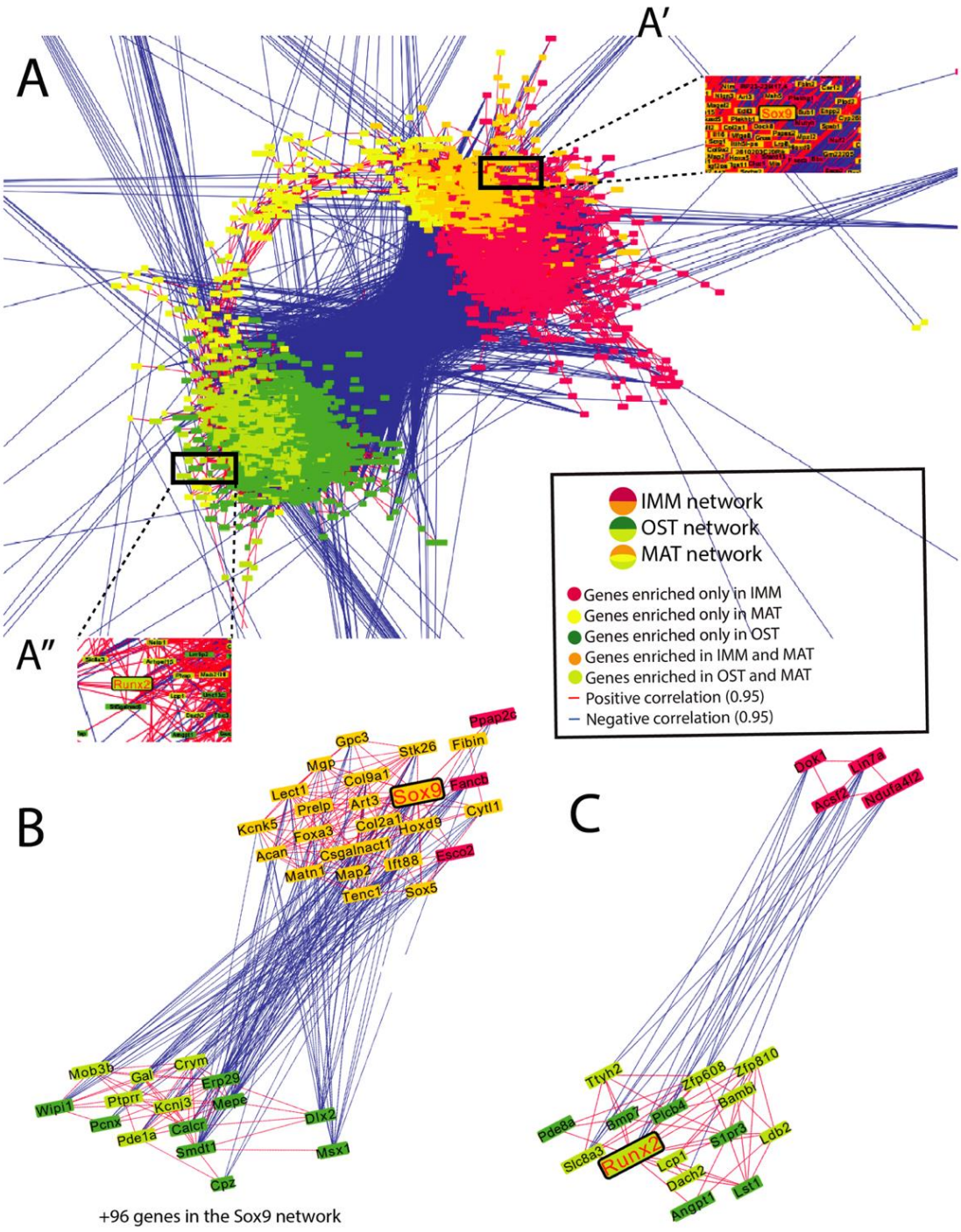



**Figure 15. Gene co-expression network (GCN) analyses in mouse skeletal cell types.** (A)Using GCN analyses, two portions of the GRN active in IMM and OST interact in MAT. One portion of the GRN contained the chondrogenic transcription factor SOX9, associated mostly with IMM genes (A'), while another portion of the GRN contained the osteogenic transcription factor RUNX2, associated mostly with OST genes (A''). Portions of the GRN active in IMM and OST operate in MAT (yellow bridging portion). (B) The isolated *Sox9* portion of the network showed a high degree of centrality (135), and many of its nearest neighbors (i.e. *Col2a1*, *Col9a1*, and *Acan*) are known to be under SOX9 transcriptional control. For visualization purposes, only a portion of the SOX9 network were included (see Fig. 16B for entire network). (C) The isolated *Runx2* portion of the network had 18 genes (Credit: Katie Ovens).

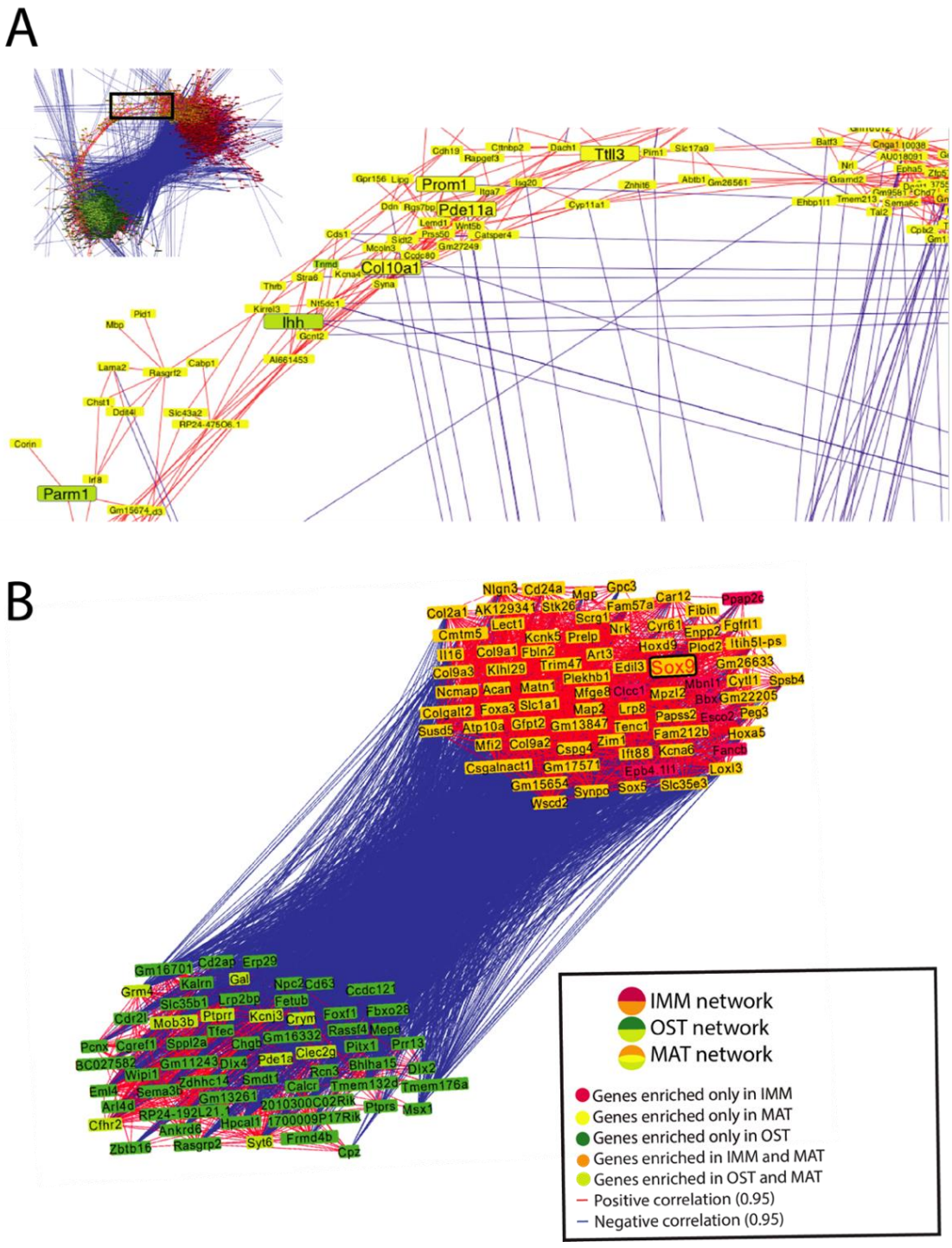



**Figure 16. Co-expression network analyses.** (A) The bridging portion of the GCN contained several mature chondrocyte genes, such as *Col10a1* and *Ihh*, as well as the novel MAT genes whose expression was validated, such as *Parm1*, *Prom1*, *Ttll3*, *Rhox8*, and *Pde11a*. (B) Genes associated with the *SOX9*-mediated portion of the GRN. Co-expression network analysis identified 135 Sox9 nearest neighbors. Two groups of positively correlated genes were identified. One side of the diagram included genes enriched in IMM and MAT, which are positively correlated with Sox9, whereas the other group included enriched genes in OST and MAT, which are negatively correlated with Sox9. Enriched genes in MAT are included in both groups of genes, meaning that this tissue expresses both chondrogenic and osteoblast markers (Credit: Katie Ovens).

**Table 2.** Pairwise comparisons showing differentially expressed transcription factors in the three skeletal tissues

### MAT vs IMM

| Gene | logFC | logCPM | PValue |
|---|---|---|---|
| ***Runx2*** | **2.95** | **7.53** | **4.86E-05** |
| *Sox5* | -2.13 | 6.93 | 1.92E-07 |
| *Sp7* | 5.42 | 7.81 | 1.12E-14 |
| *Sox8* | -2.42 | 4.12 | 3.00E-08 |

### OST vs IMM

| Gene | logFC | logCPM | PValue |
|---|---|---|---|
| ***Runx2*** | **3.91** | **7.53** | **2.44E-07** |
| ***Sox9*** | **-4.33** | **8.67** | **1.41E-16** |
| *Sox5* | -5.65 | 6.93 | 9.70E-32 |
| *Sox6* | -1.99 | 6.43 | 5.03E-06 |
| *Sp7* | 6.93 | 7.81 | 2.26E-20 |
| *Runx3* | -2.64 | 5.15 | 1.32E-06 |
| *Sox8* | -5.81 | 4.12 | 2.28E-28 |
| *Sox10* | -4.67 | 1.51 | 8.35E-06 |

### OST vs MAT

| Gene | logFC | logCPM | PValue |
|---|---|---|---|
| ***Sox9*** | **-3.48** | **8.67** | **5.05E-12** |
| *Sox5* | -3.52 | 6.93 | 1.34E-15 |
| *Runx3* | -2.94 | 5.15 | 1.04E-07 |
| *Sox8* | -3.38 | 4.12 | 1.52E-11 |

**Table 3.** Expression levels of common skeletal cell transcription factors. The table includes averaged normalized counts in the three different biological replicates of each skeletal cell type.

| TFs | IMM1 | IMM2 | IMM3 | MAT1 | MAT2 | MAT3 | OST1 | OST2 | OST3 |
|---|---|---|---|---|---|---|---|---|---|
| *Runx1* | 61 | 94 | 61 | 143 | 271 | 173 | 113 | 152 | 133 |
| ***Runx2*** | **500** | **23** | **434** | **3131** | **2606** | **1690** | **5420** | **4989** | **4315** |
| *Runx3* | 831 | 159 | 831 | 730 | 829 | 691 | 85 | 91 | 124 |
| *Sox10* | 76 | 33 | 161 | 11 | 2 | 40 | 2 | 2 | 6 |
| *Sox5* | 4120 | 3008 | 4891 | 992 | 727 | 1021 | 97 | 74 | 74 |
| *Sox6* | 2139 | 2936 | 1444 | 769 | 539 | 1168 | 571 | 509 | 585 |
| *Sox8* | 587 | 468 | 706 | 109 | 112 | 107 | 8 | 12 | 12 |
| ***Sox9*** | **11857** | **12345** | **7193** | **6401** | **2712** | **8160** | **405** | **630** | **552** |
| *Sp7* | 128 | 15 | 26 | 1808 | 2751 | 2576 | 7472 | 6787 | 6589 |

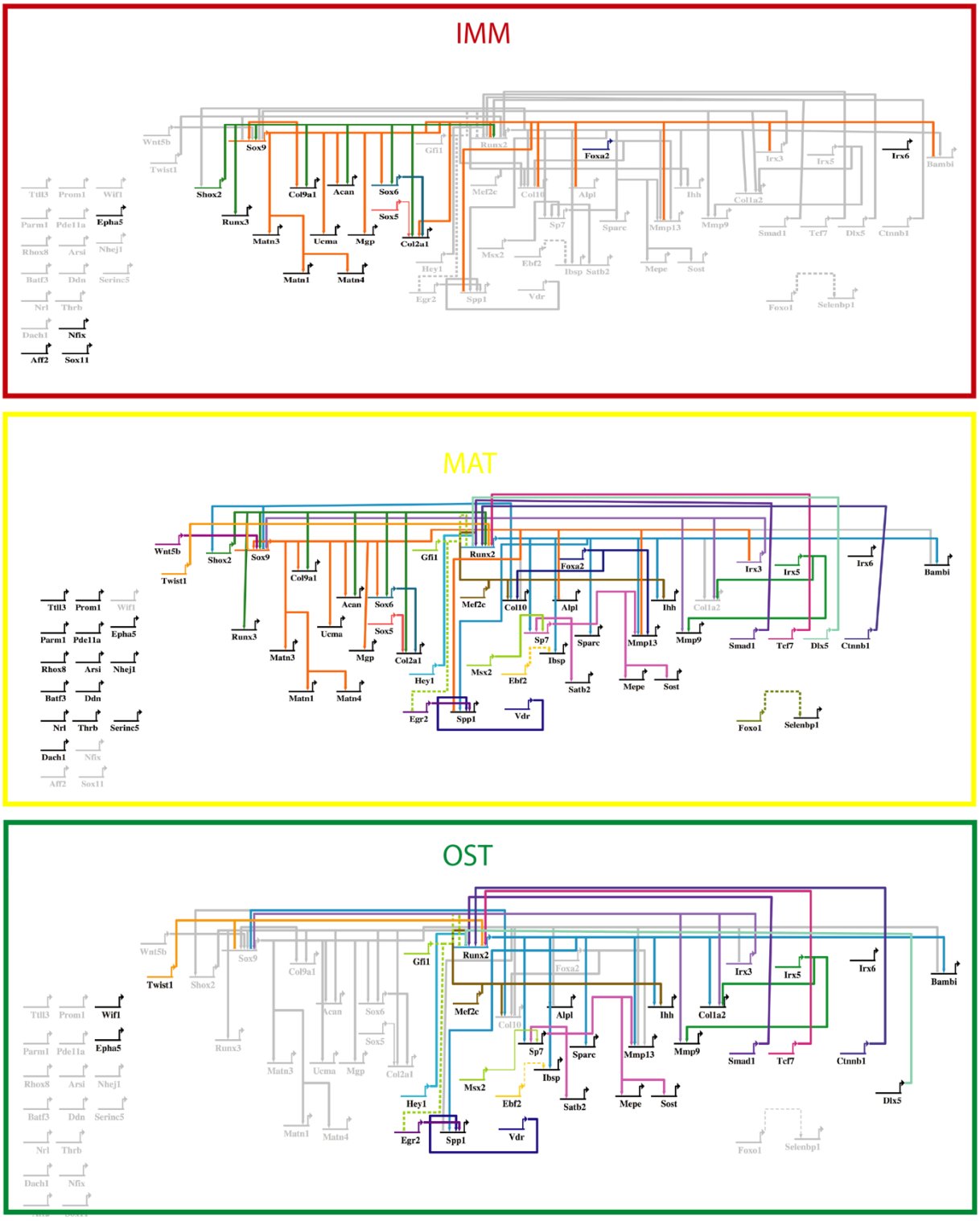



**Figure 17. Validated gene regulatory network of vertebrate skeletal cells shows that the MAT regulatory state is an overlap of gene expression between IMM and OST states.** Interaction data was compiled from the literature. Regulatory interactions are separated in three categories: genes enriched in IMM, genes enriched in MAT, and genes enriched in OST. Novel transcription factors identified by GCN analyses are included as single-nodes to the left of the network. Arrowheads represent positive interaction, whereas negative interactions are depicted as –|. Interactions included in this network could be direct or indirect. Inactive nodes are represented in gray.

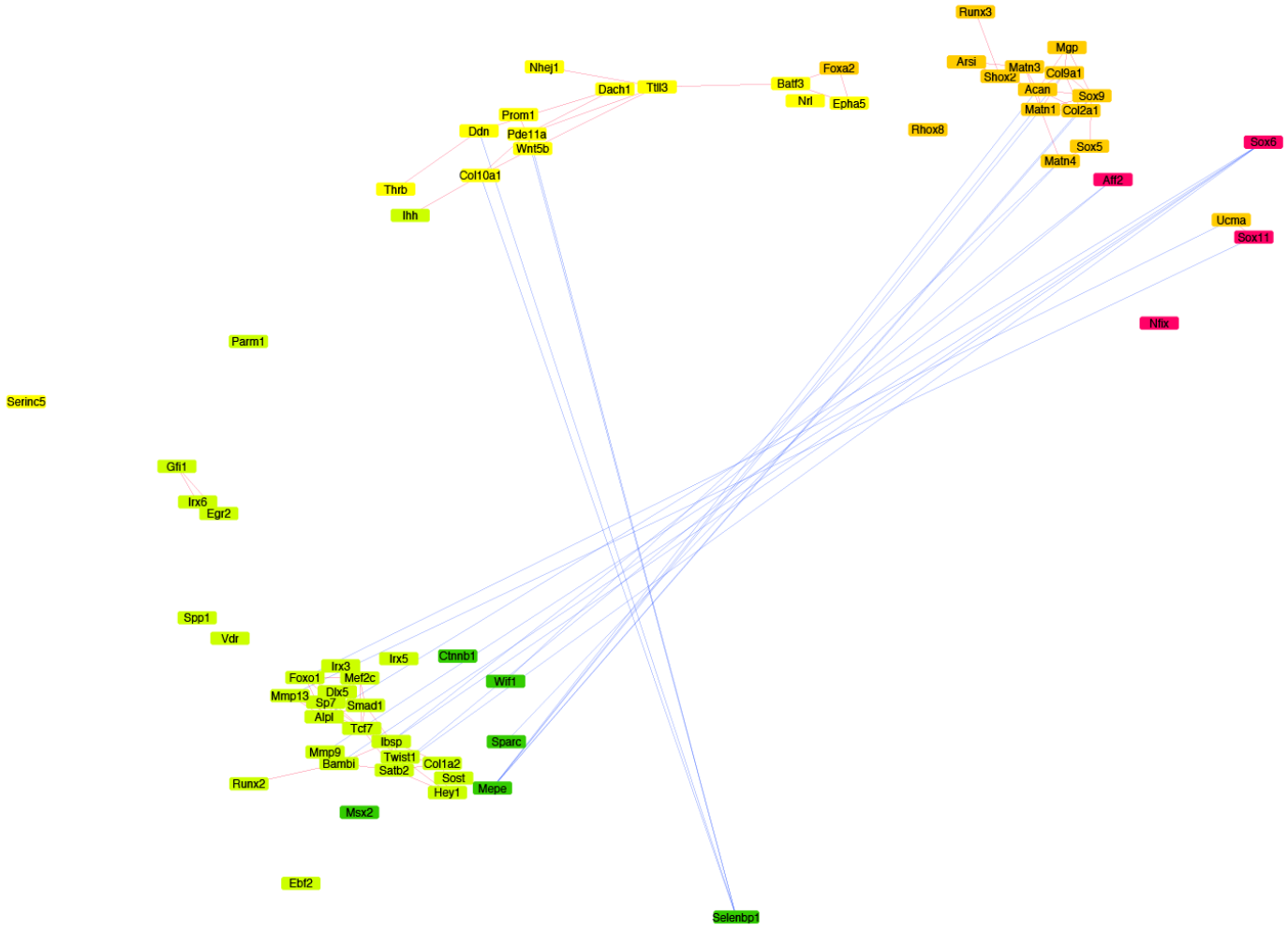



**Figure 18. GRN structure remains the same after network validation.** Network diagram in cytoscape after validating regulatory connections in BioTapestry (Credit: Katie Ovens).

### 3.3.6. Portions of the GRN active in IMM and OST were expressed independently in MAT, whereas other portions suggested interaction between distinct transcriptional programs via averaging and synergism

These analyses of *in vivo* LCM-RNA-seq data suggested that MAT somehow utilizes portions of the GRN present in IMM and OST, but the exact molecular mechanisms of this regulation are completely unknown. To test whether portions of the GRN driving IMM and OST might be

interacting in MAT, model-based cluster analysis was performed, identifying specific categories of gene expression (Fig. 19A). In some clusters, gene expression in MAT was statistically similar to one of the other cell types (IMM or OST; Fig. 19D,H). For example, cluster 7 had expression levels in MAT statistically similar to IMM, with both MAT and IMM being statistically different from OST. Cluster 3 had expression levels in MAT that were indistinguishable statistically from OST, and both MAT and OST were statistically different from IMM. Cluster 3 included *Runx2*, as well as important mineralization genes and classic osteoblast differentiation markers, such as *Mmp13* and *Spp1/Osteopontin*, that are regulated by RUNX2 (Drissi et al. 2000; Komori 2010). These clusters suggested that portions of the GRN active in IMM and OST do not interact in MAT, but rather are expressed independently.

Other clusters provided evidence that portions of the GRN operating in IMM and OST interacted in MAT. For instance, clusters 1 and 4 had gene expression in MAT that was statistically indistinguishable from the average (AVE) gene expression between IMM and OST, meanwhile being statistically different from either IMM or OST (Fig. 19B,E). Cluster 1 included *Sox9*, as well as classical genes involved in chondrogenesis, such as *Col2a1*, *Col9a1*, *Acan, Mgp,* and *Sox5*, that are regulated by SOX9 (Mead et al. 2013; Oh et al. 2014; Ohba et al. 2015). Other clusters also showed intermediate expression in MAT, compared to IMM and OST, even though they were not statistically similar to the AVE (i.e. average expression between IMM and OST; Fig. 19C,F,G,J,K). Cluster 8 had expression levels in MAT that suggested a synergistic interaction between IMM and OST GRN portions (Fig. 19I). Cluster 8 contained the hallmark chondrocyte maturation markers *Col10a1* and *Ihh* (Kielty et al. 1985; St-Jacques et al. 1999; Zheng et al. 2003), which showed enhanced expression in MAT (Table 1). In general, these data identified specific genomic loci that are predicted to have two distinct molecular mechanisms of GRN interaction: averaging and synergism. In summary, GRN portions directing IMM and OST that are associated with *Sox9* and *Runx2* expression, respectively, regulate MAT gene expression in both independent and interactive manners.

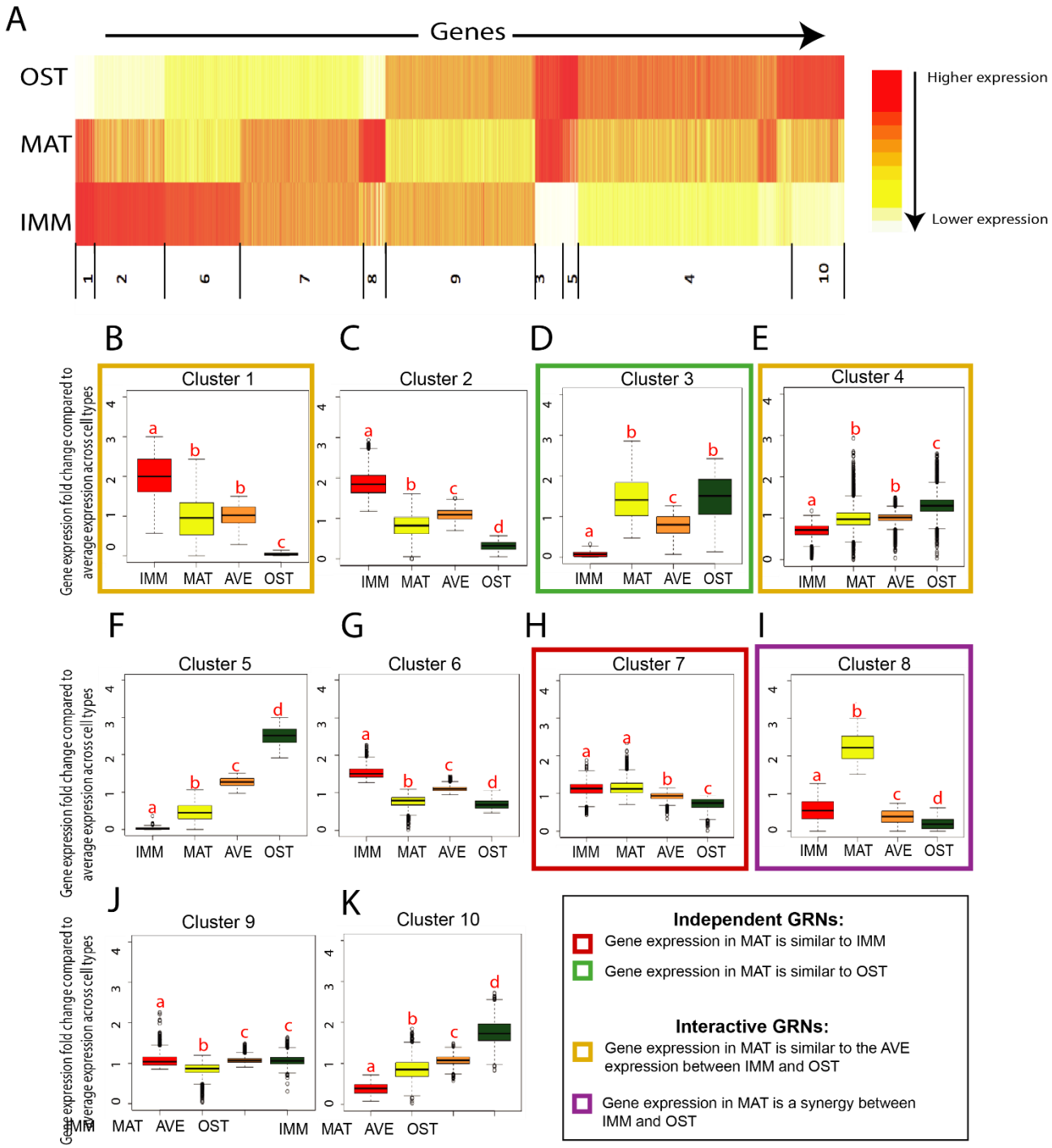



**Figure 19. Model-based clustering demonstrated that groups of IMM, MAT, and OST genes had discrete categories of expression.** (A-K) Heatmap and box plots showed the distribution of gene expression change among cell types compared to the average expression across all three cell types. AVE represented average expression of IMM and OST. MAT expression in cluster 3 (D) was significantly similar to OST, whereas MAT expression in cluster 7 (H) was significantly similar to IMM, suggesting that p IMM- and OST-enrieched portions of the GRN operate independently in MAT. (B, E) In clusters 1 and 4, gene expression in MAT was statistically indistinguishable from AVE, suggesting that portions of the GRN active in IMM and OST interact via averaging in MAT. (I) Cluster 8 had statistically higher gene expression in MAT than in either IMM or OST, suggesting that IMM- and OST-portions of the GRN interact synergistically in MAT (Credit: Katie Ovens).

### 3.3.7. EBF2 and IRX6 are putative novel regulators of chondrocyte maturation

To test whether other transcription factors besides *Sox9* and *Runx2* might be regulating mature chondrocyte differentiation and to provide functional verification of our data, cell transfection experiments were performed with two transcription factors upregulated in MAT compared to IMM, Early B cell factor 2 (*Ebf2*) and Iroquois homeobox 6 (*Irx6*). *Ebf2* has a role in regulating the osteoblast-dependent differentiation of osteoclasts (Kieslinger et al. 2005), but its role in cartilage maturation is unknown. Likewise many Iroquois genes have roles in skeletal development (Askary et al. 2015; Cain et al. 2016; Tamamura et al. 2017), but the function of *Irx6* during cartilage maturation remains to be revealed.

Micromass cultures of the ATDC5 chondrogenic cell line (Yao and Wang 2013) were transfected with *Ebf2* and *Irx6* to confirm their functional roles during chondrocyte maturation (Fig. 20). The ATDC5 cell line has been derived from mouse teratocarcinoma cells, and it is a well-established in vitro model for chondrocyte differentiation. Numerous studies have used ATDC5 cells to understand the mechanisms driving endochondral ossification since this cell line mimics the different stages involved in this type of ossification such as, mesenchymal condensation, chondrocyte differentiation, hypertrophy, and matrix mineralization (Yao and Wang 2013). Micromass cultures were analyzed at days 7 and 14 and imaged by fluorescent microscopy which showed that both EBF2 and IRX6 remain highly expressed at these time points, confirmed by the positive expression of mCherry and RT-PCR (Fig. 20A,B). Chondrogenic differentiation in micromass cultures occurs by day 7 which is characterized by an increased expression of classic immature cartilage genes, such as *Col2a1* and *Acan* (Tanoue et al. 2018). Then, by day 14 the onset of cartilage matrix mineralization occurs and expression of the hallmark mature cartilage marker *Col10a1* reaches its highest level (Newton et al. 2012; Li et al. 2016). Chondrogenic differentiation of transfected micromass cultures at day 14 was evaluated using different staining techniques. Alcian blue staining confirmed the presence of secreted sulfated proteoglycans in both GFP- transfected control (Fig. 21A,D) and EBF2- and IRX6-transfected groups (Fig. 21B,C-E,F). No differences in Alcian blue staining were observed between EBF2/IRX6 groups and GFP-controls. Alkaline Phosphatase (ALP) staining activity was also evaluated in both GFP-controls and EBF2/IRX6 experimental groups. An increase in ALP stained areas was observed in EBF2/IRX6 groups relative to GFP-controls suggesting increased mineralization, but no statistically significant difference was

found between groups ($p>0.05$; Fig. 21M). No Alizarin red staining was confirmed in any of the groups by day 14 (Fig. 21G-L). Differentiation of these cultures was also confirmed by assessing the expression of important maturation and mineralization markers, including *Col10a1*, *Ihh*, *Mmp13* and *Spp1* (Fig. 22). RT-PCRs are not quantitative, therefore *Ebf2/ Irx6* groups were not assessed for differences in gene expression relative to *GFP*-controls. Since *Ebf2* and *Irx6* showed higher fold changes in MAT compared to IMM in our data (*Ebf2*, LogFC 7.86; *Irx6*, LogFC 8.034) these transcription factors are believed to accelerate cartilage maturation perhaps by regulating the expression of some classic mature cartilage genes.

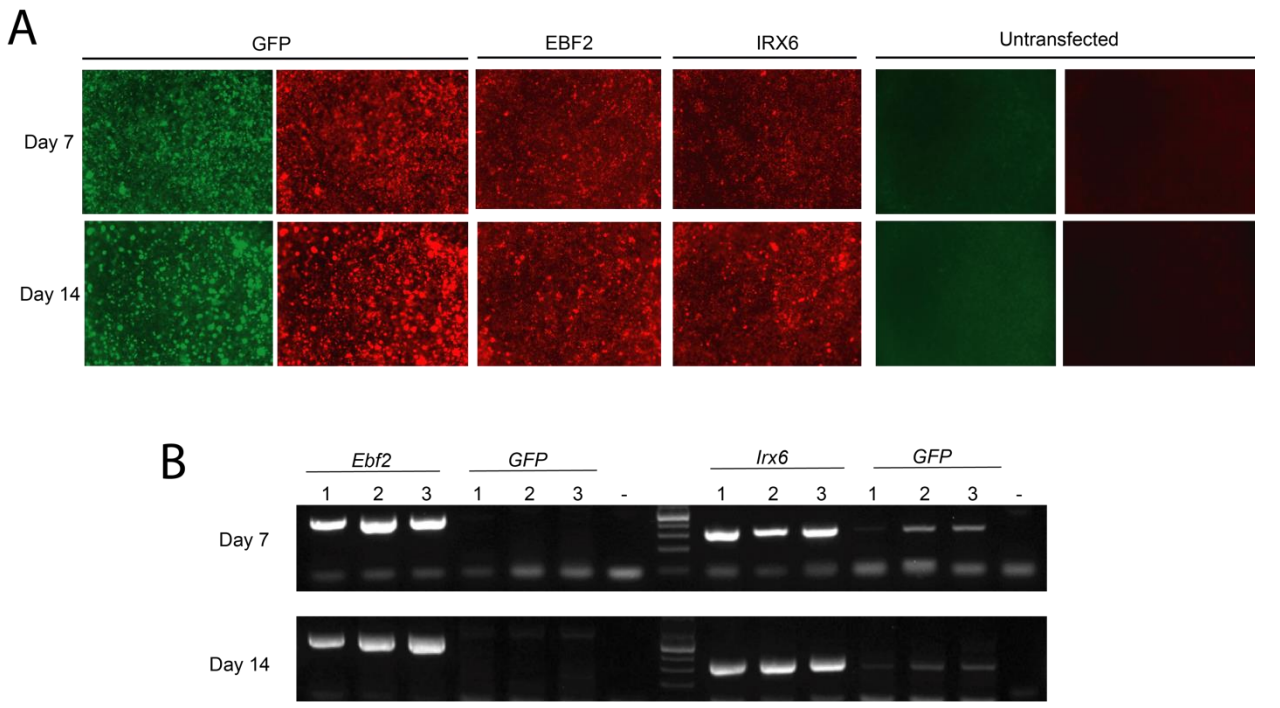



**Figure 20. Overexpression of EBF2 and IRX6 in micromass cultures of ATDC5 cells.** (A) Protein expression of GFP (control), EBF2, and IRX6 in micromass cultures of ATDC5 cells at days 7 and 14. The genes *GFP*, *Irx6* and *Ebf2* were all transfected in the pP2A-mCherry-N2 vector which enables bicistronic expression of protein of interest and mCherry (see methodology for more details). Untransfected controls do not show expression of GFP or mCherry because the vector was not overexpressed in these cell types. No fluorescence above background was detected in untransfected controls. (B) RT-PCR confirming the overexpression of Ebf2 and Irx6 in micromass cultures at days 7 and 14. Expression levels of *Ebf2* and *Irx6* in transfected cells are higher compared to the GFP- transfected control, which expresses low levels of endogenous *Ebf2* and *Irx6*.

Immunofluorescence quantitation results provided preliminary evidence that overexpression of EBF2 and IRX6 in ATDC5 micromass cultures led to an increase in COLX expression area,

compared to GFP-transfected controls (Fig. 23A,B). These results suggest that EBF2 and IRX6 might be directly or indirectly influencing COLX expression, possibly in coordination with SOX9 and RUNX2 (Fig. 24), but future *in vivo* functional studies can further confirm the role of Ebf2 and Irx6 during cartilage maturation.

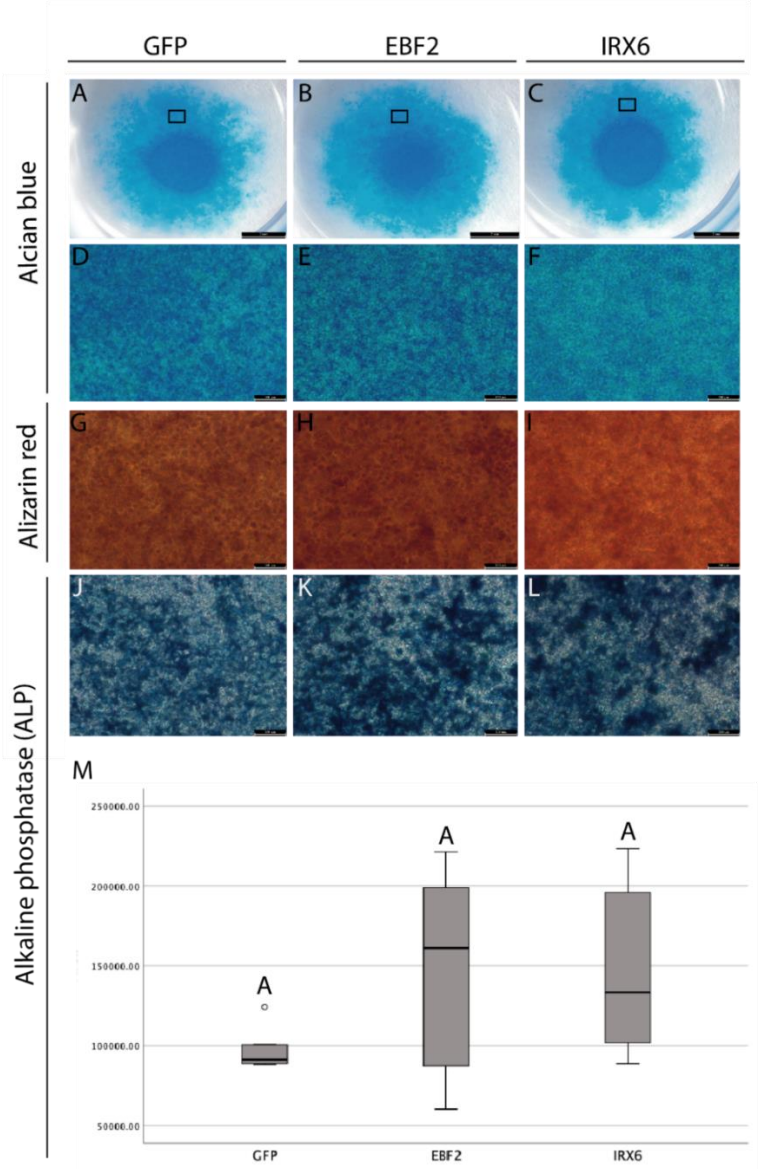



**Figure 21. Stained Alkaline phosphatase area increases in EBF2/IRX6 experimental groups relative to GFP-controls, but no differences in Alcian blue or Alizarin red staining patterns were observed between controls and experimental groups. (**A-F) Alcian blue staining of micromass cultures at day 14 confirmed the secretion of sulfated proteoglycans. (D-F) High magnification of boxes in (A-C). (G-I) Alizarin red positive staining was absent at day 14, only background staining was observed. (J-L) Alkaline Phosphatase activity (ALP) was confirmed (Scale bars: A-C, 2mm; D-L, 100 μm). (M) Box plots representing stained ALP area show an increase in stained areas in both EBF2 and IRX-transfected groups compared to the GFP-control but this trend did not reach statistical significance (p=0.05; n=6 per group).

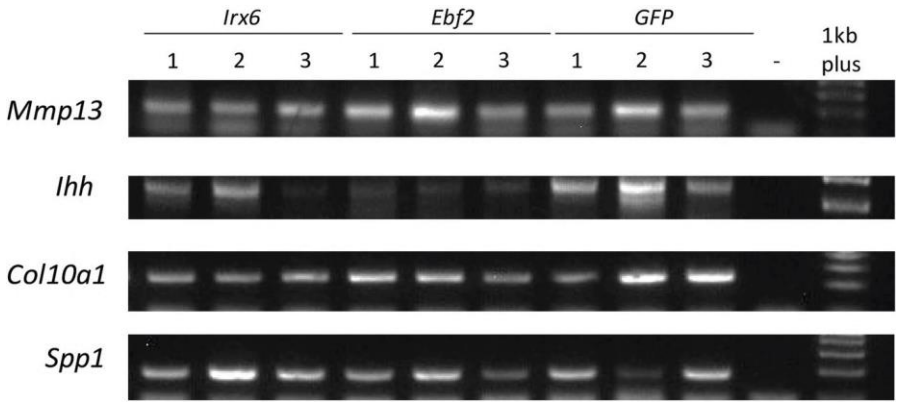



**Figure 22. RT-PCR confirmed the expression of important chondrocyte markers in transfected micromass cultures at day 14.**

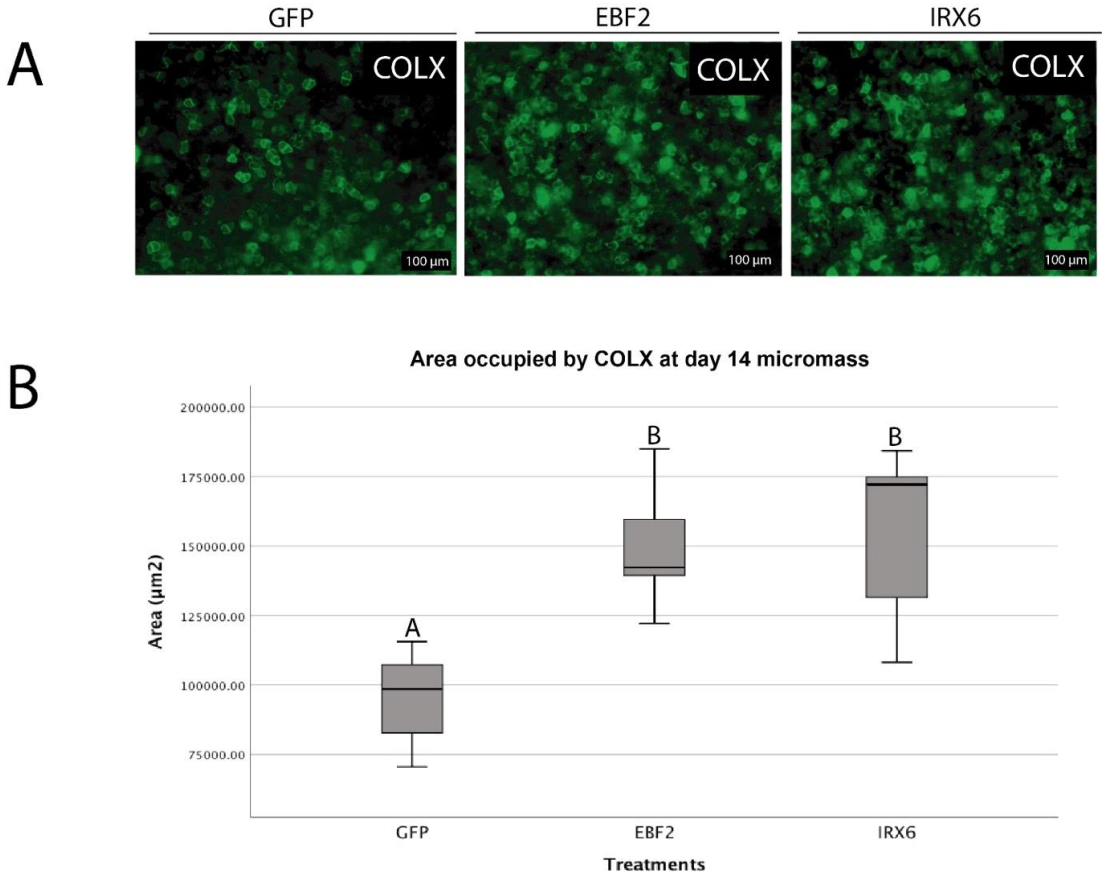



**Figure 23. Overexpression of EBF2 and IRX6 in ATDC5 micromass cultures led to statistically significant increases in COLX protein expression area.** (A) Expression of COLX in micromass cultures transfected with GFP, EBF2, and IRX6. (B) The immunostained COLX area in cells overexpressing EBF2 and IRX6 is significantly higher compared to the GFP-transfected control cells. A similar region in the micromass was quantified and compared between GFP control and experimental groups (n=5 per group).

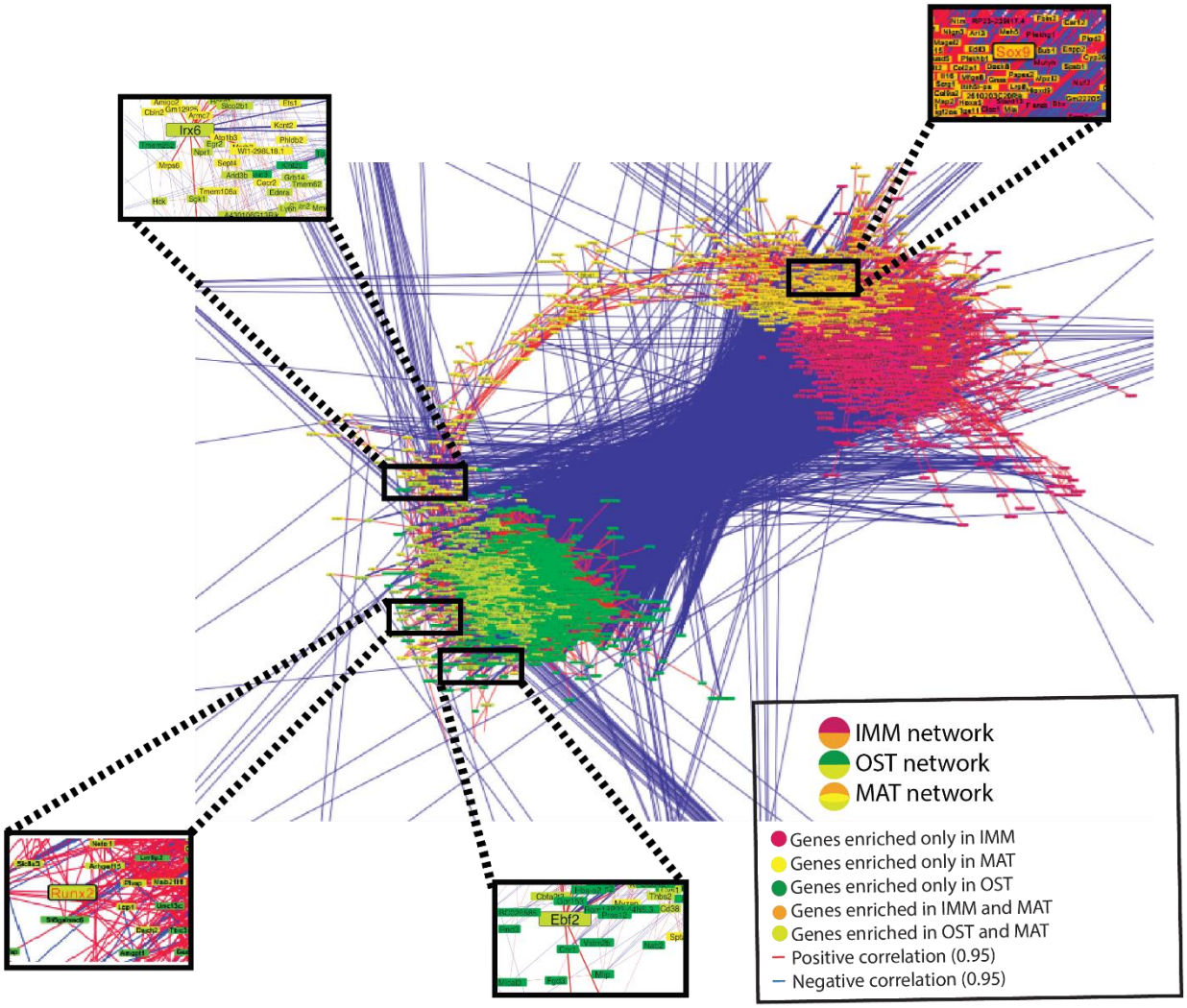



**Figure 24. EBF2 and IRX6 are potential novel transcription factors regulating gene expression in MAT in coordination with SOX9 and RUNX2.** Two transcriptional programs regulated by SOX9 and RUNX2 are present in IMM, MAT, and OST. GCN analysis indicate that the transcription factors EBF2 and IRX6 might be regulating the expression of genes in these cell types in coordination with SOX9 and/or RUNX2 (Credit: Katie Ovens).

## 3.4.DISCUSSION

A deeper understanding of biological processes can come from comparing transcriptomes of different cells. Here, LCM-RNA-seq was used to characterize the *in vivo* transcriptomic profiles of three related cell types: immature chondrocytes (IMM), mature chondrocytes (MAT), and

osteoblasts (OST). In addition to identifying a number of genes previously unassociated with skeletal cells, our findings shed light on two main issues regarding the molecular mechanisms of skeletal cell differentiation: the structure of the underlying GRN(s) and regulatory control of MAT gene expression.

Skeletal cells differentiate from common osteochondroprogenitors (Day et al. 2005), so one GRN might regulate differentiation of these progenitors along three developmental pathways. Indeed, almost 80% of genes expressed in IMM, MAT, and OST were shared (Fig. 11). Also, MAT gene expression was largely an overlap of genes expressed in IMM or OST, since it had the fewest unique genes and DEGs (Figs. 11,12, and 13). The fact that MAT expressed an overlapping subset of IMM and OST genes also was illustrated by evaluation of GRN structure only using genes with regulatory connections that had been confirmed by functional molecular genetic experiments (Figs. 17&18). GCN analyses estimated one GRN, and distinct portions active in each cell type. Two large transcriptional programs (i.e. SOX9 and RUNX2) active in distinct portions of the GRN were identified in these three cell types, with genes that were unique or differentially expressed in MAT located in a bridging portion between those two programs (Fig. 15). GCN analyses cannot identify all regulatory relationships of a GRN, so functional experiments should be carried out in the future to validate the GRN estimated for skeletal cells here (Azpeitia et al. 2010; Streit et al. 2013; He et al. 2017).

The transcriptome data presented here supported a binary code of SOX9 and RUNX2 in determining skeletal cell differentiation (Eames et al. 2004), suggesting that these chondrogenic and osteogenic transcription factors dominate the underlying GRN (Cole 2011; Kerkhofs et al. 2012; Wu et al. 2014; Gomez-Picos and Eames 2015; Ohba et al. 2015). Consistent with previous studies (Hall 1978; Bi et al. 1999; Akiyama et al. 2002; Lian and Stein 2003; Eames and Helms 2004; Eames et al. 2004; Yoshida et al. 2004; Zhou et al. 2006; Cheng and Genever 2010), *Sox9* was upregulated in IMM and MAT, compared to OST, whereas *Runx2* was upregulated in MAT and OST, compared to IMM (Tables 2,3). Using LCM-RNA-seq on *Sox9* and *Runx2* gain- and loss-of-function tissues *in vivo* would help to validate the estimated transcriptional programs active in this GRN, but SOX9 or RUNX2 regulate many genes that were highly correlated to *Sox9* and *Runx2* expression. For example, *Sox9* was positively correlated with, and SOX9 is known to bind enhancers and promote transcription of, many cartilage genes, such as *Col2a1*, *Col9a1*, and *Sox5* (Lefebvre et al. 1997; Akiyama et al. 2002; Zhang et al. 2003). Likewise *Runx2* expression was correlated with important

cartilage maturation and bone genes that RUNX2 regulates, including *Col10a1*, *Mmp13*, and *Ibsp* (Wang et al. 2004a; Komori 2010). The present study expands current understanding of the skeletal cell GRN *in vivo*, which was previously estimated using published gene interaction data (Cole 2011; Kerkhofs et al. 2012). Seven transcription factors in the previous chondrocyte GRN implicated in cartilage differentiation (*Runx3*, *Foxa2*, *Foxa3*, and *Sox8*) and hypertrophy (e.g. *Dlx6*, *Atf3*, and *Irx3*) had no known regulatory connections (Cole 2011). These transcription factors were incorporated into the GRN estimated here. Minimal overlap was found between the *in vivo* GCN reported here and published SOX9 and RUNX2 target gene datasets in skeletal cells (data not shown; Wu et al. 2014; Ohba et al. 2015). This discordance could be explained by limited statistical power of the published datasets, or the different techniques used to isolate cell types of interest.

Although both SOX9 and RUNX2 are required for cartilage maturation (Komori et al. 1997; Bi et al. 1999; Nakashima et al. 2002; Karsenty et al. 2009; Dy et al. 2012), the analyses presented here suggested distinct molecular mechanisms through which the SOX9 and RUNX2 transcriptional programs might interact in this process. Several DEGs and unique genes in MAT were highly correlated with both *Sox9* and *Runx2* expression, so SOX9 and RUNX2 might be able to regulate coordinately some of the same genes. Indeed, both transcription factors promote the expression of the hallmark chondrocyte maturation gene *Col10a1* (Zheng et al. 2003; Dy et al. 2012). In other cases, SOX9 and RUNX2 might have antagonistic effects on a given gene. For example, SOX9 inhibits, while RUNX2 promotes, the expression of MAT (and OST) genes, such as *Spp1*, *Mmp13*, *Alpl*, *and Vegfa* (Drissi et al. 2000; Zelzer et al. 2001; Zheng et al. 2003; Wang et al. 2004a; Hattori et al. 2010; Komori 2010; Leung et al. 2011; Peacock et al. 2011; Dy et al. 2012; Weng and Su 2013; Herlofsen et al. 2014).

Are there other transcription factors driving a portion of the GRN operating specifically in mature chondrocytes? Functional data overexpressing two MAT-enriched transcription factors, EBF2 and IRX6, in the chondrogenic cell line ATDC5 provided preliminary evidence that these transcription factors might be directly or indirectly regulating expression of COLX in MAT (Fig. 23). Increased ALP staining was also observed in EBF2/IRX6 groups relative to GFP-controls, suggesting that these transcription factors might also accelerate matrix mineralization, even though no statistically significant differences were found between groups (Fig. 21M). Both *Ebf2* and *Irx6* are negatively correlated to genes regulated by SOX9 and positively correlated with genes regulated by RUNX2 (Fig. 24), so these results suggest that these transcription factors do not seem to be

regulating MAT differentiation independently. IRX6 might be part of a repression circuit that enables the transition from IMM to MAT phenotypes. The general relevance of a repression circuit was demonstrated by several negative interactions GRN portions including SOX9 and RUNX2 in our GCN analyses (Fig.13). Future gain- and loss-of-function experiments coupled to quantitative protein assays can confirm whether *Ebf2* and *Irx6* (and other novel MAT-enriched transcription factors included in this GRN) are indeed altering COLX expression levels and influencing cartilage maturation.

Model-based clustering analysis identified the same exact categories of gene expression in MAT, where different portions of one GRN interact, that were described in cells of hybridized organisms, where interaction between portions of different GRNs would be expected. For example, in hybrid species of cotton, expression levels of some groups of genes were similar to either one or the other parent (Hu et al. 2016), suggesting independent GRN activity (Fig. 25A). Other groups of genes in the hybrid either had expression levels that were an average of the parent species or above parental levels (Hu et al. 2016), suggesting GRNs interact via averaging or synergism, respectively (Fig. 25B). Expanding the generality of this phenomenon, identical trends in gene expression also were described in hybrid species of hamster (Brekke et al. 2016). Importantly, the presence of synergistic genes in MAT argues strongly that, despite the heterogeneity of captured mature chondrocytes, MAT expression did not simply represent an amalgamation of portions of the GRN active in IMM and OST. An intriguing possibility for future research is that these categories of gene expression might predict specific groups of genes that should have distinct molecular mechanisms of regulation. Loci showing independent activity of GRN portions should share specific features of regulatory architecture, perhaps using binding sites of transcription factors exclusive to one portion of the GRN. In the case of interacting GRN portions, transcription factors from different portions of the GRN might antagonize each other at loci showing average levels of gene expression, or they might bind cooperatively at loci showing synergistic expression levels. The LCM-RNA-seq approach utilized here establishes a novel, unbiased system through which GCN analysis can estimate GRNs in any cell type *in vivo*, meanwhile identifying specific genomic loci predicted to have unique molecular mechanisms of interaction between distinct portions of a GRN.

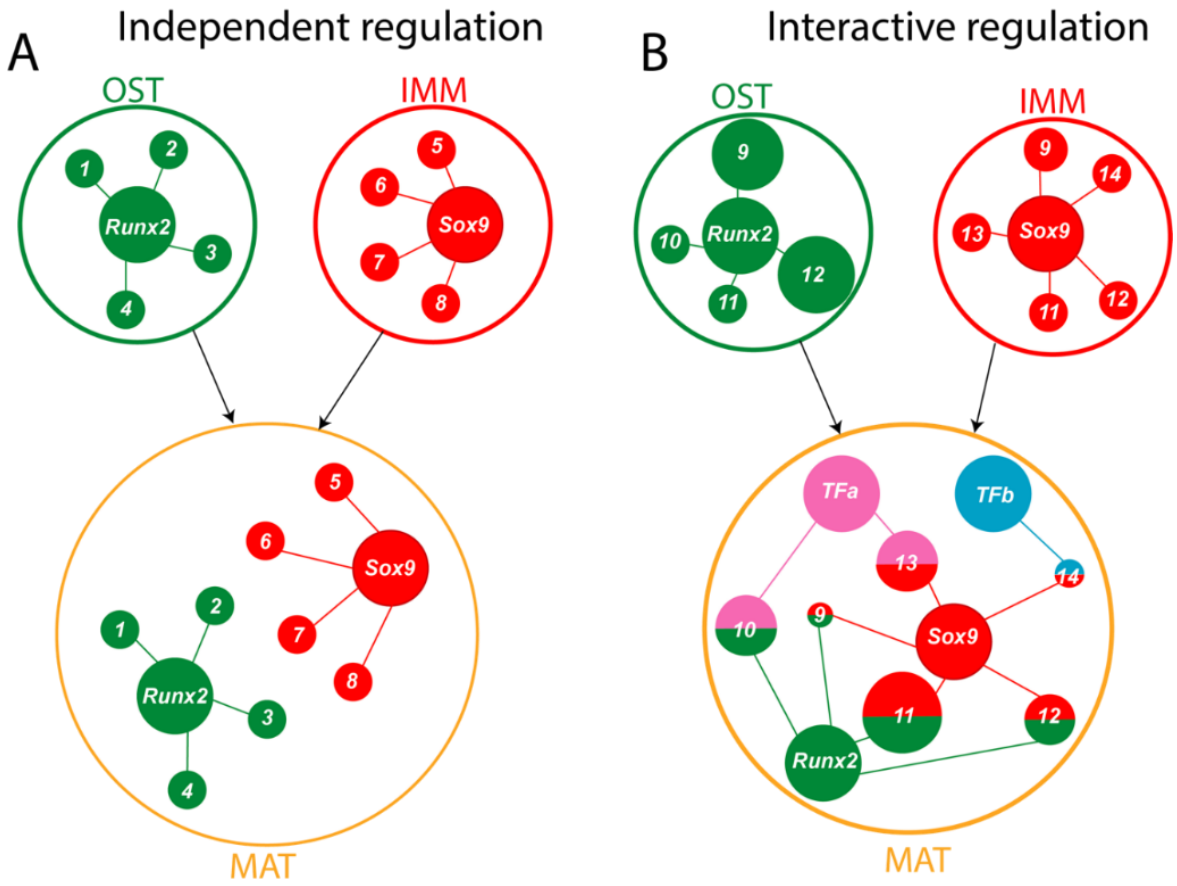



**Figure 25. Model based upon LCM-RNA-seq data showing independent and interacting portions of the GRN in MAT.** (A) Independent regulation: portions of the GRN including SOX9 from IMM and the portion of the GRN containing RUNX2 from OST do not interact in MAT, resulting in gene expression in MAT that is similar to either IMM or OST. (B) Interactive regulation: portions of the GRN including SOX9 from IMM and the GRN portion including RUNX2 from OST interact in MAT. In some cases, genes in MAT showed an average expression between levels seen in IMM and OST (e.g. *gene 12*). In other cases, genes in MAT showed synergistic effects compared to expression levels in IMM or OST (e.g., SOX9 and RUNX2 cooperatively repress or promote expression, such as *genes 9* or *11*, respectively). Other MAT transcription factors, *TFa* and *TFb*, might also interact with SOX9 and RUNX2, enhancing (e.g. *genes 10* and *13*) or repressing (e.g. *gene 14*) the expression of MAT genes. Pink indicates regulation by *TFa*, whereas blue indicates regulation by *TFb*. Abbreviation: TF, transcription factor.

## 3.5. METHODS

**Embryo Collection and tissue processing.** All animal procedures were performed according to guidelines approved by the University of Saskatchewan Animal Care and Use Committee. C57BL/6J mice were obtained from The Jackson Laboratory (Stock# 000664). Pregnant females at E14.5 were euthanized by $CO_2$ asphyxiation and embryos were collected in cold phosphate buffered saline (PBS) solution. The (right and left) humeri and lower jaw were dissected from E14.5 embryos in an RNase free environment, and then immediately placed in OCT (optimum cutting temperature) compound (Tissue-Tek; Sakura Finetek USA, Torrance, CA) using standard cryomolds (Tissue-Tek; Sakura Finetek USA, Torrance, CA). Tissues samples were snap frozen in OCT medium by using isopentane/ liquid nitrogen, and then stored at -80°C.

**Alcian blue/ Alizarin red whole mount skeletal stain.** Mouse E14.5 embryos were fixed using 4% paraformaldehyde, and then stained with Alcian blue and Alizarin red using an acid-free solution that included $MgCl_2$ to differentiate staining, and then cleared in glycerol/KOH as described elsewhere (Eames et al. 2011).

**Histology.** Safranin O/Fast Green staining was performed on 10 μm thick frozen sections of the E14.5 mouse humeri, as described previously (Ferguson et al. 1998). Trichrome staining was performed on 10μm thick frozen sections of the E14.5 mouse mandible, as described previously (Clark and Smith 1993). Humeri were sectioned longitudinally whereas the mandible was sectioned along the transverse plane.

**RNA in situ hybridization.** *In situ* hybridization was performed on mouse E14.5 humeri 10 μm thick frozen sections. *Pde11a* (NM_001081033.1; nt 1-695) was synthesized by Bio Basic (Canada). Plasmids containing the mouse genes, *Col10a1* (Lu et al. 2005), *Ihh* (Spater et al. 2006), *Parm1* (Wittler et al. 2012), *Ttll3* (Bosch Grau et al. 2013), *Prom1* (Holmberg Olausson et al. 2014), and *Rhox8* (Brown et al. 2013) were kindly provided by several colleagues (see Acknowledgements). RNA probes were prepared from linearized plasmids by *in vitro* transcription using T7, Sp6, and T3 RNA polymerases, and labeled with digoxigenin. *In situ* hybridization was performed as described previously (Zakin and De Robertis 2004), with some modifications. Briefly, slides were treated with

0.2N HCl for 10 min, followed by proteinase K (1 μg/ml) for 15 min at 37 °C, and postfixed in 4% paraformaldehyde. Hybridization was carried out with RNA probes overnight at 58°C. After hybridization, unbound probes were washed and incubated in blocking solution (Roche). Slides were then incubated with anti-DIG antibody coupled to alkaline phosphatase (Roche) followed by the addition of BM purple substrate (Roche).

**Preparation of histological sections for laser capture microdissection (LCM).** Tissue sections were cut on a cryostat (Microm GmbH; ESBE Scientific, Walldorf, Germany), and mounted on PALM MembraneSlides (P.A.L.M. Olympus - Microlaser Technologies, Bernried, Germany). Humeri were sectioned longitudinally whereas the mandible was sectioned along the transverse plane.

**Laser Capture Microdissection (LCM).** Six to ten unfixed frozen sections (10μm/each) were collected on a single MMI MembraneSlide (Prod#50102; MMI Molecular Machines & Industries). On the day of laser capture, a single slide was removed from -80°C and allowed to thaw for no longer than 30 seconds. The slide was transferred to 70% ethanol for 30 sec to allow fixation of sections and then rinsed in chilled diethylpyrocarbonate (DEPC)-treated water for 30 sec to remove residual OCT. The slide was dehydrated through ethanol series (75%, 95%, 100%, 100%) for 1 min each at room temperature and allowed to quickly air dry before proceeding immediately to laser microdissection. Since all relevant morphological features of skeletal tissues were easily identifiable under light microscopy, no staining step was included in the protocol which greatly reduced the risk of RNAse contamination. LCM was performed on a Laser Microdissection - Molecular Machines & Industries (MMI) CellCut apparatus. The capture of each tissue was done in triplicate from three distinct embryos. Immature and mature chondrocytes were captured from the developing E14.5 mouse humeri. In contrast, osteoblasts were isolated from the developing dentary in the jaw at E14.5. A larger capture area was necessary to obtain a sufficient amount of RNA from mature cartilage (both left and right humeri; Table 4). The captured cells were collected onto the inner lid of 0.5ml MMI IsolationCaps (either Diffuser caps (Prod#50202) or Transparent caps (Prod#50204; MMI Molecular Machines & Industries). Capture time per slide was restricted to 30 min in order to ensure good RNA quality from cells. Immediately following capture, 50 μL of PicoPure lysis buffer was

added to the microcentrifuge tube, which was then inverted so captured cells were submerged in lysis buffer, and then stored at -80°C.

**Table 4.** LCM Surface area obtained from mouse E14.5 skeletal tissues

| Skeletal element | Tissue | Surface Area captured ($\mu m^2$) |
|---|---|---|
| Humerus | IMM 1 | 451,440 |
| Humerus | IMM 2 | 883,909 |
| Humerus | IMM 3 | 806,649 |
| Humerus | MAT 1 | 1,483,648 |
| Humerus | MAT 2 | 1,855,732 |
| Humerus | MAT 3 | 1,741,191 |
| Mandible | OST 1 | 218,981 |
| Mandible | OST 2 | 396,441 |
| Mandible | OST 3 | 759,814 |

**RNA Isolation and Amplification**. RNA was isolated using the ARCTURUS PicoPure RNA Isolation Kit (ThermoFisher Scientific; Cat# KIT0204) according to the manufacturer's instructions and DNase treatment was done using RNase-Free DNase (Qiagen; Cat#79254). The RNA from immature cartilage ($n$ = 3), mature cartilage ($n$ = 3), and bone ($n$ = 3) were extracted in triplicate from three unique tissue samples that were obtained from three different embryos, and thus a total of 9 cDNA libraries were constructed. RNA was amplified one round using MessageAmp II aRNA Kit (ThermoFisher Scientific; Cat# AM1751). The RNA integrity was evaluated on the observation of a signature eletropherogram pattern (Bioanalyzer). A minimum of 50 ng of amplified RNA was used per sample for the construction of each library.

**Library preparation and deep RNA Sequencing.** RNASeq libraries were prepared using the Illumina TruSeq RNA Sample Prep Kit v2 with the following modification: the protocol was started at the Elute, Prime, Fragment step using 5 µl amplified mRNA (minimum amount was 50 ng mRNA as determined using Quant-iT RiboGreen RNA Assay Kit (Invitrogen)). The quality of each library was checked on a DNA 1000 chip using the 2100 Bioanalyzer (Agilent Technologies Inc.). The size range of the prepared libraries (insert+ adapters) was 200-600 bp with an average size of 290 bp. Library concentrations were determined by qPCR using the KAPA SYBR FAST ABI Prism qPCR Kit (Kapa Biosystems) and the StepOnePlus Real-Time PCR System (Applied Biosystems).

Equimolar concentrations of the libraries were pooled and a final concentration of 12 pM was used for clustering in one lane of a flowcell on the cBOT (Illumina). The samples were then sequenced (2 x 101 cycles, paired-end reads) on the HiSeq2500 (Illumina) using the TruSeq SBS Kit v3-HS 200 cycles Kit (Illumina).

**Reads preprocessing, mapping, quantitation and primary analysis of RNA-seq data.** The paired-end Illumina reads were trimmed using a Java -based tool, Trimmomatic v0.30 (Bolger et al. 2014), and the reads were then mapped to the mm10 mouse genome fom Ensembl using TopHat2 and Bowtie2 (Langmead and Salzberg 2012; Kim et al. 2013). The location of each read was matched to genome annotation using HTSeq-count (Anders et al. 2015). Differential expression analysis was performed using EdgeR after excluding genes with zero or very low counts (less than 3 counts for all cell types) across the cell type, which resulted in 16,553 genes. The distribution of average $\log_2$ expression across three replicates of each tissue produced three bimodal distributions, which were used to set the count thresholds to 25, 24, and 18 for immature cartilage, mature cartilage, and bone respectively. Pairwise comparisons between tissues were made with Fisher's exact test, and a gene was considered differentially expressed if it had an absolute $\log_2$ fold change greater than 2 ($p<0.01$). The Venn diagrams were constructed using gplots v3.0.1 for isoforms and RNA-seq expression data.

**Principal component analysis (PCA).** PCA was performed on the data using prcomp from the stats library in R to determine if the biological replicates of each cell type separated into distinct groups based on gene expression variance. The 95% confidence ellipses were included using R package car 2.1-2.

**Cytoscape Visualization using Pearson's Correlation.** Pairwise differential expression results were visualized using Cytoscape 3.4.0 prefuse force directed layout using Pearson's correlation values of the edges (Lopes et al. 2010). Edges between nodes were limited to a correlation of at least +/- 0.95 (the nodes are the two genes or conditions that are correlated). All differentially expressed genes were visualized with *Runx2* and *Sox9* highlighted, and the nearest neighbours isolated.

**Validated skeletal cell GRN.** The skeletal cell GRN was constructed using BioTapestry version 7.1.2 (www.BioTapestry.org/) following developer's protocol (Longabaugh et al. 2005; Longabaugh et al. 2009). Regulatory interactions were validated using published work on mouse skeletal tissues.

**GO analysis**. DAVID v6.8 (http://david.abcc.ncifcrf.gov/home.jsp) functional annotation analysis was performed on the list of differentially expressed genes (DEGs) in IMM, MAT, and OST, with a fold change > 2 or < -2 as well as, in all genes expressed above threshold (i.e. non-DEGs). The GO term biological process (BP) in DAVID was used to perform the gene-annotation enrichment analysis using *Mus musculus* as a background.

**Cluster Analysis.** The algorithms from MBCluster.Seq 1.0 package in R were used to cluster the genes from our RNA-seq data (Si et al. 2014). Genes were assigned to 10 clusters based on expression profiles across all three cell types.

**Cell culture and micromass experiments.** Cell experiments were performed using the chondrogenic mouse cell line ATDC5 (Yao and Wang 2013). When cells reached 70-80% confluency on a 10 cm dish (~$7x10^6$), the cell culture medium was removed, and cells were washed with 10 mL of sterile phosphate-buffered saline (PBS). PBS was then removed and 1ml of 0.25% trypsin-EDTA (Sigma) was added into the dishes. Ten mL of DMEM/F12 complete medium was then added to stop trypsin digestion. Cells were then collected and transferred to a 15 mL conical tube and centrifuged at 1000 rpm for 3-5 mins to pellet the cells at room temperature. Digested cells were then resuspended in fresh medium and they were counted using a Neubauer chamber. Based on the cell numbers, the volume of differentiation medium to make a final cell centration at $2.5x10^5$ cells/ml was calculated. A 10 µl dot (micromass) was then platted in the center of a well in 24-well plate. Micromass cultures were incubated for 90 mins to allow the cells to aggregate and attach to culture plate. Finally, 1 mL of differentiation medium was gently added to each well, and it was changed every other day.

**Alcian blue staining on micromass cultures.** Micromass cultures of ATDC5 cells were examined for secretion of sulfated proteoglycans at day 14 ($n$=3), using Alcian blue staining as previously

described (Izadifar et al. 2016). Cells were fixed for 2 hours at room temperate, and then stained overnight at room temperature in 0.5% Alcian blue in 3% acetic acid (pH=1) with gentle rocking. Cells were then de-stained in 25% ethanol in 3% acetic acid for one hour, and then washed in 50% ethanol in 3% acetic acid, and then imaged using light-microscopy.

**Alizarin Red staining on micromass cultures.** Calcium deposition in micromass cultures of ATDC5 cells was evaluated a day 14 ($n$=3) using Alizarin red, as described previously (Gregory et. al., 2004) with minor modifications. Briefly, cells were fixed in 4% paraformaldehyde and then subsequently washed in phosphate buffered saline (PBS, pH 7.4). Cells where then washed in deionized water, and 250 µl of 40 mM ARS (Sigma) solution (pH 4.2) was added. Cells were kept in this solution for 20 min at room temperature with gentle shaking. Excess die was removed, and cells were imaged by microscopy.

**Alkaline Phosphatase staining on micromass cultures.** Alkaline phosphatase activity was identified at day 14 in micromass cultures of ATDC5 cells as described by others (Eames et al. 2007). Briefly, cells were washed in phosphate-buffered saline (PBS), followed by the addition of alkaline phosphatase buffer (100 mm Tris, pH 9.5; 50 mm MgCl2; 100 mm NaCl; 0.1% Tween 20) for 5 min. BM purple (Roche, Indianapolis, IN, USA) was then added into the cells, and incubated until strong signal appeared. Cells were then washed twice in PBS and imaged by microscopy. Alkaline phosphatase stained area was then quantified using ImageJ (https://imagej.nih.gov/ij/docs/examples/stained-sections/index.html). The micromass cultures predominantly formed a uniform sheet of chondrocytes, so a similar region in each micromass was quantified and compared between both GFP-control and EBF2- and IRX6-transfected cultures for more accurate comparisons (n=6 for each group). One-way ANOVA coupled to Bonferroni for multiple comparisons were performed using IBM SPSS Statistics v 1.0.0.1072 in order to find statistical differences ($p<0.05$).

**Cell transfection.** The enriched MAT genes *Ebf2* and *Irx6*, were obtained from Bio Basic (Canada), and then subcloned into the pP2A-mCherry-N1 vector, which was obtained from Dorus Gadella (Addgene plasmid #84329). This vector enables bicistronic expression of protein of interest and mCherry. GFP subcloned into the pP2A-mCherry-N1 vector was used as a control in transfection

experiments. Transfection was performed using TransIT-LT1 Transfection Reagent (Mirus Bio, Madison, WI). After 48 hours, transfection medium was removed, cells were harvested, and micromass experiments were performed. Differentiation media (Dulbecco's Modified Eagle's Medium (DMEM-F12), Ascorbate-2-phosphate (A2P), β-glycerophosphate (BGP), and Insulin-transferrin-sodium selenite (ITS)) was added into these cultures, and cells underwent chondrogenic differentiation. Micromass cultures were analyzed at days 7 and 14 and imaged by fluorescent microscopy which showed that both EBF2 and IRX6 remain highly expressed at these time points (Fig. 20A,B). Cells were then harvested and examined for secretion of sulfated proteoglycans at day 14, using Alcian blue staining as previously described (Izadifar et al. 2016). Alcian blue staining of micromass cultures showed that cells successfully differentiated into chondrocytes (Fig.21).

**Immunohistochemistry.** Secretion of COLX in micromass cultures was examined at day 14, as described previously (Cunliffe 2003; Hutchinson et al. 2007), with some modifications. Micromass cultures were fixed in 4% paraformaldehyde overnight at 4ºC, incubated in 1 mg/ml trypsin solution for 45 min at 37ºC, followed by incubation in 5mg/ml hyaluronidase solution (Worthington, USA) for 30 mins at 37˚C. Monoclonal COLX primary antibody X-AC9 (DHSB, USA) diluted in blocking solution (1XPBS, 0.5% TritonX-100, 4% NGS 2% NSS; 1:100) was then added into the cells, followed by the addition of secondary antibody Goat Anti-Mouse IgG1-488 (1:1000; EMD Millipore, USA). Micromass cultures were imaged by microscopy (*n=5* per experimental group), and fluorescence was quantified and expressed as area ($\mu m^2$) using ImageJ software, as described by Dr. Christine Labno (University of Chicago)- Labno, C. (2007) Basic Intensity Quantification with ImageJ (2007). https://voices.uchicago.edu/confocal/image-processing/imagej-fiji-help/. Fluorescence quantification was limited only to a specific area in the micromass. The micromass cultures predominantly formed a uniform sheet of chondrocytes, so a similar region in each micromass was quantified and compared between both GFP-control and EBF2- and IRX6-transfected cultures for more accurate comparisons. One-way ANOVA coupled to Bonferroni for multiple comparisons were performed using IBM SPSS Statistics v 1.0.0.1072 in order to find statistical differences ($p<0.05$).

**RT-PCR.** Reverse transcription PCR (RT-PCR) was used to detect the presence of important maturation and mineralization markers, including *Col10a1*, *Ihh*, *Mmp13* and *Spp1* in micromass

cultures. First, RNA was extracted from micromass cultures using the RNeasy Mini Kit (Qiagen, USA), according to manufacturer's instructions. Reverse transcription was then performed at 42ºC using RevertAid H Minus First Strand cDNA Synthesis Kit (Thermo Scientific, USA). Specific primers for genes enriched in mature chondrocytes were designed based on transcripts obtained from previously reported sequences in GenBank (NCBI). RT-PCR conditions were as follows; one cycle at 95ºC for 3 min; 35 cycles at 95º for 1 min, 55-60º for 1 min, and 72ºC for 1 min, and a final cycle at 72ºC for 10 min.

## 3.6.DATA ACCESS

The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus (Edgar et al. 2002) and are accessible through GEO Series accession number GSE110051 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE110051).

## 3.7.ACKNOWLEDGEMENTS

The authors would like to acknowledge grants NSERC RGPIN 435655-201 and 2016-06172 supporting this research, and the Nucleic Acid Solutions Team, Janet Condie at the National Research Council Canada-Saskatoon who performed library construction and the RNA-seq procedure. Thanks to Dr. Bogdan Popescu for the use of his laser capture equipment, as well as Dr. Ralph Marcucio (University of California San Francisco, USA), Dr. Christine Hartmann (University Münster, Germany), Dr. Lars Wittler (Max-Planck Institut für Molekulare Genetik, Germany), Dr. Keith L. Ligon (Harvard Medical School, USA), Dr. Carsten Janken (Institut Curie, France), and Dr. James MacLean (Southern Illinois University School of Medicine, USA), who kindly provided plasmids.

## 3.8.DISCLOSURE DECLARATION

The authors have no conflicts of interest to declare.

## CHAPTER 4:

**Manuscript 2; Comparative transcriptomics reveal that a conserved molecular program underlies mesoderm- and neural crest-derived chondrocytes (*in preparation*).**

Gomez-Picos P, Ovens K, and Eames BF

# CHAPTER 4: Comparative transcriptomics reveal that a conserved molecular program underlies mesoderm- and neural crest-derived chondrocytes

## 4.1.ABSTRACT

Cartilage is composed of chondrocytes of distinct embryonic origins, mesoderm and neural crest (NC), but the degree of similarity between chondrocytes derived from the distinct embryonic lineages is still debatable. Molecular analyses using candidate genes suggest that gene expression in chondrocytes is conserved regardless of embryonic origin and location in the body. During endochondral ossification, two types of chondrocytes differentiate in the head and limb skeletons, immature chondrocytes (IMM) and mature chondrocytes (MAT), so both cell types can derive from the mesoderm or the neural crest. To test the hypothesis that the transcriptomes of mesoderm- and neural crest-derived chondrocytes are conserved we used LCM to isolate IMM and MAT from two endochondral bones in the chick limb and head, the humerus and the ceratobranchial, which are mesoderm- and neural crest-derived, respectively. In general, our findings supported the hypothesis. Venn diagram analyses revealed that the humerus and ceratobranchial transcriptomes show a high degree of conservation (70%- 80% of the genes expressed above threshold). Although they exhibit some differences in gene expression, the fundamental set of genes driving cartilage differentiation including *SOX9, SOX5, SOX6, ACAN, COL2A1, COL9A1, COL11A1, MATN1*, and *COL10A1* was generally conserved. Indeed, some enriched biological processes in genes shared between the humerus and the ceratobranchial are related to skeletal cell differentiation. In contrast, gene ontology analyses revealed that enriched biological processes in the humerus are related to limb/forelimb morphogenesis whereas enriched terms in the ceratobranchial are related to neural crest-dependent processes. Importantly, pairwise differential gene expression revealed subtle differences in IMM and MAT differentiation markers between the humerus and the ceratobranchial. For instance, *RUNX2* showed higher expression levels in the humerus compared to the ceratobranchial, but gene expression levels of other classic IMM and MAT genes were generally similar between both skeletal elements. Together these results suggest that the molecular program driving cartilage differentiation is conserved regardless of the embryonic origin. Adding more clades into these transcriptomic

comparisons can make this conclusion more robust and might provide novel insights into mechanisms of differentiation and evolutionary origins of cartilage.

## 4.2.INTRODUCTION

Vertebrate limb and head cartilage can have two distinct embryonic origins, mesoderm and neural crest, respectively. Cartilage of the limb skeleton derive from lateral plate mesoderm, whereas cartilage of a great portion of the cranial skeleton including skull, jaws, and hyoid bone are of neural crest (NC) origin (Couly et al. 1993; Knight and Schilling 2006; Fonseca et al. 2017). Previous studies in skeletal cells have suggested that the gene regulatory network (GRN) driving cartilage is conserved regardless of its embryonic origin (Schneider et al. 1999; Eames and Helms 2004; Donoghue et al. 2008; Cattell et al. 2011), but the similarity between skeletal cells of the distinct origins is still debatable.

Endochondral ossification involves the differentiation of two types of cartilage: immature cartilage (IMM) and mature cartilage (MAT; Eames et al. 2003; Eames and Helms 2004; Eames et al. 2004). Immature cartilage is composed of proliferative and resting chondrocytes that deposit COL2A1 fibers and proteoglycans in the extracellular matrix (ECM), whereas mature cartilage is characterized by pre-hypertrophic and hypertrophic chondrocytes that modify the immature cartilage ECM by depositing COL10A1 fibers, decreasing its proteoglycan sulfation, and mineralizing it (Leboy et al. 1988; Takeda et al. 2001; Farquharson et al. 1994) . IMM and MAT are present in head and limb endochondral bones, so both types of chondrocytes can derive from the mesoderm or neural crest.

Head and limb cartilages share many histological and molecular features (Eames and Helms 2004). In both head and limb, the extracellular matrix surrounding IMM and MAT stains with Safranin O and Alcian blue (Eames and Helms 2004; Eames et al. 2004; Ovchinnikov 2009), although the latter decreases during MAT matrix mineralization (Farquharson et al. 1994). In contrast, mineralized bone stains with Alizarin red (Eames and Helms 2004; Eames et al. 2004; Ovchinnikov 2009). Moreover, several skeletal markers have also been shown to express similarly in head and limb cartilage. Classic cartilage differentiation markers, such as *SOX9, SOX5, SOX6, ACAN, and COL2A1* (Lefebvre and de Crombrugghe 1998; Smits et al. 2001a; Akiyama et al. 2002;

Eames et al. 2004; Smits et al. 2004; Dale and Topczewski 2011a; Lefebvre and Dvir-Ginzberg 2017; Xiong et al. 2018) are conserved in IMM, whereas classic maturation markers, such as *RUNX2*, *COL10A1* and *IHH* (Eames et al. 2004; Yoshida et al. 2004; Young et al. 2006) are conserved in MAT. Importantly, genetic experiments have shown that if the function of any of these genes is perturbed then it could have an impact throughout the body, suggesting that the molecular program driving chondrogenesis in conserved in both head and limb, regardless of the different embryonic origins (Komori et al. 1997; Bi et al. 1999b; Smits et al. 2001a; Smits et al. 2004; Yoshida et al. 2004).

Studies in amphioxus suggest that neural crest-derived cartilage evolved from co-option of a primitive mesodermal gene regulatory network (GRN) due to the tremendous overlap in gene expression between these two lineages (Meulemans and Bronner-Fraser 2007; Hall and Gillis 2013a; Jandzik et al. 2015). Moreover, one particular study analyzed and compared the expression of a few selected cartilage markers in head vs limb chick cartilages and revealed that generally a conserved set of genes directs skeletal cell differentiation, but molecular differences were also identified (Eames and Helms 2004). More unbiased studies comparing gene expression globally between head and limb are needed to verify if the GRN underlying cartilage differentiation is conserved.

Comparative transcriptomics has revealed differences between mesenchymal precursors of the different origins, but none of these studies looked directly at skeletal cells. For instance, in a previous study in the mouse embryo, neural crest- and mesoderm-derived mesenchymal cells were isolated from the first pharyngeal arch using laser-capture microdissection (LCM), and gene expression profiles of the two cell lineages were obtained and compared revealing 140 differentially expressed genes (Bhattacherjee et al. 2007). Another study in mouse craniofacial structures also showed that the transcriptomes of neural-crest and mesoderm-derived cells differ (Fan et al. 2016). These studies highlight important differences between mesenchymal cells of the distinct embryonic lineages, but none of them has focused specifically in differentiated chondrocytes *in vivo*.

Using laser capture-microdissection (LCM) coupled to RNA-seq, we aim to test the hypothesis that the transcriptomes of mesoderm- and neural crest- derived chondrocytes are conserved. To test this hypothesis, we used LCM to isolate IMM and MAT from two skeletal elements of the chick embryo, the humerus and ceratobranchial (Fig. 26), which are of mesoderm and NC origin, respectively, and we then present a transcriptomic comparison between chondrocytes derived from the distinct lineages.

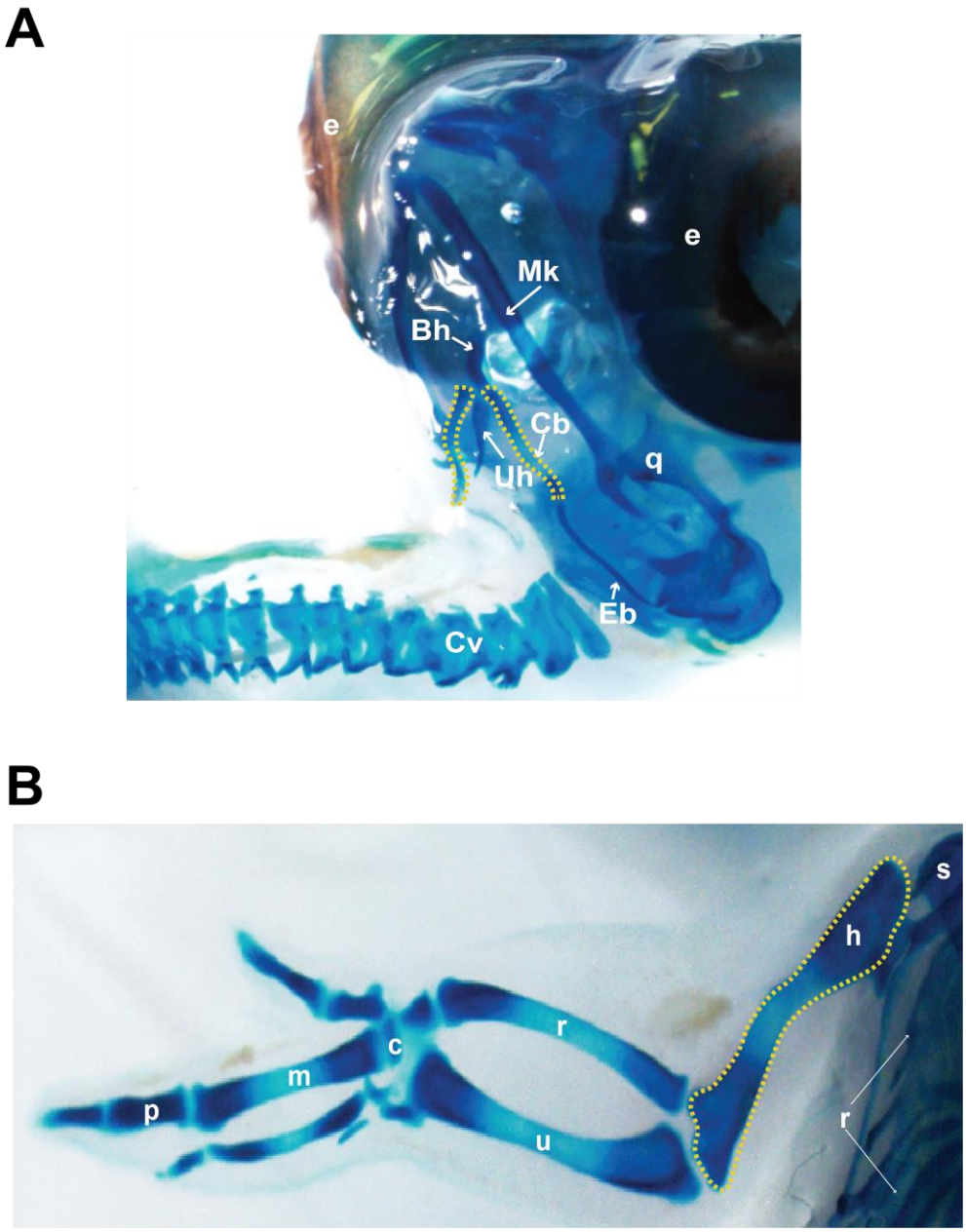



**Figure 26. Whole mount Alcian blue staining of chick HH36 head and limb skeletal elements.** (A) Chick head: The ceratobranchial is a neural crest-derived bone, and it undergoes endochondral ossification in the second pharyngeal arch (yellow dotted line). (B) Chick forelimb: The humerus (yellow dotted line) is an endochondral bone of mesodermal origin that forms in the upper arm. Abbreviations: Cb, ceratobranchial; Uh, urohyal; Bh, basihyal; Eb, epibranchial; Cv, cervical vertebrae; e, eye; h, humerus; s, scapula; r, radius; u, ulna; c, carpal; m, metacarpals; p, phalanges; r, ribcage.

## 4.3.RESULTS

### 4.3.1. Histological identification of chick HH36 skeletal tissues/cells and laser capture microdissection.

Chick IMM and MAT were obtained from two endochondral bones of distinct embryonic origin, the humerus and ceratobranchial (Figs. 27&28), which are derived from the mesoderm and neural crest, respectively. The ceratobranchial undergoes endochondral ossification in the second pharyngeal arch, and it is one of the five distinct bones that comprise the hyoid apparatus of most avian species, being the other four the epibranchial, basihyal, urohyal, and paraglossal (Figs. 26A, Fig. 28A; Homberger and Meyers 1989). The epiphyseal growth plate containing IMM, MAT, and OST in long bones is detected in HH36 chick (E10; Conen et al. 2009). The epiphyseal growth plate is a thin layer of cartilage where longitudinal growth of long bones takes place, and it is composed of cells in different stages of differentiation (Milz et al. 2002; Lui et al. 2014). By HH36, chondrocytes have already proliferated and maturation (i.e. hypertrophy) becomes obvious in the mid-diaphyseal region (i.e. shaft of a long bone) of the humerus and ceratobranchial as shown by Safranin O staining (Fig. 27C,E; Fig. 28B,D). Alcian blue identified cartilage, whereas Alizarin Red identified perichondral bone in both skeletal elements (Fig. 27A,B; Fig. 28A). At HH36 perichondral bone is obvious in the mid-diaphyseal region surrounding the mature cartilage region of both the humerus and the ceratobranchial. Once skeletal elements and cells of interest were identified, laser capture microdissection (LCM) was used to isolate IMM and MAT from the humerus (Fig. 27F-I) and the ceratobranchial (Fig. 28E-H; see methodology for more details).

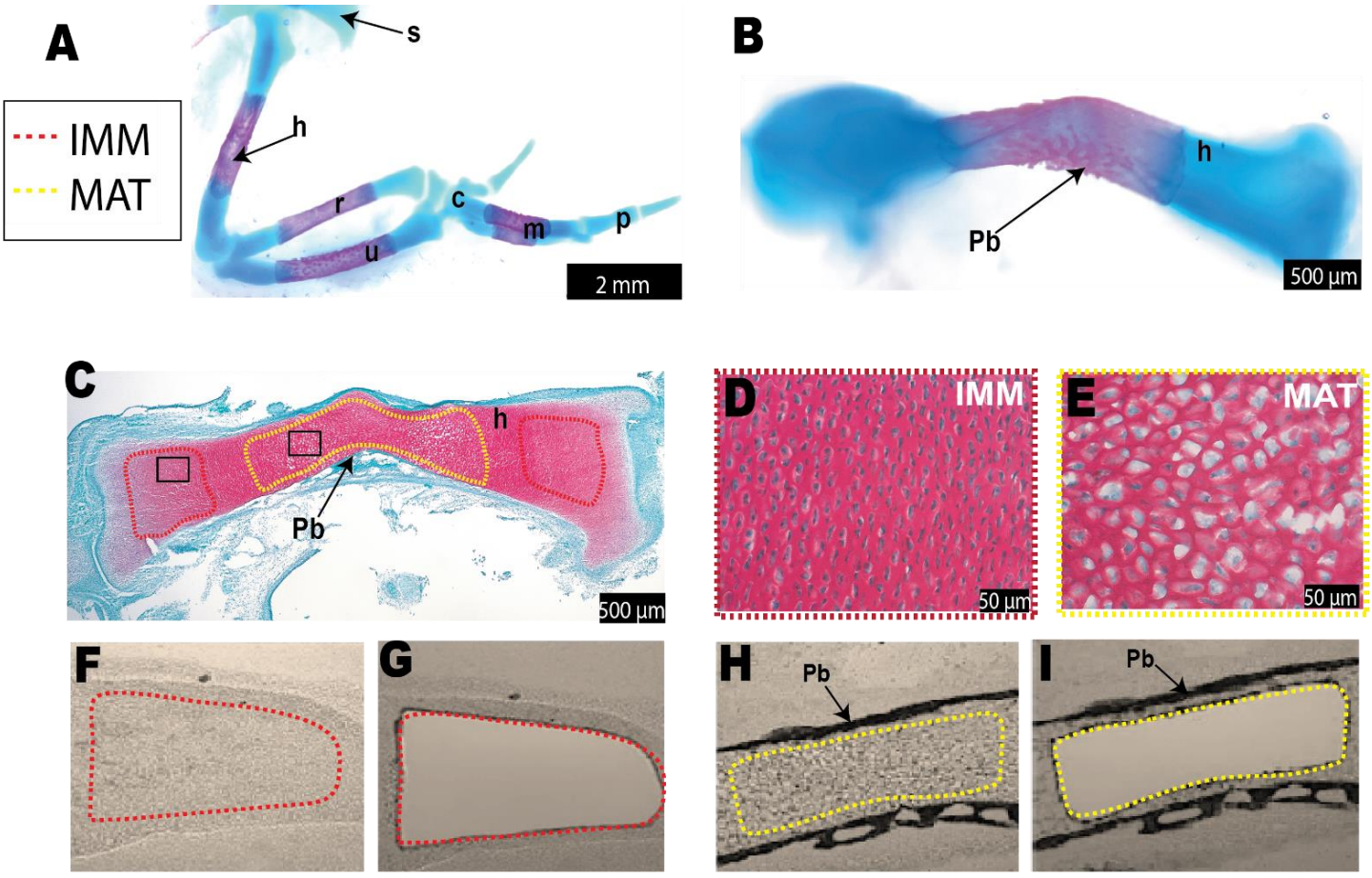



**Figure 27. Laser capture microdissection was used to isolate chondrocytes from the chick HH36 humerus.** (A,B) Whole-mount Alcian blue and Alizarin red staining identified cartilage and perichondral bone in endochondral bones of the chick forelimb. (C) Safranin O-stained section of HH36 humerus highlighted the immature cartilage (red dotted outline) and the mature cartilage (yellow dotted outline) regions. High-magnification images of immature (D) and mature chondrocytes (E) from black boxes in (C). Unstained sections of HH36 chick humerus before (F) and after (G) laser capture of immature chondrocytes, and before (H) and after (I) laser capture of mature chondrocytes. Abbreviations: IMM, immature chondrocytes; MAT, mature chondrocytes; Pb, perichondral bone; h, humerus; s, scapula; r, radius; u, ulna; c, carpal; m, metacarpals; p, phalanges.

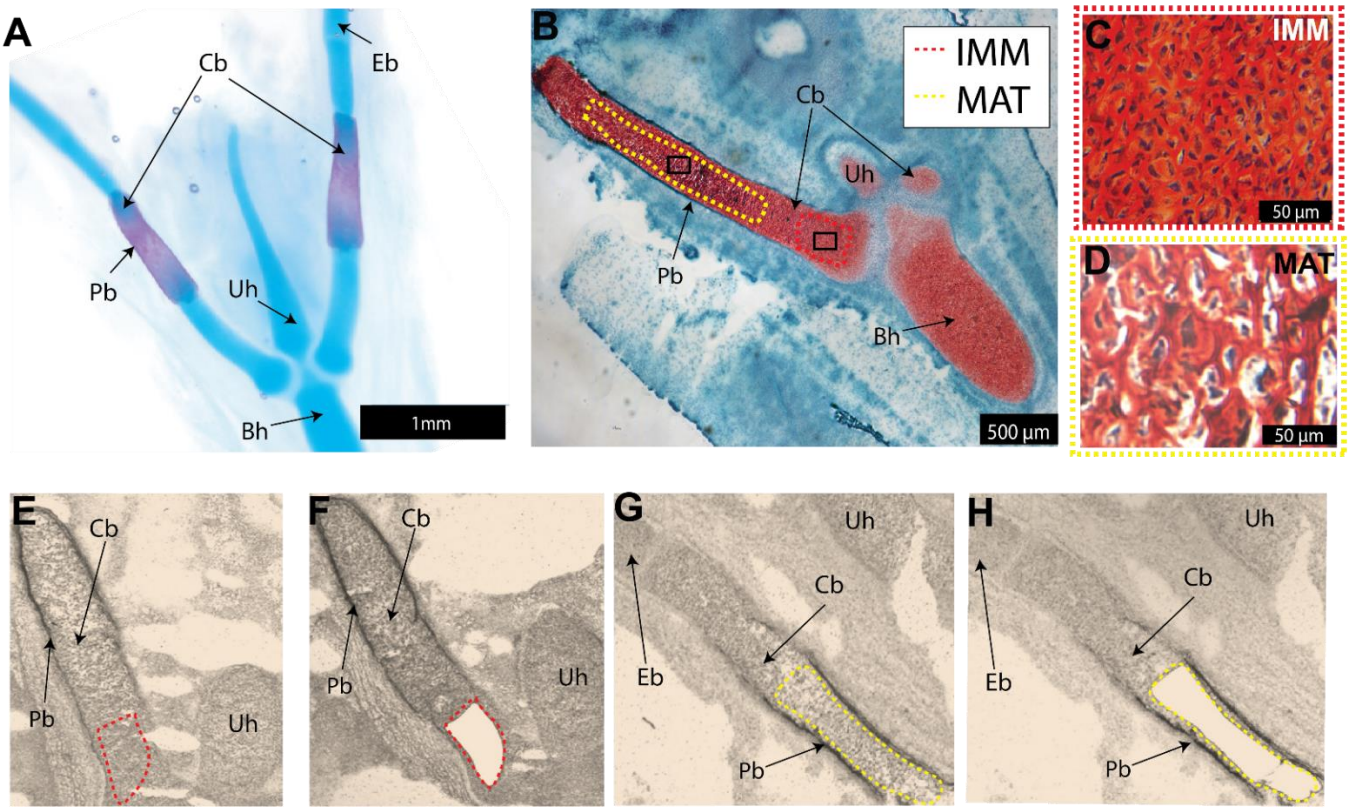



**Figure 28. Laser capture microdissection was used to isolate specific skeletal cell types from the chick HH36 ceratobranchial.** (A) Whole-mount Alcian blue and Alizarin red staining identified cartilage and perichondral bone in the chick ceratobranchial. (B) Safranin O-stained section of HH36 ceratobranchial highlighted the mature cartilage region (yellow dotted outline). High-magnification images of immature (C) and mature chondrocytes (D) from black boxes in (B). Unstained sections of HH36 chick ceratobranchial before (E) and after (F) laser capture of immature chondrocytes, and before (G) and after (H) laser capture of mature chondrocytes. Abbreviations: IMM, immature chondrocytes; MAT, mature chondrocytes; Cb, ceratobranchial; Uh, urohyal; Bh, basihyal; Eb, epibranchial; Pb, perichondral bone.

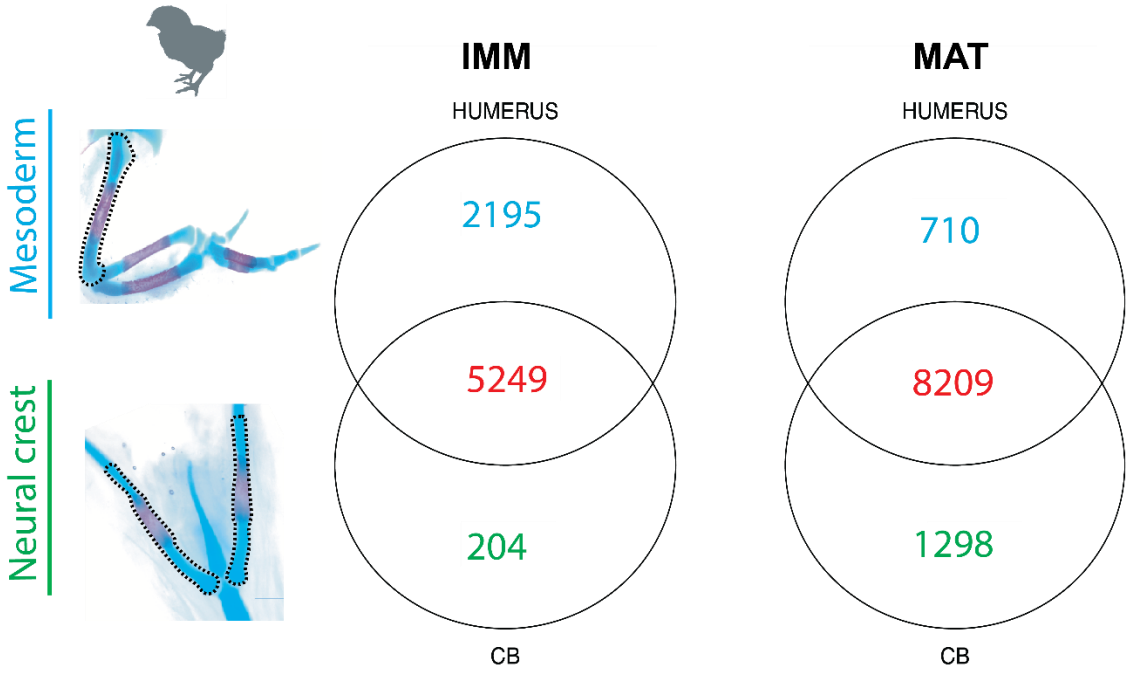
### 4.3.2. Comparative transcriptomics reveal that the transcriptomes of neural crest- and mesoderm-derived cartilage are highly conserved

To test the hypothesis that transcriptomes of mesoderm- and NC-derived chondrocytes are conserved, datasets obtained from the chick humerus and ceratobranchial were compared (Fig. 29). Comparisons of transcriptomic profiles of IMM from both embryonic lineages revealed that they shared almost 70% of the total number of genes expressed above threshold, whereas MAT shared

80% of the genes expressed above threshold (Fig. 29A). Gene ontology analyses showed that enriched biological processes related to multicellular organism development, cell cycle, cell proliferation, cell differentiation, DNA replication, transmembrane transport, transcription, translation, and protein phosphorylation were enriched in uniquely expressed genes in the humerus, uniquely expressed genes in the ceratobranchial, as well as the overlap including approximately 50% of total genes expressed above threshold. To test whether the fundamental set of genes driving cartilage differentiation was shared between the humerus and the ceratobranchial we then focused on the Venn diagram overlap (Fig. 29A; IMM, 5249 genes; MAT, 8209 genes). GO analyses revealed that biological processes related to cartilage differentiation and maturation were conserved between the distinct chondrocyte lineages including genes such as *SOX9, RUNX2, SOX5, SOX6, ACAN, COL2A1, COL9A1, COL11A1, MATN1, COL10A1, MEF2C, MMP13* and *PTHLH* (Fig. 29B; Vortkamp et al. 1996a; Zhao et al. 1997; Bridgewater et al. 1998; Watanabe et al. 1998; Bi et al. 1999b; Smits et al. 2001a; Smits et al. 2004; Arnold et al. 2007; Nicolae et al. 2007; Dy et al. 2012; Lu et al. 2014; Nakatani and Partridge 2017), but genes associated exclusively to cartilage specific processes correspond to approximately 3% of the total number of genes shared between the humerus and the ceratobranchial. Approximately 20% of the total number of IMM and MAT genes in the overlap have not been associated to any specific GO terms so the software does not include them into these analyses, thus, it is possible that more genes associated with cartilage differentiation processes were just excluded from the analysis. However due to the tremendous overlap in gene expression between chick mesoderm- and neural crest-derived cartilage transcriptomes, and the conservation of classic cartilage markers in both the humerus and the ceratobranchial these results suggest that the GRN underlying cartilage differentiation might also be conserved between chondrocytes derived from the distinct embryonic lineages.

Gene ontology analyses was then used to reveal enriched biological processes in genes uniquely expressed in the humerus and genes uniquely expressed in the ceratobranchial (Fig. 29). Enriched biological processes in cartilage isolated from the humerus included, but were not limited to, skeletal system development, ossification, cell proliferation, and embryonic limb/ forelimb morphogenesis including genes , such as *TBX5, DLX5/6, SHOX2, SALL4, HOXA10, HOXD9,* and *HOXD10* which are all known regulators of limb development (Wahba et al. 2001; Ng et al. 2002; Robledo et al. 2002; Zakany and Duboule 2007; Vieux-Rochas et al. 2013; Neufeld et al. 2014; Akiyama et al. 2015; Fig. 29B). In contrast, enriched processes in the ceratobranchial included, but

A




B

**Enriched biological processes**

| Humerus<br>Mesoderm | Overlap | Ceratobranchial<br>Neural crest |
|---|---|---|
| Skeletal system development, ossification, cell proliferation, and embryonic limb/ forelimb morphogenesis, Wnt signaling pathway. | Cartilage differentiation, chondrocyte differentiation, skeletal system development, Wnt signaling pathway. | Skeletal system development, cranial skeleton morphogenesis, middle ear morphogenesis, ectodermal cell differentiation, neuron differentiation, nervous system development, cell migration, heart development, melanocyte differentiation, Wnt signaling pathway. |

**Note**: *Biological processes related to multicellular organism development, cell cycle, cell proliferation, cell differentiation, DNA replication, transmembrane transport, transcription, translation, and protein phosphorylation were enriched in these three categories including ~50% of total genes expressed above threshold.*

**Figure 29. The transcriptomes of chondrocytes derived from the chick ceratobranchial and humerus show a high degree of conservation.** (A) IMM and MAT chondrocytes isolated from the ceratobranchial and humerus share 70-80% of the genes. The overlap includes hallmark cartilage differentiation and maturation genes such as *SOX9*, *SOX5*, *SOX6*, *ACAN*, *COL2A1*, *COL9A1*, *COL11A1*, *MATN1*, *RUNX2*, *MEF2C*, and *COL10A1*. (B) Enriched biological processes in the overlap include cartilage differentiation processes, whereas neural crest-related processes and limb morphogenesis processes are enriched in the ceratobranchial and humerus, respectively.

were not limited to, skeletal system development, cranial skeleton morphogenesis, middle ear morphogenesis, ectodermal cell differentiation, neuron differentiation, nervous system development, cell migration, heart development, melanocyte differentiation, which are all neural crest lineage-dependent processes (Fig. 29B). Again, most of the genes uniquely expressed in the humerus and the ceratobranchial were related to cell cycle processes. A portion of the GRN driving neural crest-derived cartilage was revealed (Meulemans and Bronner-Fraser 2007), and includes many orthologs that were also expressed in our chick neural crest-derived data , such as *SOX9, PAX7, ETS1, MYB* and *MYC* genes, *ID1, ID2, ID3, BARX1, BARX2, RUNX1, RUNX2,* and *RUNX3* (Fig. 29B).

### 4.3.3. Gene expression levels of classic cartilage differentiation and maturation markers were generally similar between the humerus and the ceratobranchial

Differential gene expression analyses were performed using genes included in the overlap (i.e. genes shared between humerus and ceratobranchial) of the Venn diagrams presented in Fig. 29A (IMM, 5249; MAT, 8209). These analyses revealed 150 DEGs between IMM isolated from the humerus vs IMM isolated from the ceratobranchial, whereas 271 DEGs were identified between MAT isolated from the humerus vs MAT isolated from the ceratobranchial (Fig. 30). In general, the expression levels of hallmark cartilage differentiation and maturation genes was conserved, but this study also revealed a few differences between both skeletal elements. For instance, *RUNX2* was upregulated in IMM isolated from the humerus (mesoderm-derived) compared to IMM isolated from the ceratobranchial (neural crest-derived). Enriched biological processes were also identified when gene ontology analyses was performed in differentially expressed genes (DEGs) between the humerus and the ceratobranchial. Biological processes in IMM isolated from the humerus include positive regulation of cell proliferation and steroid biosynthetic process, whereas processes in MAT isolated from this element include positive regulation of ossification, steroid biosynthetic processes, osteoblast differentiation, and cell proliferation. Biological terms enriched in IMM isolated from the ceratohyal include collagen fibril organization, extracellular matrix organization, cell migration, and skeletal muscle differentiation, whereas processes enriched in MAT of this element include several

neural crest related processes such as, muscle development, heart development, cell migration, bone mineralization, skeletal system development, ossification, Wnt signaling pathway, among others.

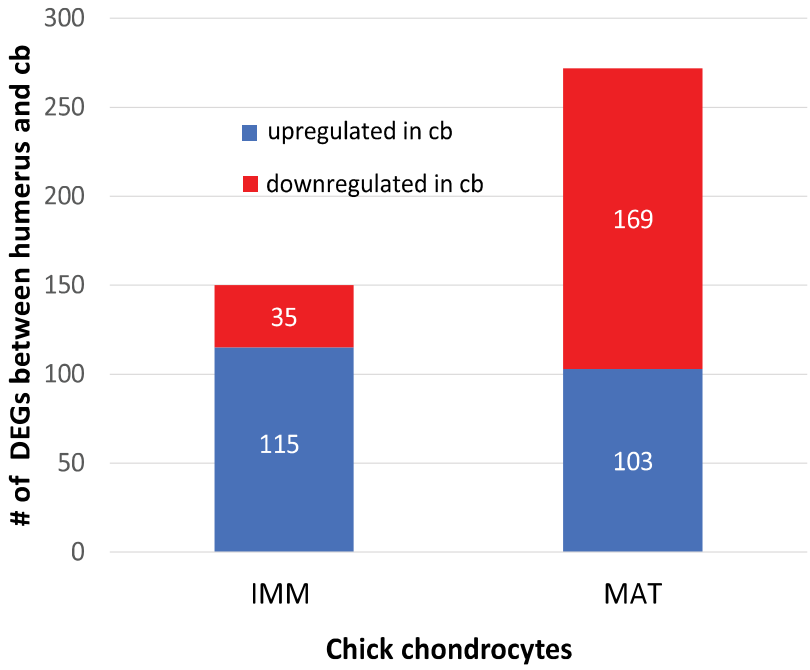



**Figure 30. Pairwise differential gene expression between immature and mature chondrocytes isolated from the humerus and the ceratobranchial (cb).** Bars represent genes upregulated (and downregulated) in the ceratobranchial compared to humerus. Genes considered for this analysis were located in the overlap between humerus and cb (Fig. 45A; IMM, 5249 genes and MAT, 8209 genes). Gene lists used for DEGs graphs were generated by Katie Ovens.

A few skeletal markers that were differentially expressed between MAT of both skeletal elements were also revealed (Inada et al. 2004; Mailhot et al. 2008). For example, *BMP5* was upregulated in the ceratobranchial, whereas *MMP13* was upregulated in the humerus. Importantly, hallmark cartilage differentiation genes, such as *SOX9*, *SOX5*, *SOX6*, *ACAN*, *COL2A1*, as well as and maturation markers including *COL10A1* and *RUNX3* were not differentially expressed (absolute $\log_2$ fold change lower than 2 ($p>0.05$)) between the humerus and the ceratobranchial since gene expression is either at comparable levels or perhaps variation is high between biological replicates that the statistical tests are unable to confidently detect differential expression.

## 4.4. DISCUSSION

Skeletal cells have two distinct embryonic origins, neural crest and mesoderm (Poole et al. 2000; Knight and Schilling 2006; Schilling and Le Pabic 2014; Tickle 2015; Prummel et al. 2020), but the degree of conservation between the transcriptomes of these two lineages is still unresolved (Schneider et al. 1999; Tucker et al. 1999; Eames and Helms 2004; Wang et al. 2019). We used comparative transcriptomics to test the hypothesis that transcriptomes of chondrocytes derived from the neural crest and the mesoderm are conserved. In general, the analyses presented here supported the hypothesis. Venn diagram analyses revealed a tremendous overlap in gene expression between the transcriptomes of IMM and MAT isolated from the humerus and the ceratobranchial (Fig. 29A). Indeed, the core set of genes directing chondrocyte differentiation is conserved regardless of embryonic origin, as previously suggested by others (Schneider et al. 1999; Eames and Helms 2004; Fig. 29A). Gene ontology analyses revealed that some biological processes enriched in genes shared between humerus and ceratobranchial were related to cartilage differentiation (Fig. 29B; Venn diagram overlap). Together these data suggest that portions of the GRN directing IMM and MAT are conserved in mesoderm- and neural crest- derived cartilage, so perhaps one might have evolved from the other one. Indeed, previous work in amphioxus supports this conclusion (Meulemans and Bronner-Fraser 2007; Hall and Gillis 2013; Jandzik et al. 2015). As expected, enriched terms in the ceratobranchial included neural crest-derived processes, whereas enriched biological processes in the humerus were generally related to limb/forelimb morphogenesis (Fig. 29B).

Molecular makers of IMM and MAT were generally conserved between the humerus and the ceratobranchial, but a few differences in gene expression levels were also identified that might have an impact in specific regulation of head and limb skeletogenesis. For example, *RUNX2* showed higher expression levels in chondrocytes isolated from the humerus (mesoderm-derived) compared to chondrocytes isolated from the ceratobranchial (neural crest-derived; Fig. 30), and these results are consistent with previous work on mouse head bones where mesoderm-derived bones showed higher expression levels of this transcription factor compared to neural crest-derived bones (Al-Amer 2017). Moreover, this analysis revealed a few skeletal markers that were differentially expressed between MAT of both skeletal elements including *BMP5* which was upregulated in the ceratobranchial and *MMP13* which was upregulated in the humerus (Fig. 30). Despite these

differences in gene expression levels, most cartilage differentiation molecular markers were not differentially expressed between the humerus and the ceratobranchial. For instance, the master cartilage differentiation gene, *SOX9*, as well as many other genes that are regulated by this transcription factor such as *SOX5*, *SOX6*, *COL2A1*, *COL9A1*, *COL10A1*, and *ACAN* (Lefebvre et al. 1997; Zhao et al. 1997; Bi et al. 1999b; Sekiya et al. 2000; Smits et al. 2001a; Zhang et al. 2003; Dy et al. 2012) showed similar expression levels between the humerus and the ceratobranchial. Together these results suggest that the molecular program (i.e. portion of the GRN) driving cartilage differentiation is highly conserved throughout the body regardless of embryonic origin.

Revealing the GRN underlying mesoderm and NC-derived skeletal cells in a single species can potentially increase understanding on skeletal cell differentiation. Besides identifying conserved cartilage markers, the analyses presented here can also identify novel genes that might be specifically involved in chondrogenesis of the head or the limb, and the expression of these novel genes can be validated using molecular techniques such as, *in situ* hybridization or immunohistochemistry. In addition, gain-of-function approaches using the retroviral system known as RCAS (Tickle 2004; Gordon et al. 2009) can be performed in chick embryos to test the function of novel genes in the head and limb. For instance, a transcription factor specifically expressed in the head can be misexpressed in the limb (and vice versa) to assess if the function of this gene during chondrogenesis is maintained throughout the body.

To increase understanding on skeletal cell GRN evolution, transcriptomes of homologous mesoderm- and NC-derived skeletal cells can be compared across animal clades. Using this approach, GRNs driving the distinct embryonic lineages can be revealed, and conserved and species-specific portions can be identified. Since distinct chondrocyte lineages were analyzed in a single species (i.e. chick), skeletal cells derived from both the mesoderm and NC should be compared within another species to confirm if the degree of similarity between skeletal cells of the distinct embryonic origins is still high. In chapter 6, the present chick data was analyzed and compared with data obtained from homologous elements in gar and mouse and these evolutionary comparisons provide insight into GRN organization and mechanisms of gene regulation in skeletal cells. In summary, identifying conserved markers between the head and limb skeletons is crucial for developing new therapies and treatments for cartilage injuries and skeletal disorders, as well as to have a more comprehensive understanding of cartilage differentiation and evolution.

## 4.5.METHODS

**Embryo Collection and tissue processing.** All animal procedures were performed according to guidelines approved by the University of Saskatchewan Animal Care and Use Committee. *Chick (Gallus gallus):* White leghorn chicken eggs were incubated in a humified incubator at a constant temperature of 37ºC. Embryos were harvested at Hamburger-Hamilton stage 36 (~E10.5; (Hamburger and Hamilton 1951). Each embryo was decapitated, and the forelimbs and lower jaws were dissected and immediately placed in 1X PBS/DEPC, followed by embedding in OCT (Tissue Tek, Torrance, CA, USA; for detailed methodology see chapter 3), and immediately flash-frozen using liquid N2 and 2-Methylbutane (isopentane).

**Alcian blue/ Alizarin red whole mount skeletal stain.** Chick HH36 embryos were fixed using 4% paraformaldehyde, and then stained with Alcian blue and Alizarin red using an acid-free solution that included $MgCl_2$ to differentiate staining, and then cleared in glycerol/KOH as described elsewhere (Eames et al. 2011).

**Histology.** Safranin O/Fast Green staining was performed on 10 μm thick frozen sections of the HH36 chick humeri and ceratobranchial, as described previously (Ferguson et al. 1998). Trichrome staining was performed on 10μm thick frozen sections of the HH36 chick mandible, as described previously (Clark and Smith 1993).

**Laser Capture Microdissection (LCM).** LCM was performed on a Laser Microdissection - Molecular Machines & Industries (MMI) CellCut apparatus. Five biological replicates for each tissue were captured. Immature and mature chondrocytes were captured from the developing chick HH36 humeri and ceratobranchial. The captured cells were collected onto the inner lid of 0.5ml MMI IsolationCaps (either Diffuser caps (Prod#50202) or Transparent caps (Prod#50204; MMI Molecular Machines & Industries). Once RNA-seq data was obtained, bioinformatic analyses identified some samples as outliers, and these were not included in our analyses. For details on bioinformatic analyses see chapter 3 methodology.

**RNA Isolation and Amplification**. RNA was isolated using the ARCTURUS PicoPure RNA Isolation Kit (ThermoFisher Scientific; Cat# KIT0204) according to the manufacturer's instructions and DNase treatment was done using RNase-Free DNase (Qiagen; Cat#79254). Immature and mature chondrocytes were captured from both chick humerus and ceratobranchial. RNA was extracted from immature cartilage ($n$ = 5 per skeletal element) and mature cartilage ($n$ = 5 per skeletal element; Table 5), and then amplified with one round using MessageAmp II aRNA Kit (ThermoFisher Scientific; Cat# AM1751). The RNA integrity was evaluated on the observation of a signature eletropherogram pattern (Bioanalyzer).

**Table 5.** LCM Surface area obtained from chick HH36 humerus and ceratobranchial

| Skeletal element | Cell type | Surface Area captured ($um^2$) |
|---|---|---|
| Humerus | IMM 1_h | 1,828,237 |
| Humerus | IMM 2_h | 1,000,000 |
| Humerus | IMM 3_h | 1,900,000 |
| Humerus | IMM 4_h | 1,281,487 |
| Humerus | IMM 5_h | 1,116,820 |
| Humerus | MAT 1_h | 3,080,490 |
| Humerus | MAT 2_h | 2, 934, 645 |
| Humerus | MAT 3_h | 2,635,045 |
| Humerus | MAT 4_h | 1,697,166 |
| Humerus | MAT 5_h | 2,218,223 |
| Ceratobranchial | IMM 1_cb | 400,000 |
| Ceratobranchial | IMM 2_cb | 930,247 |
| Ceratobranchial | IMM 3_cb | 635,000 |
| Ceratobranchial | IMM 4_cb | 680,274 |
| Ceratobranchial | IMM 5_cb | 588,083 |
| Ceratobranchial | MAT 1_cb | 1,385,585 |
| Ceratobranchial | MAT 2_cb | 1,703,189 |
| Ceratobranchial | MAT 3_cb | 1,383,251 |
| Ceratobranchial | MAT 4_cb | 940,041 |
| Ceratobranchial | MAT 5_cb | 1,600,000 |

**Library preparation and deep RNA Sequencing.** RNA-seq libraries were prepared by the National Research Council (NRC, Saskatoon) using the Illumina TruSeq RNA Sample Prep Kit v2 with the following modification: the protocol was started at the Elute, Prime, Fragment step using 5 µl amplified mRNA (minimum amount was 50 ng mRNA as determined using Quant-iT RiboGreen

RNA Assay Kit (Invitrogen)). The quality of each cDNA library was checked on a DNA 1000 chip using the 2100 Bioanalyzer (Agilent Technologies Inc.).

**Reads preprocessing, mapping, quantitation and primary analysis of RNA-seq data.** The paired-end Illumina reads were trimmed using a Java -based tool, Trimmomatic v0.30 (Bolger et al. 2014), and the reads were then mapped to the chicken genome fom Ensembl using TopHat2 and Bowtie2 (Langmead and Salzberg 2012; Kim et al. 2013). The location of each read was matched to genome annotation using HTSeq-count (Anders et al. 2015). Differential expression analysis was performed using EdgeR after excluding genes with zero or very low counts (less than 3 counts for all cell types) across the cell type. The distribution of average $\log_2$ expression across three replicates of each tissue produced three bimodal distributions, which were used to set the count thresholds to 142 and 23 for immature cartilage and mature cartilage isolated from the ceratobranchial, and 37 and 58 for IMM and MAT isolated from the humerus. Pairwise comparisons between tissues were made with Fisher's exact test, and a gene was considered differentially expressed if it had an absolute $\log_2$ fold change greater than 2 ($p<0.01$). The Venn diagrams were constructed using gplots v3.0.1 for isoforms and RNA-seq expression data.

**GO analysis**. DAVID v6.8 (http://david.abcc.ncifcrf.gov/home.jsp) functional annotation analysis was performed on the list of differentially expressed genes (DEGs) in IMM, MAT, and OST, with a fold change > 2 or < -2 as well as, in all genes expressed above threshold (i.e. non-DEGs). The GO term biological process (BP) in DAVID was used to perform the gene-annotation enrichment analysis using *Mus musculus* as a background.

# CHAPTER 5

**Manuscript 3; Comparative transcriptomics reveals the gene regulatory network driving osteoblast differentiation and identifies higher expression of chondrocyte genes in gar osteoblasts (*in preparation*)**

Gomez-Picos P, Ovens K, Nguyen JKB, McQuillan I, and Eames BF

**CHAPTER 5: Comparative transcriptomics reveals the gene regulatory network driving osteoblast differentiation and identifies higher expression of cartilage genes in gar bone**

## 5.1.ABSTRACT

Defining the gene regulatory network (GRN) underlying osteoblast development is key to understanding the differentiation and even the origins of this skeletal cell type, but current knowledge is limited to a few candidate genes. Here, comparative transcriptomics was used to reveal aspects of the GRN driving osteoblast (OST) differentiation and to make an evolutionary comparison between bones of distinct clades. Bone is a vertebrate-specific tissue, and evidence from the fossil record suggests that it first originated in the throat or skin, in tooth-like structures named odontodes, but recent molecular evidence suggests that bone might have actually evolved from cartilage. Based on this idea, it is expected that osteoblasts of earlier diverged clades would be more 'chondrogenic' compared to osteoblasts of later diverged ones. Indeed, previous work with candidate genes support this idea. Here, we aim to test the hypothesis that gar OST express higher levels of 'chondrocyte' genes compared to mouse and chick OST. To test this hypothesis in an unbiased manner, LCM-RNA-seq was used to isolate OST from the dentary bone, a homologous element between mouse, chick, and gar. First, the consensus set of genes driving OST differentiation was revealed, and the known regulatory interactions were summarized into the format of a GRN. Regulatory interactions were validated based on published genetic molecular studies. Then, mouse, chick, and gar OST transcriptomes were compared in order to identify differences in gene expression between the three clades. Mouse and chick share 67% of the total number of OST genes. Tetrapod (i.e. mouse and chick) and gar osteoblasts share 44% of the total number of OST genes, including the master osteoblast regulator *Runx2* and many known OST markers regulated by this transcription factor, such as *Col1a1*, *Col1a2*, *Spp1*, *Ibsp*, *Satb2*, *Sparc*, and *Mmp13*. Some other important OST genes were not conserved in all three species including *Sp7*, *Alpl*, *Bmp4*, and *Bmp7*, and most of these OST genes (except for *Sp7*) were exclusively expressed in mouse and chick osteoblasts. The results presented here partially supported the hypothesis. Pairwise differential gene expression analyses revealed that gar OST expressed higher levels of some chondrogenic markers including *Col2a1*, *Sox6*, *Col10a1*, and *Acan* compared to mouse or chick OST, but not compared to both.

Moreover, model-based clustering analysis revealed one cluster that showed higher expression in the gar OST and included the hallmark mature chondrocyte gene *Col10a1*, but no other classic cartilage genes were included in this cluster. Comparative transcriptomics combined with functional studies and cis-regulatory analysis in osteoblasts of distinct species will contribute in expanding and verifying connections in this GRN in order to have a more comprehensive understanding of the early evolution of the OST.

## 5.2. INTRODUCTION

The emergence of genomic and transcriptomic technologies has greatly improved the identification of cell types, as well as evolutionary mechanisms underlying cell type specification (Arendt 2008; Arendt et al. 2019). Traditionally, classification schemes for skeletal cell types were mostly focused on morphology and function (Andrews et al. 1979; Cole and Hall 2004b; Cole and Hall 2004a). However, these traditional approaches can be problematic when cell types are compared across animal clades. Comparing gene expression patterns between jawless vertebrates (i.e. lamprey and hagfish) and cephalochordates (i.e. amphioxus) has greatly benefited the identification of homologous skeletal cell types across animal clades and has shed light on the evolution of cartilage and bone (Zhang and Cohn 2006; Zhang and Cohn 2006; Hecht et al. 2008; Cattell et al. 2011). Specifically, bone is a vertebrate-specific tissue, and evidence from the fossil record suggests that it first originated in the throat or skin, in tooth-like structures named odontodes (Smith and Hall 1990; Donoghue and Sansom 2002; Wagner and Aspenberg 2011). However based on recent molecular evidence, it has been suggest that bone might have evolved from cartilage (Gomez-Picos and Eames 2015; see chapter 2 for additional details, Fig. 6). To understand the origin and subsequent evolutionary changes underlying bone, osteoblasts across vertebrate clades should be analyzed and compared using unbiased transcriptomic approaches that allow the identification of the gene regulatory network (GRN) driving osteoblast differentiation (Fisher and Franz-Odendaal 2012; Gomez-Picos and Eames 2015).

Defining the GRN underlying osteoblast differentiation (i.e. portion of the GRN active in OST) is key to understanding the differentiation and even origins of this cell type, but current

knowledge is limited to a few candidate genes. Previous studies using mouse microarray data have aimed to reveal the GRN driving bone formation but they have some limitations for evolutionary studies, since they were performed on a single species (Someren et al. 2005; Calabrese et al. 2012). Here, we use an unbiased approach to reveal the portion of the skeletal cell GRN driving bone formation and make an evolutionary comparison between bones of mouse, chick, and gar. As discussed above, we previously hypothesized that the osteoblast might have evolved from a chondrocyte (see chapter 2 for details; Gomez-Picos and Eames 2015). Here, we aim to further test this hypothesis by comparing gene expression in bones of fish vs tetrapods.

What are the fundamental differences in gene expression between gar, mouse, and chick osteoblasts? Do these differences provide insight into the evolutionary origin of bone? Previous molecular studies have identified differences in gene expression between tetrapod (i.e. mouse and chick) and fish osteoblasts. For instance, 'chondrogenic' markers including *sox9, col2a1*, and *col10a1* were found to be expressed in fish developing bones, but not in tetrapod bones (Eames et al. 2012). Other studies in fish have also revealed the expression of chondrogenic genes in bones (Benjamin 1989; Benjamin 1990; Benjamin and Ralphs 1991), while the expression of these genes appears to be reduced in bones of aquatic tetrapods (i.e. Xenopus, Aldea et al. 2013; Bertin et al. 2015; Enault et al. 2015), and almost completely absent in mammals (Hilton et al. 2007). Based on these previous studies and our data (see chapter 6, Fig. 48B), we hypothesize that the gar osteoblast expresses higher levels of chondrogenic markers than tetrapod osteoblast. This lower expression of chondrocyte genes in more recently diverged vertebrates could be the result of a gradual repression of chondrogenic genes during evolution (Nguyen and Eames 2020). To test this hypothesis, osteoblasts were isolated from the dentary bone, a homologous element between mouse, chick, and gar, and then their transcriptomes were compiled. In general, the results presented here partially supported the hypothesis that gar osteoblasts express higher levels of chondrogenic genes compared to mouse or chick osteoblasts. Adding more species into these evolutionary comparisons, such as homologous bones isolated from an aquatic tetrapod (i.e. frog) or chondrichthyan bone-like tissue can verify and expand this initial GRN and confirm whether expression of chondrocyte genes has indeed been repressed in the osteoblasts of later diverged animals.

## 5.3. RESULTS

### 5.3.1. Osteoblasts were isolated from the dentary, a homologous element between mouse chick and gar

To analyze gene expression in the osteoblast, the present analyses were focused on the dentary bone, an intramembranous bone in the mandible adjacent to Meckel's cartilage which is homologous between mouse, chick, and gar. Alcian blue/Alizarin red staining confirmed the presence of the dentary bone adjacent to Meckel's cartilage in the three vertebrates (Fig. 31). In mouse, bone matrix of the dentary bone was evident by E14.5 (Ramaesh and Bard 2003). Likewise mineralized bone matrix deposition in the chick skull begins at around HH36 (Hamburger and Hamilton 1951; Hall and Miyake 1992; Dunlop and Hall 1995; Eames and Helms 2004; Fig. 31A,E). In gar, the presence of the dentary bone was confirmed by Alizarin red staining, adjacent to Meckel's cartilage (Fig. 31I) at 13dpf. Furthermore, Aniline blue staining was evident in the bone matrix of the dentary bone of mouse, chick, and gar (Fig. 31B,F,J). Histological analyses suggested that these three species were at similar stages of bone development, but not equivalent. While in mouse and chick, extensive mineralization of the dentary was evident at the stages analyzed, in gar mineralization was not as prominent (Fig. 31A,E,I). However, 13dpf was selected for gar due to the presence of teeth adjacent to the dentary bone at later stages which could potentially introduce contamination in captured osteoblasts (see chapter 6, Fig. 47).

Once skeletal elements and cells of interest were identified in mouse, chick, and gar, osteoblasts were isolated from the dentary bone (Fig. 31). This allowed capture of osteoblasts without chondrocyte contamination (Fig. 31C,D,G,H,K,L). Transcriptomic data obtained from OST presented in this chapter, were also used for the analyses presented in chapters 3 and 6.

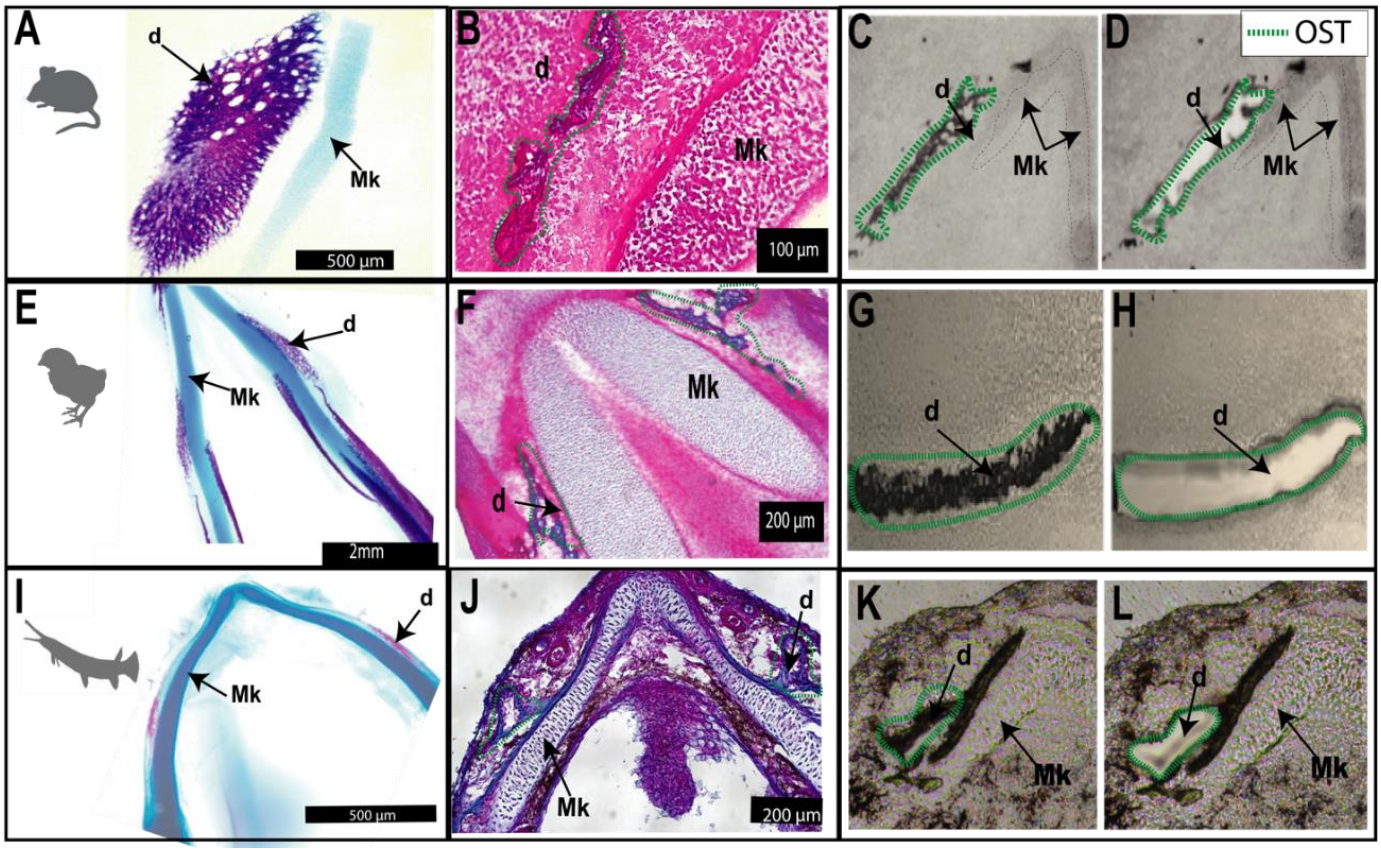



**Figure 31. Laser capture microdissection was used to isolate osteoblasts from the dentary bone, a homologous element between mouse (E14.5), chick (HH36), and gar (13dpf).** (A,E,I) Whole-mount Alcian blue and Alizarin red staining identified intramembranous bone in the mandibles of mouse, chick, and gar. (B,F,J) Trichrome-stained section of mouse, chick, and gar mandibles showed Aniline blue staining of bone matrix in the dentary (d), separated from Meckel's cartilage (Mk). Unstained sections of mouse E14.5, chick HH36, and gar 13dpf dentary bones before (C,G,D) and after (D,H,L) laser capture of osteoblasts. Abbreviations: OST, osteoblasts; Mk, Meckel's cartilage; d, dentary bone.

### 5.3.2. The consensus set of genes driving OST differentiation was revealed and regulatory interactions were summarized into the format of a GRN.

Deciphering the gene regulatory network (GRN) driving osteoblast differentiation is crucial to understand how this cell type might have evolved, but current knowledge is limited to a few candidate genes. Here, we used an unbiased approach to reveal for the first time the consensus set of genes driving osteoblast differentiation and make an evolutionary comparison between bones of three phylogenetic clades. For these inter-species comparisons, only one-to-one orthologs were

considered, so only a percentage of data is retained after normalization (mouse 65% genes retained; chick 75% genes retained; gar 70% genes retained). One-to-one orthologs are present in a single copy in the genome of each species and they are expected to conserve function after genome duplication events (Hubbard et al. 2009). Venn diagram analyses identified 9021 genes (one-to-one orthologs) expressed above threshold in the OST of all species (Fig. 32A). Of these, mouse, chick, and gar share 44% of the total number of genes expressed above threshold (one-to-one orthologs Fig. 32A, 3961 genes). Considering genes expressed in the OST of only two species, mouse and chick share the most genes (2091, Fig. 32A), and both mouse and chick share less genes with gar (Fig. 32A).

Genes shared between mouse, chick, and gar (one-to-one orthologs; 44% of total number of OST genes, Fig. 32A), were then subjected to gene ontology analyses. GO analysis revealed that 124 out of the 3961 genes were involved in processes related to bone formation, suggesting that these genes might be the core set of genes required for osteoblast differentiation (Fig. 32B, Table 5). We aimed to summarize the known regulatory interactions that underlie osteoblast differentiation into the format of a GRN using the modelling and visualization software BioTapestry (Fig. 33; Longabaugh et al. 2005). This OST GRN highlights several genes known to be involved in OST differentiation, but also genes whose role during osteogenesis still need to be elucidated. Regulatory interactions were validated based on published studies including genetic molecular experiments and cis-regulatory analyses performed in bones of mouse, chick, and zebrafish (Owen et al. 1990; Sodek et al. 1996; Ducy et al. 1997; Komori et al. 1997; Otto et al. 1997; Aslam et al. 1999; Dodig et al. 1999; Lee et al. 2000; Satokata et al. 2000; Göllner et al. 2001; Hu et al. 2001; Kundu et al. 2002; Miller et al. 2002; Nakashima et al. 2002; Yoshida et al. 2002; Kahler and Westendorf 2003; Lee et al. 2003; Kim et al. 2004; Stock et al. 2004; Wang et al. 2004; Yang et al. 2004b; Yang et al. 2004a; Burdan 2005; Guweidhi et al. 2005; Lee et al. 2005; Pratap et al. 2005; Vadlamudi et al. 2005; Xiao et al. 2005; Dobreva et al. 2006; James et al. 2006; Komori 2006; Holleville et al. 2007; Yu et al. 2007; Jensen et al. 2008; Kahler et al. 2008; Kimura et al. 2008; Luderer et al. 2008; Boumah et al. 2009; Fung Ling Chau et al. 2009; Hoeppner et al. 2009; Jensen et al. 2009; Komori 2009; Miclea et al. 2009; Sharff et al. 2009; Wu et al. 2009; Yu et al. 2009; Zhang et al. 2009; Lin et al. 2010; Sohaskey et al. 2010; Teixeira et al. 2010; Bond et al. 2011; Drabek et al. 2011; Miclea et al. 2011; Siqueira et al. 2011; Xu et al. 2011; Bonilla-Claudio et al. 2012; Chuang et al. 2012; Galli et al. 2012; Kokabu et al. 2012; Nishimura et al. 2012; Ohyama et al. 2012; Peruzzi et al. 2012; Stamper

et al. 2012; Li et al. 2013; Yang et al. 2013; Kawane et al. 2014; Wu et al. 2014; Bauer et al. 2015; Hill et al. 2015; Komori 2015; Pawaputanon Na Mahasarakham et al. 2015; Goto et al. 2016; Zhu et al. 2016; Bonyadi Rad et al. 2017; Frey et al. 2017; Nakatani and Partridge 2017; Giraud et al. 2018; Kawane et al. 2018; Kida et al. 2018; Komori 2018; Li et al. 2018; Qin et al. 2018; Robert et al. 2018; Zhu et al. 2018; Ishida et al. 2019; Komori 2019). Future functional studies including loss/gain of function genetic experiments in combination with epigenetic profiling such as, chromatin immunoprecipitation coupled to high-throughput sequencing (ChIP-Seq) and the assay for transposase-accessible chromatin with sequencing (ATAC-Seq) can validate direct genetic interactions in order to verify and expand this initial GRN. While ChIP-seq allows for the identification of specific transcription factor binding sites (Robertson et al. 2007), ATAC-Seq uncovers open chromatin regions where transcription factors can bind (Buenrostro et al. 2015).

Consistent with previous studies, the master osteoblast regulator *Runx2* was found to be highly expressed in osteoblasts of the three species and was placed at the top of the hierarchy of the GRN (see Fig. 33; Table 6). Several osteoblast markers including *Sparc*, *Mmp13*, *Dcn*, *Col1a1*, *Col1a2*, *Runx3*, *Msx2*, among others, were also conserved in the three vertebrate clades (see Fig. 33; Table 6). Many genes shared between the three species activate the expression of *Runx2* during osteoblast differentiation including *Bmp2*, *Dlx5*, *Msx2*, *Ctnnb1*, *Tcf7*, and *Foxo1* (Fig. 33; Holleville et al. 2007; Teixeira et al. 2010; Chen and Long 2013; Pawlowska et al. 2015), and *Runx2* then activates the expression of many important shared osteoblast specific genes, such as *Sp7*, *Mmp13*, *Sparc*, *Alpl*, *Spp1*, *Bglap*, *Ibsp*, *Fn1*, and *Fbn2* (Fig. 33; Golub and Boesze-Battaglia 2007; Boumah et al. 2009; Yu et al. 2009; Komori 2010; Stamper et al. 2012; Weng and Su 2013; Rosset and Bradshaw 2016). Interestingly, a few chondrogenic markers including *Sox9*, *Sox5*, and *Sox6* were also expressed in the OST of the three species. As discussed previously, during osteoblast differentiation SOX9 is downregulated by RUNX2 (Fig. 33; Cheng and Genever 2010). However SOX5 and SOX6 are necessary for osteoblast differentiation since they form an enhanceosome with SP7, CTNNB1, MEF2C, TCF7, DLX5, DLX6, SMAD1, and SMAD5 which binds to the *Runx2* enhancer and activates its expression in osteoblasts (Fig. 33; Komori 2015). Most of these genes, except for *Sp7*, *Smad1*, and *Smad5*, where expressed in mouse, gar, and chick osteoblasts, providing evidence on the importance of *Runx2* as a master regulator during bone differentiation. The sequence of this *Runx2* enhancer has been shown to be highly conserved in several vertebrates including human, mouse, and chicken (Komori 2015).

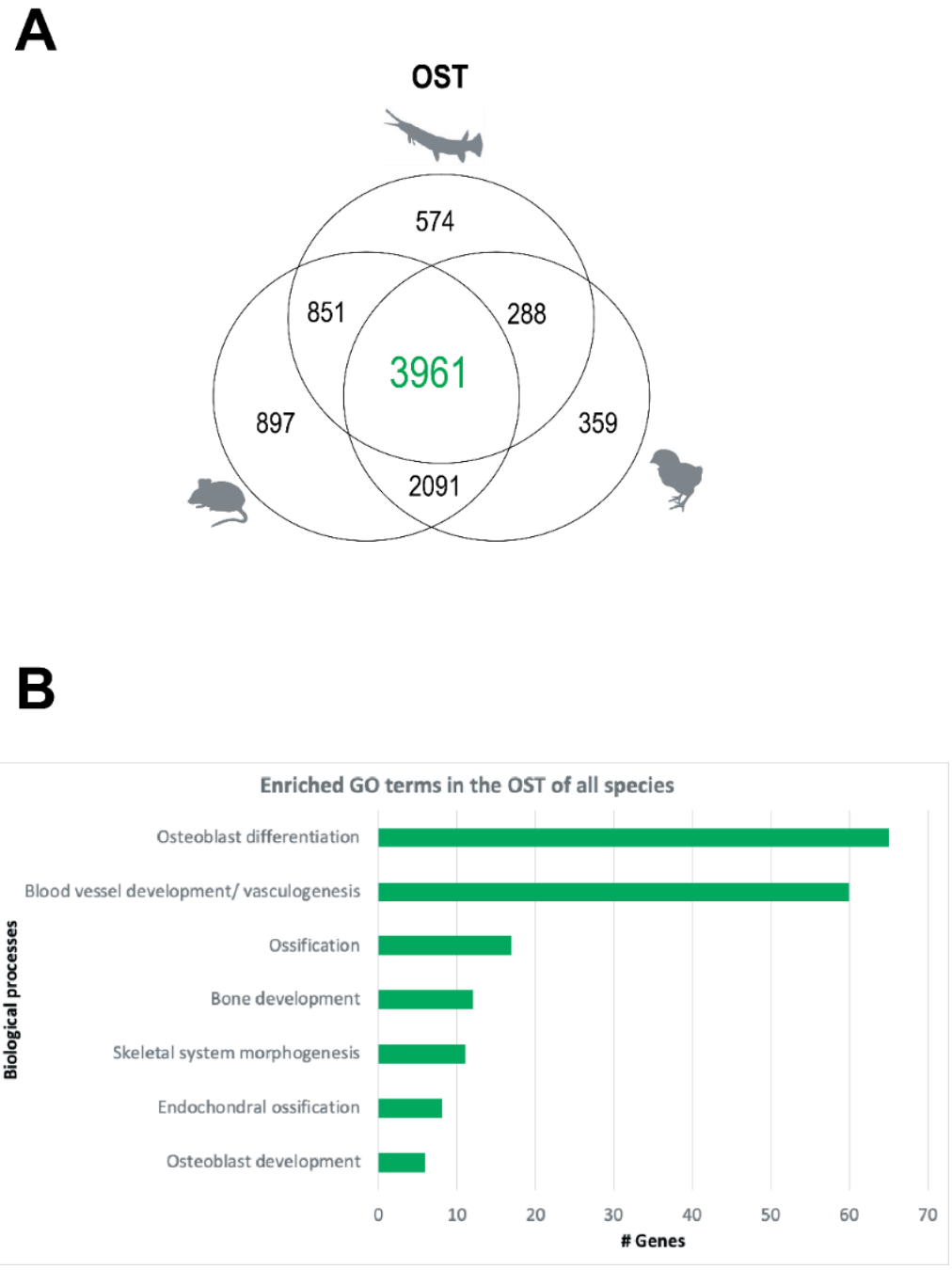



**Figure 32. Enriched GO terms in OST genes conserved between mouse, chick, and gar.** (A)Venn diagram showing genes expressed above threshold in OST of tetrapods and gar. The three species share 3961 genes (44% of total genes). (B) Gene ontology analyses of the 3961 conserved OST genes between tetrapods and gar reveals that 124 genes are involved in processes related to bone formation.

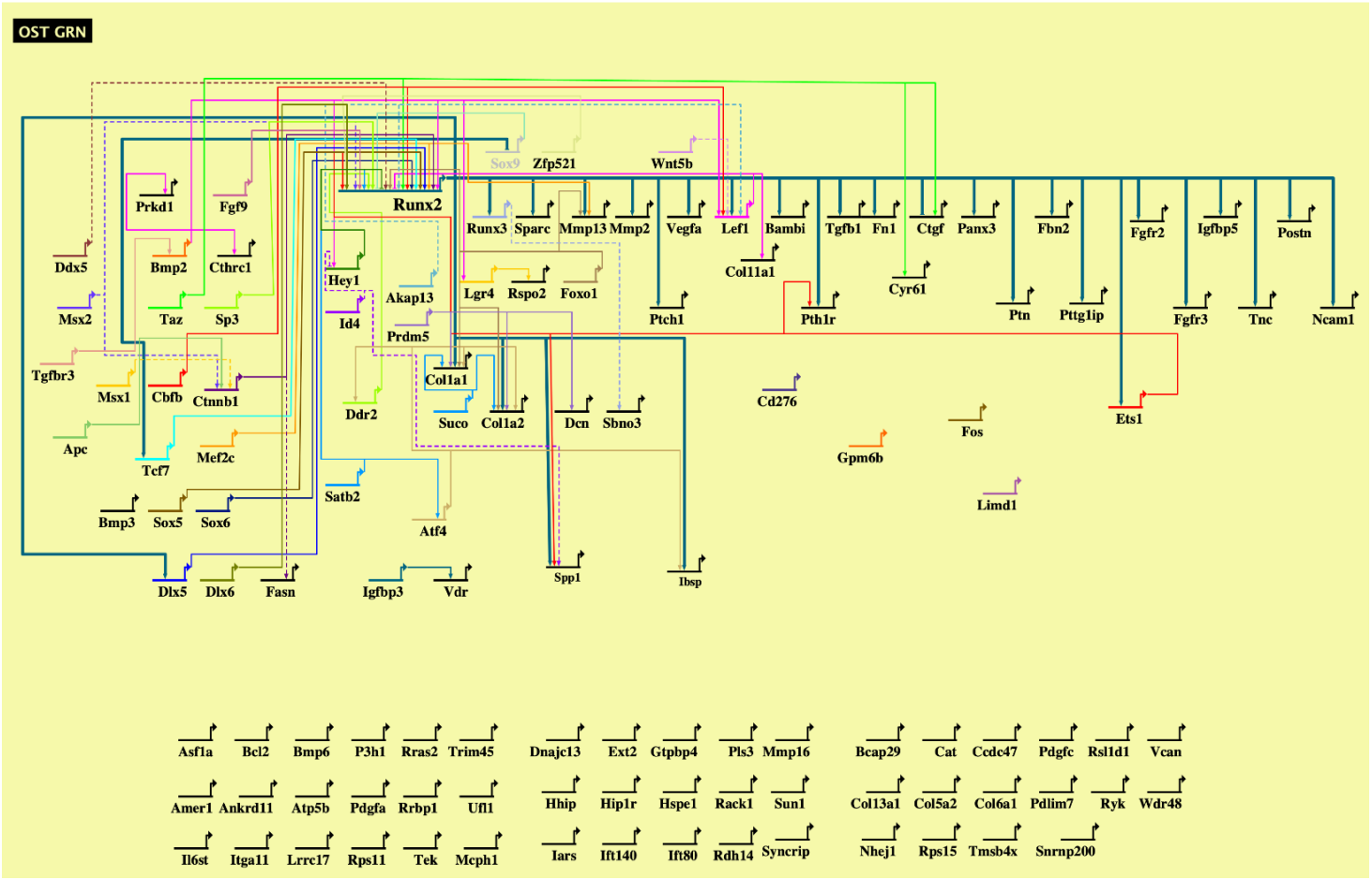



**Figure 33. Validated OST gene regulatory network included the core set of genes conserved between mouse, chick, and gar.** The 124 genes involved in bone differentiation processes were used to construct a consensus OST GRN. Interaction data was compiled from the literature. Genes that were annotated as osteoblast differentiation, but there is no further information on how these genes are regulated in the OST are included as single-nodes below the network. Arrowheads represent positive interaction, whereas negative interactions are depicted as –|. Interactions included in this network could be direct or indirect. Inactive nodes are represented in gray.

**Table 6.** Genes included in the putative OST GRN.

| Genes involved in GO terms related to OST in all species | | | | | | | |
|---|---|---|---|---|---|---|---|
| *Acvr2b* | *Ccdc47* | *Dlx6* | *Hip1r* | *Mef2c* | *Prdm5* | *Runx3* | *Tek* |
| *Akap13* | *Cd276* | *Dnajc13* | *Hspe1* | *Mmp13* | *Prkd1* | *Ryk* | *Tgfb1* |
| *Amer1* | *Col11a1* | *Ets* | *Iars* | *Mmp16* | *Ptch1* | *Satb2* | *Tgfbr3* |
| *Ankrd11* | *Col13a1* | *Ext2* | *Ibsp* | *Mmp2* | *Pth1r* | *Sbno2* | *Tmsb4x* |
| *Apc* | *Col1a1* | *Fasn* | *Id4* | *Mrc2* | *Ptk2* | *Snrnp200* | *Tnc* |
| *Asf1a* | *Col1a2* | *Fbn2* | *Ift140* | *Msx1* | *Ptn* | *Sox5* | *Trim45* |
| *Atf4* | *Col5a2* | *Fgf9* | *Ift80* | *Msx2* | *Pttg1ip* | *Sox6* | *Ufl1* |
| *Atp5b* | *Col6a1* | *Fgfr2* | *Igfbp3* | *Ncam1* | *Rack1* | *Sox9* | *Vcan* |
| *Bambi* | *Ctgf* | *Fgfr3* | *Igfbp5* | *Nhej1* | *Rdh14* | *Sp3* | *Vdr* |
| *Bcap29* | *Cthrc1* | *Fn1* | *Il6st* | *P3h1* | *Rps11* | *Sparc* | *Vegfa* |
| *Bcl2* | *Ctnnb1* | *Fos* | *Itga11* | *Panx3* | *Rps15* | *Spp1* | *Wdr48* |
| *Bmp2* | *Cyr61* | *Foxo1* | *Lef1* | *Pdgfa* | *Rras2* | *Suco* | *Wnt5b* |
| *Bmp3* | *Dcn* | *Gpm6b* | *Lgr4* | *Pdgfc* | *Rrbp1* | *Sun1* | *Zfp521* |
| *Bmp6* | *Ddr2* | *Gtpbp4* | *Limd1* | *Pdlim7* | *Rsl1d1* | *Syncrip* | |
| *Cat* | *Ddx5* | *Hey1* | *Lrrc17* | *Pls3* | *Rspo2* | *Taz* | |
| *Cbfb* | *Dlx5* | *Hhip* | *Mcph1* | *Postn* | *Runx2* | *Tcf7* | |

Some other important OST genes, such as *Spp1* and *Ibsp* did not have a clear one-to-one ortholog in all three species. Importantly, previous work identified *spp1* and *ibsp* in jaw and scales of gar (Braasch et al. 2016), but these genes are not considered to be one-to-one orthologs with mouse and chick according to Ensembl orthology quality controls (Aken et al. 2016). Mouse and chick *Spp1* and *Ibsp* are considered one-to-one orthologs according to Ensembl, but these genes are not considered to be one-to-one orthologs with any Ray-finned fish including gar (e.g. Mouse *Ibsp*-https://uswest.ensembl.org/Mus_musculus/Gene/Compara_Ortholog?db=core;g=ENSMUSG00000029306;r=5:104299171-104311469;t=ENSMUST00000031246). However due to the high expression levels of *spp1* and *ibsp* in gar OST, these genes were also included in this putative OST GRN. Other classic OST markers were not conserved in all three species including *Sp7*, *Alpl*, *Bmp4*, and *Bmp7* and most of these OST genes (except for *Sp7*) were exclusively expressed in mouse and chick osteoblasts.

### 5.3.3. Mouse and chick osteoblasts share the expression of many genes, and both share fewer genes with gar osteoblasts

Comparative transcriptomics was used to show that mouse and chick osteoblasts are more similar to each other compared to gar osteoblasts. Specifically, we aim to test the hypothesis that gar osteoblasts express higher levels of 'chondrogenic' genes than tetrapod osteoblasts. To test this hypothesis, unique gar OST genes (574 genes; Fig. 32A) were subjected to gene ontology analyses. GO analyses revealed that many neural-crest related processes were enriched in the gar OST, but none of this was related to chondrocyte differentiation, which do not support the hypothesis. The hypothesis also predicts that mouse and chick OST will have less differentially expressed genes compared to gar OST. Indeed, pairwise differential gene expression analyses show that tetrapod OST have fewer differentially expressed genes compared to gar OST (Fig. 34). A total of 1766 DEGs (Fig. 34; 522 upregulated and 1244 downregulated) were differentially expressed between mouse and chick. When gar OST was compared to chick OST, 3631 DEGs were identified (Fig. 34; 1268 upregulated and 2363 downregulated), and when gar OST was compared to mouse OST, a total of 2940 DEGs were identified (Fig. 34; 2300 upregulated and 640 downregulated). These results suggest that mouse and chick OST are more similar to each other than to gar OST (Figs. 32A, Fig. 34).

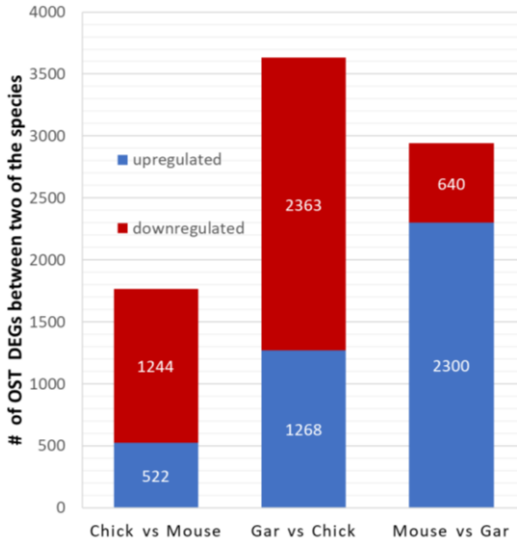



**Figure 34. Tetrapod OST have less differentially expressed genes compared to gar OST.** Mouse and chick OST have less differentially expressed genes (DEGs, absolute $\log_2$ FC greater than 2 ($p<0.05$)) compared to gar OST. The first bar shows upregulated (and downregulated) genes in chick compared to mouse, the second bar shows upregulated (and downregulated) genes in gar compared to chick, and the third bar shows upregulated (and downregulated) genes in mouse compared to gar. Gene lists used for DEGs graphs were generated by Katie Ovens.

To test the hypothesis that fish OST express higher levels of chondrogenic genes compared to tetrapod OST, we then focused on DEGs between tetrapods and gar. Interestingly, some chondrogenic markers show higher levels in gar compared to mouse or chick. For instance, differentially expressed genes between gar and chick osteoblasts include cartilage differentiation and maturation genes, such as *Col2a1*, *Sox6*, *Runx3*, and *Wnt5b*, which show higher expression in gar OST (Table 7; Zhao et al. 1997; Smits et al. 2001; Yang et al. 2003; Yoshida et al. 2004).

Gene ontology analyses revealed that genes upregulated in gar compared to chick OST are involved in processes related to cartilage differentiation. Differentially expressed genes between gar and mouse OST, with higher levels in gar, include the chondrocyte genes *Acan*, *Runx3*, *Wnt5b*, and *Pthlh* (Watanabe et al. 1998; Minina et al. 2002; Yang et al. 2003; Yoshida et al. 2004). GO analyses reveal that these and other DEGs upregulated in gar are involved in chondrocyte differentiation, cartilage development, and endochondral ossification (i.e. cartilage maturation, Table 7). These results partially support the hypothesis that gar OST express higher levels of 'chondrogenic' markers compared to tetrapods, since only a few of these genes (Table 7) were shown to be upregulated in gar compared to mouse or chick, but not compared to both species.

**Table 7.** Genes involved in cartilage differentiation processes upregulated in gar OST vs chick or mouse OST included in Fig.53B

| Upregulated 'chondrocyte' genes in gar OST vs chick OST | |
|---|---|
| *Bmp1* | *Osr2* |
| *Bmp5* | *Pax7* |
| *Bmp6* | *Runx3* |
| *Col2a1* | *Shox2* |
| *Dicer1* | *Six2* |
| *Dlx2* | *Sox6* |
| *Evc* | *Wnt5b* |
| *Hspg2* | *Wnt7a* |
| *Mapk14* | *Zbtb16* |

| Upregulated 'chondrocyte' genes in gar OST vs mouse OST | |
|---|---|
| *Acan* | *Pthlh* |
| *Hhip* | *Hspg2* |
| *Cytl1* | *Evc* |
| *Osr2* | *Dicer1* |
| *Hspg2* | *Hyal1* |
| *Runx3* | *Pax7* |
| *Wnt5b* | *Wnt7a* |
| *Wnt7a* | *Zbtb16* |

### 5.3.4. Cluster analyses reveal gene expression differences between tetrapod and gar osteoblasts

Cluster analysis was performed in OST of gar, chick, and mouse (Fig. 35). We predict that more clusters will show similar expression levels between mouse and chick OST compared to gar OST, and that enriched clusters in gar OST will include a higher number of chondrocyte genes. Out of the nine clusters included in this analysis, three clusters showed similar gene expression levels between mouse and chick, which included over 30% of the total number of genes (2944 genes, Fig. 35B,C,E). In clusters 1 and 4, gene expression was statistically lower in gar compared to mouse and chick (Fig. 35B,E), whereas in cluster 2 gene expression was higher in gar, and statistically different compared to mouse and chick (Fig. 35C). Cluster 2 included the hallmark mature chondrocyte gene *col10a1,* but no other classic cartilage genes were identified in this cluster, which do not support our hypothesis (Fig. 35C; Lu et al. 2014). Previous work and our data have shown that *col10a1* is highly expressed in the gar OST (see chapter 6; Fig. 46E), but not in tetrapod OST (Eames et al. 2012).

Other clusters showed enhanced expression in the OST of one species vs the other two. For example, cluster 3 showed higher expression in the chick OST (Fig. 35D), whereas cluster 6 showed higher expression in the mouse OST (Fig. 35G). Among the clusters discussed above, some also showed similar expression levels between only two species. For example, clusters 3 and 5 showed statistically similar expression levels between mouse and gar OST (approx. 5% genes, Fig. 35D,F). Cluster 6 showed similar low gene expression between chick and gar (approx. 3% of the total genes, Fig. 35G), and both species were significantly different from mouse. Cluster 9 included almost 50% of the total number of genes, and expression was not statistically similar between mouse, chick, and gar osteoblasts (Fig. 35J). This cluster encompassed important osteoblast markers such as *Runx2, Bmp2, Hey1, Dlx5, Sparc, Col1a2,* and *Mmp13* which are expressed in the osteoblasts of all three species, and were also included in our putative OST GRN (Fig. 33; Table 6). Enriched biological processes in cluster 9 include osteoblast differentiation, bone development, apoptosis, angiogenesis, ossification, and blood vessel development. Although cluster analysis shows that gene expression levels in mouse and chick OST are more similar to each other than compared to gar, in general this analysis only partially supported the hypothesis that gar osteoblasts express higher levels of chondrogenic genes than tetrapod osteoblasts.

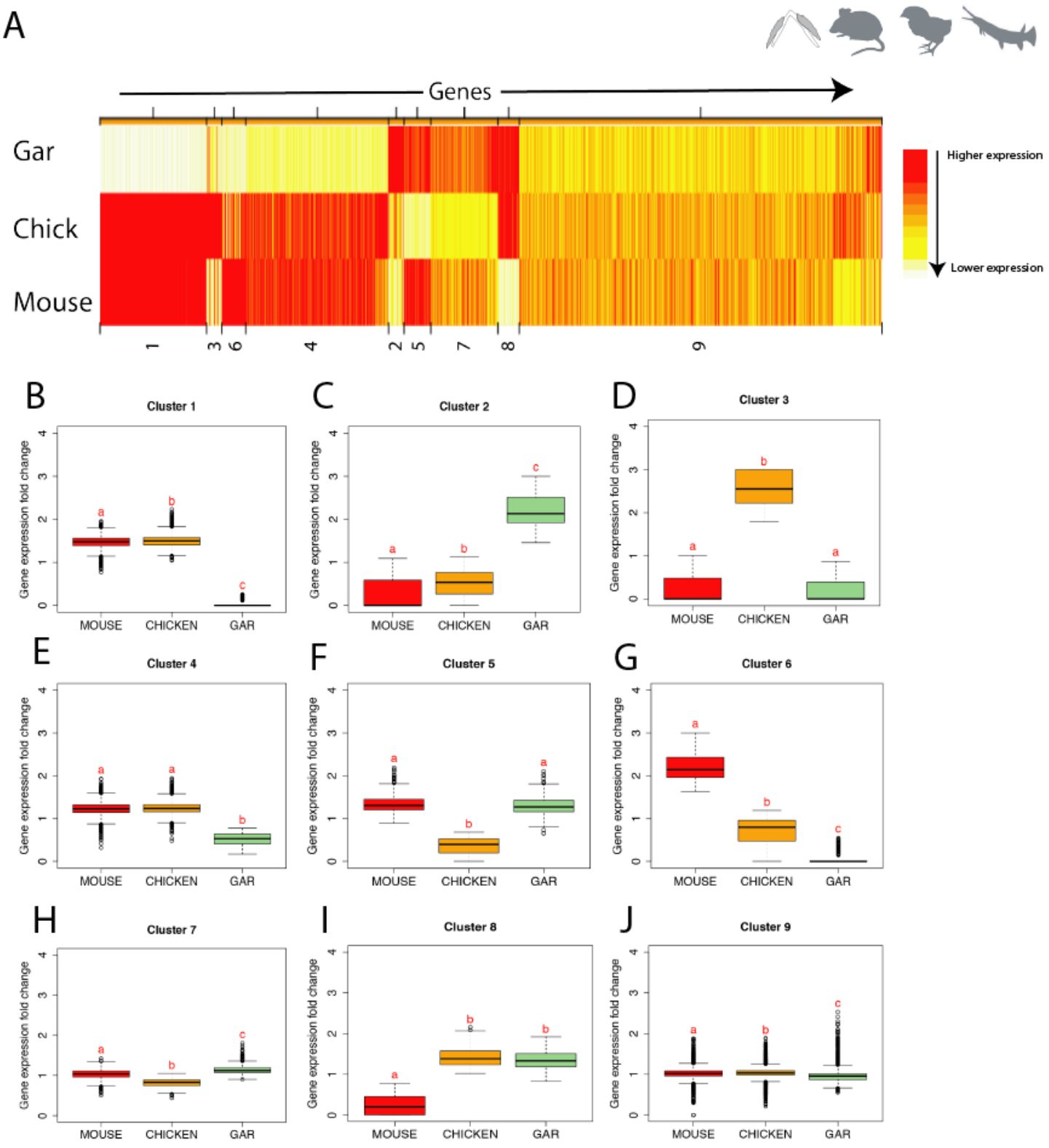



**Figure 35. Model based clustering comparing gene expression in OST of mouse, chick, and gar.** Three clusters (1, 2, 4) showed similar gene expression in tetrapods (only cluster 2 statistically significant), which included around 34% (2944 genes) of the total number of genes. In all these clusters, gene expression in mouse and chick was statistically different from gar (Credit: Katie Ovens).

## 5.4. DISCUSSION

Revealing the GRN driving bone formation can shed light into molecular mechanisms of differentiation and even evolutionary origins of this vertebrate-specific tissue, but current studies are limited to some candidate genes in only a few species (Someren et al. 2005a; Someren et al. 2005b; Calabrese et al. 2012). Here, we focused on the osteoblast and made an evolutionary comparison between bones of gar, mouse, and chick to reveal the consensus GRN underlying osteoblast differentiation, and to identify differences between bones of these vertebrates. Using Venn diagram coupled to gene ontology analyses we identified the consensus set of genes driving OST differentiation. A total of 124 genes conserved in mouse, chick, and gar were identified to be specifically involved in processes related to osteoblast differentiation (Roach 1994; Zelzer et al. 2001; Kundu et al. 2002; Lian and Stein 2003; Kim et al. 2004; Stock et al. 2004; Wang et al. 2004; Pratap et al. 2005; Dobreva et al. 2006; Nie et al. 2006; Holleville et al. 2007; Hoeppner et al. 2009; Sharff et al. 2009; Datta et al. 2010; Miclea et al. 2011; Nishimura et al. 2012; Nassif et al. 2014; Wu et al. 2014; Pawaputanon Na Mahasarakham et al. 2015; Pawlowska et al. 2015; Goto et al. 2016; Rosset and Bradshaw 2016; Zhang et al. 2017; Kawane et al. 2018; Qin et al. 2018; Komori 2019), and these were used to construct a core OST GRN (Figs. 32&33). Importantly, the master osteoblast regulator *Runx2* and many known OST markers regulated by this transcription factor, such as *Col1a1*, *Col1a2*, *Spp1*, *Ibsp, Satb2*, *Sparc*, and *Mmp13* were conserved between mouse, chick, and gar (see Table 6; Sodek et al. 1995; Drissi et al. 2000; Lian and Stein 2003; Pratap et al. 2005; Dobreva et al. 2006; James et al. 2006; Hoeppner et al. 2009; Nishimura et al. 2012; Rosset and Bradshaw 2016; Komori 2018; H. et al. 2019).

Regulatory interactions in this putative GRN were validated using published work including functional experiments in bones and cis-regulatory analyses. Some other important OST genes were not conserved in all three species including *Sp7, Alpl*, *Bmp4*, and *Bmp7* (Nie et al. 2006; Golub and Boesze-Battaglia 2007; Chang et al. 2009; Yano et al. 2014; Niu et al. 2017), and most of these OST genes (except for *Sp7*) were exclusively expressed in mouse and chick osteoblasts. *Sp7* was expressed in mouse and gar osteoblasts in our data, but no clear *SP7* ortholog has been identified in chick (Kague et al. 2016). *Alpl*, *Bmp4* and *Bmp7* were not expressed in gar osteoblasts in our data, but both genes have been identified in scales, jaw, and vertebrae of teleost (Coble 1966; Johnston et al. 1994; Hwang et al. 1997; Nie et al. 2006). Together, these data suggest that the GRN driving

OST differentiation has been modified throughout evolution perhaps conferring novel functions to the bones of more recently diverged vertebrates, such as tetrapods. Comparative transcriptomic analyses including skeletal cells isolated from species that lack mineralized bone can also help to identify the consensus set of genes underlying osteoblast differentiation and by incorporating these species, a GRN which might even resemble the original ancestral OST GRN could be revealed. For example, expression of ancestral genes to *Runx2* (i.e. *RunxA* and *RunxB)*, has been revealed in jawless vertebrates (i.e. hagfish and lamprey; Hecht et al. 2008; Cattell et al. 2011), as well as invertebrates (i.e. amphioxus; Hecht et al. 2008). Furthermore, expression of classic bone markers such as *Runx2* and *Col1* has been identified in bone-like tissue in some species of Cartilaginous fish (Eames et al. 2007; Nah et al. 2014).

This work elucidates how genes are organized into a GRN in the OST *in vivo* which is crucial to understand the function and evolution of this vertebrate-specific cell type. A previous study using gene co-expression network analyses and perturbation data obtained from microarray profiling of mouse bone samples was able to identify a group of 354 highly correlated genes that were used to construct an OST specific GRN (Calabrese et al. 2012). Genes identified in this previous network include many important genes during osteoblast differentiation such as, *Col1a1, Col1a2, Bglap1, Sp7, Ibsp, Mmp13, Alpl,* and *Phex* (Nakashima et al. 2002; Inada et al. 2004; Golub and Boesze-Battaglia 2007; Yuan et al. 2008; Nishimura et al. 2012; Yano et al. 2014; Yang et al. 2016; Niu et al. 2017), and most of these genes were conserved in the core OST GRN presented here (Fig. 33, Table 6). In another experiment, RNA extracted from an osteoblastic mouse cell line was subjected to microarray gene expression analysis (Someren et al. 2005b). In this experiment, an OST GRN was inferred *in vitro* using bioinformatic network analyses. Classic OST genes included in this inferred network such as, *Dcn*, *Postn,* and *Col1a1* (Mochida et al. 2009; Seto et al. 2017; Zhang et al. 2017) were also included in the consensus OST GRN presented here (Fig. 33, Table 6)**.** These previous studies, however, have some limitations since they were performed on a single species, whereas the OST GRN presented here, included transcriptomic data from three distinct vertebrate clades making the study more robust. Also, as discussed in chapter 2, RNA-seq produces the most robust, unbiased results, and it has the ability to detect genes that are expressed at very low levels as well as, novel genes and splice variants, which is not possible to obtain from microarrays (Wang et al. 2009; Necsulea and Kaessmann 2014).

Finally, we focused on revealing the differences in gene expression levels between gar OST and tetrapod OST. Specifically, we aimed to test the hypothesis that gar expressed higher levels of 'chondrogenic genes' compared to tetrapods. This hypothesis was only partially supported by our results. First, gar osteoblasts expressed higher levels of some chondrogenic genes compared to mouse or chick, but not compared to both species (see Table 7). However the expression levels of many other cartilage genes including the master cartilage regulator *Sox9*, were not higher in gar OST compared to mouse and/or chick OST. Second, cluster analysis comparing OST of all three species revealed that one cluster that showed enriched expression in the gar OST included the hallmark mature chondrocyte gene *Col10a1*. These results are also supported by previous findings in fish and amphibians that have revealed that *Col10a1* is not limited to chondrocyte hypertrophy, but it might also have a role during osteoblast differentiation, at least in nonamniote vertebrates (Laue et al. 2008; Albertson et al. 2010; Eames et al. 2012; Aldea et al. 2013). In clusters 1 and 4, gene expression was statistically lower in gar compared to mouse and chick (Fig. 35B,E). Water-to-land transition required the development of novel muscular and skeletal structures to withstand the forces of gravity and facilitate locomotion (Wood and Nakamura 2018). Some enriched biological processes in these clusters that showed higher expression in tetrapods compared to gar were related to smooth muscle cell proliferation and angiogenesis including genes such as, *Vegfc*, *Fgf1*, *Fgf2, Fgfr1*, and *Vash1* (Hinsbergh and Rabelink 2005; Jacob et al. 2006; Benest et al. 2007; Kimura et al. 2009; Sato 2012; Su et al. 2014; Hu and Olsen 2016; Charoenlarp et al. 2017) which could be related to higher remodeling activity in tetrapod bones as a result of increased loading and fracture healing (Hankenson et al. 2011; Herbst et al. 2019). It has also been proposed that during early tetrapod evolution, vascularization increased in dermal bones in order to buffer the respiratory acidosis caused by elevated $CO_2$ during transition to land (Janis et al. 2012).

To define how an osteoblast transcriptome can be considered more 'chondrogenic' several criteria might be used for future analyses, as suggested by others (Nguyen and Eames 2020). For instance, instead of comparing a few selected genes, chondrocyte and osteoblast transcriptomes could be compared within a single species in order to reveal the percentage of genes shared, and then compare how this percentage varies across phylogenetic clades (Nguyen and Eames 2020). In chapter 6, we further analyzed the hypothesis that gar OST express higher levels of chondrogenic genes using this approach. For example, our data shows that considering only genes expressed in immature chondrocytes (IMM) and osteoblasts (OST), the overlap between these cell types is higher

in gar (see chapter 6 for more details, Fig. 48B) compared to mouse or chick (see chapter 6, Fig. 37A,B; Fig. 41A; Fig. 48A), which gives further support to the hypothesis. From an evolutionary perspective, overall, these data support that the osteoblast might have evolved from a chondrocyte by co-opting portions of an ancestral GRN initially used to make cartilage (see chapter 2 for details, Fig. 6).

Concepts to be developed in the future include gain- and loss- of function studies in osteoblasts of distinct vertebrates coupled to cis-regulatory analysis which could contribute in expanding and verifying connections in this GRN. To further analyze OST GRN conservation across animal clades, transgenic approaches can be used to test whether transcription factors enriched in osteoblasts of one species can bind and activate the expression of genes in osteoblasts of other species. For example, previous work has provided evidence that the GRN driving chondrogenesis might be conserved in vertebrates and invertebrates, since *SoxE* from a hard-shell invertebrate was able to activate the expression of the human *COL2A1*, revealing deep conservation in the chondrocyte GRN (Tarazona et al. 2016). In summary, comparative transcriptomics coupled with cis-regulatory analysis, as well as gain- and loss-of-function studies in osteoblasts of distinct species will contribute in expanding and verifying connections in this initial GRN, and will expand our understanding on the evolutionary origin of bone.

## 5.5.METHODS

**Embryo Collection and tissue processing.** All animal procedures were performed according to guidelines approved by the University of Saskatchewan Animal Care and Use Committee. *Chick (Gallus gallus):* White leghorn chicken eggs were incubated in a humified incubator at a constant temperature of 37ºC. Embryos were harvested at Hamburger-Hamilton stage 36 (~E10.5; (Hamburger and Hamilton 1951). Each embryo was decapitated, and the lower jaws were dissected and immediately placed in 1X PBS/DEPC, followed by embedding in OCT (Tissue Tek, Torrance, CA, USA; for detailed methodology see chapter 3), and immediately flash-frozen using liquid N2 and 2-Methylbutane (isopentane). *Spotted gar (Lepisosteus oculatus):* Gar larvae were kindly provided by Dr. Allyse Ferrara (Nicholls State University Thibodeaux, LA). Larvae were placed in beakers and raised at a constant temperature of 24ºC and a pH between 6.5-8 in a water bath. Once

gar larvae reached 13 and 14 dpf they were euthanized using 0.2% tricaine. The length of these larvae varied between 1.6-2.2 cm. Heads were dissected and immediately placed in 1X PBS/DEPC, followed by embedding in OCT. For whole mount and section histology, both chick and gar dissected tissues were fixed in 4% PFA prior embedding in OCT (in the case of section histology).

**Laser Capture Microdissection (LCM).** LCM was performed on a Laser Microdissection - Molecular Machines & Industries (MMI) CellCut apparatus. For gar and chick analyses, five biological replicates for each tissue were captured. *Chick:* Osteoblasts were isolated from the developing dentary in the jaw at HH36. *Spotted gar:* Osteoblasts were isolated from the developing dentary in the jaw at 13 dpf. Similar to mouse, a larger capture area was necessary to obtain a sufficient amount of RNA from mature cartilage of both chick and gar (Tables 7&8). The captured cells were collected onto the inner lid of 0.5ml MMI IsolationCaps (either Diffuser caps (Prod#50202) or Transparent caps (Prod#50204; MMI Molecular Machines & Industries). Once RNA-seq data was obtained, bioinformatic analyses identified some samples as outliers, and these were not included in our analyses. For details on bioinformatic analyses see chapter 3 methodology.

**RNA Isolation and Amplification**. RNA was isolated using the ARCTURUS PicoPure RNA Isolation Kit (ThermoFisher Scientific; Cat# KIT0204) according to the manufacturer's instructions and DNase treatment was done using RNase-Free DNase (Qiagen; Cat#79254). Osteoblasts were captured from mouse, chick, and gar dentary bones (Table 8), and then RNA was extracted (mouse, n=3; chick, n=5; gar, n=5). RNA was then amplified with one (mouse and chick RNA) or two rounds (gar RNA) using MessageAmp II aRNA Kit (ThermoFisher Scientific; Cat# AM1751). OST data obtained from these samples was also used for some analyses presented in chapters 3 and 5. The RNA integrity was evaluated on the observation of a signature eletropherogram pattern (Bioanalyzer).

**Table 8.** LCM Surface area obtained from mouse, chick, and gar dentaries

| Species | Skeletal element | Cell type | Surface Area captured (um2) |
|---|---|---|---|
| Mouse | Dentary | OST 1 | 218,981 |
| | Dentary | OST 2 | 396,441 |
| | Dentary | OST 3 | 759,814 |
| Chick | Dentary | OST 1 | 835,931 |
| | Dentary | OST 2 | 1,817,000 |
| | Dentary | OST 3 | 1,630,000 |
| | Dentary | OST 4 | 1,005,075 |
| | Dentary | OST 5 | 1,497,360 |
| Gar | Dentary | OST 1 | 239,200 |
| | Dentary | OST 2 | 119,862 |
| | Dentary | OST 3 | 136,791 |
| | Dentary | OST 4 | 192,750 |
| | Dentary | OST 5 | 230,961 |

**Library preparation and deep RNA Sequencing.** RNA-seq libraries were prepared by the National Research Council (NRC, Saskatoon) using the Illumina TruSeq RNA Sample Prep Kit v2 with the following modification: the protocol was started at the Elute, Prime, Fragment step using 5 µl amplified mRNA (minimum amount was 50 ng mRNA as determined using Quant-iT RiboGreen RNA Assay Kit (Invitrogen)). The quality of each cDNA library was checked on a DNA 1000 chip using the 2100 Bioanalyzer (Agilent Technologies Inc.).

**GO analysis**. DAVID v6.8 (http://david.abcc.ncifcrf.gov/home.jsp) functional annotation analysis was performed on the list of differentially expressed genes (DEGs) in IMM, MAT, and OST, with a fold change > 2 or < -2 as well as, in all genes expressed above threshold (i.e. non-DEGs). The GO term biological process (BP) in DAVID was used to perform the gene-annotation enrichment analysis using *Mus musculus* as a background.

**Validated Osteoblast GRN.** The OST GRN was constructed using BioTapestry version 7.1.2 (www.BioTapestry.org/) following developer's protocol (Longabaugh et al. 2005; Longabaugh et al. 2009). Regulatory interactions were validated using published work on bones of mouse, chick, and zebrafish.

## CHAPTER 6

**Manuscript 4; Skeletal cell GRN structure and regulatory control of mature chondrocytes are highly conserved features between mouse and chick, while gar shows more variation (*in preparation*).**

Gomez-Picos P, Ovens K, McQuillan I, and Eames BF

# CHAPTER 6: Skeletal cell GRN structure and regulatory control of mature chondrocytes are highly conserved features between mouse and chick, while gar shows more variation

## 6.1.ABSTRACT

Skeletal cell evolution has been little explored since most evolutionary studies focus on morphological differences between homologous skeletal elements in various phylogenetic clades rather than differences in skeletal cell histogenesis. The present chapter focuses on the evolution of skeletal cells since molecular studies have revealed clade specific differences between homologous skeletal cell types. Specifically, this work aims to shed light into the evolution of skeletal cell gene regulatory networks (GRNs). Most of the vertebrate skeleton is composed of three skeletal cell types, immature chondrocytes (IMM), mature chondrocytes (MAT), and osteoblasts (OST). We have previously shown in mammals that the differentiation of IMM, MAT, and OST is directed by one GRN, and distinct portions of this GRN are active in each cell type. In mouse, interactions between portions of the GRN regulating IMM and OST regulate gene expression in MAT. Model-based clustering analyses revealed specific mechanisms of gene regulation in mouse MAT: independent and interacting (via averaging and synergism). We hypothesize that GRN structure and mechanisms of gene regulation in mature chondrocytes are conserved features across vertebrates. To test this hypothesis, homologous skeletal cells (i.e. IMM, MAT, and OST) in mouse, chick, and gar were isolated and their transcriptomes were revealed. Venn diagram and differential gene expression analyses revealed that in mouse and chick, MAT and OST shared the most genes, whereas IMM and OST were the most different cell types. In gar, however IMM and OST shared the most genes and MAT was the most different cell type. Furthermore, gene co-expression network (GCN) analyses in mouse and chick skeletal cells revealed that two independent portions of the GRN active in IMM and OST show strong cross-inhibition with each other, and they positively interacted in MAT. In contrast, GCN analysis in gar showed increased positive correlations between genes enriched in IMM and OST. Model-based clustering analysis identified specific categories of gene expression that were generally conserved between mouse, chick, and gar. Some genes in MAT had expression levels similar to IMM or OST, suggesting that portions of the GRN active in IMM or OST are

independently regulating the expression of genes in MAT in a similar manner as in these two cell types. Other genes had expression levels that represented an averaging or synergy between levels in the IMM and OST, suggesting that portions of the GRN active in these cell types are negatively or positively, respectively, interacting in MAT to regulate the expression of a specific set of genes. These mechanisms of gene regulation were conserved across the three species, but when gar skeletal cells were analyzed more clusters showed a dramatic downregulation of gene expression in MAT. These results in gar suggest that in earlier diverged clades, MAT was more different to IMM and OST, and this cell type acquired more similarities in gene expression over time. In general, these data highlight the complexity of skeletal cell GRN organization and evolution, and proposes a novel experimental approach to compare GRN structure across clades, as well as to identify evolutionary changes in regulatory control of skeletal cell differentiation that could ultimately shed light into evolutionary origins of cartilage and bone.

## 6.2.INTRODUCTION

The recent advances in comparative transcriptomics and genomics have illuminated in great extent many events in skeletal evolution, but the organization and evolution of the gene regulatory network (GRN) underlying skeletal cell types are not well understood. A GRN describes regulatory interactions between genes in a given biological process (Davidson and Erwin 2006; Peter and Davidson 2011). Despite the current lack of data, further testing might support that GRN driving skeletal cells includes a core portion of conserved genes that is homologous in distinct animal clades, while other species-specific portions of the GRN show more variation perhaps in response to specific selective pressures (Fig. 2A). These modifications in specific portions of the GRN can give rise to evolutionary novelties (Wagner and Lynch 2010; Purnell 2020). Comparative transcriptomics can be used to identify conserved portions of the GRN and novelties by revealing skeletal cell GRNs in different phylogenetic clades.

The majority of the vertebrate skeleton is composed of three main cell types: immature chondrocytes (IMM), mature chondrocytes (MAT), and osteoblasts (OST; Eames et al. 2004; Gomez-Picos and Eames 2015; Li and Dong 2016; Aghajanian and Mohan 2018). These three cell

types are distinct, yet they also share many histological and molecular features (Hardingham 1981; Fang and Hall 1997; Kronenberg 2003; Eames and Helms 2004; Eames et al. 2004; Day et al. 2005; Hill et al. 2005; Gentili and Cancedda 2009; Gomez-Picos and Eames 2015), suggesting that their underlying GRNs are also related. Since previous work has shown that skeletal cells evolve (Volkmann and Baluska 2006; Zhang and Cohn 2006b; Zhang et al. 2006a; Cattell et al. 2011; Eames et al. 2012; Jandzik et al. 2015; Tarazona et al. 2016), the present work aims to compare skeletal cell transcriptomes globally within and between different clades in order to test how evolutionary changes have impacted skeletal cell GRN structure. We have previously shown in mouse that interactions between portions of the GRN regulating IMM and OST regulate gene expression in MAT (see chapter 3 for more details). Using comparative transcriptomics, we aim to test the hypothesis that GRN structure and mechanisms of gene regulation in mature chondrocytes are conserved features across vertebrates. To test this hypothesis, laser capture microdissection was used to isolate IMM, MAT, and OST from homologous skeletal elements in mouse, chick, and gar embryos/ larvae, and then their transcriptomes were compiled (Fig. 36).

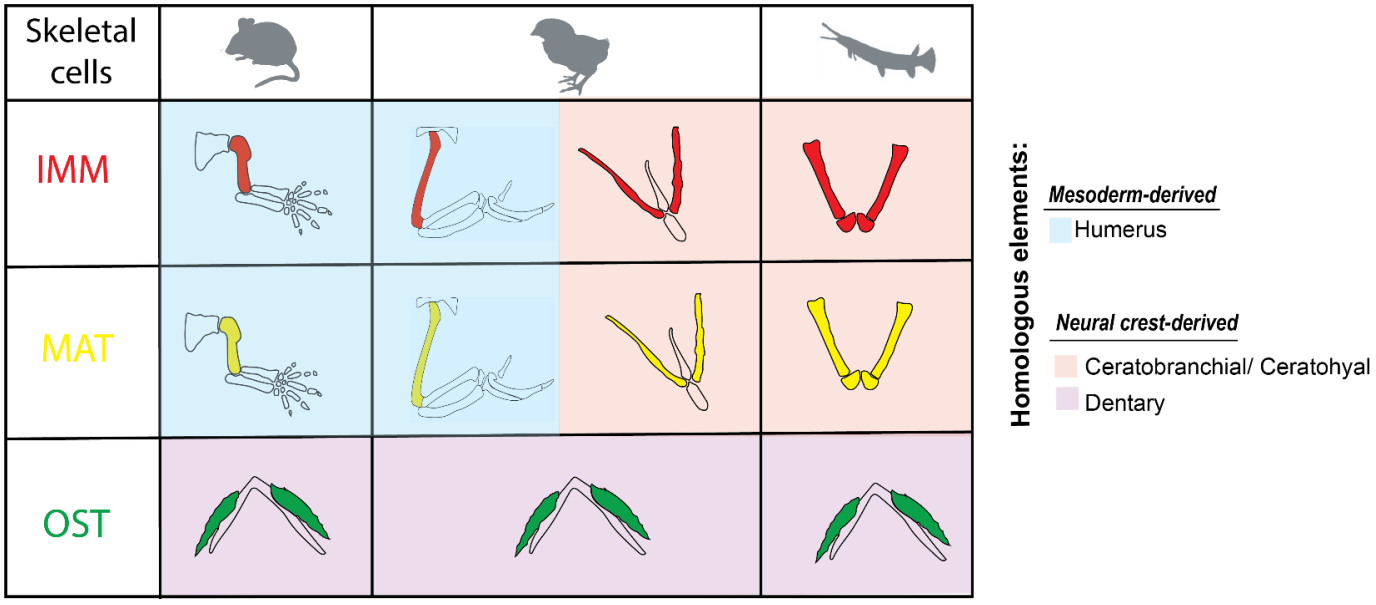



**Figure 36. Mouse, chick, and gar skeletal elements used for capturing IMM, MAT, and OST. In** mouse IMM and MAT were captured from the humerus. In chick IMM and MAT were captured from the humerus and the ceratobranchial. In gar, IMM and MAT were captured from the ceratohyal which is homologous to the chick ceratobranchial. In all three species, OST were isolated from the dentary bone, a homologous element between the three vertebrates.

Revealing the structure of skeletal cell GRNs *in vivo* in distinct vertebrates can shed light into regulatory control of chondrocytes and osteoblasts during evolution. Comparing GRN structure rather than individual gene interactions can be more useful to understand function and evolution of any cell type (Sun et al. 2012; Winterbach et al. 2013), and to identify evolutionary novelties, such as synapomorphies (i.e. shared derived traits that are common between an ancestor and its descendants; Choudhuri 2014). Previous work has analyzed synapomorphies comparing a few molecular markers (Sanderson and Doyle 1992; Venkatesh et al. 2001; Cameron et al. 2012; Shadwick and Ruiz-Trillo 2012), but comparing GRN structure globally between distinct vertebrate clades can reveal molecular synapomorphies at the transcriptome level (McCune and Schimenti 2012).

As discussed in chapter 3, gene co-expression network (GCN) analysis in mouse showed that MAT share the expression of many genes with both IMM and OST and the overlapping actions of portions of the GRN active in IMM and OST regulate gene expression in MAT (Fig. 15). On the other hand, portions of the GRN active in IMM and OST seem to act antagonistically, illustrated by the cross-inhibition between genes enriched in these two cell types (see chapter 3, Fig. 15, blue lines separating IMM and OST portions of the GRN). Two general mechanisms of gene regulation were identified in mouse MAT: independent and interactive (see chapter 3, Fig. 23). Independent regulation occurs when portions of the GRN active in IMM or OST are also active in MAT, and this type of regulation results in similar expression levels between MAT and one of the other two cell types (see chapter 3; Fig. 23A). Interactive regulation occurs when portions of the GRN active in IMM and OST interact in MAT. Together, this results in synergistic or average levels of expression in this cell type compared to IMM and OST (see Fig. 23B). In summary, this approach can reveal skeletal cell GRNs in *vivo* and accurate comparisons of skeletal cell GRN structure can be performed across clades to uncover how the molecular relationship between portions of the GRN has been modified over time. Furthermore, molecular mechanisms of gene regulation in skeletal cells can be compared across animal clades in order to have a more comprehensive understanding on the development and evolution of cartilage and bone.

## 6.3. RESULTS

### 6.3.1. Histological identification of chick HH36 skeletal tissues/cells and laser capture microdissection.

Chick IMM and MAT were obtained from two endochondral bones of distinct embryonic origin, the humerus and ceratobranchial, which are mesoderm- and neural crest-derived, respectively (see chapter 4 for more details, Figs. 27&28). These same chick chondrocytes were used in transcriptomic analyses presented in chapter 4. Osteoblasts were obtained from the HH36 dentary, an intramembranous bone of neural crest origin located in the mandible (see chapter 5 for more details, Fig. 31E-H). The same chick osteoblasts were used in analyses presented in chapter 5.

Despite data presented in chapter 4 suggesting that chick chondrocyte transcriptomes derived from the neural crest and mesoderm are highly similar (see chapter 4, Fig. 29A), after analyzing these datasets separately (and together) in the present chapter, some suggestive differences were found with these additional comparisons. More analysis comparing mesoderm vs NC within another species are needed to further confirm the level of similarity between mesoderm and NC skeletal datasets. Since chondrocytes of two different embryological origins (i.e. mesoderm and neural crest) were isolated from chick skeletal elements, analyses were performed in three different dataset combinations depending on the vertebrates analyzed, so analyses were mostly limited to homologous elements allowing for more direct transcriptomic comparisons (Fig. 36): (1) mouse vs chick: only humerus vs dentary datasets, (2) mouse vs chick: combined humerus and ceratobranchial vs dentary datasets, and (3) chick vs gar: only ceratobranchial/ceratohyal vs dentary datasets. The results of the different comparisons are presented in the following sections. Mouse data presented here was also used for analyses in chapter 3, chick data was also used for analyses in chapter 4 and 5, and gar OST data was also used for analyses in chapter 5. For most analyses presented here, IMM and MAT data were always compared to OST data in all three species.

### 6.3.2. Evolutionary comparisons of mouse and chick skeletal cells (including only humerus and dentary datasets)

#### 6.3.2.1. In mouse and chick, MAT share the expression of many genes with both IMM and OST, but IMM and OST showed the least similarity

Transcriptomic analyses of genes expressed above threshold in chick identified genes that were unique to each skeletal cell. The numbers of genes expressed in chick IMM, MAT, and OST were 7671, 9147, and 9054, respectively (Fig. 37B). Once genes expressed above threshold were identified in chick skeletal cells, these data were compared to mouse data obtained from homologous elements (i.e. humerus and dentary, Fig. 37). For these analyses, data compared within mouse IMM, MAT, and OST were then compared to data within chick IMM, MAT, and OST. Similarities between mouse and chick skeletal cell transcriptomes include the following: (1) IMM, MAT, and OST share over 70% of the genes (Fig. 37A,B); (2) MAT shared the expression of many genes with only IMM or OST (Fig. 37A,B); (3) IMM and OST share the least genes (228 genes in mouse and 70 genes in chick; Fig. 37A,B) whereas MAT and OST shared the most (857 in mouse and 1091 in chick; Fig. 37A,B). Due to these similarities in gene expression, these data suggest that skeletal cell GRN structure between mouse and chick might also be conserved.

Unique genes (i.e. genes expressed exclusively above threshold values in only one cell type) for each skeletal cell type were also identified in mouse and chick (Fig. 37A, B). We predict that MAT will express the fewest unique genes since its transcriptome would be mostly a mixture of gene expression between IMM and OST transcriptomes. As explained in chapter 3, mouse MAT had less unique genes compared to IMM and MAT (Fig. 11A, Fig. 37A), but these results were not consistent in chick (Fig. 37B). Our analyses show that in chick, IMM had the fewest uniquely expressed genes compared to the other two cell types (105 genes, Fig. 37B). Importantly, the presence of unique genes in mouse and chick mature chondrocytes support the idea that MAT is not merely a combination of genes expressed in IMM or OST.

Consistent to the unique gene analyses, we also predict that MAT would have fewer differentially expressed genes (DEGs) compared to IMM or OST. Similar to mouse (Fig. 37C), of those DEGs specific to one cell type in chick, MAT expressed far fewer (124) than IMM (513) and

OST (482; Fig. 37D). Among these DEGs specific to one cell type, mouse and chick MAT showed up-regulated expression of the hallmark mature chondrocyte gene *Col10a1*.

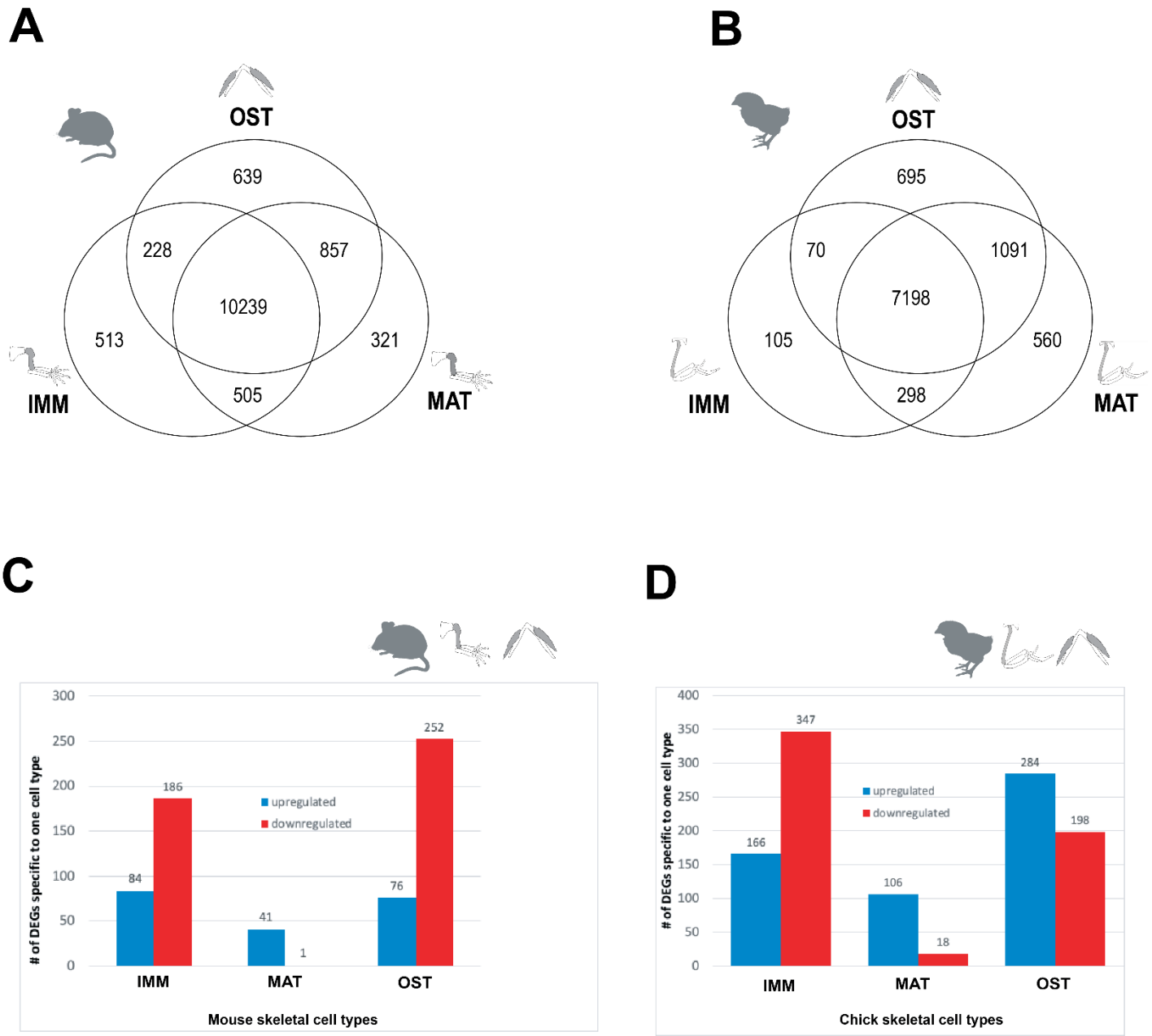



**Figure 37. Gene expression distribution in mouse and chick using humerus and dentary datasets.** (A,B) In tetrapods, IMM and OST share the least genes, whereas MAT and OST share the most. (B) In both mouse and chick, MAT had the least differentially expressed genes compared to IMM and OST. The chick skeletal element from which IMM and MAT were isolated is indicated in the figure (i.e. humerus). Gene lists used for Venn diagram, analyses and DEGs graphs were generated by Katie Ovens.

Furthermore, pairwise differential gene expression comparisons between two of the cell types in both mouse and chick showed that many genes are differentially expressed in IMM compared to OST (Figs. 13&38), supporting the idea that independent portions of the GRN are active in IMM and OST (see Fig. 1). Many DEGs in mouse and chick skeletal cells show the same trend of expression. For instance, MAT and OST markers are upregulated in either of these cell types compared to IMM including genes, such as *SPP1, IBSP, SATB2, MMP9, and MMP13* (Minina et al. 2001; Zhang et al. 2003; Inada et al. 2004; Dobreva et al. 2006; Zhang et al. 2014; Liang et al. 2016).

Other genes involved in cartilage differentiation are upregulated in IMM and MAT compared to OST, such as *SOX5, SOX9, COL2A*1, and *ACAN* (Zhao et al. 1997; Watanabe et al. 1998; Bi et al. 1999; Smits et al. 2001).

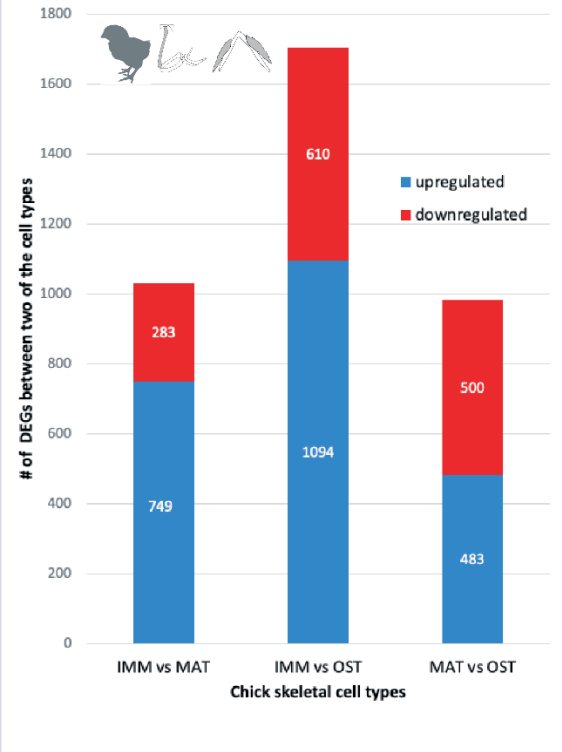



**Figure 38. Pairwise comparisons in chick skeletal cells isolated from the humerus and dentary showed that many genes are differentially expressed in IMM compared to OST.** A gene was considered differentially expressed if it had an absolute log2 fold change greater than 2. A total of 1032 genes are differentially expressed between IMM and MAT, 1704 genes are differentially expressed between IMM and OST, and 983 genes are differentially expressed between MAT and OST. Gene lists used for DEGs graphs were generated by Katie Ovens.

#### 6.3.2.2.Skeletal cell GRN organization is conserved in mouse and chick

To test the hypothesis that skeletal cell GRN structure is conserved in distinct vertebrates, gene co-expression network analyses was also performed on chick skeletal datasets. For these first comparisons, we focused on the chick humerus and dentary datasets, the same skeletal elements analyzed for mouse (Fig. 15). A total of 2693 DEGs were included in this GCN analysis. Overall, GCN analyses showed that GRN structure is conserved between chick and mouse (Fig. 39), but genes included in each portion of the GRN vary between species showing evidence of adaptive evolution. In general, two independent portions of the GRN which are active in IMM and OST, show strong cross-inhibition with each other, and they positively interact in MAT (Fig. 39A,B). In contrast to mouse GRN, positive correlations are also present to some degree between the IMM- and OST- portions of the chick GRN. Positive correlations between IMM and OST were never seen in the mouse skeletal cell GRN (Fig. 39A). A total of 338,805 positive correlations (edges) were included in the Cytoscape estimated GRN. 395 positive correlations between IMM and OST out of 338,805 total positive correlations in the network were present in chick (0.11% of total positive correlations between genes enriched in different cell types, Fig. 39B), but positive correlations between MAT and OST were higher (47,620/338,805= 14% of total positive correlations between genes enriched in different cell types, Fig. 39B). Most of the positive correlations included in this estimated GRN were between genes enriched within the same cell type, predominantly OST-OST and IMM-IMM. Despite minor structural differences between mammalian and avian skeletal GRNs, in general these findings suggest that GRN structure is conserved within tetrapods, at least in amniotes, suggesting that regulatory control of skeletal cell types is also conserved during evolution.

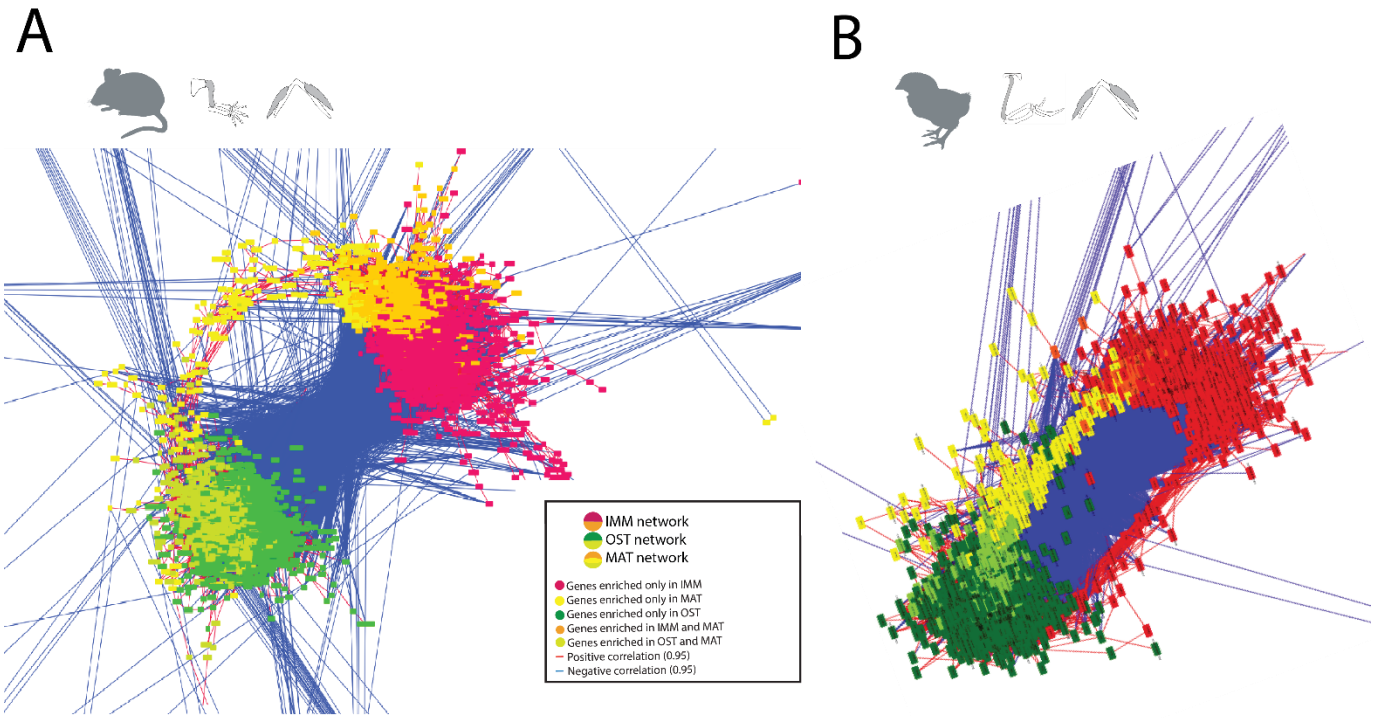



**Figure 39. Skeletal cell GRN structure is conserved in mouse and chick when only humerus and dentary datasets are analyzed.** (A,B) GCN analysis in both tetrapods showed that MAT contributed several genes to both IMM- and OST-enriched portions of the GRN, whereas genes enriched in IMM and OST were located in opposite portions of the GRN. (B) In chick, however, there is also some positive correlation between IMM- and OST-enriched portions of the GRN. Chick GCN was constructed using only humerus cartilage datasets and comparing them to dentary bone datasets (same datasets as in mouse). For a more detailed view of the mouse GCN (A) see Fig. 15. The chick skeletal element from which IMM and MAT were isolated is indicated in the figure (i.e. humerus). In these analyses, IMM and MAT datasets were always compared to OST datasets (Credit: Katie Ovens).

### 6.3.2.3.Molecular mechanisms of gene regulation in MAT are conserved in mouse and chick

As discussed in chapter 3, mouse skeletal cells showed specific categories of gene expression (Fig. 19), providing insight into specific mechanisms of gene regulation in MAT. To test whether regulatory control in mature chondrocytes is conserved in other vertebrates, homologous skeletal cell types were analyzed in chick. Interestingly, when the same skeletal datasets (i.e. humerus and dentary) were analyzed in chick, mechanisms of gene regulation were conserved (Fig. 40). Independent regulation was observed when portions of the GRN active in IMM or OST do not interact in MAT, resulting in similar expression levels in MAT compared to IMM or OST. For

example, clusters 3 and 9 showed similar expression levels between IMM and MAT (Fig. 40D,J). In cluster 3, gene expression between IMM and MAT was similarly low and significantly different compared to OST, which showed higher levels (Fig. 40D). Cluster 3 includes genes involved in ossification such as, *BMP2, BMP3, PTH1R,* and *DCN* (Minina et al. 2001; Lee et al. 2003; Nie et al. 2006; Datta et al. 2010). Other clusters showed similar expression levels between MAT and OST. For example, cluster 5 showed statistically significant lower expression levels in MAT and OST compared to IMM (Fig. 40F,J). On the other hand, interactive regulation in MAT was observed when portions of the GRN active in IMM and OST showed averaging (AVE) or synergistic activity in MAT. For example, the chondrogenic transcription factor *SOX9* was included in cluster 2 were MAT had expression statistically similar to AVE between IMM and OST (Fig. 40C). Cluster 2 also included genes , such as *SOX5, COL2A1, COL9A1, and ACAN,* which are all regulated and/or work together with *SOX9* (Fig. 40C; Smits et al. 2001; Oh et al. 2014; Liu and Lefebvre 2015; Ohba et al. 2015). All these 'cartilage' genes are expected to have intermediate levels in MAT compared to IMM and OST. Cluster 10 showed enhanced expression in MAT including genes, such as *COL10A, MEF2C, WNT5B,* and *VEGFA,* which have been shown to have important roles during chondrocyte maturation (Fig. 40K; Yang et al. 2003; Zelzer et al. 2004; Arnold et al. 2007; Lu et al. 2014; Nakatani and Partridge 2017). Importantly, all the genes discussed above were also included in clusters that showed the same categories of gene expression in mouse (Fig. 19). In general, these results suggest that mechanisms of GRN regulation in mature chondrocytes are a conserved feature within tetrapods (i.e. mouse and chick).

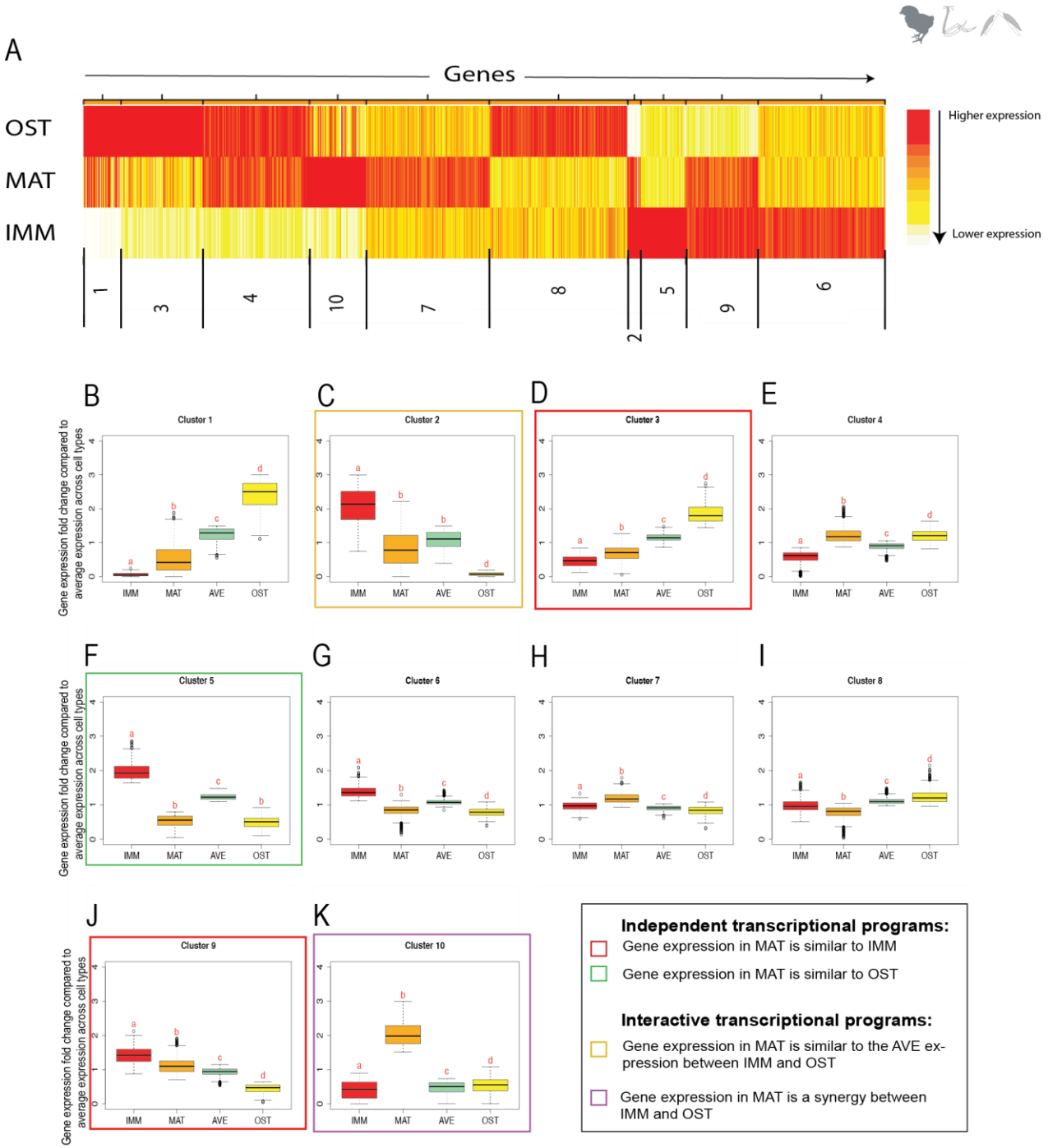



**Figure 40. Model based clustering using only humerus and dentary datasets in the chick.** (A) Model-based clustering demonstrated that groups of IMM, MAT, and OST genes had discrete categories of expression. (B-K) Box plots showed the distribution of gene expression change among cell types compared to the average expression across all three cell types. AVE represented average expression of IMM and OST. MAT expression in clusters 3 and 9 (D,J) was similar between IMM and MAT, and statistically different from OST. MAT expression in cluster 5 (F) was significantly similar to OST. (C) In cluster 2, gene expression in MAT was statistically indistinguishable from AVE. (K) Cluster 10 had statistically higher gene expression in MAT than in either IMM or OST. The chick cartilage datasets used for these comparisons are indicated in the figure. The chick skeletal element from which IMM and MAT were isolated is indicated in the figure (i.e. humerus). In these analyses, IMM and MAT datasets were always compared to OST datasets (Credit: Katie Ovens).

### 6.3.3. Evolutionary comparisons of mouse and chick skeletal cells (combined humerus and ceratobranchial cartilage datasets vs dentary datasets)

#### 6.3.3.1. MAT expresses less differentially expressed genes than IMM and OST

Previous work in chick using candidate genes (Eames and Helms 2004) and our data (see chapter 4, Fig. 29A) have revealed a tremendous overlap in gene expression between chondrocytes isolated from distinct locations (and embryological origin) in the body. Hence for these analyses transcriptomic data obtained from chick chondrocytes isolated from the humerus (mesoderm, appendicular skeleton) and ceratobranchial (neural crest, cranial skeleton) were combined and compared. Again, for these analyses, data compared within mouse IMM, MAT, and OST were then compared to data within chick IMM, MAT, and OST. When data obtained from chondrocytes isolated from the ceratobranchial were also included into these comparisons, many similarities were observed with the analyses discussed above (i.e. using only humerus and dentary datasets, Fig. 37). These results were also generally conserved with mouse (Fig. 37A,C). Again, MAT shared expression of several genes with only one of the other cell types (866 with OST and 268 with IMM, Fig. 41A) suggesting that portions of the GRN active in IMM and OST are also active in MAT. Importantly, IMM and OST still share the least genes (64 genes, Fig. 41A), supporting the idea that independent portions of the GRN regulate the differentiation of IMM and OST. Uniquely expressed genes were also analyzed in these combined datasets. Similar to the results discussed above for only chick humerus and dentary datasets (Fig. 37B), IMM again had fewer uniquely expressed genes (Fig. 41A), and these results also differ from mouse skeletal cells (i.e. MAT had fewer uniquely expressed genes, Fig. 37A).

When differential gene expression analyses were performed using these combined hum+cb vs dentary datasets, the same patterns of gene expression were identified as when only humerus vs dentary datasets were analyzed in chick (Fig. 37D), and these patterns of gene expression were also conserved compared to mouse (Fig. 37C). Again, MAT had fewer differentially expressed genes (61) compared to IMM (122) and OST (423; Fig 41B). Of the 61 DEGs in MAT, only 29% were downregulated, compared to 43% in IMM and 46% in OST. Importantly, many of the genes known to be important during the differentiation of IMM, MAT, and OST were conserved between mouse and chick. For example, important genes during cartilage differentiation such as *SOX9*, *COL2A*1,

*COL9A1*, *ACAN*, *SOX5* (Lefebvre and de Crombrugghe 1998; Bi et al. 1999; Lefebvre et al. 2001; Akiyama et al. 2002; Mead et al. 2013; Oh et al. 2014; Liu and Lefebvre 2015; Ohba et al. 2015) were upregulated in IMM and MAT and downregulated in OST. Moreover, many important mineralization and osteoblast markers such as *MMP9*, *MMP13*, *IBSP*, *SPP1*, *BAMBI*, *MEF2C*, *SATB1*, *BMP4*, *DLX6*, and *VDR* (Drissi et al. 2000; Wang et al. 2004; Li et al. 2008; Chang et al. 2009; Komori 2010; Nyman et al. 2011; Yamamoto et al. 2013; Wu et al. 2014; Nakatani and Partridge 2017) were up-regulated in MAT and OST and downregulated in IMM.

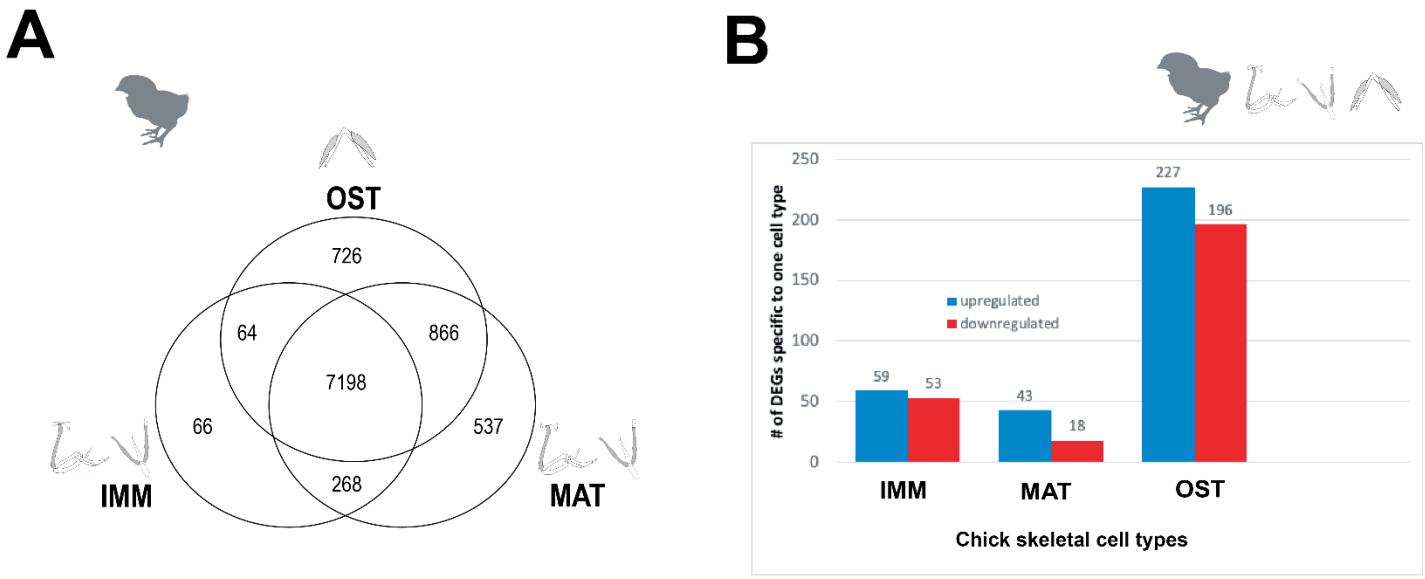



**Figure 41. Gene expression distribution in chick using combined humerus and ceratobranchial datasets compared to dentary datasets.** (A) MAT and OST share the most genes, whereas IMM and OST share the least. IMM has the fewest uniquely expressed genes (B) MAT had less differentially expressed genes (DEGs) than the other two cell types. Genes included in the graphs were differentially expressed in one cell type compared to the other two cell types. Of the 61 DEGs in MAT in chick, only 18 were downregulated, compared to IMM (53) and OST (196). The chick skeletal element(s) from which IMM and MAT were isolated is indicated in the figure (i.e. humerus + cb). Gene lists used for Venn diagram, analyses and DEGs graphs were generated by Katie Ovens.

Some key differences between classic OST markers between mouse and chick were the following. In chick, *RUNX2* was expressed highly in both MAT and OST, but it was not differentially expressed when compared to IMM. Consistent with previous work (Kague et al. 2016), another important difference between the present mouse and chick skeletal datasets was the absence of a clear *SP7* orthologue in chick. These results suggest that *SP7* might not be required for

avian bone formation, as it is during mammalian and fish osteogenesis (Niu et al. 2017; Yu et al. 2017).

#### 6.3.3.2. GRN structure is conserved between mouse and chick (combined hum+cb chick cartilage datasets)

Gene co-expression network analyses (GCN) were also performed in chick combined datasets (hum+cb vs dentary). A total of 1074 DEGs were included in this analysis. GRN structure was highly conserved in chick regardless of the dataset combinations used (Figs. 39&42). Moreover, GRN structure obtained from combined chick datasets was very similar to mouse (Fig. 39A), yet again some genes included in each portion of the network were different between mouse and chick providing evidence that each species GRN has undergone evolutionary changes. Again, two distinct portions of the GRN active in IMM and OST, showed strong cross-inhibition with each other, and they interacted positively in MAT (Figs. 39&42). Consistent with the chick analyses presented above (i.e. only humerus vs dentary datasets, Fig. 39) positive correlation between IMM- and OST-enriched portions of the GRN is still observed when chick chondrocyte hum+cb datasets are combined. A total of 73 positive correlations between IMM and OST out of 16,269 total positive correlations in the network were observed in chick (0.45% of total positive correlations between genes enriched in different cell types, Fig 42), but as in the analyses in chick discussed above (i.e. only humerus and dentary datasets, Fig. 39B) the positive correlation between MAT and OST is still stronger (557/16,269=3.4% of total positive correlations between genes enriched in different cell types; Fig. 42). Most of the positive correlations included in this estimated GRN were between genes enriched within the same cell type, predominantly OST-OST and IMM-IMM. Negative correlations between IMM and OST were dominant in both mouse and chick skeletal cell GRNs (Figs. 39&42).

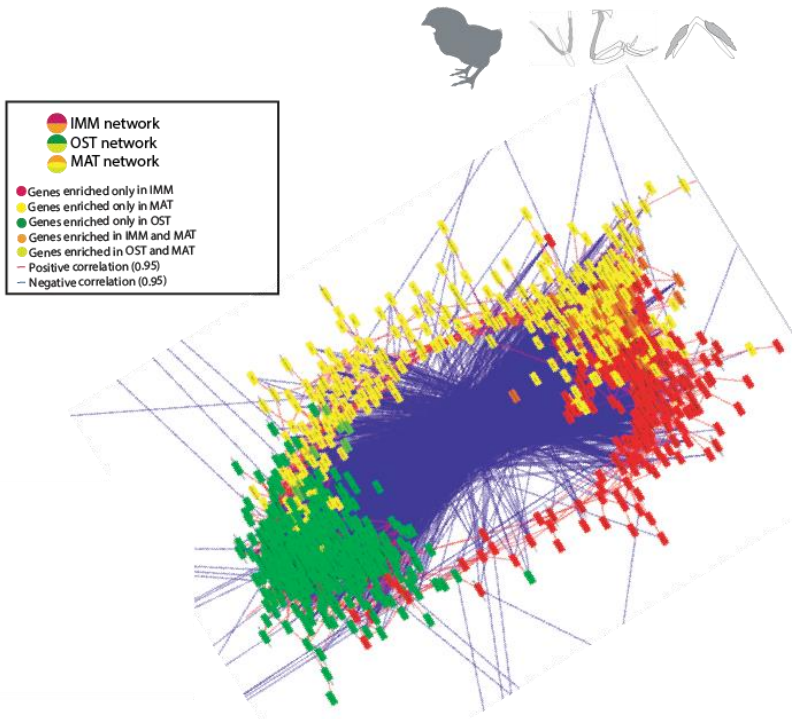



**Figure 42. Skeletal cell GRN structure is still conserved when chick humerus and ceratobranchial datasets are combined and compared to dentary datasets.** GCN analysis revealed two independent portions of the GRN active in IMM and OST, which are negatively correlated with each other. Both IMM- and OST-enriched portions of the GRN are positively correlated when they overlap in MAT. Again, in this chick skeletal cell GRN there is also some positive correlation between IMM- and OST-enriched portions of the GRN. Chick GCN was constructed combining humerus and ceratobranchial cartilage datasets and comparing them to dentary bone datasets. The chick skeletal element(s) from which IMM and MAT were isolated is indicated in the figure (i.e. humerus + cb). In these analyses, IMM and MAT datasets were always compared to OST datasets (Credit: Katie Ovens).

### 6.3.3.3.Molecular mechanisms of gene regulation in MAT are conserved in mouse and chick (combined hum+cb chick cartilage datasets)

Cluster analysis was also performed using combined hum+cb datasets vs dentary datasets (Fig. 43), and categories of gene expression were also conserved compared to results discussed above in chick (i.e. only humerus and dentary; Fig. 40) and mouse (Fig. 19). Gene expression patterns in some clusters suggest that portions of the GRN operating in IMM and OST operate independently in MAT. For instance, cluster 7 includes genes with high expression levels in MAT and OST compared to IMM, such as *RUNX2*, *FOXO1*, *POSTN*, and *MMP13*, *DLX5*, *DLX6* and *WNT5A*, which are crucial markers for osteoblast differentiation and mineralization (Nakashima and de Crombrugghe 2003; Zaragoza et al. 2006; Holleville et al. 2007; Komori 2010; Siqueira et al.

2011; Zhang et al. 2017). In clusters 1, 2, and 6, gene expression was similar between IMM and MAT, and significantly different from OST (Fig. 43B,C,G). While cluster 6 showed higher expression levels in both IMM and MAT compared to OST (Fig. 43G), clusters 1 and 2 showed lower expression levels in these cell types compared to OST (Fig. 43B,C). Indeed, cluster 1 includes classic bone markers, such as *BAMBI*, *COL1A2*, *IBSP*, *ALPL*, and *SPARC* (Golub and Boesze-Battaglia 2007; Yano et al. 2014; Rosset and Bradshaw 2016; Komori 2019), which are all expected to have higher expression levels in OST, intermediate to high expression levels in MAT, and lower levels or no expression in IMM. Cluster 2 also includes genes involved in ossification such as, *BMP3, BMP4, BMP6, BMP7, MMP9, MSX2*, and *PTH1R* (Satokata et al. 2000; Minina et al. 2001; Lee et al. 2003; Nie et al. 2006; Datta et al. 2010; Liang et al. 2016) which are all expected to have higher levels in OST (Fig. 43C).

Again, other clusters provided evidence that portions of the GRN operating in IMM and OST interacted in MAT. For instance, cluster 5 showed enhanced expression levels in MAT compared to IMM and OST, suggesting that portions of the GRN active in IMM and OST act in a synergistic manner in MAT (Fig. 43F). As expected, this cluster includes genes known to be important during cartilage maturation and osteoblast differentiation such as *COL10A1, MEF2C, WNT5B, WNT11,* and *VEGFA* (Yang et al. 2003; Zelzer et al. 2004; Arnold et al. 2007; Lu et al. 2014; Nakatani and Partridge 2017; Boyan et al. 2018). In contrast, cluster 8 showed average expression levels in MAT compared to IMM and OST (Fig. 43I) including genes, such as *SOX9, SOX5, SOX6, COL2A1, COL9A1*, and *ACAN*, which are crucial cartilage differentiation markers (Bell et al. 1997; Bi et al. 1999; Lefebvre et al. 2001; Kiani et al. 2002; Carlsen et al. 2006; Liu and Lefebvre 2015b), and are expected to have slightly lower levels in MAT compared to IMM, but higher compared to OST. Together these results in mouse and chick suggest that GRN structure and mechanisms of GRN regulation in MAT are conserved features within tetrapods, at least in amniotes.

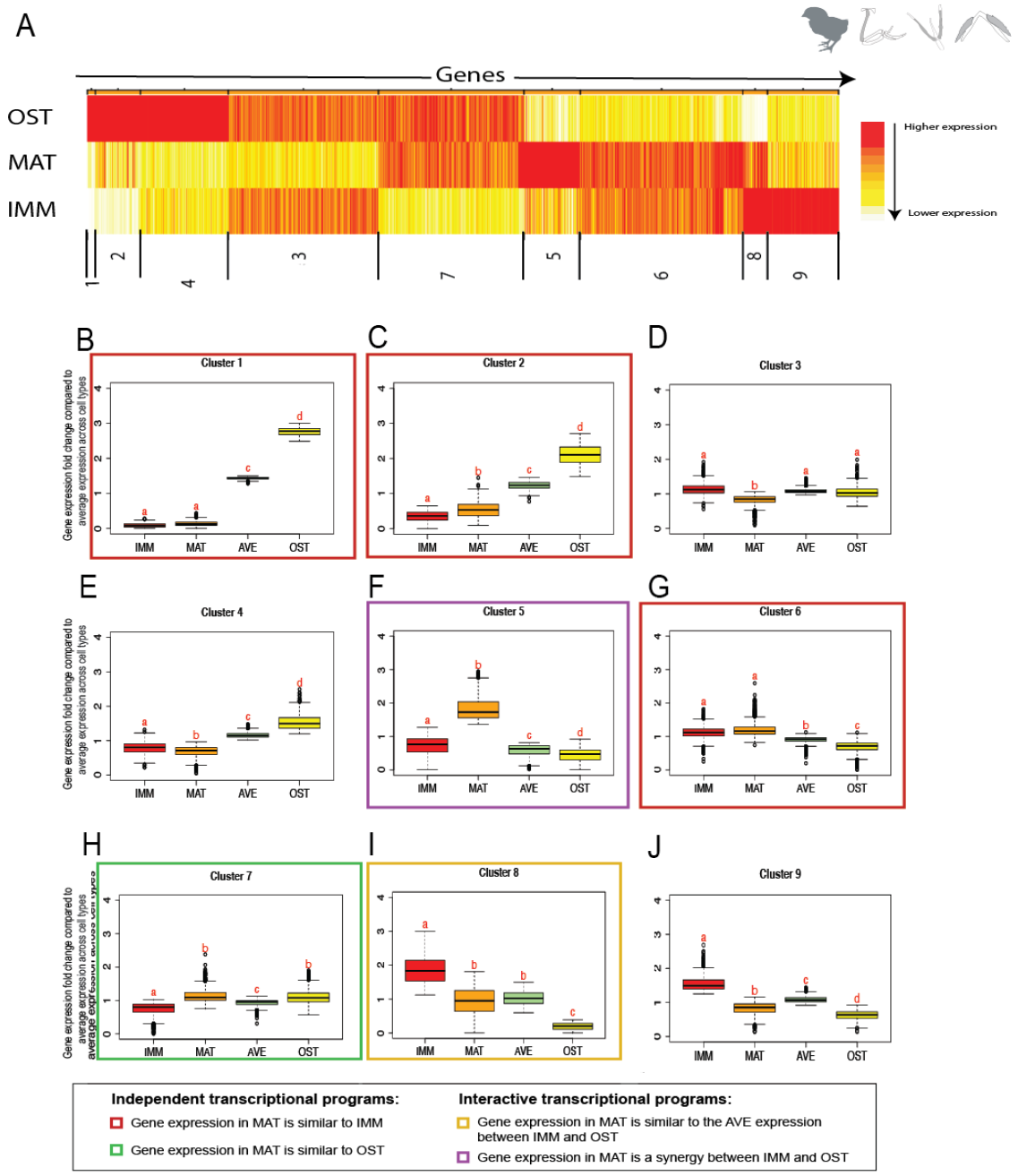



**Figure 43. Model based clustering using combined humerus and ceratobrancial datasets compared to the dentary datasets in the chick.** (A) Model-based clustering demonstrated that groups of IMM, MAT, and OST genes had discrete categories of expression. (B-J) Box plots showed the distribution of gene expression change among cell types compared to the average expression across all three cell types. AVE represented average expression of IMM and OST. MAT expression in clusters 1, 2, and 6 (B,C,G) was similar to IMM, and significantly different from OST. MAT expression in cluster 7 (H) was significantly similar to OST, suggesting that the IMM- and OST-enriched portions of the GRN operate independently in MAT. (F) Cluster 5 had statistically higher gene expression in MAT than in either IMM or OST, suggesting that IMM- and OST-enriched portions of the GRN interact synergistically in MAT. (I) In cluster 8, gene expression in MAT was statistically indistinguishable from AVE, suggesting that portions of the GRN in IMM and OST interact via averaging in MAT. The chick skeletal element(s) from which IMM and MAT were isolated is indicated in the figure (i.e. humerus + cb). In these analyses, IMM and MAT datasets were always compared to OST datasets (Credit: Katie Ovens).

#### 6.3.3.4.Gene expression of candidate MAT genes is conserved between mouse and chick

RNA in situ hybridization was used to confirm and compare the expression of the hallmark mature chondrocyte genes *COL10A1* and *SPP1* (Pullig et al. 2000; Lu et al. 2014), between mouse and chick. Expression domains of these genes were highly conserved in mouse and chick MAT (Fig. 44A-C, E-G). However in chick HH36 expression of *COL10A1* was more restricted to the pre-hypertrophic zone as previously reported by others (Provot et al. 2006). RNA *in situ* hybridization was also used to validate the expression of Non-Homologous End Joining Factor 1 (*NHEJ1*), a novel evolutionarily conserved gene not previously associated with mature chondrocytes. Importantly, this gene was differentially expressed (upregulated) in MAT from both mouse and chick, suggesting that it might have a conserved role during cartilage maturation. Expression of *NHEJ1*, was also confirmed to be expressed in mature chondrocytes, predominantly in the pre-hypertrophic region (Fig. 44D,H). Interestingly, *NHEJ1* was expressed in a similar domain as *COL10A1* in chick, suggesting that one gene might affect the expression of the other during cartilage maturation.

*In situ* hybridization in mouse MAT shows that *Nhej1* is expressed in the same domain as *Ihh*, although the former is more restricted to pre-hypertrohic chondrocytes (see Fig. 11D, Fig. 44D). In chick, however, only a few transcripts of *IHH* in some MAT samples were found to be expressed in the data presented here. Since there is a spatial relationship between mature cartilage and perichondral bone, the captured area for MAT in both mouse and chick was restricted to chondrocytes located in the mid-diaphyseal region (i.e. shaft of the long bone) surrounded by this collar (see chapter 3, Fig. 9I,J; see chapter 4, Fig. 27H,I and Fig. 28G,H). *In situ* hybridization in chick HH36 cartilage confirmed the expression of *IHH* closer to the epiphyses (i.e. heads of long bones) of the bone and outside the perichondral bone area (data not shown; Provot et al. 2006), but this area was not captured in the chick samples analyzed here (see chapter 4, Fig. 27H,I and Fig. 28G,H).

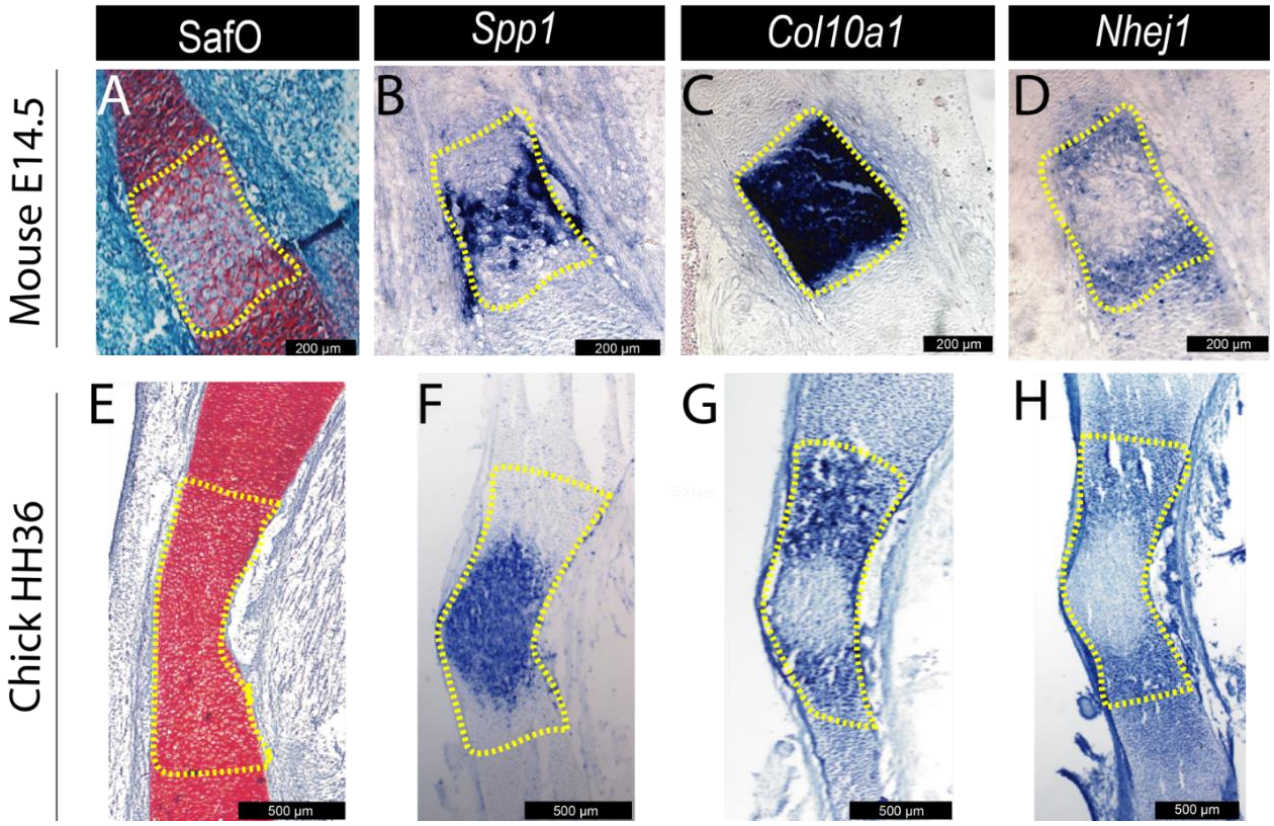



**Figure 44. Gene expression of MAT genes in mouse and chick is conserved.** Differentially expressed genes conserved in mouse and chick MAT. (A,E) Safranin O-stained section of E14.5 humerus highlighted the mature cartilage region (yellow dotted outline). RNA in situ hybridization validated (B,F) *Spp1* and (C,G) *Col10a1* as a positive controls for MAT, and also (D,H) *Nhej1* as novel evolutionarily conserved gene enriched in MAT. Holes are artifacts and do not represent vascular invasion. Abbreviation: SafO=Safranin O.

### 6.3.4. Evolutionary comparisons between chick and gar neural crest-derived skeletal elements

#### 6.3.4.1. Histological identification of gar 13dpf skeletal cells

As stated at the beginning of this chapter, evolutionary comparisons presented here were mostly limited to homologous elements (Fig. 36). Hence, in this section homologous neural crest-derived skeletal elements were compared between chick and gar: the ceratobranchial/ ceratohyal and the dentary. The dentary is an intramembranous bone of neural crest-origin that is homologous between mouse, chick, and gar. Despite mouse also having a homologous element to the chick

ceratobranchial/ gar ceratohyal, this element is very small, making it difficult to capture enough cells of interest to obtain appropriate amounts of RNA for sequencing. For some comparisons between chick and gar, however, mouse is occasionally brought into the picture since bioinformatic analysis of mouse skeletal cells showed high conservation with chick skeletal cells, regardless of the chick dataset combinations used (humerus vs dentary, section 6.3.2 and humerus+cb vs dentary, section 6.3.3).

To make an evolutionary comparison between gar and tetrapods, a histological analysis to identify skeletal elements and cell types of interest in the gar larvae was performed (Fig. 45). Alcian blue/ Alizarin red staining at 13dpf demonstrated well defined cartilage in the ceratohyal, but no presence of mineralized perichondral bone characterized by the lack of Alizarin red staining in this region (Fig. 45A). On the other hand, the presence of the dentary bone was confirmed by intense Alizarin red staining, adjacent to Meckel's cartilage (see chapter 5, Fig. 31I). Safranin O staining identified immature chondrocytes in the epiphyses, and mature chondrocytes in the diaphysis region of the 13dpf ceratohyal (Fig. 45C-E, Fig. 46A), but this was not so conclusive since at the stage analyzed hypertrophy was minimal (Fig. 45D). On the other hand, aniline blue staining was evident in the bone matrix of the dentary bone and perichondrium (Fig. 45E; see chapter 5, Fig. 31J; Fig. 46D). *In situ* hybridization of hallmark chondrocyte genes, *col2a1* and *col10a1*, in the gar ceratohyal revealed immature and mature cartilage regions, respectively (Fig. 46B,C). Although chondrocyte hypertrophy was not so evident in the 13dpf ceratohyal (Fig. 45D), expression of the hallmark mature chondrocyte gene *col10a1* was already high and the expression domain expanded towards the epiphyses (Fig. 46C). One key difference between gar and tetrapods is that expression of *col10a1* in gar is not limited to mature chondrocytes, but it is also highly expressed in osteoblasts of the dentary bone (Fig.46E).

Once skeletal cell types of interest were identified, laser capture microdissection (LCM) was performed to isolate skeletal cells from gar 13dpf cranial skeletal elements (Fig. 45F-I). Gar immature chondrocytes and mature chondrocytes were isolated from the 13dpf ceratohyal (Fig. 45F-I), whereas osteoblasts were isolated from the dentary bone (see chapter 5, Fig. 31K,I). Gar larvae were analyzed at both 13 and 14 dpf (Fig. 47), but analyses were focused on 13dpf since little or no teeth were observed adjacent to the dentary (Fig. 47A-C). The presence of teeth nearby the dentary bone at 14dpf could lead to a higher risk of cross-contamination when capturing osteoblasts (Fig. 47D-F). Although some genes expressed in the gar OST have been linked to both bone and tooth

formation, no dentin or enamel specific genes were found to be expressed in the gar OST data (data not shown) which validates the purity of the osteoblasts captured.

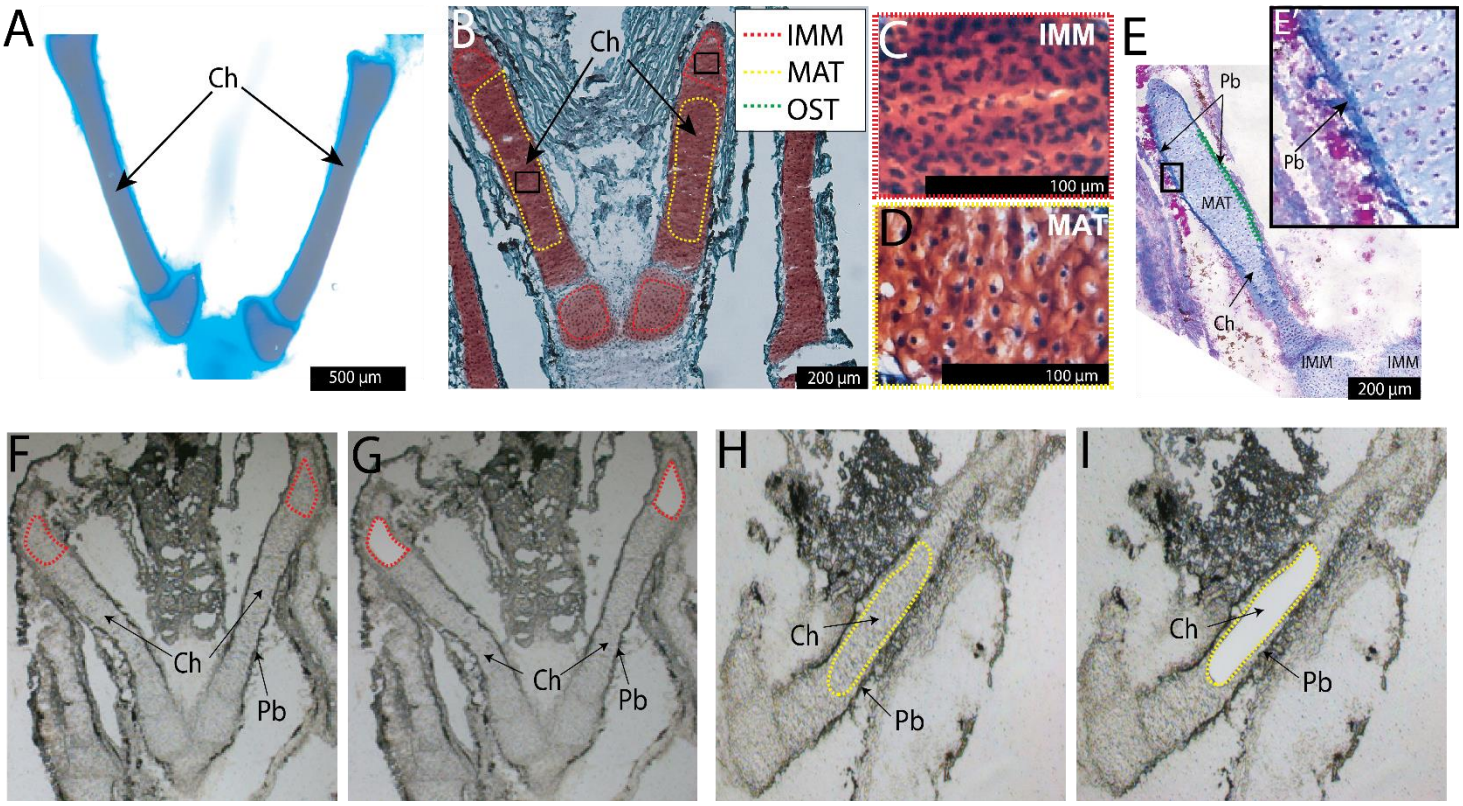



**Figure 45. Laser capture microdissection was used to isolate specific skeletal cell types from the gar 13 dpf ceratohyal.** (A) Whole-mount Alcian blue and Alizarin red staining identified cartilage and perichondral bone in endochondral bones of the gar ceratohyal, and intramembranous bone in the mandible. (B) Safranin O-stained coronal section of 13 dpf ceratohyal highlighted the mature cartilage region (yellow dotted outline). High-magnification images of immature (C) and mature chondrocytes (D) from black boxes in (B). (E) Trichrome staining identified perichondral bone in the 13dpf ceratohyal.Unstained sections of HH36 gar ceratohyal before (F) and after (G) laser capture of immature chondrocytes, and before (G) and after (H) laser capture of mature chondrocytes. Abbreviations: IMM, immature chondrocytes; MAT, mature chondrocytes; OST, osteoblasts; ch, ceratohyal; Pb, perichondral bone.

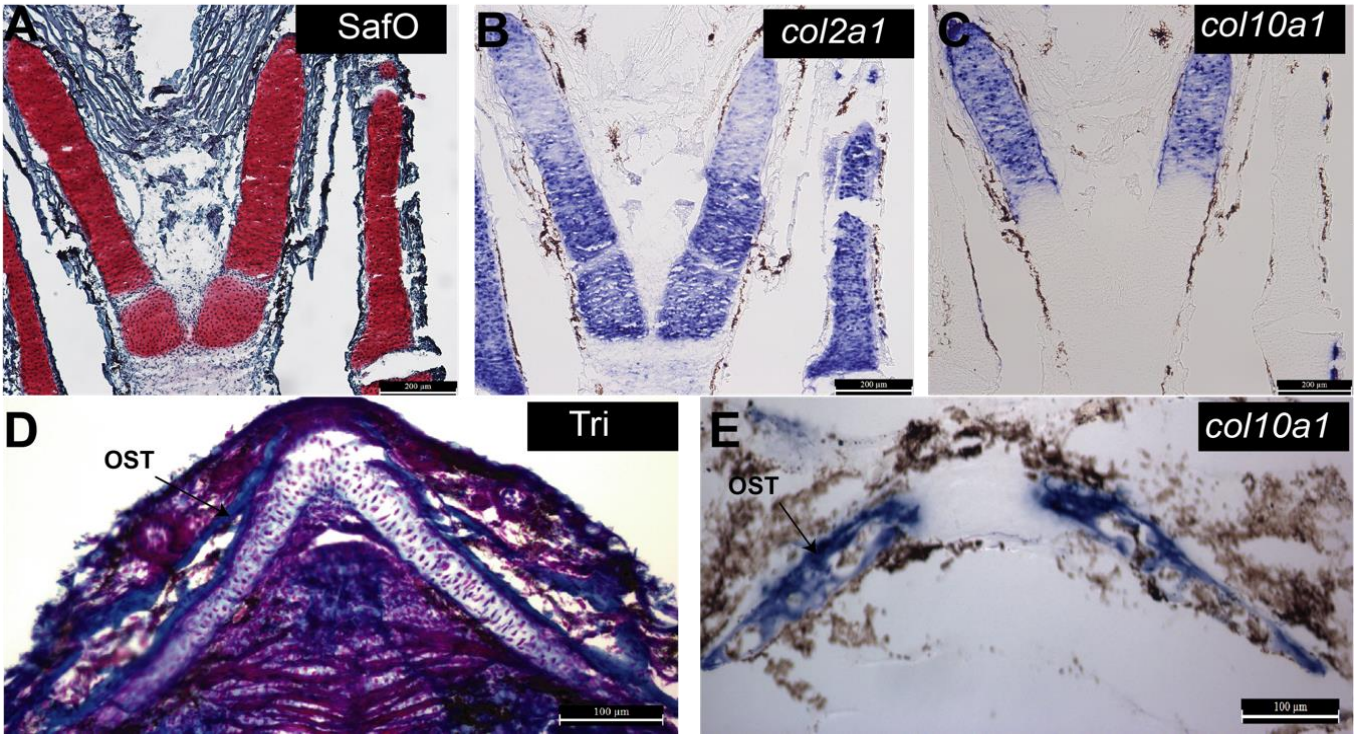



**Figure 46. Osteoblasts in the gar express chondrocyte genes, such as *col10a1*.** Safranin O/Fast green staining and in situ hybridizations on gar ceratohyal. A) Safranin O binds to cartilage glycosaminoglycans, staining the extracellular matrix red. B) *In situ* hybridization identified the domains of expression of (B) *col2a1*, a characteristic marker of immature cartilage, and (C) *col10a1*, a characteristic marker of mature cartilage. (D) Trichrome staining stains tightly wound collagen fibers, such as *col10a1* dark blue. (E) *col10a1* is not restricted the mature chondrocytes in the gar, since it is also expressed in osteoblasts.

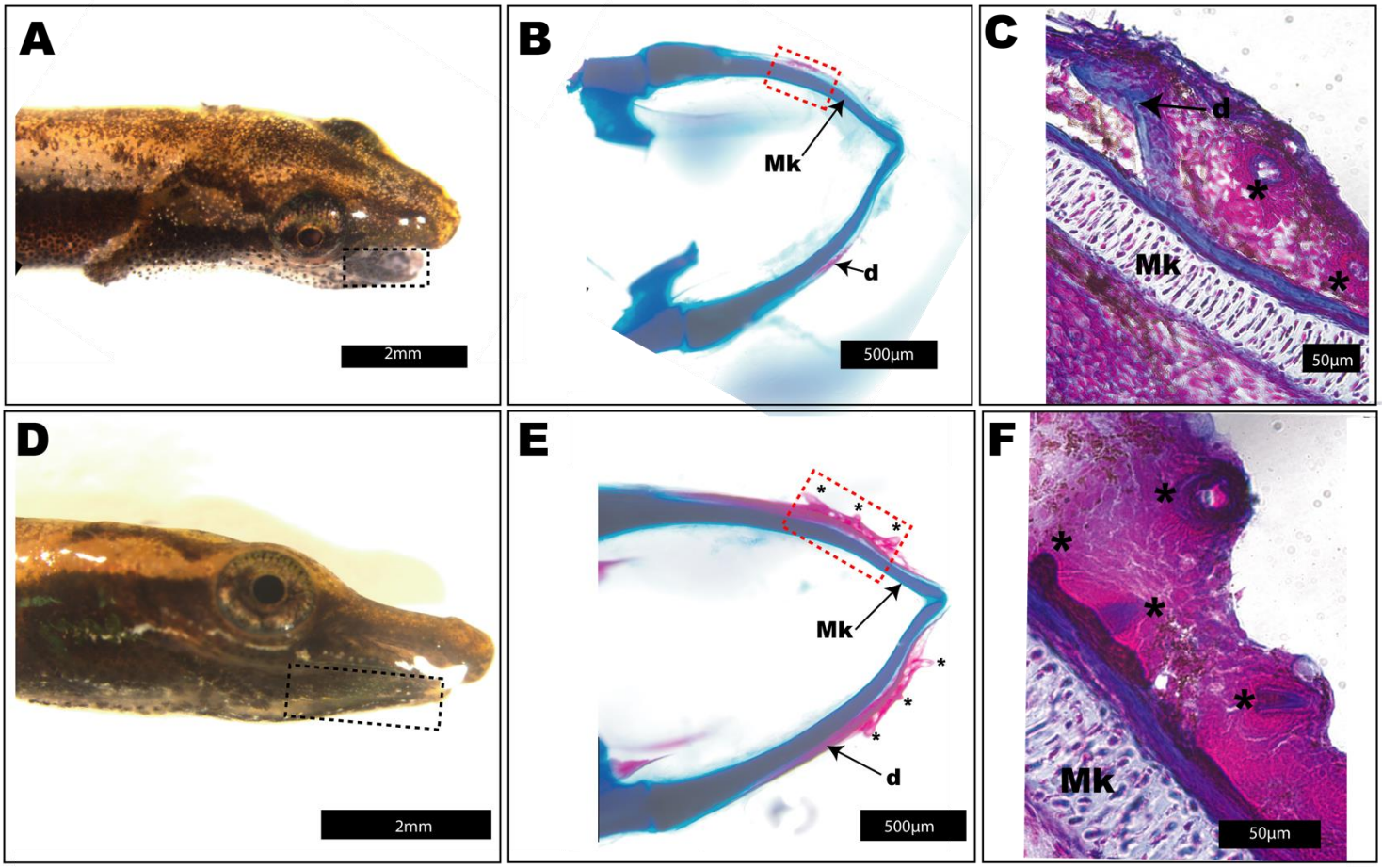



**Figure 47. Histological characterization of gar skeletal tissues at 13 and 14dpf.** (A) In 13dpf gar larvae, dentary bone is observed adjacent to Meckel's cartilage (B). A few teeth are observed close to the dentary bone at 13dpf (C). (D) In 14 dpf gar larvae, extensive teeth adjacent to the dentary were identified (E,F). Asterisks indicate the location of teeth.

### 6.3.4.2.When only NC-derived skeletal elements are analyzed, gar IMM and OST are the most similar cell types, but chick IMM and OST are still the least similar cell types

Chick datasets obtained from neural crest-derived skeletal tissues (i.e. ceratobranchial and dentary) were compared to datasets obtained from homologous NC-derived skeletal elements in gar (i.e. ceratohyal and dentary). In chapter 4, analyses were only limited to chick IMM and MAT, but the present analyses in chick compared IMM and MAT in context of OST data. Data compared within chick IMM, MAT, and OST were then compared to data within gar IMM, MAT, and OST (and in some cases both chick and gar were also compared to data within mouse IMM, MAT, and

OST). Venn diagram analysis in chick comparing only NC-derived skeletal elements show that IMM and OST still share the least genes (Fig. 48A, 139 genes), whereas MAT and OST still share the most (Fig. 48A, 681 genes). These results are also consistent with chick results discussed above using the different dataset combinations including mesoderm-derived samples (Figs. 37B&41A). An important difference between chick and gar is the number of genes shared between IMM and OST. In chick, MAT and OST shared the most genes (higher overlap in Venn diagram), whereas IMM and OST shared the least genes (Fig. 37A,B, Fig. 41, Fig. 48A). This high overlap of gene expression between MAT and OST is also conserved in mouse, while IMM and OST also share the least genes. In contrast to tetrapods, gar IMM and OST shared the most genes (2135, Fig. 48B). Importantly, the results in chick show that there is still a high degree of overlap in gene expression between MAT and OST regardless of embryonic origin (NC vs mesoderm) while in IMM and OST the overlap is still low. This strongly supports the idea that regardless of embryonic origin, tetrapod IMM and OST are always the least similar cell types. Finally, analyses of transcripts expressed above threshold identified genes that were unique to each skeletal cell. While in chick IMM had the least uniquely expressed genes (51) compared to MAT (1126) and OST (692), in gar, MAT had the least uniquely expressed genes (365) compared to IMM (867) and OST (3058, Fig. 48B).

#### 6.3.4.3. Differential gene expression analyses in gar show that MAT express many downregulated genes compared to IMM and OST, while in chick IMM is the cell type that had less DEGs

Differential gene expression analyses in chick and gar were performed in transcriptomic data obtained from IMM, MAT, and OST derived from the neural crest. In mouse, OST was also derived from the neural crest, but IMM and MAT were derived from the mesoderm. Despite analyses in chapter 4 suggesting that chick mesoderm- and neural crest-derived cartilage transcriptomes are highly conserved (see chapter 4, Fig. 29A), analyses presented here were limited exclusively to skeletal cells derived from the neural crest since subtle differences in gene expression levels were also identified between the different chondrocyte lineages (see chapter 4, Fig. 30). Also, in chapter 4, differential gene expression analyses in chick were limited to IMM and MAT (see chapter 4 for

more details, Fig. 30), but in this section both IMM and MAT are also compared to OST. In general, differential gene expression analyses in chick showed that IMM had less differentially expressed genes compared to MAT and OST (Fig. 48C). In contrast to chick, gar OST had the least differentially expressed genes, and MAT had several downregulated genes compared to IMM and OST (Fig. 48D, Fig. 49). Gene ontology analysis showed that most of these downregulated genes in MAT are involved mainly in cell cycle processes and chondrocyte proliferation. An important difference between gar and tetrapods, is that in gar the hallmark mature chondrocyte gene *col10a1* is differentially expressed between IMM and MAT and between IMM and OST (downregulated in IMM in both cases), but not between MAT and OST (in mouse and chick *Col10a1* is upregulated in MAT compared to both IMM and OST). Other differentially expressed genes between gar IMM and OST include the osteoblast markers *mbp* and *col1a2* with higher expression levels in OST, and chondrogenic markers, such as *sox5*, *can*, and *col9a2* with higher expression levels in IMM. Unlike tetrapods, other chondrogenic genes such as *sox9* and *col2a1* were not differentially expressed between IMM and OST in gar. This result is consistent with a previous study in zebrafish and gar where the expression of these 'chondrogenic' markers was confirmed in developing osteoblasts of these fish, but not in tetrapods (Eames et al. 2012; also see chapter 5 for more details).

Differentially expressed genes between MAT and OST include classic mineralization and osteoblast markers, such as *msx1, msx2, satb2, mmp9, mmp13, hey1*, members of the Irx family (*irx1* and *irx5*) *vegfa, pthlh, sparc, col1a2* (Satokata and Maas 1994a; Weir et al. 1996; Satokata et al. 2000; Zelzer et al. 2004; Salie et al. 2010; Nyman et al. 2011; Nishimura et al. 2012; Conner and Hornick 2013; Nassif et al. 2014; Yano et al. 2014; Cain et al. 2016; Rosset and Bradshaw 2016) which showed higher expression levels in the OST. This result differs from analyses in tetrapods since many of these genes including *Satb2, Mmp9, Mmp13, Pthlh,* and *Sparc* show upregulated expression levels in both MAT and OST compared to IMM. Differentially expressed genes between MAT and OST also include cartilage markers , such as *sox5*, *acan*, *matn1*, *col11a1*, *col9a1* and *col9a2* (Li et al. 1995; Watanabe et al. 1998; Lefebvre et al. 2001; Zhang et al. 2003; Nicolae et al. 2007), which showed higher expression levels in MAT.

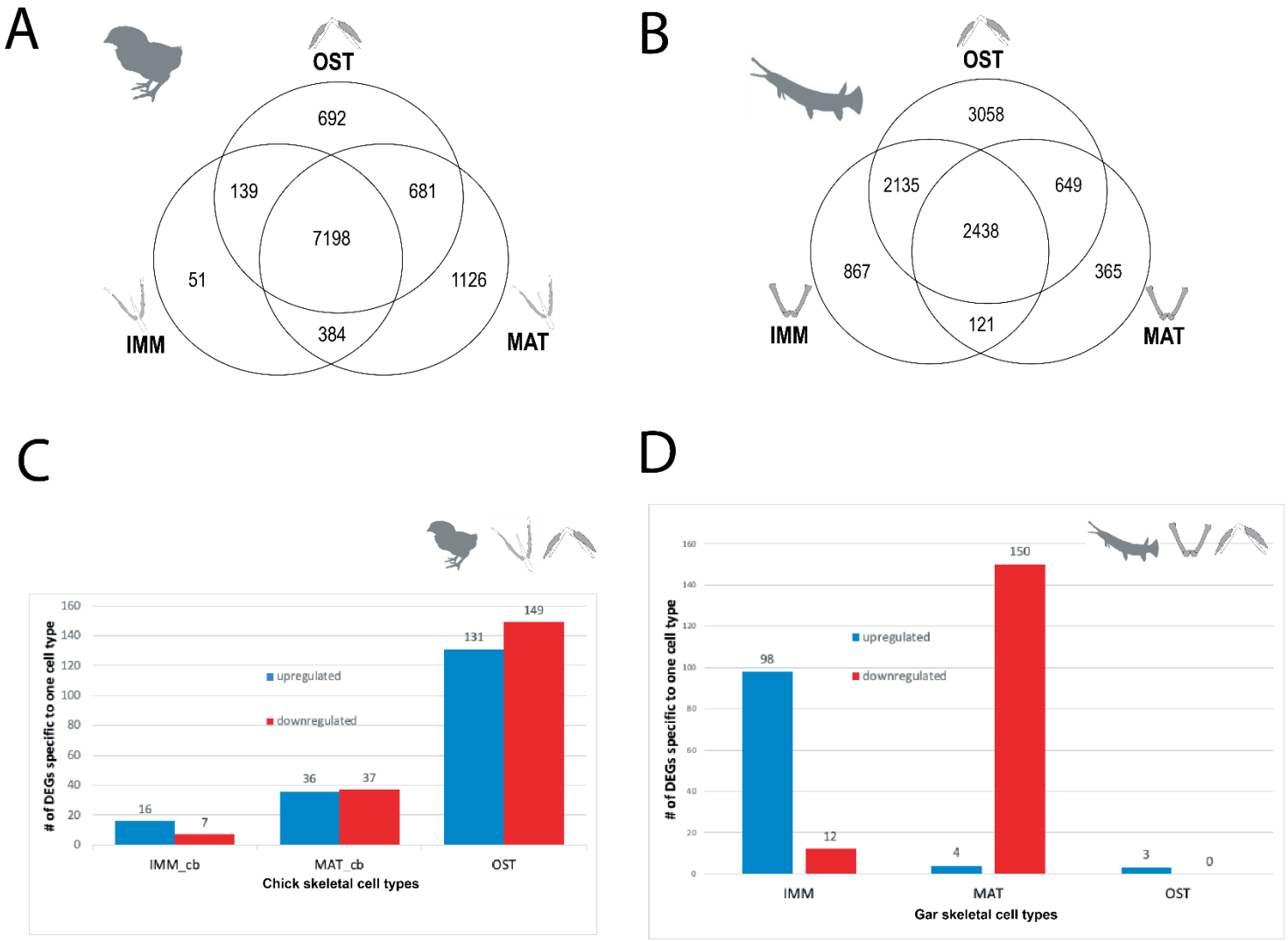



**Figure 48. Gene expression distribution using only neural crest-derived skeletal datasets in chick and gar.** A) Venn diagram analysis in neural crest-derived elements in chick (cb and dentary) shows that MAT and OST still share the most genes, whereas IMM and OST share the least. These results are consistent with chick datasets analyzed above when only humerus and combined hum+cb datasets are compared to dentary dentasets. (B) Venn diagram analysis in gar shows that IMM and OST share the most genes, whereas IMM and MAT share the least. (C) Differential gene expression analysis in chick show that IMM has the least DEGs, whereas the OST has the fewest DEGs in the gar (D). There are many downregulated genes in gar MAT compared to IMM and OST. The chick skeletal element(s) from which IMM and MAT were isolated is indicated in the figure (i.e. cb). Gene lists used for Venn diagram, analyses and DEGs graphs were generated by Katie Ovens.

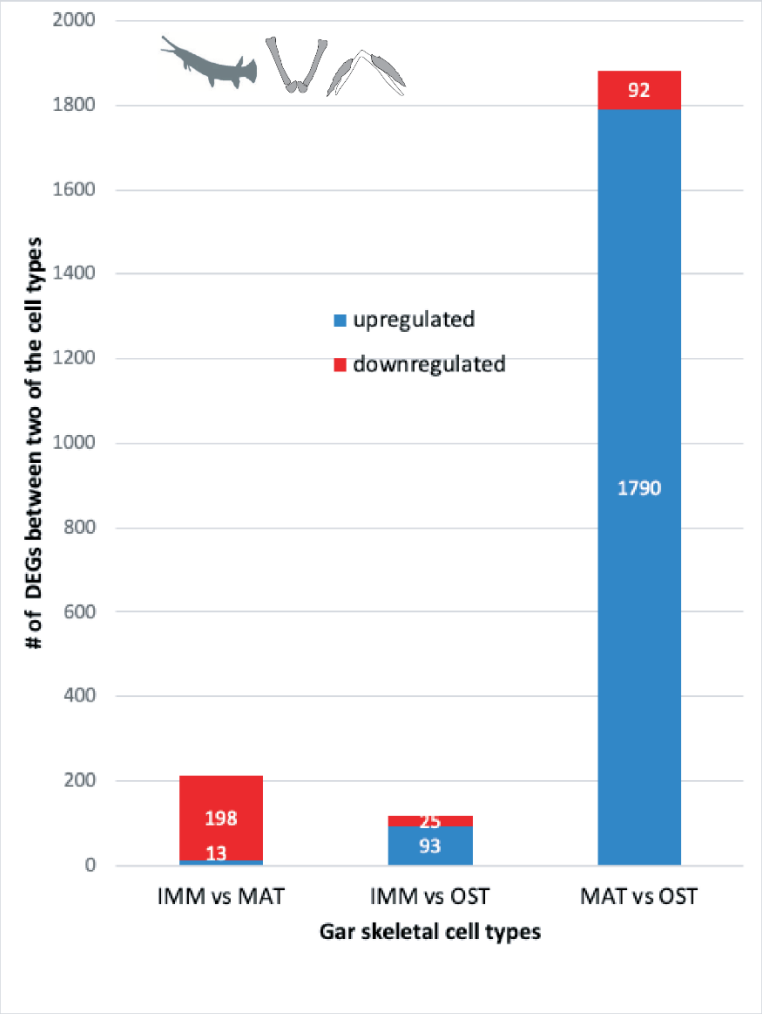



**Figure 49. Pairwise differentially expressed genes between gar skeletal cell types.** Unlike tetrapods, IMM and OST in the gar have the least DEGs (118 genes), whereas MAT and OST have the most (1882 genes). The graph shows upregulated genes in the second cell type compared to the first cell type. Gene lists used for DEGs graphs were generated by Katie Ovens.

#### 6.3.4.4.NC-derived skeletal GRN shows strong positive correlation between IMM and OST in chick and gar

Gene co-expression network analyses described in the previous sections (sections 6.3.2.2. and 6.3.3.2, Figs. 39&42) showed that in mouse and chick that there is a strong cross-inhibition between IMM- and OST-enriched portions of the GRN, and these portions interact positively in MAT (Figs. 39&42). To test whether GRN structure is also conserved in gar, GCN analysis was performed in gar NC-derived skeletal cells (Fig. 50A). A total of 1882 differentially expressed genes were included in the gar GCN analysis. In contrast to tetrapods, genes enriched in gar IMM show a strong positive correlation with genes enriched in OST, and this big portion of positively correlated genes in IMM and OST were negatively correlated with genes enriched in MAT (Fig. 50A,A'). A total of 4156 positive correlations between IMM and OST out of 127,179 total positive correlations in the network were observed in gar (3.3% of total positive correlations between genes enriched in different cell types, Fig. 50A). Most of the positive correlations included in this estimated GRN were between genes enriched within the same cell type, in this case OST-OST. When GCN analysis was performed in chick homologous NC-derived skeletal cells, GRN structure between gar and chick was remarkably similar, although this conclusion in chick was based on an extremely limited dataset (Fig. 50B; only 490 DEGs included in the GRN). Interestingly, in gar, an interacting bridging portion between two distinct cell types is absent (Fig. 50A), so these results suggest that cross-inhibition between IMM and OST and interacting bridging portion between two distinct portions of the skeletal cell GRN is a tetrapod synapomorphy.

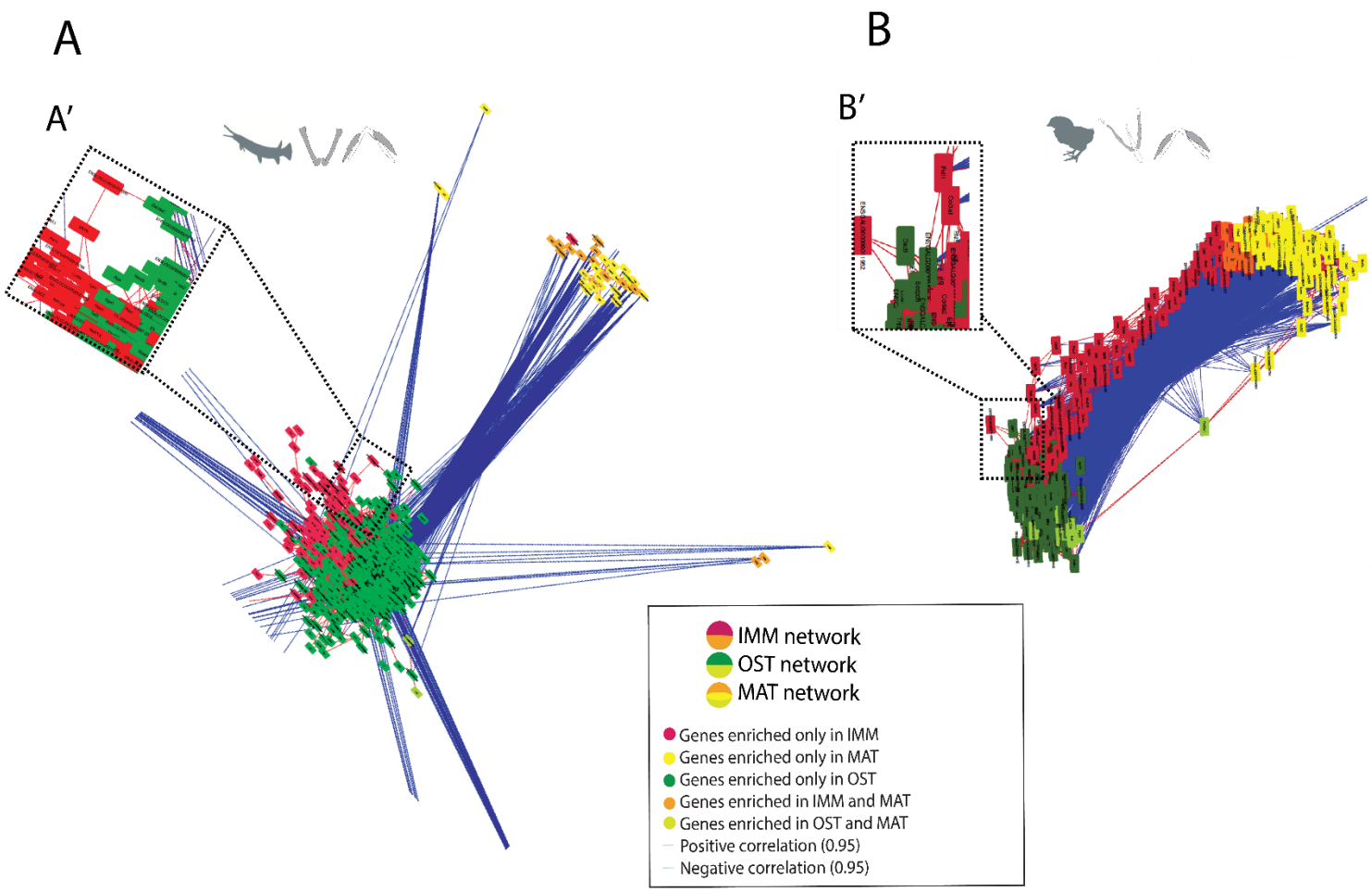



**Figure 50. GRN structure of neural crest-derived elements in gar and chick shows strong positive correlation between IMM and OST.** (A) GCN analysis in gar revealed strong positive correlation between genes enriched in IMM and OST, which are located in the same portion of the GRN (A'). These IMM-OST enriched genes are negatively correlated with genes enriched in MAT. (B) Gene co-expression network analysis in chick using only NC-derived skeletal elements revealed two independent GRN portions enriched in OST and MAT that are negatively correlated with each other. In this case the bridging interacting portion is composed of genes enriched in IMM instead of MAT. In the ceratobranchial, there is strong cross-inhibition between MAT and OST enriched portions of the GRN (B'). The chick skeletal element(s) from which IMM and MAT were isolated is indicated in the figure (i.e. cb). In these analyses, IMM and MAT datasets were always compared to OST datasets (Credit: Katie Ovens).

Since an increase in positive correlations between IMM and OST was seen in NC-derived skeletal cell GRNs of both gar and chick (Fig. 50A',B'), we aimed to test whether NC origin had an influence in GRN structure. To test this, gene ontology analyses were performed in genes positively correlated between IMM and OST of chick (187/490 total genes in GRN=38%) and gar (510/1882 total genes in GRN=28%) GRNs (Fig. 50A,B). GO analyses revealed that some genes positively

correlated between IMM and OST were related to neural crest-dependent processes in chick and gar (Fig. 52A,B), but the number of genes involved in NC-related processes was low (less than 20%) in both species. Furthermore, when chick mesoderm-derived cartilage datasets are included in these GRN structure analyses, positive correlations between IMM and OST portions of the GRN are still observed but to a lesser extent (Fig. 39B, Fig. 42).

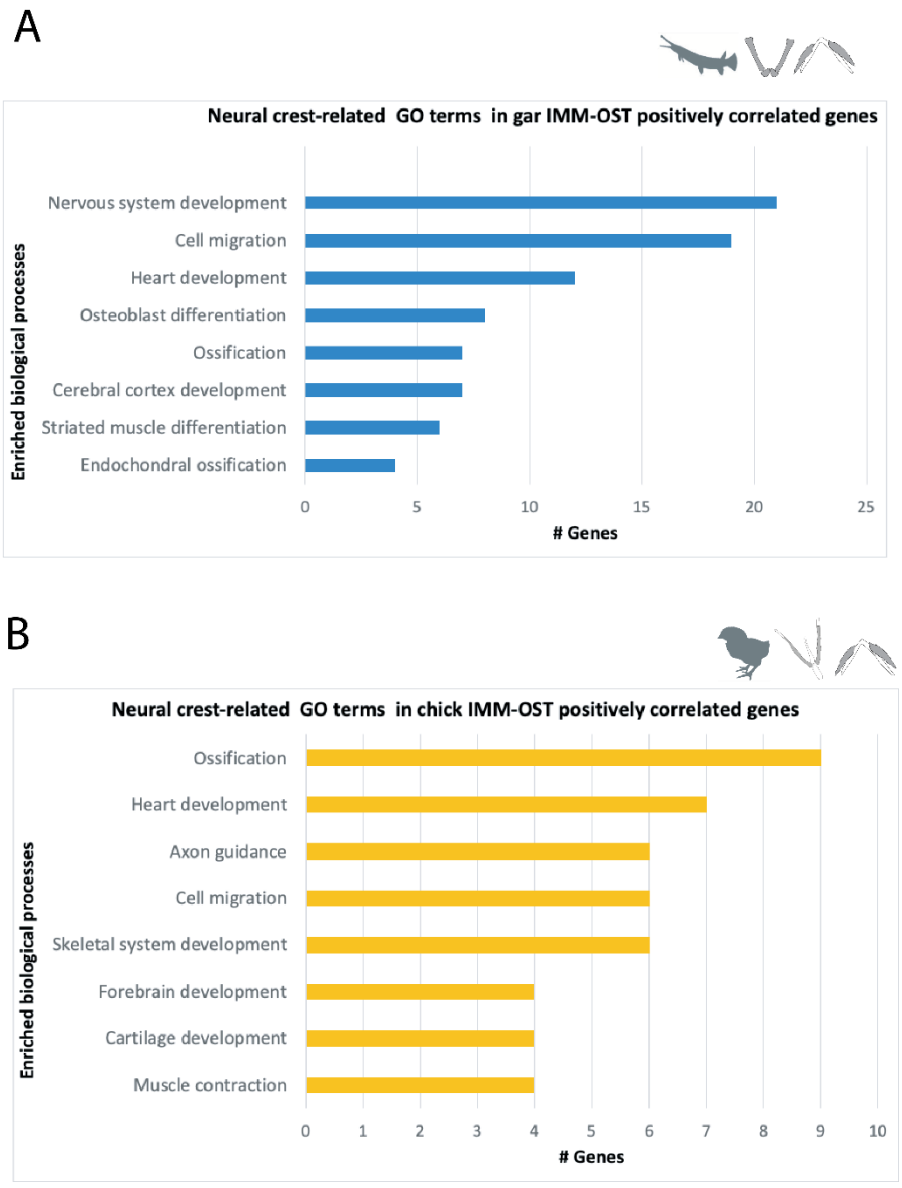



**Figure 51. Gene ontology analyses reveal that positively correlated genes between IMM and OST in both chick and gar are involved in several neural crest related processes.** GO analyses of positively correlated genes between IMM and OST in chick and gar (included in Fig.46A,B) reveal several enriched neural crest-related biological processes.

#### 6.3.4.5. Cluster analysis in chick and gar reveal that mechanisms of gene regulation in MAT are conserved regardless of the embryonic origin, but in gar more clusters show downregulated gene expression in MAT

To test whether regulatory control of gene expression in MAT is conserved in the different vertebrates (tetrapods vs fish), model-based cluster analysis was also performed in NC-derived skeletal datasets of chick and gar (Figs. 52&53). Similar to the previous cluster analyses in chick (Figs. 40&43), when only NC-derived elements were analyzed, categories of gene expression were conserved (Fig. 52A). Some clusters show evidence of independent GRN regulation in MAT. For instance, in cluster 2, expression in IMM and MAT is similar and significantly lower compared to OST (Fig. 52C), and this cluster includes genes such as *PTH1R*, *DLX5, DLX6*, and *HEY1* which are known to be involved during osteoblast differentiation (Robledo et al. 2002; Holleville et al. 2007; Samee et al. 2008; Datta and Abou-Samra 2009; Sharff et al. 2009; Datta et al. 2010; Kawane et al. 2014). An important discrepancy compared to the previous cluster analyses in mouse and chick (see chapter 3, Fig. 19; Figs. 40&43) is that no cluster showed evidence of similar expression between MAT and OST (Fig. 52). Although different numbers of genes are included in GCN and cluster analyses, the fact that no cluster shows similar expression between chick MAT and OST is consistent with chick GCN analysis which showed increased negative correlations between MAT and OST (Fig. 50B, blue lines between MAT and OST portions of the GRN), causing antagonistic expression levels in these cell types. Other clusters showed evidence of GRN interaction in MAT. For example, in cluster 1 expression is enhanced in chick MAT, suggesting specific portions of the GRN driving IMM and OST are acting synergistically in MAT (Fig. 52B,I). These clusters include genes involved in cartilage differentiation and maturation, such as *SOX5, COL10A1*, and *MEF2C* (Smits et al. 2001; Lu et al. 2014; Nakatani and Partridge 2017). *SOX5* has a well-known role during early cartilage differentiation (Smits et al. 2001), but previous work has shown that this gene together with *SOX6* are also needed to develop and maintain hypertrophic chondrocytes (Smits et al. 2004), which could explain the high levels of expression in MAT.

Other clusters showed that expression levels in MAT were an average of expression between IMM and OST. For example, cluster 10 includes the master chondrocyte differentiation marker *SOX9*, as well as, many genes are under the control of this transcription factor, such as *COL2A1*, *SOX6*, *ACAN*, *COL9A2*, and *BARX2* which are crucial during cartilage development and are

expected to show intermediate expression levels between IMM and OST (Bell et al. 1997; Bi et al. 1999b; Lefebvre et al. 2001; Kiani et al. 2002; Meech et al. 2005; Carlsen et al. 2006).

Cluster analysis in gar shows that regulation of gene expression in gar MAT was conserved to some extent, regardless of the dramatic differences of GRN structure in gar (i.e. genes enriched in IMM and OST grouped together and positively correlated + no interacting bridging portion in MAT; Fig. 50A). Again, some clusters show evidence of independent regulation in gar MAT. For instance, in cluster 1 and 5, gene expression in IMM and MAT was similar and significantly lower compared to OST (Fig. 53B,F) including genes involved in bone related processes such as, osteoblast differentiation, angiogenesis, vasculogenesis, and mineralization. In cluster 1, gene expression in IMM and MAT was significantly similar, and lower compared to the OST (Fig. 53B) including genes such as, *mmp13, mmp9, hey1,* and *ctnnb1 (β-catenin;* Inada et al. 2004; Sharff et al. 2009; Chen and Long 2013; Liang et al. 2016). In cluster 5, gene expression in IMM and MAT was similarly low and statistically different from OST (Fig. 53F) including the classic osteoblast markers *sp7* and *spp1* (Sodek et al. 1995; Nakashima et al. 2002; Nishimura et al. 2012; Niu et al. 2017). In contrast to mouse and chick MAT, many of these genes highlighted above (e.g. *mmp9, mmp13,* and *spp1*) showed little to no expression in gar MAT.

Other clusters showed similar expression between MAT and OST in gar. For example, clusters 3 and 7 showed similarly high expression in MAT and OST, and both were statistically significant different from OST (Fig. 53D). Cluster 3 includes *col10a1* which is highly expressed in both MAT and OST suggesting that this gene is not limited to chondrocyte maturation in gar. Strikingly, in gar, more clusters (particularly clusters 2, 4, and 6) showed dramatic downregulation in MAT compared to IMM and OST (Fig. 53C,E,G). These results in gar suggest that gene expression in MAT of earlier diverged vertebrates was different to IMM and OST, and this cell type acquired more similarities with IMM and OST over time. Similar to tetrapods, other clusters in gar show evidence of GRN interaction in MAT. For example, cluster 9 shows that gene expression in MAT is similar to the average expression in IMM and OST. As expected, this cluster includes genes that are involved in both cartilage and bone differentiation such as *sox9* and *runx2, bmp5, bmp6, dlx5, runx3,* and *sparc,* which are also crucial for proper MAT formation (Fig. 53J; Minina et al. 2001; Yoshida et al. 2004; Holleville et al. 2007; Rotllant et al. 2008; Samee et al. 2008; Rosset and Bradshaw 2016). These results suggest that regulatory control of MAT is somewhat conserved across vertebrates, but unlike tetrapods, gar exhibits a dramatic downregulation in MAT.

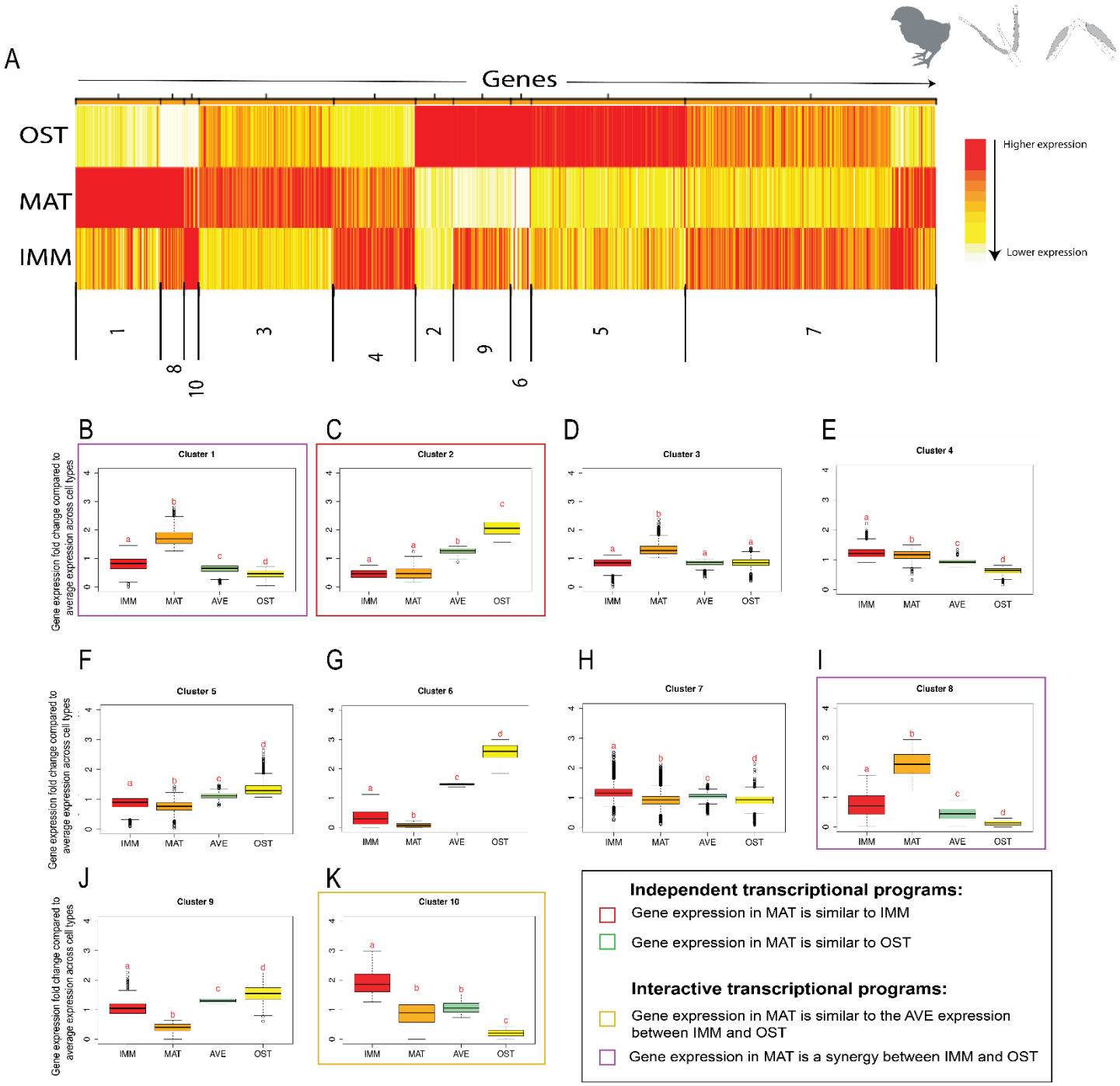



**Figure 52. Model based clustering using only ceratobranchial datasets compared to the dentary datasets in the chick.** (A) Model-based clustering demonstrated that groups of IMM, MAT, and OST genes had discrete categories of expression. (B-K) Box plots showed the distribution of gene expression change among cell types compared to the average expression across all three cell types. AVE represented average expression of IMM and OST. MAT expression in cluster 2 (C) was significantly similar to IMM. (I) Cluster 8 had statistically higher gene expression in MAT than in either IMM or OST. (I) In cluster 10, gene expression in MAT was statistically indistinguishable from AVE. Some clusters show a trend in which IMM and OST have similar expression levels (D,J), but only in cluster 3 gene expression between IMM and OST was significantly similar (D; $P<0.05$). The chick skeletal element(s) from which IMM and MAT were isolated is indicated in the figure (i.e. cb). In these analyses, IMM and MAT datasets were always compared to OST datasets (Credit: Katie Ovens).

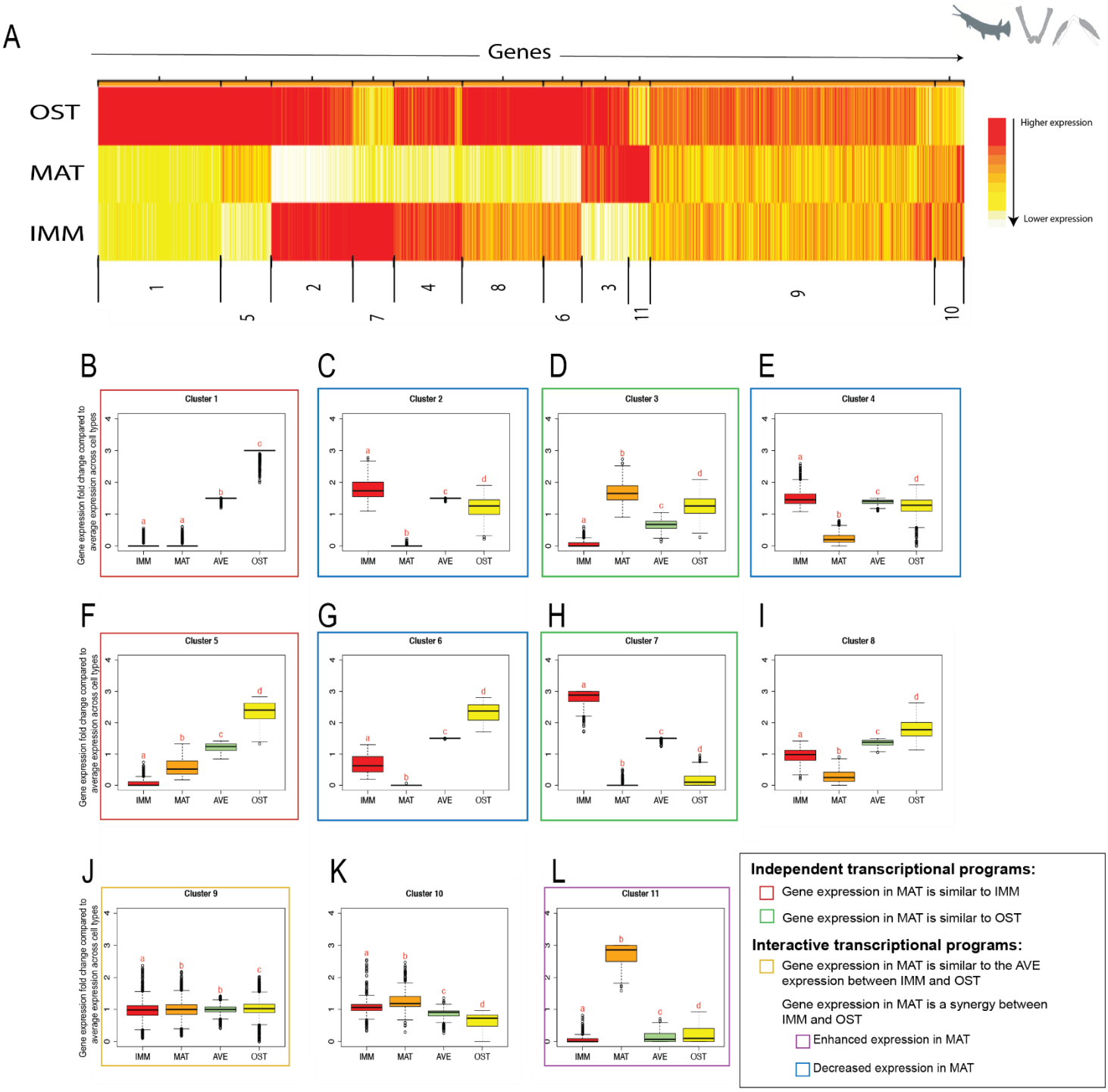



**Figure 53. Model based clustering using gar skeletal datasets.** (A) Model-based clustering demonstrated that groups of IMM, MAT, and OST genes had discrete categories of expression. (B-L) Box plots showed the distribution of gene expression change among cell types compared to the average expression across all three cell types. AVE represented average expression of IMM and OST. MAT expression in cluster 1 and 5 (B,F) was similar to IMM, and significantly different from OST. In clusters 3 and 7 (D,H), gene expression in MAT was similar to OST, and statistically different from IMM. (J) In cluster 9, gene expression in MAT was statistically indistinguishable from AVE. (L) MAT showed enhanced expression compared to IMM and OST. In clusters 2, 4, and 6 (C,E,G), MAT showed significantly lower expression compared to IMM and OST ($P<0.05$; Credit: Katie Ovens).

### 6.3.4.6.The overlap between transcriptomes of chondrocytes isolated from the humerus is higher compared to the ones isolated from the ceratobranchial/ceratohyal

Skeletal tissues of distinct vertebrate clades have evolved in response to different selective pressures. To reaffirm the evolutionary relatedness between gar and tetrapods, we compared the transcriptomes of the different homologous skeletal elements used in the present study. While chick and mouse humeri share 65-70% of the genes expressed above threshold (Fig. 54C,D), gar and chick ceratohyal/ceratobranchial share only 27-38% of the genes (Fig. 54A,B). This could be the result of a larger phylogenetic distance between common ancestors of gar, chick, and mouse (~430 MYA; Kumar et al. 2017; Fig. 2B) compared to the ancestor of only mouse and chick (~312 MYA; Kumar et al. 2017; Fig. 2B).

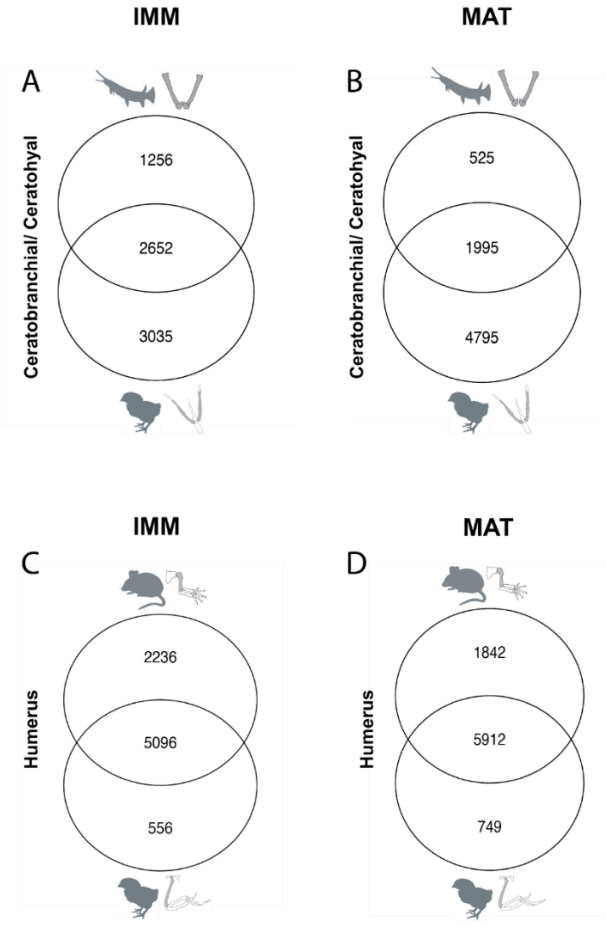



**Figure 54. Venn diagrams comparing homologous skeletal elements.** Comparative transcriptomics showed that IMM and MAT isolated from the humerus of mouse and chick share 65% and 70% of the total number of genes, respectively. IMM and MAT isolated from the ceratohyal/ceratobranchial of gar and chick only share 38% and 27% of the genes, respectively. Gene lists used for Venn diagram, analyses and DEGs graphs were generated by Katie Ovens.

### 6.3.4.7.The osteoblast shows a higher degree of conservation compared to the chondrocyte

In chapter 2, it was hypothesized that earlier diverged cell types are more conserved between distinct clades, than later diverged ones (see chapter 2, Fig 8). This idea is based on the concept of phyletic constraint where a cell type appearing later in evolution is more free to vary compared to a cell type appearing earlier which retains more ancestral features (Gould et al. 1979; Hanken and Thomson 1990; McKitrick 1993). Thus, when comparing IMM, MAT, and OST of mouse, chick, and gar, it is predicted that IMM would be more similar (i.e. share more genes) between the three species, since this cell type evolved first (see chapter 2 for details). In contrast, it is predicted that MAT and OST would show more variation across species. Previous studies with candidate genes support this hypothesis. When comparing gene expression in chondrocytes and osteoblasts of mouse, chick, zebrafish, and gar, gene expression was conserved in chondrocytes of these species, while in osteoblasts expression showed some variation (Eames et al. 2012). For example, chondrogenic markers were found to be expressed in osteoblasts of zebrafish and gar, but not in tetrapods (Eames et al. 2012; our data, see chapter 5 for a more detailed OST evolutionary comparison).

In order to test which skeletal cell type shows a higher degree of conservation across clades, IMM, MAT, and OST isolated from homologous elements were compared between mouse, chick, and gar (Fig. 55). Only one-to-one orthologs were used for these comparisons (IMM, 8387 genes; MAT, 8679 genes; OST, 9021 genes). Using an unbiased transcriptomic approach, these analyses revealed that OST was the most similar cell type between mouse, chick, and gar since all three species shared 44% of the genes expressed above threshold (Fig. 55C, 3961 shared genes). Gene ontology analysis revealed that some genes shared between the OST of the three species are involved in bone related processes such as, osteoblast differentiation, blood vessel development, and ossification including classic bone genes such as the master osteoblast regulator *Runx2*, and many genes that are under the control of this transcription factor such as, *Col1a1*, *Col1a2*, *Mmp13*, and *Sparc* (see chapter 5 for more details regarding evolution of the OST GRN). The three species shared a small percentage of IMM and MAT genes (Fig. 55A,B). While mouse, chick, and gar shared 30% of IMM genes expressed above threshold (2520 genes; Fig. 55A), MAT was the least similar cell type across the three species since they only shared 21% of MAT genes expressed above

threshold (1841 genes; Fig. 55B). However, when comparisons were limited to only two species, MAT was the most similar cell type between mouse and chick (4017 genes, Fig. 55B).

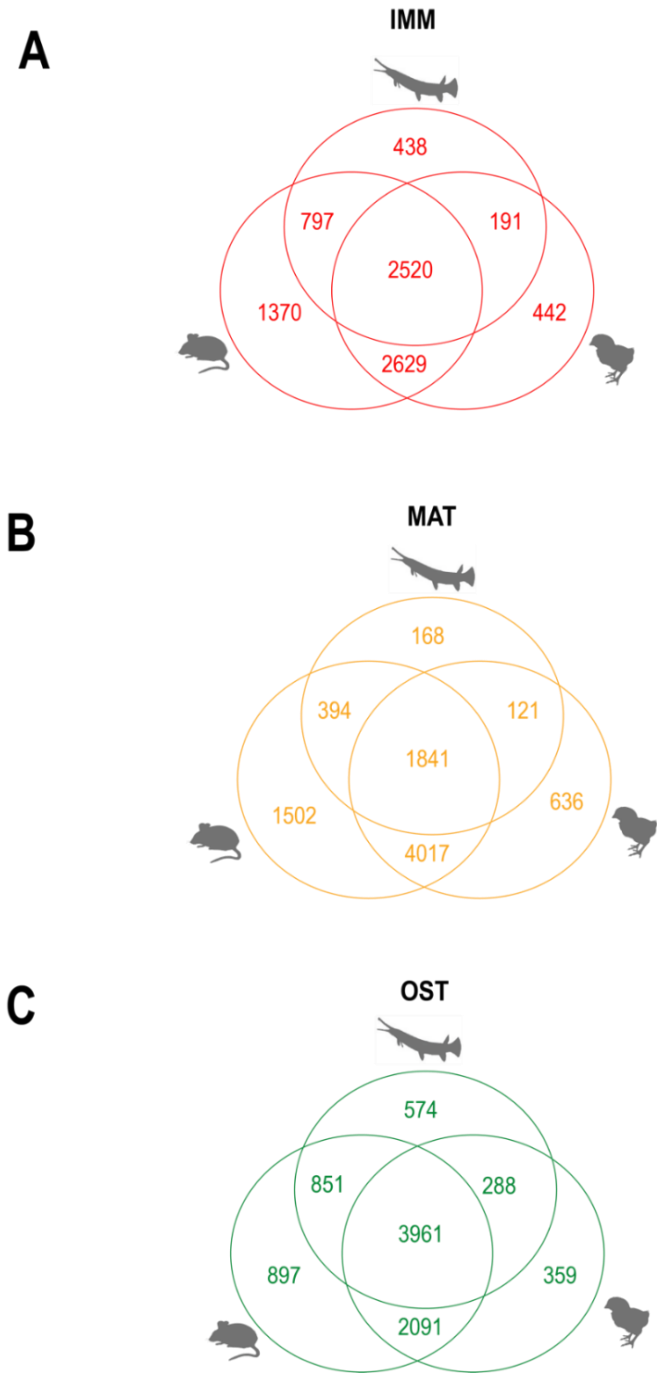



**Figure 55. OST is the most similar cell type between mouse, chick, and gar, and MAT is the least similar.** Venn Diagrams showing genes expressed above threshold in IMM, MAT, and OST of land tetrapods (mouse and chick) and gar. The three species share a total of 2520 IMM genes, 1841 MAT genes, and 3961 OST genes. Gene lists used for Venn diagram, analyses and DEGs graphs were generated by Katie Ovens.

Gene ontology analysis revealed that IMM and MAT genes shared across all three species were related to chondrocyte differentiation, cartilage development, ossification, osteoblast differentiation, Wnt signaling pathway, bone mineralization, bone development, and apoptosis, among others. Genes shared in IMM and MAT of all three species include hallmark cartilage differentiation markers such as, *Sox9*, *Sox5*, *Sox6*, *Acan*, and *Col9a1* (Watanabe et al. 1998; Bi et al. 1999; Smits et al. 2001; Zhang et al. 2003), as well as classic cartilage maturation genes such as, *Runx2*, *Col10a1*, *Mef2c*, *Runx3*, and *Wnt5b* (Yang et al. 2003; Yoshida et al. 2004; Arnold et al. 2007; Lu et al. 2014). In all three IMM, MAT, and OST comparisons, mouse and chick shared a higher percentage of genes (IMM, 61% of shared genes; MAT, 67% of shared genes; OST, 67% of shared genes), and both tetrapods shared a smaller percentage with gar (Fig. 55).

## 6.4. DISCUSSION

Skeletal cell evolution has been little explored since most evolutionary studies focus on morphological differences between homologous skeletal elements in various phylogenetic clades rather than differences in skeletal cell histogenesis. Here, LCM was used to characterize the transcriptomic profiles of IMM, MAT, and OST in three distinct vertebrate clades, and skeletal cell gene regulatory networks (GRNs) were revealed. Previous studies have provided insight into how the GRNs driving cartilage and bone might be organized during their differentiation (Cole 2011; Calabrese et al. 2012; Kerkhofs et al. 2012; Oh et al. 2014; Liu and Lefebvre 2015; Ohba et al. 2015; Hojo and Ohba 2019; Kang et al. 2019), but the evolution of skeletal cell GRNs needs to be further explored.

As discussed in chapter 3, analyses in mouse suggest that one GRN regulates the differentiation of IMM, MAT, and OST, and specific portions of this GRN are active in each cell type (Fig. 1; see chapter 3, Figs. 15&17), but is GRN structure conserved in distinct animal clades? To understand GRN organization and evolution, we compiled transcriptomes of homologous skeletal cells in mouse, chick, and gar, and constructed skeletal cell GRNs. We hypothesize that skeletal cell GRN structure and mechanisms of gene regulation in MAT are conserved features across vertebrates. Gene expression analyses in mouse and chick showed that MAT was largely an

overlap of gene expression between IMM or OST, since it shared many genes with both cell types, and it also had the fewest differentially expressed genes (see chapter 3, Fig. 15; Figs. 37&41). This overlapping gene expression in MAT was also supported by gene co-expression network (GCN) analyses. GCN analyses in mouse and chick identified two independent portions of the GRN enriched in IMM and OST, which showed strong cross-inhibition with each other, and they were positively interacting through enriched genes in MAT (see chapter 3, Fig. 12; Figs.39&42). In chick, however, positive interactions were also observed between IMM and OST, but negative interactions between these two cell types were dominant (Figs. 39B&42). Although a few exceptions exist (Blumer et al. 2004; Ytrehus et al. 2004; Blumer et al. 2005; Hall 2015), cartilage is typically avascular, whereas bone is highly vascularized. A particular study in embryonic chick femur showed the presence of 'communicating canals' in the different zones of the growth plate, including resting and proliferative zones (i.e. IMM in the present study), and confirmed the expression of *Vascular endothelial growth factor* (*VEGF*) in these canals (Blumer et al. 2004), so expression of *VEGF* was not limited to blood vessels in bone. During endochondral ossification, MAT in tetrapods often is invaded by vasculature, forming the bone marrow cavity (Stricker et al. 2002; Ortega et al. 2004; Moriishi et al. 2005). Although blood vessels were not observed at the timepoint chondrocytes were collected from chick humerus and ceratobranchial (see chapter 4, Figs. 27&28), it is possible that these cells have already expressed 'OST' genes that would later induce vessels and were influencing the expression of chick IMM genes at this early timepoint. These results in chick suggest that this network structure (i.e. positive correlation between IMM and OST portions of the GRN) might be a conserved feature in avian species but not in mammals (Figs.39B&42), but skeletal cells from other animal clades need to be analyzed to support this assumption.

Despite minor structural differences between the mouse and chick GRNs being identified, important cartilage and bone markers were conserved in both species. For instance, the expression of classic MAT markers, such as *Col10a1* and *Spp1* was verified in mature chondrocytes of both mouse and chick (Fig.44B,C-F,G). Importantly, the expression of a novel conserved MAT enriched gene, *Nhej1*, was also confirmed in mature chondrocytes of mouse and chick (Fig.44D,H). NHEJ1 is part of a protein complex that includes XRCC4, XRCC5, XRCC6, PRKDC, LIG4, PAXX, and NAA15 (Berman et al. 2000). Specifically, NHEJ1 interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining, and may serve as a bridge between XRCC4 and other factors located in DNA ends (Ahnesorg et al. 2006). The NHEJ complex activates the

expression of important MAT and OST differentiation genes including *RUNX2, MSX2, DLSX5* and *BGLAP* (Berman et al. 2000). Interestingly, most of the proteins involved in the NHEJ complex were expressed in tetrapod MAT and OST in our data (data not shown). Also, studies in mutant mice and chicken have shown that complete or partial deletion of *Nhej1* can inhibit expression of *Ihh* in certain tissues, leading to short limbs, reduced skull ossification and lack of trabecular bone (Will et al. 2017; Kinoshita et al. 2020). The effects that NHEJ1 has in chondrocytes are still uncertain, however due to the importance of the NHEJ1 complex in activating important MAT and OST genes, the idea that NHEJ1 might influence cartilage maturation needs to be further explored. In the present work, we attempted to study *Nhej1* function, but overexpression of this gene in ATDC5 cells resulted in significant cell death (data not shown). When comparing chick and mouse, subtle gene expression differences were also identified. In chick, *RUNX2* was expressed highly in both MAT and OST, but it was not differentially expressed when compared to IMM. A possible explanation for this is that pre-hypertrophic and hypertrophic zones in the avian growth plate are disorganized, so cells at different stages of differentiation can be located next to each other (Leach Jr and Gay 1987; Pines and Hurwitz 1991; see chapter 4, Figs. 27&28). It is possible that there was more variation in the cells captured from chick skeletal elements compared to mouse. Furthermore, chicks were raised in farms, so they become more susceptible to variation, compared to the mouse that is an inbred individual raised in a controlled laboratory environment which shows little phenotypic variation. Consistent with previous work (Kague et al. 2016), another important difference between the mouse and chick skeletal datasets presented here was the absence of a clear *SP7* orthologue in chick. These results suggest that *SP7* might not be required for avian bone formation, as it is during mammalian and fish osteogenesis (Niu et al. 2017; Yu et al. 2017).

Gene expression in gar MAT showed important differences when compared to tetrapods. In contrast to mouse and chick, gar MAT has several downregulated genes compared to IMM and OST (Fig. 48D, Fig. 49, Fig. 53). In gar, differentially expressed genes between MAT and OST include classic mineralization and osteoblast markers, such as *msx1, msx2, satb2, mmp9, mmp13, hey1*, members of the Irx family (*irx1, irx5*) *vegfa, pthlh, sparc, col1a2* (Satokata and Maas 1994a; Weir et al. 1996; Satokata et al. 2000; Zelzer et al. 2004; Salie et al. 2010; Nyman et al. 2011; Nishimura et al. 2012; Conner and Hornick 2013; Nassif et al. 2014; Yano et al. 2014; Cain et al. 2016; Rosset and Bradshaw 2016) which showed higher expression levels in the OST. In mouse and chick, however, many of the genes highlighted above were highly expressed in both MAT and OST such

as, *Satb2, Mmp9, Mmp13, Pthlh,* and *Sparc*, whereas in gar MAT they exhibited little to no expression. These results suggest that gar MAT was perhaps at a relatively younger stage than chick and mouse MAT. Despite the presence of bone matrix in the perichondrium of the 13dpf ceratohyal was revealed by Trichrome staining (Fig. 45E), histological analysis in gar ceratohyal revealed that chondrocyte hypertrophy was not so evident in the 13dpf ceratohyal (Fig. 45D). However expression of the hallmark mature chondrocyte gene *col10a1* was already high in the 13dpf ceratohyal and the expression domain was already expanding towards the epiphyses (Fig. 46C). These results in gar suggest *col10a1* is only a hypertrophy marker, and should not be considered alone to identify MAT, since there are other factors potentially driving cartilage hypertrophy.

The mechanisms that trigger chondrocyte hypertrophy have been largely unexplored, but previous studies suggest that this increase in chondrocyte size could be associated with accumulation of organic osmolytes (i.e. betaine, aminoacids, and sugars), an increase in organelle size, and water accumulation in the cytoplasm and nucleoplasm (Buckwalter et al. 1986; Farnum et al. 2002). In a more recent study in mouse and jerboa, Cooper *et al.* identified three phases of chondrocyte enlargement, and the last of these three phases is regulated by the Insulin-like Growth Factor (IGF) signaling (Cooper et al. 2013). Previous work in mutant mice has also revealed that *Igf1* increases chondrocyte hypertrophy, and hypertrophic chondrocytes of $Igf1^{-/-}$express normal levels of MAT markers such as, *Col10a1, Ibsp*, and *Alpl,* but they are smaller in size compared to hypertrophic chondrocytes of wild-type mice (Wang et al. 1999). The effects of *Igf2* during cartilage hypertrophy have also been assessed (Murray et al. 2013; Uchimura et al. 2017). For instance, a particular study in mutant mice revealed a role for *Igf2* during cartilage hypertrophy. Bones of *Igf2* mutant mice grew slower compared to wild type mice, exhibiting a shorter prehypertrophic zone and a disproportionally larger hypertrophic zone (Uchimura et al. 2017). Likewise, in *Igf1* mutant mice, the expression of classic MAT markers such as, *Col10a1, Mmp13,* and *Alpl* was not significantly altered in *Igf2* mutant mice, although subtle gene expression differences were detected (Uchimura et al. 2017). Despite current studies analyzing the role of *Igf1* and *Igf2* have mostly focused in mammals, it is possible that these genes might also have an effect during chondrocyte hypertrophy of nonmammalian vertebrates. Two previous studies analyzed *Igf1* and *Igf2* genes in different nonmammalian vertebrates in order to identify regulatory changes during evolution, and the results obtained were then compared to mammals. Despite structural and functional features of *Igf1* and *Igf2* appear to be generally conserved between nonmammalian vertebrates and mammals, the spotted

gar was one of the few nonmammalian species that exhibited more variation in *Igf1* and *Igf2* gene organization (Rotwein 2018b; Rotwein 2018a). Currently it is not well understood how alterations in gene structure may affect the function of *Igf1* and *Igf2* during cartilage hypertrophy in gar, so a more detail histological and molecular study analyzing expression patterns of both genes at different developmental stages during MAT differentiation, as well as expression of classic MAT markers (e.g. *col10a1*, *spp1*, *ibsp,* and *mmp13*), would provide more insight into how cartilage hypertrophy might be regulated in gar.

GRN structure in gar also exhibited important differences compared to tetrapods (Fig. 50A). In contrast to mouse and chick, the antagonistic relationship between the IMM and OST portions of the GRN was dramatically reduced in gar (Fig. 50A). A very limited dataset of chick NC-only genes suggests that gar results might be related to neural crest origin (Fig. 50B), but additional neural crest-derived skeletal datasets from other vertebrates should be included into these comparisons to confirm whether embryonic origin is indeed affecting GRN structure (e.g. increased positive correlations between IMM and OST). While in earlier diverged vertebrates (i.e. gar) portions of the GRN underlying IMM and OST overlapped with each other and shared the expression of many genes (Fig. 50A, Fig. 56A), in later diverged vertebrates (i.e. mouse and chick), the increased number of negative correlations between IMM and OST (Figs. 39,42&56B) suggest that portions of the GRN directing these cell types became distinct during evolution. In terms of GRN structure, the gar OST appears to be more 'chondrogenic' due to the reduced number of negative correlations (and increased positive correlations) between IMM and OST portions of the GRN in this species (Fig. 50A, Fig. 56A). These results are consistent with Venn diagram analysis which revealed that gar IMM and OST transcriptomes showed the highest similarity (Fig. 48B) compared to tetrapods (Fig. 37A,B; Fig. 41A; Fig. 48A). In chapter 2, we hypothesized that the osteoblast might have evolved from a chondrocyte by co-opting portions of an ancestral GRN initially used to make cartilage (see chapter 2, Fig. 6). Then in chapter 5, we provided evidence that supported this hypothesis since the OST of an earlier diverged vertebrate (i.e. gar) expresses higher levels of classic chondrogenic markers compared to OST of later diverged vertebrates (i.e. mouse and chick). These results suggest that in earlier diverged vertebrate clades (i.e. gar) the transcriptional programs driving IMM and OST overlapped (Fig. 56A), and then they became more distinct in later diverged vertebrates (i.e. mouse and chick). Then, in later diverged animal clades new genes were incorporated into the

osteoblast GRN conferring novel functions and causing the IMM and OST portions of the GRN to become more and more distinct over time (Fig. 56B).

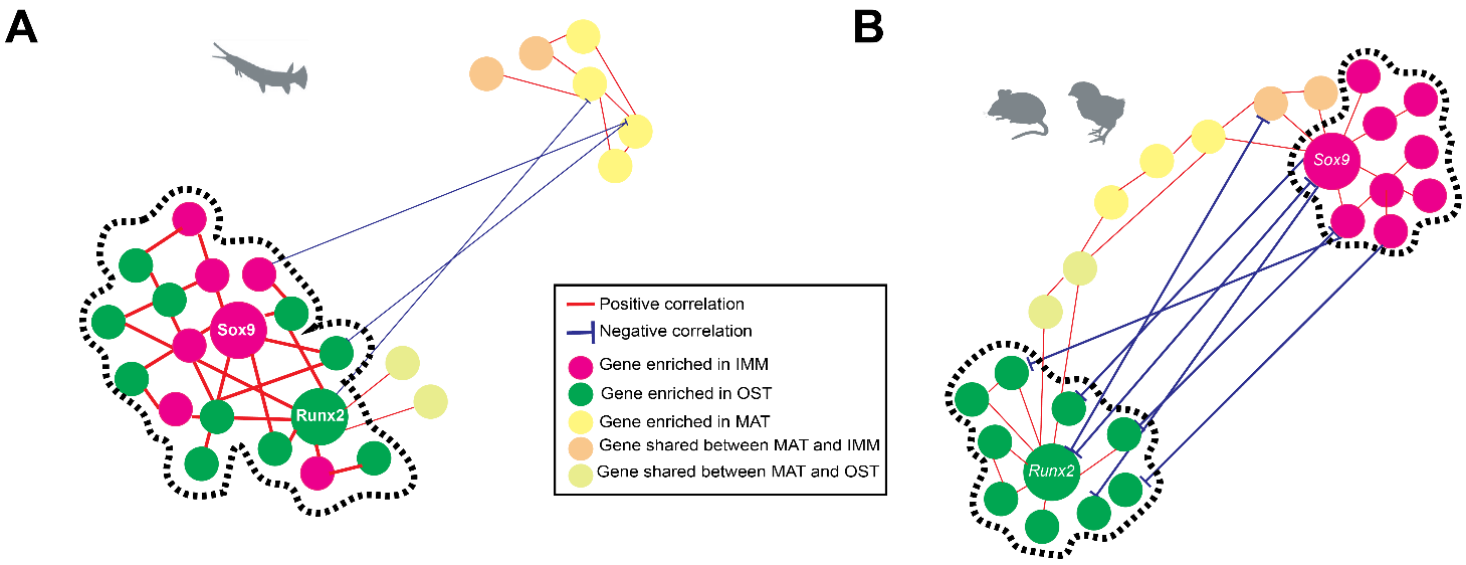



**Figure 56. Portions of the GRN directing IMM and OST became more distinct over time.** Schematic representation of estimated skeletal cell GRNs for gar, chick, and mouse. (A)While in earlier diverged vertebrates (i.e. gar) portions of the GRN underlying IMM and OST overlapped with each other and shared the expression of many genes. (B) In more recently diverged vertebrates (i.e. mouse and chick), the increased number of negative correlations between IMM and OST, suggest that portions of the GRN directing these cell types became distinct during evolution.

Despite differences in GRN structure were identified between tetrapods and gar, cluster analyses revealed that regulatory control of mature chondrocytes is generally conserved among vertebrates (see chapter 3, Fig. 19; Figs. 40,43,52 and 53). Some clusters showed independent regulation in MAT, suggesting that portions of the GRN active in IMM and OST were also active in MAT, regulating gene expression in this cell type in a similar manner as in IMM or OST (see chapter 3, Fig. 25A). Other clusters showed evidence of interaction in MAT via averaging and synergism (see chapter 3, Fig. 25B). Strikingly, when cluster analysis was performed in gar skeletal cell types, more clusters showed significant downregulated gene expression in MAT compared to IMM and OST (Fig. 53C,E,G). These results are consistent with differential gene expression analyses in gar which revealed many downregulated genes in MAT compared to the other two cell

types (Fig. 48D, Fig. 49). As discussed above, it is possible that gar was not at an equivalent stage compared to mouse and chick, since hypertrophy in MAT was not as evident at the staged analyzed (Fig. 45D) which could also explain this result. In general, these findings suggest that in earlier diverged vertebrates there was a repression circuit activated in a neural crest specific GRN to make MAT, and later this GRN was modified and transcriptional repressors in this circuit were inhibited. Evidence of evolution in a repression circuit has been identified in fungi, where changes in the coding sequence of the protein *Matα2*, and sub-sequent modification of cis-regulatory sequences in its future target genes, gave rise to a new transcriptional program that involved repression of these novel genes (Britton et al. 2020). Similar changes in regulatory proteins and its target genes in an early MAT (e.g. gar) could have led to the formation of novel transcriptional programs (and by extension activation of novel genes) in MAT of more recently diverged clades (e.g. tetrapods) that resulted in inhibition of MAT repressors that were active in early MAT. Since gar MAT evolved before mouse and chick MAT, it is likely that this 'repression state' in gar MAT closely resembled the ancestral MAT. These data recalls to the concept of phyletic constraint were earlier diverged lineages retain more ancestral features (Gould et al. 1979; Hanken and Thomson 1990; McKitrick 1993).

Skeletal tissues of distinct vertebrate clades have evolved perhaps in response to different selective pressures. In order to test gene expression similarities between skeletal elements in mouse, chick, and gar, transcriptomes of the different homologous skeletal elements used in the present study were compared. While chick and mouse humeri share 65-70% of the genes expressed above threshold (Fig. 54C,D), gar and chick ceratohyal/ceratobranchial share only 27-38% of the genes expressed above threshold (Fig. 54A,B). A possible explanation for this is that tetrapod limbs were exposed to similar selective pressures to withstand the effects of gravity which ultimately favored locomotion in the open air. Thus, the genes employed to achieve this are expected to be highly conserved between both clades. In contrast, avian and fish ceratobranchial/ceratohyal were subjected to different evolutionary pressures related to diet and food processing, so the gene expression overlap between skeletal elements between these species is not as high. Furtheremore, there is a larger phylogenetic distance between common ancestors of gar, chick, and mouse (~430 MYA; Kumar et al. 2017; Fig. 2B) compared to the ancestor of only mouse and chick (~312 MYA; Kumar et al. 2017; Fig. 2B).

Based on the concept of phyletic constraint that posits that earlier evolved traits are more conserved (Gould et al. 1979; Hanken and Thomson 1990; McKitrick 1993), we predicted that when comparing IMM, MAT, and OST transcriptomes across vertebrate clades, IMM would be more conserved than MAT and OST, since IMM evolved first (see chapter 2 for details). However the analyses presented here revealed that the OST transcriptome was actually more conserved between mouse, chick, and gar (Fig. 55C, 44% shared genes), which do not support the hypothesis. The OST was the only cell type that was isolated from exactly the same element in all three species (i.e. dentary bone), whereas IMM and MAT were isolated from two elements of different embryological origin (i.e. humerus in mouse and chick, ceratobranchial/ceratohyal in chick and gar), and these chondrocyte data were then combined to make evolutionary comparisons of the three species possible. Although, combining datasets could have added a little more variation into the chondrocyte data, analysis in chick presented in chapter 4 show that the transcriptomes of chondrocytes derived from different embryonic lineages (i.e. mesoderm vs neural crest) are highly similar (see chapter 4, Fig. 29A) which supports the reliability and validity of these results. Moreover, as discussed above, gar appeared to be in an earlier developmental stage compared to mouse and chick (i.e. gar MAT hypertrophy not as evident as in tetrapods, see Fig. 45D) which could have also added more variation into the chondrocyte data. Despite these differences, when comparing gene expression between IMM, MAT, and OST of the three clades, essential genes during cartilage (e.g. *Sox9, Sox5, Sox6,* and *Col10a1*) and bone differentiation (i.e. *Runx2, Col1a1,* and *Sparc*) were highly conserved between mouse, chick, and gar. These results suggest that these core set of genes were likely included in the ancestral GRN to make cartilage and bone. In summary, our results generally supported our hypotheses. First, mechanisms of gene regulation and GRN structure are conserved to some extent in the three vertebrate clades analyzed, however, some differences were noted when these analyses were performed in gar neural crest-derived elements. Adding more phylogenetic clades into these evolutionary comparisons will provide more insight into the GRN driving skeletal cell differentiation and evolution. Together these results highlight the complexity of skeletal cell GRN organization and propose a novel unbiased approach through which to understand differentiation and evolutionary origins of cartilage and bone.

## 6.5.METHODS

**Embryo Collection and tissue processing.** All animal procedures were performed according to guidelines approved by the University of Saskatchewan Animal Care and Use Committee. *Chick (Gallus gallus):* White leghorn chicken eggs were incubated in a humified incubator at a constant temperature of 37ºC. Embryos were harvested at Hamburger-Hamilton stage 36 (~E10.5; (Hamburger and Hamilton 1951). Each embryo was decapitated, and the forelimbs and lower jaws were dissected and immediately placed in 1X PBS/DEPC, followed by embedding in OCT (Tissue Tek, Torrance, CA, USA; for detailed methodology see chapter 3), and immediately flash-frozen using liquid N2 and 2-Methylbutane (isopentane). *Spotted gar (Lepisosteus oculatus):* Gar larvae were kindly provided by Dr. Allyse Ferrara (Nicholls State University Thibodeaux, LA). Larvae were placed in beakers and raised at a constant temperature of 24ºC and a pH between 6.5-8 in a water bath. Once gar larvae reached 13 and 14 dpf they were euthanized using 0.2% tricaine. The length of these larvae varied between 1.6-2.2 cm. Heads were dissected and immediately placed in 1X PBS/DEPC, followed by embedding in OCT. For whole mount and section histology, both chick and gar dissected tissues were fixed in 4% PFA prior embedding in OCT (in the case of section histology).

**Alcian blue/ Alizarin red whole mount skeletal stain.** Mouse E14.5 embryos, chick HH36 embryos, and 13dpf gar larvae were fixed using 4% paraformaldehyde, and then stained with Alcian blue and Alizarin red using an acid-free solution that included $MgCl_2$ to differentiate staining, and then cleared in glycerol/KOH as described elsewhere (Eames et al. 2011).

**Histology.** Safranin O/Fast Green staining was performed on 10 μm thick frozen sections of the E14.5 mouse humeri, the HH36 chick humeri and ceratobranchial, and the 13dpf gar ceratohyal as described previously (Ferguson et al. 1998). Trichrome staining was performed on 10μm thick frozen sections of the E14.5 mouse, HH36 chick, and 13dpf gar mandibles, as described previously (Clark and Smith 1993).

**Laser Capture Microdissection (LCM).** LCM was performed on a Laser Microdissection - Molecular Machines & Industries (MMI) CellCut apparatus. For gar and chick analyses, five

biological replicates for each tissue were captured. *Chick:* Immature and mature chondrocytes were captured from the developing HH36 humeri and ceratobranchial. In contrast, osteoblasts were isolated from the developing dentary in the jaw at HH36. *Spotted gar:* Immature and mature chondrocytes were captured from 13 dpf ceratohyal whereas osteoblasts were isolated from the developing dentary in the jaw at 13 dpf. Similar to mouse, a larger capture area was necessary to obtain a sufficient amount of RNA from mature cartilage of both chick and gar (Tables 7&8). The captured cells were collected onto the inner lid of 0.5ml MMI IsolationCaps (either Diffuser caps (Prod#50202) or Transparent caps (Prod#50204; MMI Molecular Machines & Industries). Once RNA-seq data was obtained, bioinformatic analyses identified some samples as outliers, and these were not included in our analyses. For details on bioinformatic analyses see chapter 3 methodology.

**RNA Isolation and Amplification**. RNA was isolated using the ARCTURUS PicoPure RNA Isolation Kit (ThermoFisher Scientific; Cat# KIT0204) according to the manufacturer's instructions and DNase treatment was done using RNase-Free DNase (Qiagen; Cat#79254). For gar and chick, the RNA from immature cartilage ($n$ = 5), mature cartilage ($n$ = 5), and bone ($n$ = 5) were extracted. Chick chondrocytes isolated from the humerus and ceratobranchial were the same ones presented in chapter 4 (Table 5). Chick and gar osteoblasts were the same ones presented in chapter 5 (Table 8) RNA was then amplified with one or two rounds using MessageAmp II aRNA Kit (ThermoFisher Scientific; Cat# AM1751). The RNA integrity was evaluated on the observation of a signature eletropherogram pattern (Bioanalyzer).

**Table 9.** LCM Surface area obtained from 13dpf gar skeletal tissues

| Skeletal element | Cell type | Surface Area captured ($um^2$) |
|---|---|---|
| Ceratohyal | IMM 1_ch | 462,704 |
| Ceratohyal | IMM 2_ch | 458,932 |
| Ceratohyal | IMM 3_ch | 225,184 |
| Ceratohyal | IMM 4_ch | 398,958 |
| Ceratohyal | IMM 5_ch | 458,283 |
| Ceratohyal | MAT 1_ch | 276,732 |
| Ceratohyal | MAT 2_ch | 333,678 |
| Ceratohyal | MAT 3_ch | 501,978 |
| Ceratohyal | MAT 4_ch | 707,613 |
| Ceratohyal | MAT 5_ch | 564,122 |

**Library preparation and deep RNA Sequencing.** RNASeq libraries were prepared using the Illumina TruSeq RNA Sample Prep Kit v2 with the following modification: the protocol was started at the Elute, Prime, Fragment step using 5 µl amplified mRNA (minimum amount was 50 ng mRNA as determined using Quant-iT RiboGreen RNA Assay Kit (Invitrogen)). The quality of each library was checked on a DNA 1000 chip using the 2100 Bioanalyzer (Agilent Technologies Inc.). The size range of the prepared libraries (insert+ adapters) was 200-600 bp with an average size of 290 bp. Library concentrations were determined by qPCR using the KAPA SYBR FAST ABI Prism qPCR Kit (Kapa Biosystems) and the StepOnePlus Real-Time PCR System (Applied Biosystems). Equimolar concentrations of the libraries were pooled and a final concentration of 12 pM was used for clustering in one lane of a flowcell on the cBOT (Illumina). The samples were then sequenced (2 x 101 cycles, paired-end reads) on the HiSeq2500 (Illumina) using the TruSeq SBS Kit v3-HS 200 cycles Kit (Illumina).

***In situ* hybridization (ISH).** Skeletal elements were dissected and fixed overnight in 4% paraformaldehyde in DEPC- phosphate buffered saline (DEPC-PBS) at 4°C. Tissues were then washed twice with PBS and embedded in OCT using liquid N2 and 2-methylbutane. Ten µm sections were collected from frozen tissues on Superfrost plus slides (Fisher scientific) for *in situ* hybridization analysis. Chick plasmids containing the *SPP1* and *COL10A1* genes were kindly provided by Dr. Rich Schneider (University of California, San Francisco) and Dr. Ralph Marcucio (University of California, San Francisco). Rat plasmid containing the *Nhej1* gene was kindly provided by Dr. Laurent Villard (Aix-Marseille Université, France). The mouse and the rat NHEJ1 protein sequences are 85% similar. Gar plasmids containing the *col2a1* and *col10a1* genes were kindly provided by Dr. John H. Postlethwait (University of Oregon). Plasmids were transformed and then sequence was confirmed in order to make RNA antisense probes. For details on methodology see chapter 3 (pp. 89,90).

# CHAPTER 7

## Limitations and future directions

# CHAPTER 7. Limitations and future directions

The recent advent in comparative transcriptomics and genomics have greatly aided in revealing the gene regulatory network (GRN) underlying skeletal cell differentiation, but the organization and evolution of skeletal cell GRNs needs to be further explored. To increase understanding in these relatively new areas of research, immature chondrocytes (IMM), mature chondrocytes (MAT), and osteoblast (OST) were isolated from homologous elements of distinct vertebrate clades and their transcriptomes were compiled. The analyses presented in this thesis provided insight into two unexplored areas in skeletal cell differentiation and evolution: the structure of GRNs driving skeletal cells in three phylogenetic clades and regulatory control of MAT.

Comparisons between mouse, chick, and gar data revealed the complexity of the organization and evolution of the GRNs underlying skeletal cells. Overall, the data presented here suggest that GRN structure and mechanisms of regulation in skeletal cells are generally conserved across tetrapods (i.e. mouse and chick), but important differences were observed in gar. While in mouse and chick genes in portions of the GRN active in IMM and OST had an antagonistic relationship (see chapter 3, Fig. 15; see chapter 6, Figs. 39&42), in gar, genes in IMM and OST showed an increase in positive correlations (i.e. positive interactions; see chapter 6, Fig. 50A). These results suggest that in earlier diverged vertebrates (i.e. gar) portions of the GRN active in IMM and OST shared the expression/ regulation of many genes, but then these portions of the GRN became more independent over time resulting in mutual-repressive relationship (see Fig. 56). Most importantly, while the portion of the GRN directing MAT in mouse and chick exhibits overlapping gene expression patterns with portions active in IMM and OST, gar MAT exhibits an antagonistic relationship with IMM and OST. Venn diagram, differential gene expression, and model-based clustering analyses in skeletal cells of mouse, chick, and gar also supported these conclusions. Transcriptomic analyses in mouse and chick revealed that MAT share many genes with both IMM and OST (see chapter 3, Figs. 11A, 12A, 13 and 19; see chapter 6, Figs. 37, 38, 40, 41 and 43), and IMM and OST are the least similar cell types. Strikingly, in gar, when comparing only two cell types, IMM and OST are the most similar cell types, whereas MAT shows more variation (Fig. 48, 49, and 53). For example, when gar skeletal cell transcriptomes were analyzed, model-based clustering analyses revealed that more clusters showed downregulation of gene expression levels in MAT compared to IMM and OST, while IMM and OST exhibited similar expression levels. This

result in gar was particularly unexpected since model-based clustering analyses in mouse and chick showed that IMM and OST generally show opposite gene expression patterns, and MAT is mostly a combination of gene expression levels between both cell types. In general, these findings suggest that in earlier diverged vertebrates there was a repression circuit activated in a neural crest specific GRN to make MAT, and later this GRN was modified and transcriptional repressors in this circuit were inhibited. Evolution in a repression circuit can occur when the coding sequence of a transcription factor and sub-sequent modification of the cis-regulatory sequences of its target genes occur, which can lead to topological changes in GRNs (Britton et al. 2020). This basic mechanism of change in GRN structure is usually caused by co-option, which is defined as a cis-regulatory re-deployment of a regulatory gene to create genomic novelty (Erwin 2019). Adding a non-amniote tetrapod into these comparisons, such as an aquatic frog (i.e. Xenopus; see Appendix I for more details), which is also a water-to-land transition animal, may fill a gap in the evolutionary puzzle of how GRN structure has been modified.

Despite all clades analyzed having homologous elements, in some cases it was not possible to isolate skeletal cells from the same elements for all species analyzed (see chapter 6, Fig. 36). Osteoblasts were isolated from the dentary bone (neural crest-derived; see chapter 5, Fig. 31), a homologous element in mouse, chick, and gar, facilitating evolutionary comparisons. Chondrocytes were isolated from skeletal elements of different embryonic origin: mesoderm and neural crest. In the case of mouse, IMM and MAT were obtained from the humerus (mesoderm-derived; see chapter 3, Fig. 9). For chick, IMM and MAT were obtained from the humerus (mesoderm-derived; see chapter 4, Fig. 27) and the ceratobranchial (neural crest-derived; see chapter 4, Fig. 28). Finally, gar IMM and MAT were obtained from the ceratohyal (neural crest-derived; see chapter 6, Fig. 45), a homologous element to the chick ceratobranchial. Mouse also has a ceratohyal/ceratobranchial homolog but capturing cells from this element would be incredibly challenging due to its small size. Transcriptomic analyses in chick such as, Venn diagram and gene ontology analyses revealed that different chondrocyte lineages have highly conserved transcriptomes (see chapter 4, Fig. 29A), but embryonic origin might influence GRN structure. However GCN analyses performed in chick skeletal cells, suggest that perhaps combining neural crest- and mesoderm-derived data might be a good approach to make evolutionary comparisons. When chick mesoderm- and NC-derived chondrocytes were combined in GCN analyses, network structure remains conserved as when only mesoderm-derived chondrocytes are included (Figs. 39B&42). Since chick was the only species

where chondrocytes of distinct lineages were analyzed, additional mesoderm- and neural-crest comparisons within another clade should be performed test whether the degree of similarity is still high between skeletal cells derived from the two embryonic lineages. If the degree of conservation between mesoderm- and NC-derived chondrocyte lineages in other vertebrates is low, then evolutionary comparisons perhaps should be limited to skeletal cell types isolated from homologous elements or at least derived from the same lineage.

As discussed above, osteoblasts were only isolated from a neural crest-derived bone (i.e. dentary), so osteoblasts isolated from perichondral bone derived from the mesoderm (i.e. humerus; Fig. 57) should also be included into these mesoderm and neural crest comparisons to further test if the portion of the GRN driving OST differentiation is also conserved throughout the body and regardless of embryonic origin. The technical reason OST were isolated from the dentary bone was to avoid potential cross-contamination with mature chondrocytes (MAT, Fig. 57). With the MMI CellCut laser equipment used it was incredibly challenging to cut through the bone samples, so cartilage would have been inevitably captured as well. However Dr. Brian Eames and a current colleague from the lab, Joseph Atake, successfully captured hard mineralized tissue using the Leica LMD 7000 laser equipment, so perhaps future captures of mineralized tissue should be performed using this LCM machine or a laser with a similar scope.

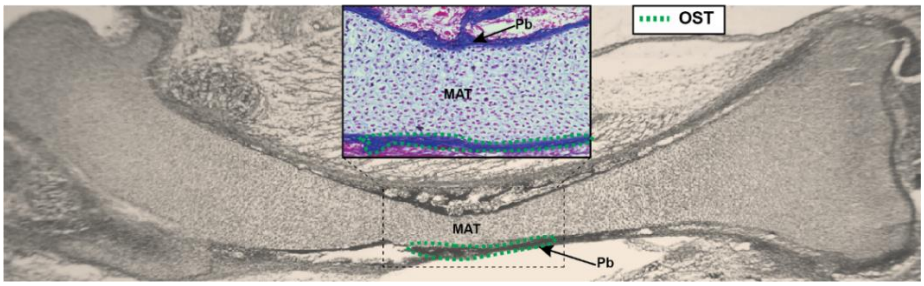



**Figure. 57. Osteoblasts could also be captured from perichondral bone using a laser with a higher scope.** Chick HH36 humerus. There is a spatial distribution between mature chondrocytes (MAT) and perichondral bone (Pb). If a laser with a higher power that allows cutting through the bone is used, osteoblasts (OST) could be captured from perichondral bone and cross-contamination with MAT could be avoided. Capturing OST from perichondral bone would also facilitate comparisons of OST derived from different embryonic origins (i.e. mesoderm vs neural crest).

An important limitation of this study is that both biological and technical factors might impact the results. For instance, there are still many bioinformatic challenges to overcome regarding data analysis. Poor genome annotation in some species is a quite common issue. A well-annotated

mouse genome sequence is available, but annotation problems were experienced with chick and gar genomes, making it difficult to perform some evolutionary comparisons. Thus, data from multiple species needed to be converted to the most well annotated species (i.e. mouse genome). To do this, reads obtained from RNA-seq were aligned to their own species' genome, and then genes in chick and gar were translated to mouse gene names based on sequence similarity data. However during this conversion there could be a loss of genes which can make species comparison difficult because many putative genes may not have an actual ortholog or it is not annotated correctly. Another common issue when dealing with multiple species is that since they are not sequenced at the exact same time and it can be challenging to know whether the effect of a group/sample is due to technical factors when processing biological samples (i.e. RNA extraction, amplification, library preparation, etc.) or technical factors due to RNA-seq. Since many normalization procedures are used during data analysis and each of these have limitations, the amount of expression data retained can be impacted.

There were also some technical challenges in the techniques used to isolate and amplify RNA. The samples submitted for RNA-seq were flash frozen unfixed tissues, thus, the RNA extracted from these tissues can be highly compromised if laser-capture microdissection is not performed within a 30 min interval. Another technical issue was that for some samples one round amplification was enough to obtain the desired RNA concentration, but for other samples two rounds amplification were needed. Specifically, for gar two rounds of amplification were required to get enough concentration, whereas for mouse and chick samples only one round amplification was needed. The MessageAmp kit protocol suggests to include an extra round of amplification in order to increase mRNA concentration, and it has also been performed by others (Choesmel et al. 2004), but since the product of the first round amplification is used as a template for the second round it was concerning that the extra round amplification could potentially add bias to the mRNAs amplified. However in a previous study Feldman *et al.* demonstrated that the fidelity in microarrays was preserved using one or two rounds of amplification, and they concluded that the quality of the data was even improved after two rounds of mRNA amplification (Feldman et al. 2002). Importantly, comparison between the same chick OST samples with one vs two round amplification from our data showed that adding an additional round of amplification does not affect the data and the results are comparable to those obtained with only a first-round amplification (Fig. 58), but this analysis was not performed in all cell types or species.

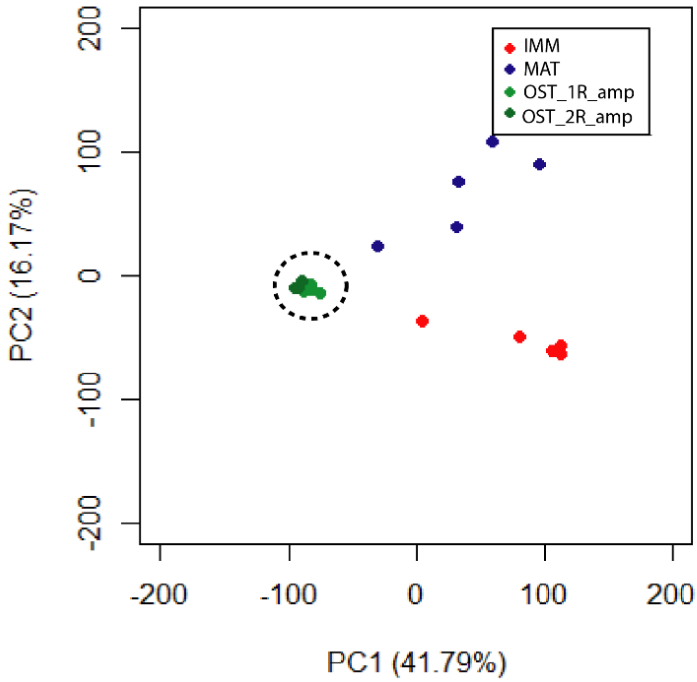



**Figure 58. Principal component analysis demonstrate that all chick OST samples amplified with one or two rounds are clustered together and are very consistent.** Chick OST samples amplified once (1R) are highlighted in light green, whereas OST samples amplified with 2 rounds (2R) are highlighted in dark green. Chick IMM samples are highlighted in red, and MAT are highlighted in blue. Both chick IMM and MAT were only amplified once (PCA Credit: Katie Ovens).

An important biological issue was the high variation between replicates. The epiphyseal growth plate contains cells in different stages of differentiation (Lui et al. 2014). Particularly in chick, the chondrocyte columns were disorganized since cells in different stages of differentiation were located next to each other in the distinct zones (Leach Jr and Gay 1987; Pines and Hurwitz 1991; see chapter 4, Figs. 27&28). Due to the lack of organization of chondrocyte columns within the avian growth plate it is likely that more variation was introduced when capturing chondrocytes. To isolate more homogeneous populations of cells, pre-hypertrophic and hypertrophic chondrocytes could be treated separately. For example, in all species the mature cartilage region was delimited to perichondral bone, since there is a spatial relationship between these two tissues (see Fig. 57). Captured MAT included some pre-hypertrophic and hypertrophic chondrocytes, but in chick the pre-hypertrophic region extended closer to the epiphyses of the bones, so some important maturation genes such as *IHH* that were expressed highly in this region that was not captured were lost during normalization due to the few counts. Another potential limitation was sample population

heterogeneity. C57BL/6J mice are an inbred strain raised in a controlled laboratory environment which shows little phenotypic variation. On the other hand, White Leghorn chickens were raised in farms and showed higher phenotypic variation compared to mice, since 20-22 hens are housed with 2 roosters (which are from the same hatch). In poultry genetics, a male parent line and a female parent line are used to produce commercial laying hens (Bell and WeaverJr 2002), and the male chicks of this commercial generation are generally discarded since they have no use in a commercial setting. The male chicks used in the present study, however, were not discarded, and instead they were crossed with Lohmann Selected Leghorn (LSL-Lite) hens, so white leghorn males used in this study did not longer have the same genetics as those used in commercial production. In contrast to mouse and chick breeding strategies, for gar there was no control of gene pool since they were caught from the wild and fertilized *in vitro*. Thus, this species is expected to show even more genetic heterogeneity compared to mouse or chick which may result in more variation between biological replicates.

Although, the overall chondrogenesis and osteogenesis pattern is conserved, the timing of ossification events can vary from species to species. Skeletal heterochronies, which are defined as a difference in the timing of skeletal development (McNamara 2012), were observed when comparing species at apparently similar stages of skeletal development. For example, mouse, chick, and gar epiphyseal growth plates are histologically similar, but terminal chondrocyte differentiation seems to occur at different rates. Our results in chick show that terminal differentiation of chondrocytes occurs at a faster rate compared to mouse characterized by a reduction in *COL10A1* expression in hypertrophic chondrocytes (Fig. 44G). A down-regulation in *Col10a1* expression in terminally differentiated chondrocytes has been shown in mouse and chick (Provot et al. 2006; Yang et al. 2014), but studies in gar show that expression of *Col10a1* is still high at later stages of chondrocyte maturation (i.e. 28dpf; Eames et al. 2012). Furthermore, while in tetrapods cartilage hypertrophy and perichondral bone mineralization occur simultaneously (see chapter 3, Fig. 9; see chapter 4, Fig. 27&28), in gar, these processes do not seem to occur at the same time (Fig. 45). The results presented here showed that in gar there is a relatively late perichondral ossification in the ceratohyal which was confirmed by a lack of alizarin red staining in this element, so perhaps gar was not at an equivalent developmental stage compared to mouse and chick. Indeed, gene expression in gar MAT showed important differences when compared to tetrapods. In contrast to mouse and chick, gar MAT had several downregulated genes compared to IMM and OST (Fig. 48D, Fig. 49, Fig. 53). Gene

ontology analysis showed that most of these downregulated genes in MAT, and upregulated in IMM and OST, are involved mainly in cell cycle processes and chondrocyte proliferation. IMM isolated from all three species included both proliferative and resting chondrocytes, but in gar and chick IMM data were mostly obtained from proliferative zones at epiphysis (see chapter 4, Fig. 27F,G and Fig. 28E,F; see chapter 6, Fig. 45F,G), while mouse IMM was mostly obtained from resting zones between diaphysis and epiphysis (see chapter 3, Fig. 9G,H). This could partially explain the difference in gene expression observed in gar MAT (i.e. more downregulated genes compared to IMM and OST), but chick chondrocytes isolated from the exact same zones in the growth plate did not follow this pattern of expression in MAT. Classic MAT and OST markers such as, *Mmp13, Spp1, Satb2, Mmp9, Mmp13, Pthlh,* and *Sparc* were highly expressed in both cell types in mouse and chick, but these genes showed little to no expression in gar MAT. These results further support that gar was at a relatively younger stage than chick and mouse MAT. Indeed, histological analysis in gar ceratohyal revealed that chondrocyte hypertrophy was not so evident in the 13dpf ceratohyal (Fig. 45D). However expression of the hallmark mature chondrocyte gene *col10a1* was already high in the 13dpf ceratohyal and the expression domain was already expanding towards the epiphyses (Fig. 46C). *In situ* hybridization in the *X.tropicalis* NF57 humerus revealed that *col10a1* is not expressed in MAT (credit: Jason Nguyen, data not shown), even when this tissue is overtly hypertrophic (see Appendix I, Fig. 61E,I). This lack of expression of *col10a1* in *X.tropicalis* MAT is also consistent with RNAseq data obtained from the NF57 humerus (Credit: Jason Nguyen, data not shown). These results in gar and Xenopus suggest *col10a1* is not a hypertrophy marker, at least in some species, and should not be considered alone to identify MAT, since there are other factors potentially driving cartilage hypertrophy. The mechanisms driving chondrocyte hypertrophy need to be further explored, but previous studies suggest that an increase in volume due to fluid accumulation and Insulin growth factor (IGF) signaling are major triggers of hypertrophy (see chapter 6 discussion for more details; Buckwalter et al. 1986; Farnum et al. 2002; Cooper et al. 2013; Wang et al. 1999; Murray et al. 2013; Uchimura et al. 2017). Since only one particular study has analyzed skeletogenesis in gar (Eames et al. 2012), a more detailed histological and molecular study analyzing expression patterns of *Igf1* and *Igf2* as well as, classic MAT and OST genes at different developmental stages during cartilage maturation in gar would provide more insight into how chondrocyte hypertrophy might be regulated in this species.

Finally, some regulatory interactions of the GRNs presented here were validated using published molecular genetic experiments and cis-regulatory analyses, but different biological and technical parameters were employed in each of these published experiments so *in vivo* functional studies must be performed to validate and expand this initial GRN. For instance, mis-expression of *Sox9* and *Runx2* orthologs in chondrogenic and osteoblastic murine cell lines, such as ATDC5 and MC3T3-E1, coupled with integrative epigenome analysis (i.e. ChIP-seq and ATAC-seq; Robertson et al. 2007; Buenrostro et al. 2015) could be performed in order to validate if the function and regulatory targets of these transcription factors are conserved during skeletal differentiation across animal clades. Moreover, cis-regulatory changes in enhancers of *Sox9* and *Runx2* targets and mechanism of regulation of these transcription factors should also be analyzed across vertebrate clades since major evolutionary changes occur when regulatory DNA sequences or regulatory function of crucial transcription factors are modified (Indjeian et al. 2016; Yamashita et al. 2019). For example, an evolutionary change in transcriptional regulation between mouse and chick occur during heart development. *Nkx2.5* autoregulation is important for cardiac differentiation in both mouse and chick, but the mechanism of autoregulation is different for each species (Clark et al. 2013). In mouse, *Nkx2.5* first binds to *Mef2c* and then both genes bind *Nkx2.5* enhancer, whereas in chick *Nlx2.5* binds directly to its enhancer to maintain its own expression (Clark et al. 2013). Similar changes in the regulatory function of *Sox9* and *Runx2* might occur during cartilage and bone differentiation across vertebrate clades, and these changes could ultimately modify the underlying GRN.

To further examine GRN conservation across animal clades, transgenic approaches can be performed to test whether hallmark chondrocyte and osteoblast transcription factors, such as *Sox9/Runx2*, as well as novel transcription factors identified by RNA-seq enriched in skeletal cells of one species can bind and activate the expression of genes in skeletal cells of other species. For example, previous work has provided evidence that the GRN driving chondrogenesis might be conserved in vertebrates and invertebrates, since *SoxE* from a hard-shell invertebrate was able to activate the expression of the human *COL2A1*, revealing deep conservation in the chondrocyte GRN (Tarazona et al. 2016). In summary, comparative transcriptomics coupled with cis-regulatory analysis, as well as gain- and loss-of-function studies in skeletal cells of distinct species will contribute in expanding and verifying connections in this initial GRN, and will increase our understanding on the evolutionary origin of cartilage and bone.

## Appendix I:

### Histological characterization of frog skeletal tissues

Gomez-Picos P, Nguyen JKB, Liu Y, and Eames BF

## Appendix I. Histological characterization of frog skeletal tissues

### I. INTRODUCTION

Skeletogenesis in *Xenopus tropicalis* might occur in a similar manner compared to other tetrapods, such as mouse and chick, but only a few studies have analyzed skeletal development in this anuran (Miura et al. 2008; Espinoza et al. 2010; Enault et al. 2015). Despite the lack of studies, previous work suggests that skeletal development in *X. tropicalis* exhibits some unique features (Miura et al. 2008). For instance, the frog limb skeleton undergoes endochondral ossification, but this process is characterized by poor growth plates that contain some hypertrophic chondrocytes close to the marrow cavity that are not very enlarged and rarely mineralize their matrix, as well as little trabecular bone (Miura et al. 2008). These differences in Xenopus skeletal development can be in part to the fact that this species undergo metamorphosis, which involves a dramatic change in shape and size, and it is under hormonal control (Fig. 59; Buchholz 2015).

Likewise in mammals and birds, in Xenopus, *sox9* and *runx2*, are involved in chondrogenesis and osteogenesis, respectively (Miura et al. 2008). However one key difference between Xenopus and other animals is that expression *runx2* in mature chondrocytes of this amphibian is weak, whereas expression of this transcription factor is high in mature chondrocytes of mouse and chick (Miura et al. 2008). Furthermore, expression of *runx2 in Xenopus* is required during cartilage differentiation of the hyobranchial skeleton, so this transcription factor does not have an exclusive function in cartilage hypetrophy and bone differentiation in the frog as in amniote tetrapods (Kerney et al. 2007).

Since little is known about the skeletogenic process in *X. tropicalis*, and most of the work in Xenopus has been done in *X. laevis*, we aimed to study skeletal development in this extant anuran. In order to gain insight into the development of *X.tropicalis* skeletal tissues and identify the exact timing of events of cartilage differentiation, hypertrophy and bone mineralization, tadpoles from Nieuwkoop and Faber stage NF41 to NF66 were collected. We hypothesize that cartilage and bone differentiation in *Xenopus tropicalis* is conserved compared to other tetrapods. To test this hypothesis, we have employed several molecular and histological techniques during the past few years in our lab. During my PhD, I had the chance to train my first student, Yiwen Liu, and together

we started the *X. tropicalis* project. I taught her sample collection procedures, histological techniques (whole mount and section histology), cryosectioning and helped her setting up the crossing system. Then, I had the fortune to supervise and train Jason Nguyen during his Masters project, who greatly improved the crossing system that we had previously established. He performed laser capture microdissection in *X.tropicalis* skeletal tissues. His samples have been already submitted for RNA-seq, and we are planning to include these data into our OST evolution paper.

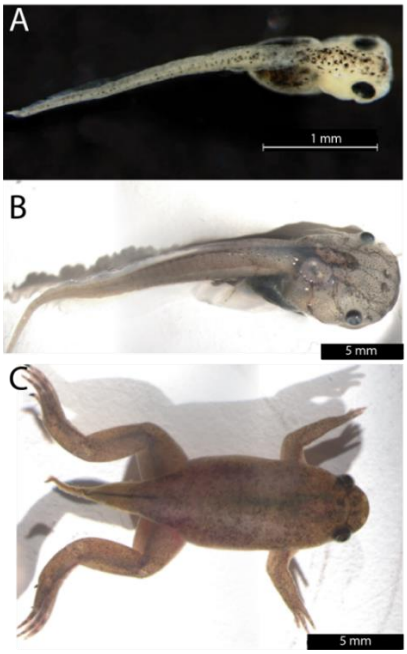



**Figure 59. Metamorphosis in X.tropicalis tadpoles.** A) Premetamorphosis in a NF42 tadpole. (B) Prometamorphosis in a NF56 tadpole. (C) Metamorphosis climax in a NF64 tadpole.

## II. METHODS

During pre-priming, Xenopus *tropicalis* males and females were injected with 10 units (0.1 mL 100U/mL) of human chorionic gonadotropin (hCG) approximately 20 hours before priming. Next day, both males and females were primed by injecting 200 units (0.2 mL 1000U/mL) of hCG. After injections, each pair of males and females were placed in plastic containers filled with water obtained from the frog tanks, which were kept in a water bath at a constant temperature of 26-27ºC (Fig. 60A). The duration of amplexus was 5-6 hours, and eggs were laid within a few hours (Fig.

60B). Adult frogs were then transferred back into the tanks, leaving the eggs in the mating plastic containers with fresh water obtained from the frog tanks. Embryos hatched within 24 hours of amplexus. The hatched tadpoles were then transferred to Modified Barth's saline (1X MBS; stock solution 10X MBS, 880 mM NaCl, 10 mM KCl, 10 mM $MgSO_4$, and 25 mM $NaHCO_3$). Developmental stages of *X. tropicalis* were determined based on Nieuwkoop and Faber staging system for *X. laevis* (Nieuwkoop and Faber 1956). Once tadpoles had reached NF 41, they were transferred into a tadpole nursery system with recycling water and a constant temperature of 26ºC (Fig. 60C,D). Tadpoles were initially fed with Sera Micron growth food, and then at around 14 days of age they were fed with HBH Pisces Pros Frog and Tadpole Bites to provide extra nutrients for growth. Tanks were cleaned twice a day to avoid yeast and bacteria growth. To monitor growth and stage of development, tadpoles were placed in a dish containing 0.02% tricaine under a dissecting microscope. Once they had reached the desired stage, tadpoles were euthanized using 0.2% tricaine and immediately transferred to 4% PFA. They were kept overnight at 4ºC in this solution in a rocker. Tadpoles were collected at stages NF41-66.

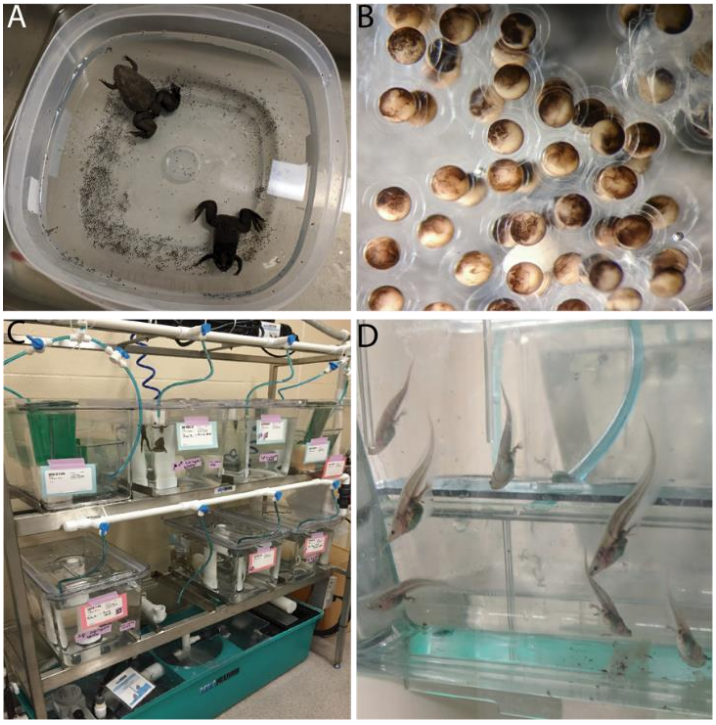



**Figure 60. Xenopus tropicalis mating and nursery system.** (A) Frogs were crossed by natural mating and underwent amplexus for 5-6 hours. (B) Eggs were laid within a few hours after the beginning of amplexus. (C) Nursery system tanks with water recirculation and a constant temperature of 26ºC were placed, and tadpoles were transferred at stage NF41. (D) Tadpoles in prometamorphorphic stage swimming inside the nursery tank.

## III. RESULTS

The skeletogenic process in X. tropicalis exhibits some conserved features compared to other animals, but some important differences were also noted. During endochondral ossification in the humerus, chondrocytes proliferate, undergo maturation, and then perichondral bone is formed in the mid-diaphyseal region (Fig. 61A,D,E,H,I). Alcian blue was evident at stage NF55 (and earlier timepoints), but Alizarin red was not obvious until stage 57, where perichondral bone was evident and early mineralization in the jaw (Fig. 61A,B,D). In contrast, endochondral ossification in the ceratohyal did exhibit some important differences compared to other animals. At NF57 the ceratohyal is already overtly hypertrophic, but it did not show evidence of mineralization (Fig. 61C, Fig. 62A). Subsequent staining of later stages confirmed that the ceratohyal in *X. tropicalis* remains cartilaginous (Fig 62B). Safranin O and Trichrome staining were perfomed in both frozen and paraffin tissue sections of the humerus, ceratohyal and dentary/ angulosplenial in order to identify immature chondrocytes, mature chondrocytes, and osteoblasts (Fig. 61E,G-I). While Safranin O staining revealed mature chondrocytes at NF57 in the humerus and ceratohyal (Fig. 61E,H), trichrome staining revealed the presence of bone in the lower jaw (Fig. 61F), as well as perichondral bone in the humerus (data not shown).

Stage NF57 was selected to perform laser capture microdissection (LCM) due to the presence of immature chondrocytes and mature chondrocytes (surrounded by perichondral bone) in the humerus (Fig. 61). Jason Nguyen collected RNA from three samples for each cell type, and then they were submitted for RNA-seq. We are currently working on analyzing these data to make an evolutionary comparison with mouse, chick, and gar in order to provide more insight into the evolution of the osteoblast, a vertebrate innovation. In the future, we also plan to collect cells from the ceratohyal and submit RNA for sequencing, but in this case, it will be impossible to collect immature and mature chondrocytes from the same stage (NF57) as the humerus and angulosplenial. Unlike the humerus, in Xenopus ceratohyal, chondrocytes mature rapidly and early. By NF57 the ceratohyal is overtly hypertrophic, making it impossible to collect immature chondrocytes. Jason performed extensive histological and molecular analysis of earlier timepoints and concluded that immature chondrocytes could be collected at NF41.

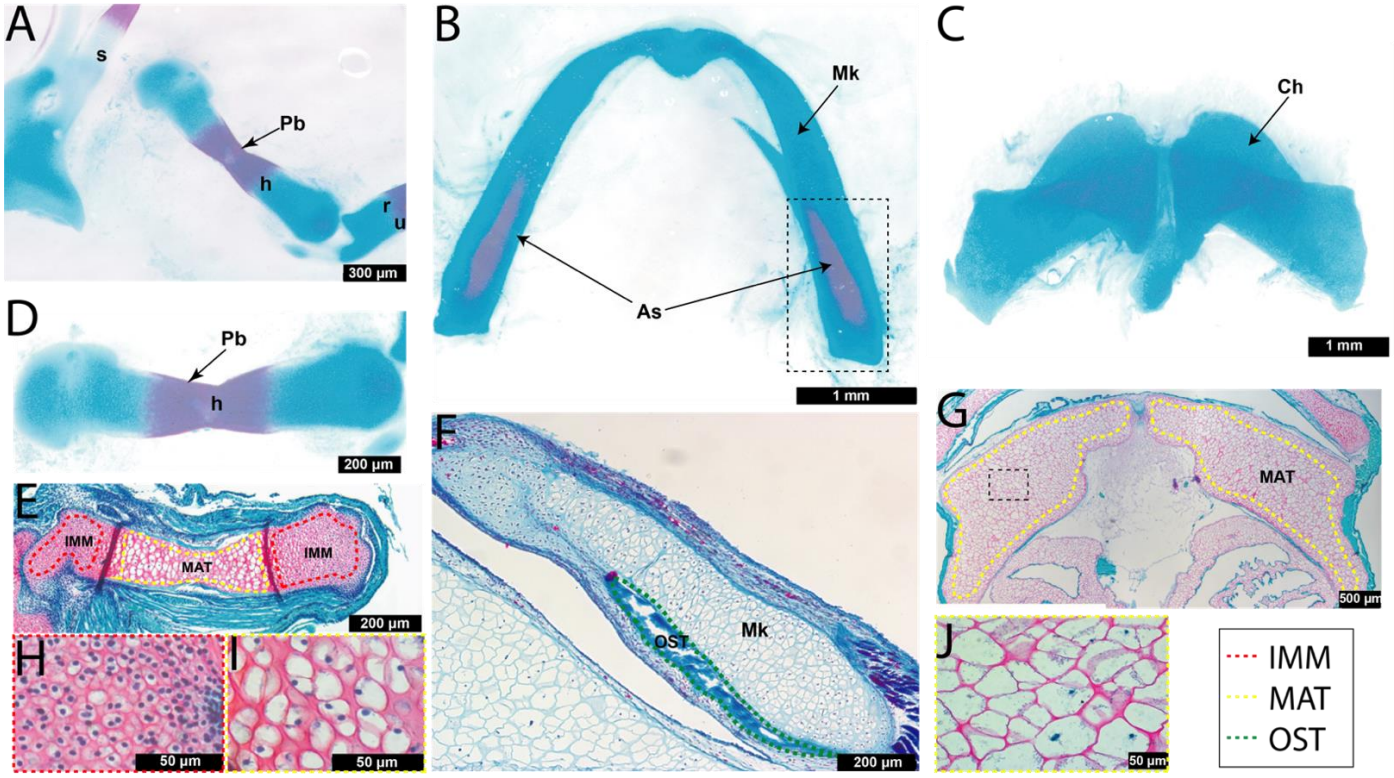



**Figure 61. Characterization of X.tropicalis skeletal elements at stage NF57.** (A-D) Whole-mount Alcian blue and Alizarin red staining identified cartilage and perichondral bone in endochondral bones of the frog forelimb, and intramembranous bone in the mandible. (E) Safranin O-stained section of NF57 humerus highlighted the mature cartilage region and mature cartilage of the ceratohyal (yellow dotted outline). (F) Trichrome-stained section of NF57 mandible showed Aniline blue staining of bone matrix (OST) in the dentary/angulosplenial, separated from Meckel's cartilage (Mk). High-magnification images of immature (H) and mature chondrocytes (I) of the humerus, and mature cartilage (J) of the ceratohyal. Abbreviations: IMM, immature chondrocytes; MAT, mature chondrocytes; OST, osteoblasts; Mk, Meckel's cartilage; Pb, perichondral bone; h, humerus; s, scapula; r, radius; u, ulna; ch, ceratohyal (Credit: Samples obtained from Jason Nguyen).

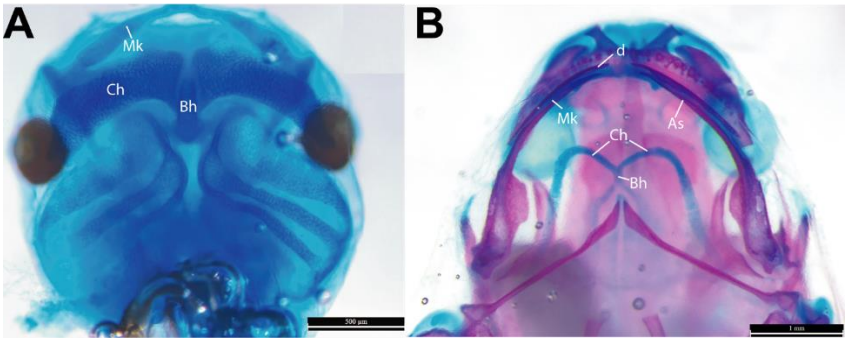



**Figure 62. In X.tropicalis, adult ceratohyal emerges from the resorption of the tadpole ceratohyal.** (A) Ceratohyal at stage NF48 shows no evidence of mineralization. (B) At the end of metamorphosis (NF66) the ceratohyal undergoes a drastic change in shape and size, emerging from deep within the tadpole ceratohyal. Abbreviations: Ch, ceratohyal; Bh, basihyal; Mk, Meckel's cartilage; As, angulosplenial; d, dentary (Credit: Jason Nguyen).

## IV. DISCUSSION

Amphibians provide great insight into the skeletogenic process since they transitioned from water to land, thus, they were probably the first tetrapods that existed (Miura et al. 2008). By comparing *X.tropicalis* with other tetrapods, such as mouse and chick, key events during cartilage and bone formation were largely similar, as reported by others (Hall 2005). Chondrocytes proliferated, and cells in the diaphysis became hypertrophic. In the humerus, perichondral bone was evident around the same time as hypertrophy, and vascular invasion occurred later. Furthermore, endochondral ossification in the humerus of mouse, chick, and frog began in the diaphysis and extended to the epiphysis. However when comparing *X.tropicalis* ceratohyal to ceratohyal/ceratobranchial of gar and chick, a few differences were noted. Unlike chick and gar, *X.tropicalis* ceratohyal never ossifies, but hypertrophy in this element occurs in earlier developmental stages in the frog (around NF42) compared to equivalent stages in chick and gar. Similar to mouse and chick, the master regulators *sox9* and *runx2*, have important roles during cartilage and bone differentiation in Xenopus, but the mechanisms of regulation of these transcription factors and many other molecules known to be involved in skeletal differentiation in anurans are still unclear. High-throughput sequencing technologies, such as RNAseq can be used to identify skeletal cell transcriptomes in anurans, in order to reveal conserved and novel genes that can be used to construct skeletal cell GRNs.

## Appendix II

**Book chapter; On the Evolution of Skeletal Cells before and after Neural Crest. In Eames BF, Meulemans Medeiros D, and Adameyko I (Eds.), Evolving Neural Crest Cells (pp. 185-218). Boca Raton: CRC Press, https://doi.org/10.1201/b22096**

Eames BF, Gomez-Picos P, and Jandzik D (2020)

## Appendix II. Book chapter: On the evolution of skeletal cells before and after neural crest

B. Frank Eames[1], Patsy Gómez-Picos[1], and David Jandzik[2,3]

[1]Anatomy, Physiology, and Pharmacology, University of Saskatchewan, Saskatoon, SK, S7N 5E5, Canada

[2]Ecology and Evolutionary Biology, University of Colorado, Boulder, CO, 80309, USA

[3]Zoology, Comenius University, Bratislava, 84215, Slovakia

**Abstract**

Along with neural crest, the skeletal tissues of cartilage and bone were considered defining features (synapomorphies) of vertebrates. Therefore, the appearance of these skeletal tissues during evolution often was conflated with the appearance of neural crest in the ancestor to all vertebrates. Evidence from the past few decades, however has demonstrated conclusively that cartilage (and a related tissue, the notochord) evolved prior to neural crest, whereas bone probably evolved after neural crest. In this chapter, we use the fossil record, comparative anatomy, and molecular genetics to summarize current understanding of how skeletal cells (and their underlying gene regulatory networks, or GRNs) evolved relative to the appearance of neural crest cells. Neural crest cells likely co-opted a GRN for chondrocyte differentiation that was employed previously by other cell lineages in other deuterostomes and even, perhaps, protostomes. Vertebrates merely modified this GRN in neural crest cells to produce multiple chondrocyte types. The evolution of bone in vertebrates seems to be linked with the appearance of other mineralized skeletal tissues, such as enamel and dentin. Current data suggest that two very similar tissues (bone and dentin) appeared in a part of the skeleton that is derived from neural crest, in which case the underlying GRNs actually evolved within neural crest cells. Comparative GRN analyses of skeletal tissues in various embryonic cell lineages of different taxonomic groups will illuminate how neural crest cells impacted the evolution of cartilage and bone within vertebrates.

## I. One of these three is not like the others: Neural crest, cartilage, or bone as a vertebrate synapomorphy?

Our understanding of how skeletal tissues evolved relative to neural crest cells has changed a lot in the past few decades. Before then, both cartilage and bone were considered synapomorphies, or defining features, of vertebrates (Gans and Northcutt 1983). From a comparative anatomist's point of view, the distribution of cartilage among extant vertebrates spanned "down" the phylogenetic tree from bony fish (osteichthyans) to cartilaginous fish (chondrichthyans) to even jawless fish (agnathan), while bone was observed only in osteichthyans (Fig. 54). These observations led to a traditional phylogenetic model whereby cartilage evolved within the ancestors to all vertebrates, and bone appeared later in the ancestor to osteichthyans. Discovery of a sufficient fossil record, however re-wrote the false story that bone evolved in the ancestral osteichthyans. Not only did primitive chondrichthyans have abundant bone, but primitive agnathans did also (Janvier 1996b; Coates et al. 1998; Coates et al. 2018). Therefore, the revised model was that cartilage and bone evolved in the ancestral vertebrates.

Embryologists recently added neural crest cells to the list of vertebrate synapomorphies. These migratory cells were identified over 150 years ago (His 1868), but the demonstration of neural crest in extant agnathans in the past twenty years solidified the understanding that neural crest cells are a vertebrate synapomorphy (Horigome et al. 1999; McCauley and Bronner-Fraser 2003; Ota et al. 2007). The contributions of neural crest to skeletal tissues of the vertebrate head were demonstrated through cell labelling and transplant experiments (Noden 1988; Couly et al. 1992; Nikitina et al. 2009). Thus, it was tempting to speculate that cartilage and bone evolved along with the appearance of neural crest in the ancestral vertebrates, but recent data refute this hypothesis.

In this chapter, we clarify the evolutionary relationship between the appearance of neural crest and the ability to form cartilage and bone. Recent studies confirm that the ability to make cartilage (and a related tissue, the notochord) preceded the appearance of neural crest (Cole and Hall 2004b; Rychel et al. 2006; Annona et al. 2015). Neural crest might have increased the number of cartilage types in vertebrates. Interestingly, the ability to make bone might have evolved slightly after the appearance of neural crest, if cladistics supports the notion that cyclostomes (such as lamprey and hagfish, whose ancestors have never shown evidence of bone formation) are one of the

earliest clades of agnathans (Janvier 1996b). Since bone appears to have formed first in a region of the skeleton (exoskeleton) that was demonstrated recently to be derived from neural crest, neural crest might have played a key role in the evolution of bone (Gillis et al. 2017).

Given that cartilage was recently confirmed to precede the appearance of neural crest, which, in turn, preceded bone during chordate evolution (Cole and Hall 2004b; Rychel et al. 2006; Annona et al. 2015), this chapter also presents likely molecular genetic scenarios by which neural crest co-opted or otherwise evolved the ability to form skeletal tissues. Gene regulatory networks (GRNs) are discrete sets of genes that interact with environmental cues to impart biological traits, such as the formation of cartilage and bone (Levine and Davidson 2005; Davidson and Levine 2008; Gomez-Picos and Eames 2015). Here, we argue that neural crest first co-opted a GRN that was used to make cartilage in other cell lineages in the ancestors to vertebrates, and then neural crest played a significant role in evolving various vertebrate-specific cartilage types through modifications to this ancestral cartilage GRN. The appearance of a bone GRN within neural crest turns out to be more complicated to model with current data. The fossil record suggests that bone evolved from dermal armor in primitive jawless vertebrates (Smith and Hall 1990b; Janvier 1996b). Since there is little molecular genetic data to support this, however and the current fossil record might not be representative of all early mineralized tissues, we present three hypotheses for the evolution of a GRN underlying bone formation.

In the first hypothesis, neural crest cells co-opted a mineralization GRN that was present in the last common ancestor of all deuterostomes. In the latter two hypotheses, a GRN driving mineralization of skeletal tissues evolved within neural crest. In the second hypothesis, the bone GRN was co-opted from a mineralization program underlying formation of related skeletal tissues, enamel and dentin, in the ancestor to non-cyclostome vertebrates. All three of these tissues are vertebrate novelties. While enamel is made by ectodermally-derived cells, both bone and dentin are generated from neural crest cells (Chai et al. 2000; Gillis et al. 2017), so GRN co-option between bone and dentin would not involve different germ layers. In the third hypothesis, the first osteoblast co-opted a GRN that was used to mineralize cartilage in the ancestor to non-cyclostome vertebrates. Intelligent application of modern technologies that characterize GRN structure can be used to test the molecular mechanism through which neural crest evolved the ability to make cartilage and bone.

## II. The players in this evolutionary game: Cartilage, notochord, bone, and dentin/enamel

One of the challenges in re-building the evolution of skeletogenesis in chordates is that most data on skeletal tissues derive from studies in mammals, who at times might represent a relatively recent vertebrate lineage's adaptations to selective pressures on the skeleton. However after summarizing these data, we provide some perspective to minimize this potential limitation for a robust understanding of skeletal tissue evolution.

### *IIA. Cartilage*

Cartilage typically functions as a flexible structural support, and the molecules that comprise cartilage extracellular matrix (ECM) impart these roles. Collagen type 2 (Col2) fibers are the most abundant protein in the ECM of cartilage, providing this tissue with the structural stability to withstand some tension while maintaining flexibility (Gray and Williams 1989; Aumailley and Gayraud 1998). These loosely-wound Col2 fibers bind Aniline blue of Milligan's Trichrome histological protocol, giving cartilage a light blue staining pattern (Eames et al. 2007). Other collagens, such as Col9 and Col11, are also abundant in cartilage (Eames et al. 2003). Sulfated proteoglycans (PGs) in the ECM lend cartilage the property of compressive resistance, due to the massive amounts of water absorbed by sulfated PGs (Gray and Williams 1989; Aumailley and Gayraud 1998). In fact, the swelling of hydrated PGs accounts for the majority of the volume of cartilage ECM (Ham and Cormack 1987). Sulfated PGs bind to Alcian blue and Safranin O, staining cartilage blue and red in these respective histological protocols (Eames et al. 2007). The most abundant PG expressed in cartilage is Aggrecan (Acan), which is a chondroitin sulfate PG, due to the repeating disaccharide glucuronic acid and N-acetylgalactosamine extending from the Acan core protein (Watanabe et al. 1998). Many other proteins present in cartilage ECM at lower levels, along with hyaluronic acid repeating disaccharide, supplement Col2 and sulfated PGs in providing cartilage with the mechanical properties that determine its function (Ham and Cormack 1987; Aumailley and Gayraud 1998).

Establishment and maintenance of cartilage ECM depends upon the action of transcription factors that regulate expression of genes encoding these molecules. Sox9 is the main transcription factor driving cartilage formation. Loss of Sox9 function abrogates cartilage differentiation, while gain of Sox9 function promotes ectopic cartilage(Bi et al. 1999a; Eames et al. 2004). Other transcription factors, such as Sox5, Sox6, C/EBP's, and FoxO's, are involved in the expression of cartilage genes, but none of these has as central a role as Sox9 in chondrocyte differentiation (Smits et al. 2001b; Okuma et al. 2015; Kurakazu et al. 2019). Sox9 binds to regulatory elements and promotes the expression of many genes that are highly expressed in cartilage, including *Col2a1, Col9a1, Col11a2,* and *Acan* (Ng et al. 1997; Bridgewater et al. 2003; Zhang et al. 2003; Hu et al. 2012). A Sox9 gene regulatory network (GRN) likely dictates cartilage formation (Cole 2011; Gomez-Picos and Eames 2015).

The two previous paragraphs focus on hyaline cartilage, which in addition to fibrocartilage and elastic cartilage, are the three main cartilage types in mammals. All cartilage types share the basic molecular features outlined for hyaline cartilage, but they have significant differences related to their specific functions. Hyaline cartilage occurs predominantly in articulating surfaces of skeletal joints, absorbing compressive forces between the bones (Gray and Williams 1989; Naumann et al. 2002). Fibrocartilage has much more tensional resistance than hyaline cartilage, so the collagen fibers are increased, while the PGs are decreased (Ham and Cormack 1987; Naumann et al. 2002). In addition to an increase in fiber quantity, fibrocartilage contains large amounts of Col1, in addition to Col2 (Wachsmuth et al. 2006). Fibrocartilage is found predominantly in the intervertebral discs of the spine and joint menisci (Gray and Williams 1989). Some parts of the anatomy, such as the external ear, epiglottis, Eustachian tube, and nose, have elastic cartilage, in which abundant elastin fibers impart more flexibility to the Col2-positive tissue (Ham and Cormack 1987; Gray and Williams 1989; Naumann et al. 2002). While the mammalian classification system recognizes these three cartilage types, many cartilage tissues exhibit intermediate characteristics among the three types. Therefore, and this becomes very important when expanding an understanding of cartilage types beyond mammals, each skeletal tissue should be considered as merely one instance along a spectrum of possibilities.

Even within the hyaline cartilage type, two versions commonly occur in mammals: immature and mature cartilage. After secretion of the hyaline cartilage matrix described above, some developing chondrocytes (usually in the middle of a developing skeletal element) progress through

a series of differentiation steps collectively termed maturation (de Crombrugghe et al. 2000; Eames et al. 2003b). During cartilage maturation, many chondrocytes undergo hypertrophy and mineralize their matrix, forming mineralized cartilage. During endochondral ossification, some mature chondrocytes also degrade their matrix and die whereas others can even trans-differentiate into osteoblasts as their ultimate fate (Hammond and Schulte-Merker 2009; Zhou et al. 2014; Park et al. 2015). By logic, hyaline cartilage that does not mature can be described as immature cartilage (Gomez-Picos and Eames 2015). Many signals, including Hedgehog, Parathyroid hormone, FGF, BMP, and Wnt, regulate cartilage maturation, and mature chondrocytes are characterized by expression of such genes as *Col10a1* and *Indian hedgehog* (Vortkamp et al. 1996b; St-Jacques et al. 1999; Vortkamp 2001; Yoon and Lyons 2004; Mak et al. 2006). A GRN under control of the transcription factor Runx2 (Cbfa1) likely drives cartilage maturation (Gomez-Picos and Eames 2015). Loss of Runx2 function abrogates cartilage maturation, and gain of Runx2 function promotes cartilage maturation (Enomoto et al. 2000; Ueta et al. 2001; Eames et al. 2004). Mature and immature hyaline cartilage are not merely transient embryonic tissues; they occur in different zones of adult articular cartilage (Gray and Williams 1989).

*IIB. Notochord*

The chordate notochord is a tissue with many similarities to cartilage, but also some interesting differences. The notochord ECM includes Col2, Col9, and Acan (Stemple 2005). At a cellular level, the notochord is formed by chordoblasts, which secrete the notochord ECM and undergo vacuole formation intracellularly; both of these features impart the mechanical properties of stiffness to the notochord. Sequestration of some PGs in chordoblast vacuoles also might play a major role in notochord mechanical properties (Stemple 2005). While Sox9 is required for cartilage formation, it appears to be dispensable for notochord formation, but is required for notochord maintenance (Bi et al. 1999a; Barrionuevo et al. 2006). Many regulatory elements drive expression of genes in both notochord and cartilage, indicating that these tissues might rely upon similar GRNs for their differentiation (Bagheri-Fam et al. 2006; Dale and Topczewski 2011b).

*IIC. Bone*

Bone serves as a rigid structure also involved in mineral homeostasis, and molecules in bone ECM impart these features. Similar to cartilage, bone ECM has abundant collagens and PGs, but bone has a much higher collagen-to-PG ratio (Ham and Cormack 1987). Bone collagen fibers are largely Col1 fibers, which are tightly wound, giving bone a dark blue stain with Aniline blue in Trichrome histological protocols (Eames et al. 2007). These fibers also provide tensile resistance that is characteristic of bone (Fyhrie and Christiansen 2015). The rigidity of bone derives from abundant biomineral, hydroxyapatite, embedded within the ECM (Fyhrie and Christiansen 2015). Alizarin red, which binds to calcium deposits, is a common histological stain for bone. The massive amount of mineral in bone also contributes to calcium and phosphorus homeostasis in the body (Oldknow et al. 2015).

The formation and maintenance of bone ECM depends upon expression of genes encoding these molecules that are regulated by transcription factors in a bone GRN. With interesting evolutionary implications (see below), Runx2, mentioned above as the main driver of cartilage maturation, is the main transcription factor involved in bone formation. Loss of Runx2 function abrogates bone differentiation, and gain of Runx2 function promotes ectopic bone formation (Ducy et al. 1997; Komori et al. 1997; Otto et al. 1997; Eames et al. 2004). Other transcription factors, such as Sp7 (Osterix), Msx1, Msx2, Twist1, and Twist2 influence the expression of bone genes, but none of these has as central role as Runx2 in osteoblast differentiation (Satokata and Maas 1994b; Dodig et al. 1999; Nakashima et al. 2002; Bialek et al. 2004). Sp7 also is required for bone formation in mice, but the arrest in osteoblast differentiation in *Sp7* null mice occurs downstream of Runx2, and Runx2 promotes *Sp7* expression (Nakashima et al. 2002). Also, as we discuss below, ancestral Sp7 function might not be required for bone formation. Runx2 binds to regulatory elements and drives high levels of expression of many bone genes, such as *Col1a1*, *Col1a2*, *Mgp*, and *Bglap* (Ducy et al. 1997; Sato et al. 1998; Harada et al. 1999; Kern et al. 2001). A Runx2 GRN might dictate bone formation (Gomez-Picos and Eames 2015).

*IID. Dentin/Enamel*

Dentin and enamel comprise the main skeletal tissues of mammalian teeth, where their hard ECM functions in mastication and predation. Dentin shares far more features with bone than enamel, including an ECM with tightly-wound Col1 fibers for tensile strength and a high degree of mineralization by hydroxyapatite for rigidity (Ham and Cormack 1987). Even non-collagenous proteins that are highly enriched in dentin ECM are expressed in bone, including Dentin sialophosphotein, Dentin matrix acidic phosphoprotein 1, and Matrix extracellular phosphoglycoprotein (Kawasaki et al. 2004). Dentin is formed by mesenchymal cells, similar to bone, while enamel is secreted by overlying epithelial cells. Similar to bone, fibers in dentin bind Aniline blue during Trichrome staining, and the high mineral content of dentin and enamel is reflected by Alizarin red staining. Enamel does not contain large collagen fibers and has more tissue-specific ECM components than dentin with respect to bone (Kawasaki et al. 2004; Moradian-Oldak 2012). In addition to Amelogenin, for example, enamel contains Ameloblastin and Enamelin.

Critical transcription factors regulating expression of the ECM molecules in dentin and enamel are currently unknown. Several transcription factors affect formation of these tissues, such as Runx2, NF-kappaB, or Pax9, but none of these appear to abrogate dentin or enamel formation completely (D'Souza et al. 1999; Ohazama and Sharpe 2004; Bonczek et al. 2017). Therefore, a transcription factor that dominates a GRN for odontoblast or ameloblast formation remains to be identified.

*IIE. Phylogenetic considerations for skeletal tissue types*

The dogma of discrete categories of skeletal tissue that has prevailed from classic studies of mammalian systems is misleading. As more information accumulates on mouse skeletal tissues, examples of intermediate tissues abound. Chondroid bone, for example, which forms in the roof of the mammalian skull, among other places, contains features intermediate between cartilage and bone, similar to, but distinct from, fibrocartilage (Beresford 1981). While the field of mammalian skeletal biology has begun to appreciate this perspective in the past decade, it has been the accepted view for many decades for skeletal biologists that study non-mammalian species. For example, instead of the three mammalian cartilage types, about ten cartilage types in teleosts were classified

almost 30 years ago (Benjamin 1990). Extant agnathans produce at least two types of cartilage, mucocartilage and (cell-rich) hyaline cartilage, and no clear homolog of mucocartilage exists in other vertebrates (Cattell et al. 2011). Considering the cases of chondroid bone and fibrocartilage, many tissues that are intermediate to "classic mammalian" cartilage and bone were described in teleosts (Benjamin 1989; Benjamin and Ralphs 1991; Benjamin et al. 1992). These observations lead to an important perspective on skeletal tissue identity that informs our understanding of how these tissues evolved in the first place: Instead of discrete entities, skeletal tissues across phylogeny often serve as examples along a spectrum of possibilities among the archetypal cartilage, notochord, bone, dentin, and enamel.

## I. Location, location, location

Two major factors impacting discussion of how cartilage, bone, and neural crest evolved are the location in the body where a tissue forms, and how these locations have evolved over time. A major theme in this chapter is that, once the ability to make a tissue is encoded in the genome as a GRN, different populations of cells can co-opt expression of this GRN, adding another location in the body where this tissue forms. For example, if cells in the tail evolved the ability to make muscle, then for muscle to appear in the head, head cells must simply co-opt expression of the muscle GRN that was employed by tail cells. How could a different population of cells co-opt expression of an entire GRN? It might not be as difficult as it seems, since GRNs typically have one to few transcription factors that operate near the top of the GRN hierarchy (Levine and Davidson 2005; Davidson and Levine 2008). In this case, so co-option of a GRN might entail simply co-opting expression of one or two genes, and the downstream gene expression would follow (Halfon 2017). In the case of cartilage and bone, Sox9 and Runx2 are prime candidates to sit at the top of their GRN hierarchies, respectively (Gomez-Picos and Eames 2015).

What evidence exists to support the contention that changing expression of one or two transcription factors can change skeletal tissue location? In threespine stickleback fish, the pelvic fin skeleton of freshwater populations is reduced dramatically, compared to oceanic populations (Peichel et al. 2001). Genetic analyses suggested that non-coding polymorphisms in enhancers near one transcription factor, Pitx1, drove these natural evolutionary differences in skeletal tissues

(Thompson et al. 2018). Also, comparison of candidate gene expression across homologous cell types in different regions of the body demonstrate few differences. For example, immature cartilage, mature cartilage, and bone express a core set of genes, including distinct Sox9 and Runx2 patterns, in the limb and in the head (Eames and Helms 2004). Since skeletal tissues occurred in the head of vertebrates before paired appendages evolved (Janvier 1996b), these data support the idea that limb mesenchyme co-opted a GRN dictating skeletal tissue differentiation in head mesenchyme. Other transcription factors, such as those in Hox, Dlx, Pax, and Nkx families, have been proposed as major players in skeletal tissue formation, but this influence is only region-specific, suggesting that these factors are more involved in patterning than differentiation *per se*. For example, Dlx5 and Nkx3.2 are referred to commonly as important transcription factors for bone and cartilage differentiation (Ferrari and Kosher 2002; Tadic et al. 2002; Zeng et al. 2002). However neither Dlx5 nor Nkx3.2 loss of function mutations abrogate bone or cartilage throughout the body, which rules these transcription factors out as major players in skeletal cell GRNs. On the other hand, evolution of regulatory elements of SoxE (ancestral to Sox9) might explain the appearance of cellular cartilage in the neural crest-derived vertebrate head (Jandzik et al. 2015).

## II. Skeletal tissues from extant animals and fossil record before the emergence of NCCs

### *IVA. Invertebrate Cartilage*

Cartilage is known from several lineages of bilaterian metazoans (Fig. 63). Generally, two cartilage categories with unclear mutual relationships are distinguished: 1) acellular cartilage, and 2) cellular cartilage. Acellular cartilage can be found in the pharynx of two deuterostome groups, hemichordates and cephalochordates (Rychel et al. 2006). It consists of cartilaginous rods (gill bars) with ECM rich in fibrillar collagen, similar to vertebrate Col1 and Col2, and proteoglycans. In contrast to other invertebrate and vertebrate cartilages, the chondrocytes are not embedded in the ECM that they produce, but rather they surround it. Acellular cartilage is presumed to be of endodermal origin in hemichordates, while all three germ layers (endoderm, ectoderm, and mesoderm) might contribute to the formation of acellular cartilage in cephalochordates (Rychel and Swalla 2007).

Cellular cartilage occurs in six somewhat distantly-related metazoan clades (Person and Philpott 1969; Cole and Hall 2004b). Cellular cartilage can be either cell-rich or matrix-rich depending on relative abundance of chondrocytes versus ECM (Benjamin 1990). Cellular cartilages of metazoans without NCCs are likely of mesodermal or, less often, putatively endodermal origin. Among protostomes, cellular cartilage occurs in chelicerates (arthropod ecdysozoans), sabellid polychaetes (annelid lophotrochozoans), gastropods, and cephalopods (latter two being mollusc lophotrochozoans). Among chordate deuterostomes, cellular cartilage can be found in two clades, cephalochordates and vertebrates (vertebrate cartilages are detailed in the next section of this chapter). Besides the acellular cartilage of pharyngeal bars, the cephalochordate amphioxus has been reported to have a cartilaginous skeleton supporting oral cirri (tentacles) that protect the mouth of this sand-burrowing chordate (Jandzik et al. 2015). This cellular cartilage is cell-rich, with chondrocytes organized in a stack-of-coins configuration, and contains collagen and proteoglycans (as suggested by Alcian blue staining).

The cellular cartilages of three protostome groups are not uniform, rather showing considerable variation, even within a single taxon. For example, the horseshoe crab (*Limulus polyphemus*), a representative of chelicerate arthropods, has four different types of cartilaginous structures: endosternite, opistosomatic endplates, branchial cartilages, and chilarial cartilages (Cole and Hall 2004b). The endosternite cartilage is dominated by extracellular matrix, with scattered small spherical or irregularly shaped chondrocytes interconnected via plasmatic processes. The ECM contains irregularly organized collagen fibers and is rich in acidic glycosaminoglycans, and generally resembles vertebrate fibrocartilage. The other cartilaginous structures of horseshoe crabs are cell-rich with vacuolated chondrocytes and minimal ECM that contains elastin, collagen, and mucopolysaccharides, specifically chondroitin-4-sulfate and chondroitin sulfate K (Cole and Hall 2004b; Hall 2015).

Cellular cartilage of at least 30 species and 18 genera of sabellid polychaetes, or feather duster worms, supports their branching feeding tentacles (Capa et al. 2011). The cartilaginous tissue is similar to non-endosternite cartilages of the horseshoe crab, being composed of large chondrocytes with vacuoles and scant vascularized ECM containing elastin and collagenous fibers and chondroitin sulfate mucopolysaccharides (Hall 2015). In distal parts of the tentacle, the cartilage could form acellular rods comprised solely of ECM (Capa et al. 2011).

Cephalopod cartilage is probably the best-known invertebrate cartilage. It can be found supporting or protecting various parts of squid, octopus, and cuttlefish bodies, being mostly associated with brain, eye, mantle, and dermis (Cole and Hall 2004b). Cephalopod cartilage resembles "typical" hyaline cartilage with sparsely distributed chondrocytes embedded in abundant ECM. The chondrocytes are connected by processes through canaliculi in matrix, however which is reminiscent of vertebrate osteocytes. Besides these cell-cell connections, the chondrocytes also form junctions with the ECM (Bairati et al. 1998). In contrast to most other known cartilages, cephalopod cartilage is commonly vascularized (Hall 2015). Besides elastin, the fibrillar part of cephalopod cartilage ECM contains several types of collagens, some resembling vertebrate Col1, Col2, and Col4, along with an additional type that appears to be cephalopod-specific (Hall 2015). The mucopolysaccharides in cephalopod cartilage ECM are surprisingly diverse, although their occurrence varies across cephalopod taxa. They include hyaluronan and various chondroitin sulfates, such as chondroitin-4-sulfate and chondroitin sulfate E (Hall 2015).

Most gastropods have either odontophore or subradular cartilages that support their radulae, a rasping toothed organ (Guralnick and Smith 1999). Subradular cartilage is typical for some basal forms and might be homologous to the dorsal odontophore cartilage of more derived forms (Golding et al. 2009). The best-studied gastropod cartilage is odontophore cartilage of the sea snail, *Busycon*. It is composed of large vacuolated chondrocytes with very little ECM that is rich in collagen (Hall 2015), thus morphologically very similar to the branchial cartilage of the horseshoe crab. The ECM does not contain chondroitin sulfate, but is instead composed of a chitin-like polyglucose sulfate. Interestingly, gastropod chondrocytes contain some amount of myoglobin, and cartilages of some other snails, aside from chondrocytes, are also formed by muscle cells (Person and Philpott 1969; Hall 2015).

Besides acellular and cellular cartilages, some invertebrates, including brachiopods, polychaetes, and urochordates, have tissues with cartilage-like properties that vary in the structure and organization, or even presence, of chondrocytes. These tissues are usually collectively referred to as chondroid connective tissue. Simpler structure, phylogenetic distribution, and variable relative contents of cells, mucopolysaccharides, and fibers indicate that chondroid connective tissue is an ancient tissue type (Cole and Hall 2004b).

The fossil record does not further expand the number of large bilaterian clades with cartilage. Cartilaginous structures, particularly those that mineralize, can be found in fossil remnants of

various chordate groups and are usually of strong evolutionary importance. Perhaps most notably, the stem chordate *Haikouella* demonstrates cartilage very similar in appearance to lamprey cartilage (Chen et al. 1999; Mallatt and Chen 2003). Other fossil examples include cephalochordates of the genus *Cathaymyrus* and stem vertebrates such as *Myllokunmingia*, *Metaspriggina*, and *Haikouichthys* (Morris and Caron 2014; Hall 2015). Despite the fact that a mineralized endoskeleton can be found among protostomes, their cartilage is not known to be capable of mineralization *in vivo* (Hall 2015), which may explain why protostome cartilage has not been well-described in the fossil record to date.

*IVB. Notochord*

The notochord, or *chorda dorsalis*, is a defining structure of chordates, a deuterostome group comprising three main lineages: cephalochordates, tunicates, and vertebrates (vertebrate notochord is discussed in section VB; Fig. 63). It forms an axial hydrostatic skeleton against which longitudinal musculature contracts during bilateral, undulating movement typical of early chordate forms and their larvae (Annona et al. 2015). The main portion of the notochord is formed by vacuolated cells, called chordocytes, connected by desmosomes and surrounded by a thick acellular perinotochordal sheath (Hall 2015). Chordocytes are filled with liquid, and their pressure against the stiff sheath facilitates the mechanical function of the notochord. The notochord persists throughout life in cephalochordates and appendicularian tunicates, but forms only transiently in the remaining tunicate lineages (ascidians and thaliaceans; Hall 2015). Ascidians and thaliaceans lose their notochord during regressive metamorphosis. Embryonically, the notochord develops from chordamesoderm. This embryonic tissue is morphologically and molecularly distinct from the remaining mesoderm, and its origin in cephalochordates lies in the dorsal organizer that arises during gastrulation (Stemple 2005). Aside from a mechanical role in chordates, the notochord also plays an important role as an embryonic signaling center, strongly affecting all three germ layers, most notably in initiating neural tube formation and axial patterning (Corallo et al. 2015; Hall 2015).

There are some significant differences in notochord structure among chordates. Antero-posteriorly, the notochord runs along almost the entire body in cephalochordates. In tunicates, the notochord does not extend past the hindbrain posteriorly (Annona et al. 2015). Chordocytes of

cephalochordates and tunicates are discoidal and stacked like coins. Unique to the cephalochordate notochord, chordocytes contain transverse myofilaments. These large vacuolated chordocytes are surrounded by a layer of thin epithelial cells. In tunicates, the perinotochordal sheath is relatively simple, formed by a basal (also called external) lamina, which is supplemented by two additional collagenous layers in cephalochordates. The notochord of late tunicate larvae, mainly its posterior portion, has a different structure with cells surrounding the lumen filled with liquid (Annona et al. 2015).

The notochord shares several characteristics with cartilage. It contains structural proteins and glycosaminoglycans typical for cartilage, such as Aggrecan, Chondromodulin-1, hyaluronic acid, chondroitin sulfate, and SPARC (Osteonectin), and the perinotochordal sheaths are predominantly formed by collagenous fibers composed of Col2 (Vasan 1987; Welsch et al. 1991; Stemple 2005; Hall 2015). Cells of the forming notochord express cartilage collagen genes and transcription factors *Tbxt (Brachyury)* and *Sox9*, which are also involved in cartilage development.

## III. Skeletal tissues from extant animals and fossil record after the emergence of NCCs

### *VA. Cartilage*

Both vertebrates and invertebrates have cartilage (Fig. 63). Consistent with a GRN expression co-option story, vertebrate head cartilage has most of the same molecular characteristics as invertebrate cartilage, but they are formed by a different embryonic population of cells. Most head and pharyngeal cartilages in vertebrates are neural crest-derived, cellular cartilage (Couly et al. 1993). In fact, neural crest cells have contributed to the formation of many structurally distinct cartilages in different vertebrate clades from cyclostomes to jawed vertebrates. These cartilages display some variations in gene expression patterns, but the basic molecular machinery composed of SoxE and Runt (ancestor to Runx2) transcription factors, as well as collagenous and proteoglycan-rich extracellular matrix is conserved among vertebrates (McCauley and Bronner-Fraser 2006; Zhang et al. 2006; Hecht et al. 2008; Cattell et al. 2011; Jandzik et al. 2015). Indeed, cyclostome cartilage is composed of Col2, sharing this feature with extant gnathostomes.

Cyclostome cartilages are found in the head, fin rays, and dorso-anterior axial skeleton (Ota et al. 2011). Cartilage types described in hagfish and lamprey include mucocartilage, soft cartilage, and hard cartilage (Zhang et al. 2006; Zhang et al. 2009; Cattell et al. 2011). Mucocartilage is composed of elastin-like molecules and fibroblasts surrounded by a perichondrium (Wright and Youson 1982; Wright et al. 2001). Agnathan soft cartilage is composed of hypertrophic chondrocytes surrounded by a thin extracellular matrix (i.e., cell-rich), whereas hard cartilage contains smaller chondrocytes surrounded by abundant extracellular matrix (i.e., matrix-rich; (Zhang et al. 2009). A potentially lamprey-specific extracellular matrix protein, Lamprin, also can be found in many cartilages (Robson et al. 1993). Although mineralized cartilage has been described in some specimens of extant lamprey of presumed older age and some agnathan fossils of unknown affinity to cyclostomes, such as *Euphanerops longaevus* and *Palaeospondylus gunni*, clear proof of mineralization in cyclostome cartilages has only come from *in vitro* experiments (Bardack and Zangerl 1968; Langille and Hall 1993b; Janvier and Arsenault 2002; Johanson et al. 2010; Hirasawa et al. 2016).

A wealth of information is coming out recently about skeletal tissues in cartilaginous fish, or chondrichthyans. Three main extant lineages of chondrichthyans include chimaera (holocephalans), sharks (selachians), and skates and rays (batoids; selachians and batoids compose elasmobranchs) (Daniel 1934). Chondrichthyan skeletons are thought to be entirely cartilaginous, mostly composed of matrix-rich hyaline cartilage (Hall 2005a; Dean and Summers 2006; Eames et al. 2007). While most vertebrates have a limited amount of mineralized cartilage as adults, it serves as a major structure that provides support and acts as a mineral reservoir in cartilaginous fish (Daniel 1934). Most of the volume of cartilage in the chondrichthyan skeleton is unmineralized, but most cartilages typically contain a surface layer of mineralized blocks, called tesserae, surrounded by a fibrous perichondrium (Eames et al. 2007; Seidel et al. 2016). Tesserae are composed of two main components: the cap (more superficial) and body (more deep) zones (Kemp and Westrin 1979). The molecular components of chondrichthyan cartilage are only becoming known with recent work. The cap zone actually contains a bone-like tissue with abundant Col1 extracellular matrix (Seidel et al. 2017). On the other hand, the body zone appears to be mineralized cartilage with an extracellular matrix rich in Col2 and Col10 (Seidel et al. 2017). The mineralized centrum of the vertebrae appears like a mineralized fibrocartilage, whereas there is an odd bone-like tissue in the neural arches of the vertebrae, reminiscent of perichondral bone (Peignoux-Deville et al. 1982; Bordat 1987; Eames et

al. 2007; Enault et al. 2015; Criswell et al. 2017a). While sharks received much focus of previous research, within the past ten years, a number of groups has recently expanded considerably what is known about chondrichthyan skeletal tissues. Tesseral features of rays are similar to sharks (Dean and Summers 2006; Seidel et al. 2016; Seidel et al. 2017). Indeed, the exact same mineralized tissues (tesserae, centra, and neural arches), have been described in both shark and skate species, arguing that these skeletal features were present in the last common ancestor to elasmobranchs (Atake et al. 2019). Almost no published data exists on holocephalan skeletal tissues, but many researchers are currently addressing this shortcoming.

Many lineages of early-derived ray-finned bony fish (i.e., actinopterygians) have living members today, and their skeletal tissues, like that of chondrichthyans, could provide a glimpse into the variety of skeletal tissues that were present in stem vertebrates. However this is also an area for future research, as there are limited publications on skeletal tissues in these species. The publication of the longnose gar genome has helped push forward interest in this clade (Braasch et al. 2016). Gar matrix-rich hyaline cartilage is abundant throughout the head and expresses genes typical of mammalian cartilage, including *Col2a1*, *Col11a2*, and *Col10a1* (Eames et al. 2012). These extremely limited published data illustrate the unexploited goldmine of molecular studies into the skeletal tissues of gar, sturgeon, paddlefish, bichir, and bowfin fishes.

Teleost fish have a very rich diversity of cartilage types, perhaps due to the teleost-specific genome duplication that resulted in more cartilage differentiation genes, including the Sox family, although no current data support this speculation (Voldoire et al. 2017). Many different cell-rich and matrix-rich cartilages occur in these fish (Benjamin 1989; Benjamin 1990; Benjamin and Ralphs 1991; Benjamin et al. 1992), but very little is known about their molecular composition. Unlike cell-rich cartilages, matrix-rich cartilages are composed of cells occupying less than 50% of total volume. So far, five types of cell-rich cartilage have been described in teleosts, based upon histology and some fiber type analyses: hyaline-cell rich cartilage, cell-rich hyaline cartilage, fibro/cell cartilage, elastin/cell-rich cartilage, and the cell-rich cartilage Schaffer's Zellknorpel. Hyaline-cell cartilage is composed of small chondrocytes and little extracellular matrix surrounding them. This cartilage sub-type is often present in lips and the head (Benjamin 1989). Distribution of all of these cartilage sub-types has only been described in the head, but all of these tissues likely occur throughout the body. In contrast, cell-rich hyaline cartilage, which is often present in the neurocranium and gill arches, is composed of large lacunae that occupy a significant percentage of the cell volume

(Benjamin 1989; Benjamin 1990). Fibro/cell cartilage is a non-hyaline type of cartilage composed of large cells surrounded by a matrix rich in collagen, which often forms articular tissues (Benjamin et al. 1992; Kapoor and Khanna 2004). Elastic/cell-rich cartilage matrix, another type of non-hyaline cartilage, is rich in elastin fibers surrounded by a thick perichondrium, and it is commonly found in the barbels (Benjamin 1990; Benjamin et al. 1992; Kapoor and Khanna 2004). The last type of cell-rich cartilage in teleost fish is Schaffer's Zellknorpel, which has a rigid structure and is composed of shrunken cells laying in lacunae (Benjamin 1989; Benjamin 1990). In addition to these five cell-rich cartilages, a familiar three matrix-rich cartilages compose the teleost skeleton: matrix-rich hyaline cartilage, fibrocartilage, and elastic cartilage (see section IIA). Matrix-rich hyaline cartilage is found in the neurocranium and gill arches of many species and is structurally similar to hyaline cartilage in tetrapods, providing the main source of support and growth (Benjamin 1990). Teleost fibrocartilage lacks a perichondrium, contains irregularly arranged cells, and can be found in the oro-mandibular region of teleosts (Benjamin and Evans 1990). The matrix-rich elastic cartilage is composed mostly of elastin fibers. Most fibrocartilages and elastic cartilages, however are cell-rich in teleosts (Kapoor and Khanna 2004). In extant tetrapods, three types of matrix-rich cartilages occur: hyaline cartilage, elastic cartilage, and fibrocartilage, as outlined in section IIA.

*VB. Notochord*

Chordoid (notochord-like) tissue is present in all chordates (Fig. 63), and the vertebrate notochord shares most features with cephalochordates and tunicates described above (section IVB). Similar to tunicates, the notochord is a transient structure in most vertebrates (exceptions include cyclostomes and sturgeons; Hall 2005a). In contrast to tunicates, in which the notochord does not extend past the hindbrain posteriorly, the vertebrate notochord does not extend past the hindbrain anteriorly (Annona et al. 2015). In various groups of vertebrates, the notochord may mineralize or be replaced by cartilage and/or bone, or some portions may become parts of intervertebral discs. Similar to cephalochordates, the embryonic origin of the notochord in vertebrates lies in the dorsal organizer that arises during gastrulation (Stemple 2005). In contrast to the discoidal and stacked-like-coins chordocytes of cephalochordates and tunicates, chordocytes of the vertebrate notochord do not show obvious stacking. Similar to cephalochordates, the vertebrate notochord has multiple epithelial layers surrounding the vacuolated chordocytes (Annona et al. 2015).

A major difference between vertebrate and invertebrate notochords is that the vertebrate notochord can mineralize. A common feature of many vertebrates is the chordacentrum, which is a typically thin mineralization in the fibrous sheet of the notochord (Arratia et al. 2001). Despite the widespread distribution of chordacentra among various vertebrate clades, they are not considered strictly homologous. The centrum of tetrapods, for example, derives from migrating somitic mesoderm, the centrum of teleosts appears to derive from notochordal sheath cells directly, while the centrum of chondrichthyans is also somite-derived (Criswell et al. 2017b). Phylogenetically, the distribution of various forms of the centrum around the notochord is patchy, further weakening an argument for strict homology. Many genes involved in bone formation, such as *runx2* and *sp7*, might regulate chordacentrum formation in teleosts (Renn and Winkler 2014). However the fossil record and sparse phylogenetic distribution does not support the idea that the bone GRN was co-opted from a centrum GRN.

*VC. Bone*

Only vertebrates make bone, but its exact anatomical location varies (Fig. 63). Based upon fossils, bone is thought to have originated on the mesenchymal side of the basement membrane in the pharynx or epidermis, in tooth-like structures termed odontodes (Smith and Hall 1990b; Donoghue and Sansom 2002). Odontodes have been argued to appear first in the fossil record in pharyngeal regions of agnathan conodonts (~515 Mya), while epidermal odontodes appear later as protective head shields in other agnathan clades, who likely had a soft cartilage endoskeleton. Perhaps the first experiment of NC making mineralized tissues, conodonts might represent the earliest form of craniate vertebrates, although more recent hypotheses place conodonts closer to gnathostomes phylogenetically (Sansom et al. 1992; Donoghue and Rucklin 2016). The feeding apparatus in these fossil species are comparable with those present in vertebrates, and their soft tissues suggest a relationship to hagfish (Donoghue et al. 2000). Odontodes were composed of dentin and bone, thus becoming the first example of bone in the fossil record (Reif 1982; Smith and Hall 1990b; Huysseune and Sire 1998). The presence of bone in tooth-like structures of conodonts, however is debated (Donoghue and Rucklin 2016). The first unequivocal evidence of bone in the

fossil record is in the agnathan clade Pteraspidormorphi (~480 Mya), historically considered a superclass of the ostracoderm group (Janvier 1996b; Donoghue et al. 2006).

Due to the association of odontodes with epithelia and the absence of adjacent cartilage, bone is argued to have formed first via intramembranous ossification, whereas endochondral ossification evolved later and gradually (Smith and Hall 1990b; Smith and Hall 1993; Janvier 1996b; Mundlos and Olsen 1997; Wagner and Aspenberg 2011). In the endoskeleton, perichondral bone formation occurred first, perhaps using the same genetic programs used to produce the dermal skeleton, whereas cartilage degradation and endochondral bone deposition occurred later (Donoghue and Sansom 2002; Wagner and Aspenberg 2011). While pteraspidomorphi were heavily armored by external odontodes, some fossils show both mineralized cartilage and perichondral bone in their endoskeleton, reflecting evolutionary progress leading to endochondral ossification (Janvier 1996b; Zhang et al. 2009).

The location of bone in vertebrates evolved. Critically for understanding the role of NC during evolution of bone, extant and fossil Cyclostomata (i.e., lamprey and hagfish) never appeared to have mineralized cartilage or bone. As outlined above, fossilized bone first appeared in exoskeletal head shields of the pteraspidomorphs, presumably formed by NC. Current data suggest that perichondral bone later appeared in the braincase of ostracoderms, and then perichondral (and perhaps endochondral) bone appeared in the vertebrae of some placoderms (e.g. petalichthyids and arthodires) (Janvier 1996b; Trinajstic et al. 2015). This would imply that after the evolution of NC, cephalic mesoderm and then somitic mesoderm acquired the ability to make bone. Whether extant chondrichthyans generate bone is currently under investigation (Eames et al. 2007; Enault et al. 2015; Atake et al. 2019), but fossil primitive chondrichthyans and especially their acanthodian ancestors had abundant perichondral and perhaps endochondral bone (Zangerl 1966; Coates et al. 1998; Long et al. 2015). Together, these studies suggest that endochondral bone is not an osteichthyan synapomorphy. Whether all examples of endochondral bone in vertebrates are homologous is complicated by another possible scenario, in which early osteichthyan chondrocytes acquired the ability to trans-differentiate into osteoblasts in response to higher mechanical forces and to provide a new osteoblast reservoir (Cervantes-Diaz et al. 2017).

Histologically, all examples of bone in vertebrates, whether it be dermal, perichondral, or endochondral, appears very similar, if not identical, when animal groups are inter-compared. Certainly, some tissues, such as chondroid bone or the mineralizing chondrichthyan tissues, share

some, but not all, of these histological features, making their designation as bone difficult. Molecular studies might reveal significant differences (Arendt 2008; Gomez-Picos and Eames 2015).

*VD. Dentin/Enamel*

Dentin and enamel are vertebrate-specific tissues (Fig. 63). Dentin and enamel/enameloid, together with bone, composed the primitive skeleton in early jawless vertebrates around 500 Mya (Gans and Northcutt 1983). Similar to bone, it is generally accepted that dentin and enamel/enameloid first appeared in odontodes of Ordovician agnathans as a protective shield (Smith and Hall 1990b; Janvier 1996b). An alternative theory is that the earliest hard skeletal tissue was dentin, forming the most superficial dermal tissue, only later covered by a hypermineralized layer of enamel/enameloid. Conodonts (~515 Mya) potentially represent the earliest form of craniate vertebrates whose skeleton was composed of dentin- and enamel-like tissues (Sansom et al. 1992). Although neural crest-like cells likely existed prior to vertebrates (Hall and Gillis 2013b), it is still unknown whether conodonts possessed neural crest cells. After conodonts, heterostracans, an extinct group of pteraspidomorphs, show traces of dentin tubercles capped with enamel/enameloid in their exoskeleton (Janvier 1996b; Keating et al. 2015). Dentin and enamel/enameloid are present in placoid scales of living chondrichthyans, and in teeth of all other extant vertebrates, including sharks, mammals, teleosts, and reptiles (Kawasaki et al. 2004). Enameloid in osteichthyans appears to serve the same function as mammalian enamel, but enameloid and enamel might actually be convergent tissues. Enameloid appears to be deposited by mesenchymal, not epithelial, cells, and the major proteins in mammalian enamel have not been identified in osteichthyans (Kawasaki et al. 2004). These and other features argue that enameloid in fish and enamel in mammals are a result of convergent evolution (Kawasaki et al. 2004). Although experimental data in the trunk region of vertebrates are limited to one paper, dentin in teeth and dermal denticles are NC-derived (Chai et al. 2000; Gillis et al. 2017). Enamel is produced by ectodermal or pharyngeal epithelia, while enameloid is presumably NC-derived (Kawasaki et al. 2004).

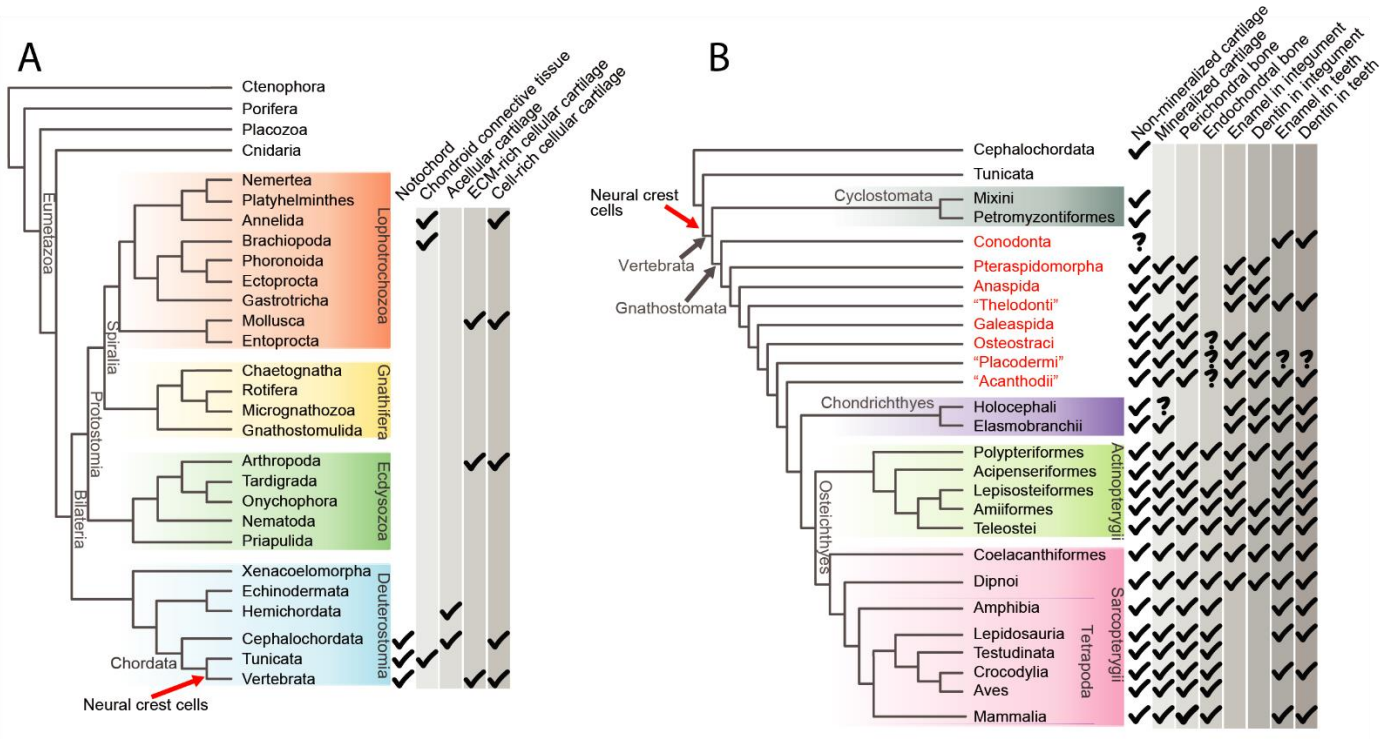



**Figure 63. Phylogenetic distribution of skeletal tissues in metazoans (A) and chordates (B).** Names of the fossil clades are in red, clades with names in "" are likely paraphyletic, though traditionally recognized, fossil groups. The term "enamel" is defined broadly here and includes enameloid as well as ganoine, while the term dentin includes both orthodentin and mesodentin. Characters in extant lineages do not reflect occurrence in fossil members of those lineages. "?" represents a character with contradictory evidence in the literature. The phylogenetic trees are based on (Dunn et al. 2014; Donoghue and Rucklin 2016; Irisarri et al. 2017; Whelan et al. 2017; Marletaz et al. 2019). The characters mapped on the phylogenetic trees are based on (Cole and Hall 2004b; Hall 2005a; Donoghue et al. 2006; Sire et al. 2009; Goudemand et al. 2011; Keating et al. 2015; Donoghue and Rucklin 2016; Leprevost et al. 2017).

## VI. Evolution of GRN underlying cartilage

### *VIA. Cellular cartilage evolved from an acellular cartilage GRN*

The cellular cartilage of vertebrates might have evolved by modifying a GRN that directed acellular cartilage formation (Fig. 64A). Similar features of agnathan and gnathostome cellular cartilage suggested that a cartilage GRN (presumably driven by SoxE, the ancestor to Sox9) is a vertebrate synapomorphy (Zhang et al. 2006; Ohtani et al. 2008b), but this concept has been revised substantially. Acellular cartilage of hemichordates might be homologous to acellular cartilage of

cephalochordates, given their general resemblance and close phylogenetic distribution, in which case the tissue was subsequently lost in echinoderms, urochordates, and vertebrates (Hall and Gillis 2013). Structures similar to acellular cartilaginous rods of hemichordates and cephalochordates were proposed in hypothetical stem vertebrates as antecedents of vertebrate cellular cartilage of the pharyngeal arches (Rychel and Swalla 2007). According to this hypothesis, NCCs that migrate to the pharynx might have gradually replaced the acellular cartilage by co-opting and modifying the SoxE GRN that presumably operated in the endodermal epithelia of ancestors to hemichordates and cephalochordates. To be clear, this would involve two steps (of uncertain order): 1) acquiring expression of the acellular cartilage SoxE GRN in NCCs; and 2) modifying that GRN to form cellular cartilage. Of course, the acellular cartilage GRN operating in ancestors to amphioxus might have acquired expression first in presumably cephalic mesoderm of larval forms of ancestors to amphioxus, then was modified to form cellular cartilage, and finally expression of this cellular cartilage GRN was co-opted by vertebrate NCCs (Jandzik et al. 2015). However acellular cartilage might be less similar to vertebrate cartilage than other protostome cellular cartilages. Cellular cartilage appears in many disparate phylogenetic groups and presumably evolved independently at least three times (Hall and Gillis 2013b). The striking similarities in morphology, function, development, and gene expression patterns, though, suggest the existence of underlying homologous mechanisms with deep metazoan origins (i.e. deep homology). For example, the endosternite cartilage of horsheshoe crab and the funnel cartilage of cuttlefish share significant parts of the GRN driving hyaline cartilage in vertebrates (Tarazona et al. 2016). Interestingly, the chondrocytes in some portions of the sabellid polychaete tentacles are arranged in a stack-of-coins fashion, reminiscent of cellular cartilage in pharyngeal arches of amphioxus, lamprey, and zebrafish (Kimmel et al. 1998; Cole and Hall 2004b; Cattell et al. 2011; Hall 2015; Jandzik et al. 2015). In summary, acellular cartilage exists in hemichordates, cephalochordates, and agnathans, while cellular cartilage is found in cephalochordates, vertebrates, and various protostomes. Therefore, if vertebrate cellular cartilage evolved from an acellular cartilage GRN, then the independence of this event from the appearance of cellular cartilage in various protostomes would have to be addressed.

*VIB. Cellular cartilage evolved from a notochord GRN*

Some fossils and other features of extant non-vertebrate chordates provide the basis for an evolutionary hypothesis that the GRN for cellular cartilage evolved from a notochord GRN (Fig. 64B). Since the notochord is usually a diagnostic character of fossil chordates, it is present in all stem chordates in which cellular cartilage has been described, but importantly for this hypothesis, the notochord is apparent in samples with no identifiable cartilage (e.g. *Pikaia*; Morris and Caron 2012). Of living animals, enteropneust hemichordates have a diverticulum of the gut called a stomochord, which has some functional and histologic similarities to the chordate notochord (Satoh et al. 2014). These animals only have acellular cartilage, so they might represent an evolutionary snapshot where a notochord GRN was being established prior to cellular cartilage. However the homology of the stomochord and notochord is questionable (Annona et al. 2015; Lowe et al. 2015; Minarik et al. 2017). The phylogenetically closest group of animals to vertebrates are urochordates, such as sea squirts, some of whose larval forms have a true notochord (Delsuc et al. 2006). However urochordates lack cartilage, and they are thought to be very derived, which led to them being designated as the sister taxon to vertebrates only relatively recently. The cephalochordate amphioxus has a notochord, but also contains both acellular and cellular cartilage, so no living animals represent a transition from notochord only, to notochord with cellular cartilage.

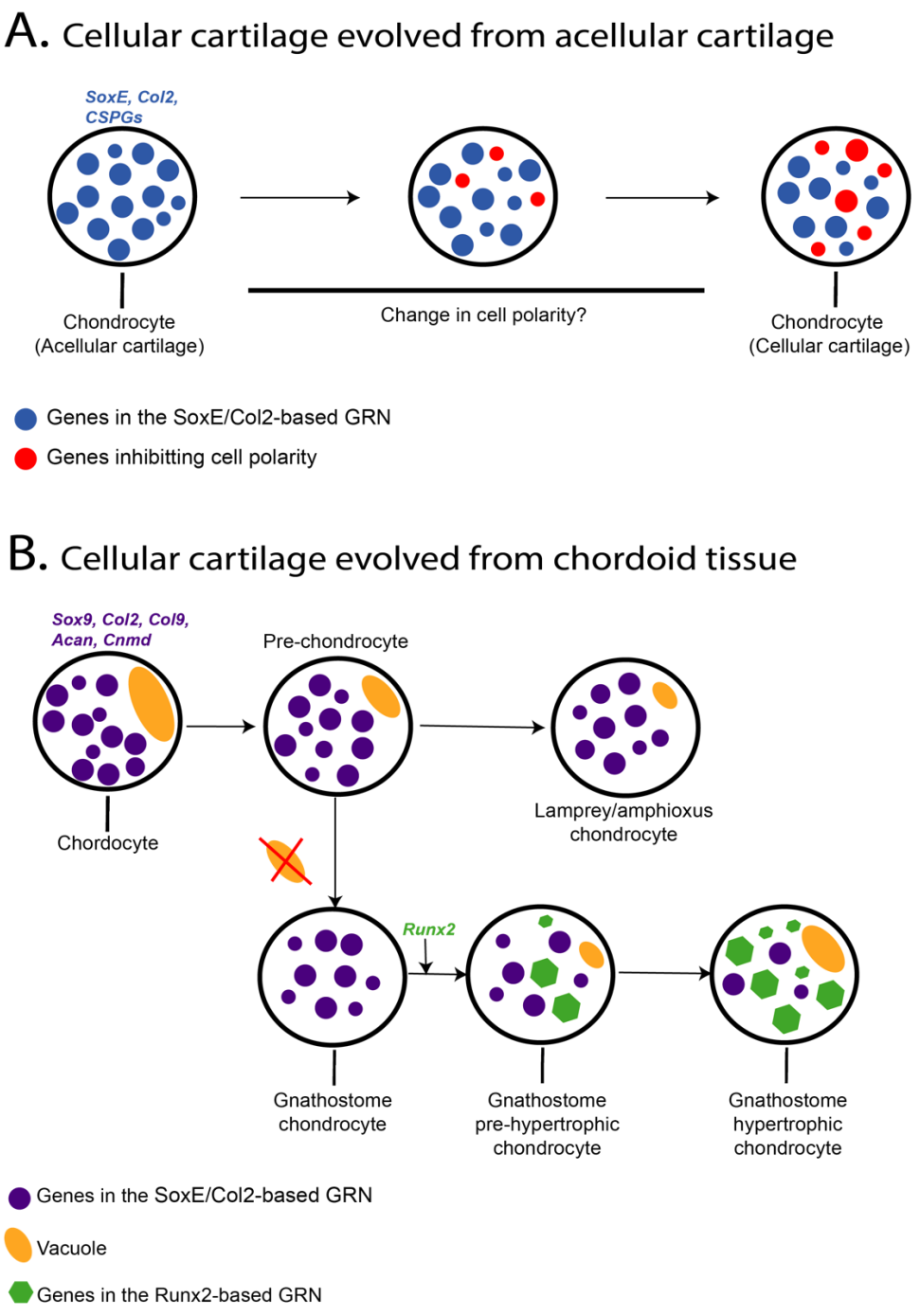



**Figure 64. Evolution of the GRN underlying cartilage.** A) Cellular cartilage evolved from acellular cartilage. NCCs acquired expression of the acellular cartilage SoxE/Collagen type II-based cartilage GRN, and subsequently modified that GRN to make cellular cartilage. Runx and Barx genes were later co-opted by these cells consolidating the cellular cartilage GRN (Cattell et al. 2011). B) Cartilage evolved from chordoid tissue. NCCs acquired expression of the notochord fibrillar collagen based GRN, and was then modified to make cellular cartilage. Chordocytes have a single large vacuole and low amounts of glycosaminoglycan- rich ECM. While chondrocytes in amphioxus and lamprey are vacuolated, chondrocytes in vertebrates are avacuolated and have abundant ECM. However when vertebrate cartilage becomes hypertrophic and degenerates (under the influence of a Runx2 GRN), the ECM frequently becomes vacuolated. Cellular cartilage and notochord might share a common origin since they share the expression of several genes including Sox9, Col2, Acan, Col9, and Cnmd (Jonasson et al. 2012).

Despite the lack of an animal phylogenetically positioned to directly support the hypothesis that the cellular cartilage GRN evolved from the notochord GRN, functional, structural, and molecular similarities supported the hypothesis that the notochord was a primitive type of cartilage. For example, chondrocytes secrete a highly hydrated ECM, which gives cartilage its structural properties, while chordocytes (notochord cells) retain liquid in their vacuoles (Stemple 2005). Regarding the vacuolar chordocyte, the mature chondrocyte also undergoes vacuolarization during hypertrophy, which under the evolutionary scenario of chordocyte to chondrocyte, suggests a rather parsimonious recovery of vacuoles, perhaps under Runx2 control (Fig. 64B). In addition to molecular and developmental similarities between vertebrate and cephalochordate notochords, amphioxus has cellular cartilage that shares properties with vertebrate cellular cartilage. This evolutionary model proposes co-option of the developmental program of originally mesodermal cartilage of an amphioxus-like vertebrate ancestor into the developmental repertoire of the cranial NCCs of vertebrates (Jandzik et al. 2015). Stemple (2015) actually proposed the opposite relationship between the notochord and cartilage, arguing that the notochord co-opted cartilage properties. Of course, the 'cellular cartilage evolved from the notochord GRN' hypothesis also would have to be reconciled with the appearance of cellular cartilage in various protostomes.

## VII. Evolution of GRN underlying bone

How could bone have evolved within vertebrates, since there is no obvious antecedent tissue in cyclostomes or vertebrate sister groups? Here, we explore three parsimonious explanations, which involve simple co-option by NCCs of a pre-existing GRN that directed formation of: 1) mineralized tissues in non-vertebrates; 2) dentin/enamel; or 3) mature cartilage. In principle, option #1 could have preceded options #2 or 3.

### *VIIA. Bone evolved from a deuterostome mineralization GRN*

The first hypothesis claims that a mineralization GRN operating in mesenchymal cells of non-vertebrate deuterostomes was co-opted by NCCs and modified to form bone in the ancestor to non-

cyclostome agnathan vertebrates (Fig. 65A). However many issues complicate this scenario. First, outside of vertebrates within deuterostomes, only echinoderms have extensive mineralization of skeletal tissues, calling into question whether a mineralization GRN could be homologous among deuterostomes. Second, the form of mineral deposited in vertebrate skeletal tissues is mostly hydroxyapatite, which is a calcium phosphate crystal, while most mineral in non-vertebrate skeletal tissues is calcium carbonate-based (Wilt et al. 2003; Matsushiro and Miyashita 2004). Therefore, at least the portions of the mineralization GRN that drive biochemical synthesis of mineral would be different among echinoderms and vertebrates. Finally, there are few clear orthologs among protein components of the mineralized matrix in echinoderms and vertebrates. Matrix proteins in sea urchin teeth cross-reacted with antibodies to a variety of bone and tooth proteins in mammals (Veis et al. 2002). However most of the proteins associated with spicules (mineralized structures) in sea urchin are in the spicule matrix (SM) of the mesenchyme-specific cell surface glycoprotein (MSP) gene family, which have no orthologs in vertebrates (Livingston et al. 2006).

Despite these obstacles, proteins in the secretory calcium-binding phosphoprotein (SCPP) family provide an interesting story for the evolution of a mineralization GRN associated with bone formation. SCPP proteins bind to $Ca^{++}$ (when phosphorylated) and facilitate hydroxyapatite crystal nucleation or modulate its growth (Kawasaki et al. 2004). The SCPP proteins Matrix gla protein (MGP) and Bone gamma-carboxyglutamate protein (BGLAP, previously referred to as Osteocalcin) are proposed to play important roles in vertebrate mineralization, although loss-of-function models indicate that they function as inhibitors (Ducy et al. 1996; Luo et al. 1997). An SCPP protein, Secreted protein acidic and rich in cysteine (SPARC, previously called Osteonectin), is the most abundant non-collagenous protein in bone ECM (Termine et al. 1981). As opposed to MGP and BGLAP, SPARC loss-of-function models have osteopenia (Delany et al. 2000), suggesting perhaps SPARC is actually a positive regulator of mineralization. Regarding the significance to this hypothesis, SPARC is found in both protostomes and deuterostomes. The family of SCPP proteins is thought to have originated from tandem duplication of SPARC-like 1 (*SPARCL1*) in the ancestor to osteichthyans following their divergence from chondrichthyans (Kawasaki et al. 2004). Interestingly, a group of putative mineralization (spicule)-associated genes in sea urchin also underwent tandem duplications (Livingston et al. 2006). SPARC and a SPARCL were identified in sea urchin, although this SPARCL was not a clear ortholog to vertebrate SPARCL (Livingston et al.

2006). Amphioxus express SPARC/SPARCL in skeletogenic cells, even though they don't mineralize their tissues (Yong and Yu 2016).

Perhaps a remaining hope for identifying conservation among vertebrate and non-vertebrate mineralization GRNs focusses on regulatory genes upstream of the biochemical and secreted components (Livingston et al. 2006). Genes involved in sea urchin spicule differentiation include *pmar1* (represses *hairy*), *delta*, *ets*, *b-catenin/tcf*, and *cart1/alx3/alx4*(*alx1*; Wilt et al. 2003). Orthologs of these genes are expressed in bone, and mutations in many of which lead to skeletal patterning/mineralization defects in mouse (Beverdam et al. 2001; Mavrogiannis et al. 2001; Itoh et al. 2012; Li et al. 2018; Shao et al. 2018).

*VIIB. Bone evolved from a dentin/enamel GRN*

Dentin, enamel, and bone are closely related tissues, so they may share evolutionary histories. For example, an ancestral GRN gave rise to one of these tissues first, and then (perhaps after genome duplication events) this ancestral GRN was modified, giving rise to other mineralized tissues (Fisher and Franz-Odendaal 2012). According to one scenario, dentin- and enamel-like tissues first originated around the basement membrane of odontodes, and subsequent spread of mineralization deeper into the dermis gave rise to bone (Donoghue et al. 2006). Both dentin and bone are formed by NCCs (presumably even in odontodes of primitive vertebrates), so these cells could have expanded the spatial domain of the dentin GRN expression (deeper in the mesenchyme) within an odontode and modified that GRN to produce bone. Many facts support the hypothesis that a bone GRN evolved by modifications to a dentin GRN (Fig. 65B). Dentin and bone share many features, including a histologically similar structure and a dense, Col1-enriched matrix that is heavy calcified by hydroxyapatite (Goldberg et al. 2011). Runx2 is essential for both osteoblast and odontoblast differentiation, regulating the expression of many bone and tooth related genes (Chen et al. 2005). In transgenic mice that manipulate Runx2 expression levels specifically in odontoblasts, which normally down-regulate Runx2 after early stages of differentiation, osteoblast-like cells embedded in a bone-like matrix were seen instead of normal odontoblasts (Miyazaki et al. 2008; Komori 2010; Li et al. 2011). By demonstrating the similarity of the GRNs driving odontoblast and

osteoblast formation, these studies support feasibility of the hypothesis that one GRN evolved from the other.

Apart from embryological and histological differences, the bone GRN also might have evolved from an enamel GRN. Enamel is formed by ectoderm-derived ameloblasts, which secrete the enamel-specific matrix proteins Ameloblastin, Enamelin, and Amelogenin (Sire et al. 2007). Therefore, the hypothesis that the bone GRN evolved from the enamel GRN requires that NCCs co-opted expression of an enamel GRN operating in ectodermal cells and modified it to form bone. Runx2 is required for the proper differentiation and function of ameloblasts, and it is a key regulator of ameloblast-specific genes (Gaikwad et al. 2001; Chu et al. 2018). Tooth development in $Runx2^{-/-}$ mice is arrested before odontoblast and ameloblast formation in the developing tooth, perhaps due to a role for Runx2 during enamel knot formation (D'Souza et al. 1999; Camilleri and McDonald 2006). In later stages of enamel mineralization, Runx2 regulates odontogenic ameloblast-associated protein (ODAM) expression and amelotin promoter activity (Lee et al. 2010; Liu et al. 2018). Runx2 down-regulation might be involved in late stages of all mineralizing skeletal cells. Due to the number of enamel-specific proteins present in the enamel GRN, however the modifications required to transform it into the bone GRN presumably were more significant than the modifications required to transform the dentin GRN to one making bone (Fig. 65B).

Whether bone evolved from dentin or enamel (or *vice versa*) is still debated, but the two hypothetical scenarios on the evolution of the osteoblast (i.e. bone evolved from either dentin or enamel) presented above support the idea that a Runx2 GRN appeared *de novo* to mineralize early vertebrate tissues, independent of cartilage. These models are consistent with saltational evolution, in which large-scale genomic changes facilitate the evolution of novelty over a short period of time (Gould 2002).

*VIIC. Bone evolved from mature cartilage*

Recently, we proposed an alternative scenario for the appearance of bone that takes advantage of molecular genetic and developmental biology principles to offer a more parsimonious explanation than the saltational models outlined above (Gomez-Picos and Eames 2015). The fossil record generally supports the idea that bone evolved before mature/mineralized cartilage, but some

exceptions suggest that mineralized cartilage might have preceded bone (Smith and Hall 1990b; Janvier 1996b; Donoghue et al. 2006). For example, the fossil *Palaeospondylus gunni* (~385 Mya) has an entire endoskeleton composed of hypertrophic, mineralized cartilage but there is no evidence of bone in its skeleton (Johanson et al. 2010). In addition, it is possible that trace amounts of perichondral bone were lost during traditional fossil preparation techniques, in which case, only the more robust odontodes persisted as the first example of bone in vertebrate ancestors.

We hypothesized that the bone GRN evolved from a cartilage maturation GRN, and this assertion complements the previous saltational models by providing a gradualist evolutionary model (Gomez-Picos and Eames 2015; Fig. 65C). The many stages of cartilage maturation occurring today during endochondral ossification likely evolved gradually over time. Hypertrophy and mineralization appeared first, followed by matrix degradation, and finally vasculogenesis, fat formation, and endochondral bone deposition occurred (Hall 1975; Smith and Hall 1990b). In our evolutionary scenario, a Runx2-dependent GRN slowly evolved these various additional maturation stages to the Sox9 GRN that drives immature cartilage. Later, ectopic expression of this Runx2 GRN outside of cartilage resulted in the evolution of the bone GRN. For example, expansion of the Runx2 GRN expression domain slightly outside of the Sox9 GRN expression domain of cartilage might produce perichondral bone (Gomez-Picos and Eames 2015). In support of this hypothesis, some osteoblasts actually trans-differentiate from chondrocytes, showing the overlapping nature of the underlying GRNs (Hammond and Schulte-Merker 2009; Zhou et al. 2014; Park et al. 2015). In addition to providing strong molecular genetic support for the hypothesis that bone evolved from a mature cartilage GRN, these recent findings also fall in line with the old adage that ontogeny recapitulates phylogeny. Those osteoblasts that trans-differentiate from mature chondrocytes during development might reflect those exact events during evolution. Alternatively, expression of the cartilage maturation Runx2 GRN in dermal tissues might have produced odontode-associated bone.

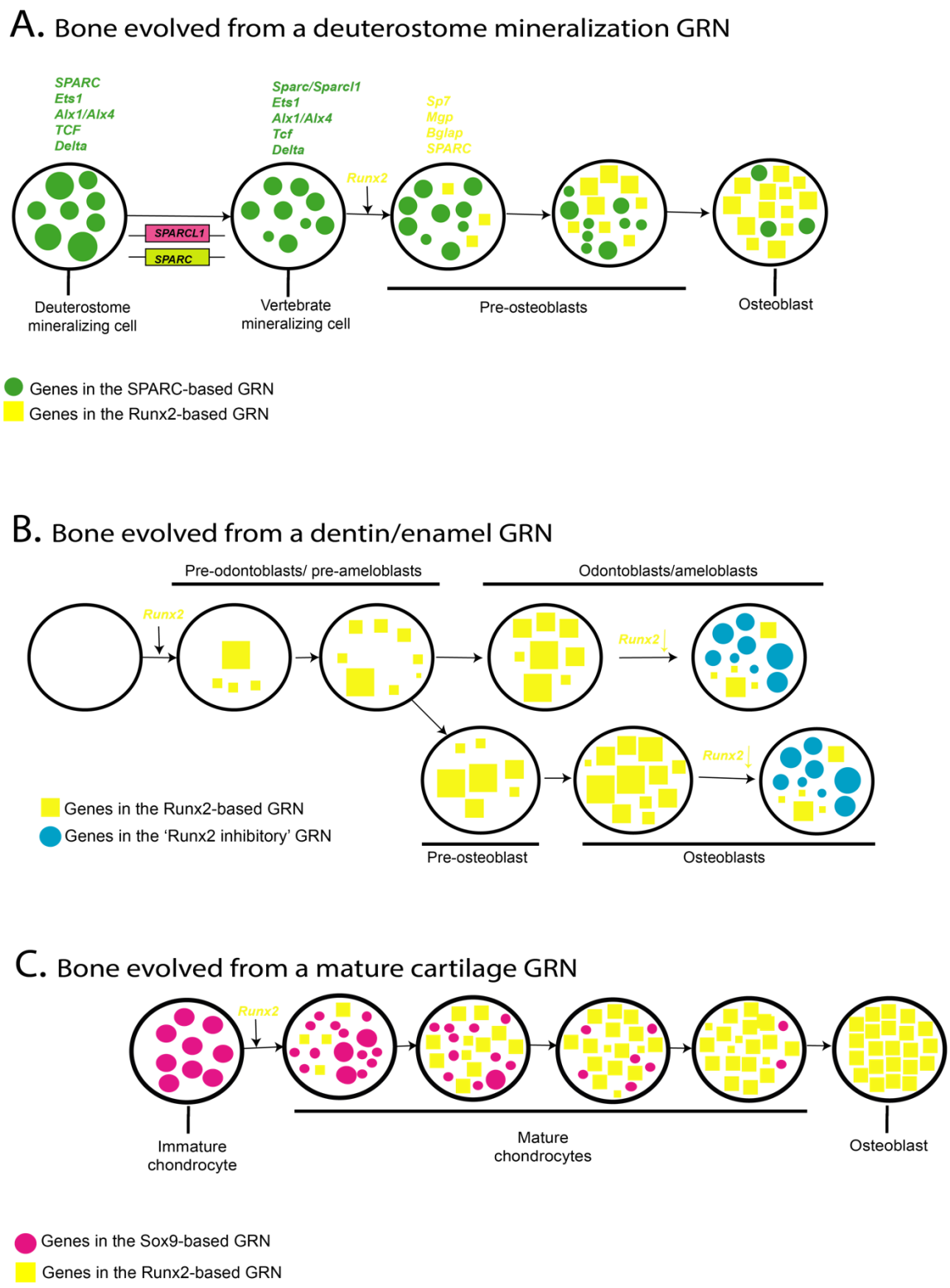



**Figure 65. Evolution of the GRN underlying bone.** A) Bone evolved from a deuterostome mineralization GRN. A mineralization GRN, likely involving *SPARC*, operating in mesenchymal cells of a deuterostome ancestor was co-opted and modified in NCCs of an ancestral vertebrate. This ancestral GRN contained many mineralization genes conserved within deuterostomes (i.e., *SPARC*, *Ets1*, *Alx1/Alx4*, *Tcf*, and *Delta*). During vertebrate evolution, a duplication event

of the ancestral *SPARC* gene generated two paralogs and gave rise to the *SCPP* gene family. Since Runx2 drives the expression of many mineralization genes in vertebrates (i.e. *Sp7*, *Bglap*, and *Mgp*), it is likely that a Runx2-mediated GRN was later established in these NCC populations to give rise to the osteoblast. B) Odontoblasts, ameloblasts, and osteoblasts are closely related cell types that may have derived one from another or from a common precursor. Runx2 is essential for odontoblast, ameloblast, and osteoblast differentiation, by regulating the expression of many bone- and tooth-related genes. During early differentiation of dentin/enamel, Runx2 is highly expressed in odontoblasts/ameloblasts, and then this transcription factor subsequently has to be downregulated at later stages of differentiation. According to this evolutionary scenario for the appearance of bone, another population of neural crest cells co-opted the dentin/enamel GRN (regulated by Runx2) to produce bone, the later stages of which also include Runx2 down-regulation. C) Immature chondrocytes, developing under a Sox9 GRN, provided a structural and molecular "buffer" for the gradual development of a Runx2 GRN. After establishment of the Runx2 GRN in mature chondrocytes, the osteoblast appeared when another population of cells co-opted this GRN.

Again, the philosophical advantages of this evolutionary hypothesis for the origin of bone is its gradualistic nature. Instead of requiring the *de novo,* saltational appearance of a Runx2 GRN driving bone formation, cartilage provided a "buffer" tissue in which a novel Runx2 GRN could evolve that drives production of a heavily mineralized ECM. Once this Runx2 GRN was established, the mis-expression of Runx2 in another tissue (NC-derived?) would have resulted in the first instance of bone in agnathan vertebrates. Perhaps the many intermediate forms of cartilage and bone highlighted earlier, such as chondroid bone and fibrocartilage, are remnants of the gradual elaboration of the Runx2 GRN as it modified immature cartilage. Interestingly, we note that osteoblasts of earlier-derived vertebrates, such as fish and frog, appear to express more "cartilage" genes than osteoblasts of later-derived vertebrates, such as chick and mouse (Eames and Helms 2004; Eames et al. 2012; Aldea et al. 2013).

## VIII. Conclusions

Traditional studies on non-model organisms, along with the advent of the modern molecular era, have revised considerably previous hypotheses on the origins of cartilage and bone, as well as the relationship of those events to the appearance of vertebrates. Studies of hemichordates and amphioxus have clearly demonstrated that an ability to form cartilage preceded vertebrates in the chordate ancestor. Bone and dentin are vertebrate-specific tissues, and their formation is tied intimately to evolved features of neural crest cells. Molecular identification of the GRNs underlying

the formation of skeletal tissues offers tantalizing clues about the molecular mechanisms underlying the evolution of cartilage and bone. The common theme for the evolutionary scenarios highlighted here is co-option of GRN expression. Once a GRN for differentiating a particular cell type was established, additional cell populations, sometimes from different germ layers, could co-opt expression of one or two genes sitting at the top of the GRN hierarchy. As a result, new cell types could evolve, or new modifications to that cell type could evolve. We are very excited that comparative transcriptomics, combined with functional genetics, can resolve among the evolutionary scenarios highlighted here, ultimately providing a clear story of how cartilage and bone evolved.

## REFERENCES


Abzhanov A, Rodda SJ, McMahon AP, Tabin CJ. 2007a. Regulation of skeletogenic differentiation in cranial dermal bone. *Development* **134**: 3133-3144.

Achim K, Arendt D. 2014. Structural evolution of cell types by step-wise assembly of cellular modules. *Curr Opin Genet Dev* **27**: 102-108.

Aghajanian P, Mohan S. 2018. The art of building bone: emerging role of chondrocyte-to-osteoblast transdifferentiation in endochondral ossification. *Bone Research* **6**: 19.

Aken BL, Ayling S, Barrell D, Clarke L, Curwen V, Fairley S, Fernandez Banet J, Billis K, García Girón C, Hourlier T et al. 2016. The Ensembl gene annotation system. *Database (Oxford)* **2016**: baw093.

Akiyama H, Chaboissier MC, Martin JF, 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 Dev* **16**: 2813-2828.

Akiyama R, Kawakami H, Wong J, Oishi I, Nishinakamura R, Kawakami Y. 2015. Sall4-Gli3 system in early limb progenitors is essential for the development of limb skeletal elements. *Proceedings of the National Academy of Sciences* **112**: 5075-5080.

Al-Amer O. 2017. Bone marker gene expression in calvarial bones: different bone microenvironments. *Journal of Biological Research-Thessaloniki* **24**: 9.

Albertson RC, Yan YL, Titus TA, Pisano E, Vacchi M, Yelick PC, Detrich HW, 3rd, Postlethwait JH. 2010. Molecular pedomorphism underlies craniofacial skeletal evolution in Antarctic notothenioid fishes. *BMC evolutionary biology* **10**: 4.

Aldea D, Hanna P, Munoz D, Espinoza J, Torrejon M, Sachs L, Buisine N, Oulion S, Escriva H, Marcellini S. 2013. Evolution of the vertebrate bone matrix: an expression analysis of the network forming collagen paralogues in amphibian osteoblasts. *J Exp Zool B Mol Dev Evol* **320**: 375-384.

Anders S, Pyl PT, Huber W. 2015. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**: 166-169.

Andrews SM, Miles RS, Walker AD. 1979. *Problems in Vertebrate Evolution: Essays Presented to Professor TS Westoll*. Academic Press for the Linnean Society of London.

Ahnesorg P, Smith P, Jackson SP. 2006. XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining. *Cell* **124**: 301-313.

Annona G, Holland ND, D'Aniello S. 2015. Evolution of the notochord. *Evodevo* **6**: 30.

Arendt D. 2003. Evolution of eyes and photoreceptor cell types. *Int J Dev Biol* **47**: 563-571.

Arendt D. 2005. Genes and homology in nervous system evolution: comparing gene functions, expression patterns, and cell type molecular fingerprints. *Theory Biosci* **124**: 185-197.

Arendt D. 2008. The evolution of cell types in animals: emerging principles from molecular studies. *Nat Rev Genet* **9**: 868-882.

Arendt D, Bertucci PY, Achim K, Musser JM. 2019. Evolution of neuronal types and families. *Current Opinion in Neurobiology* **56**: 144-152.

Arnold MA, Kim Y, Czubryt MP, Phan D, McAnally J, Qi X, Shelton JM, Richardson JA, Bassel-Duby R, Olson EN. 2007. MEF2C transcription factor controls chondrocyte hypertrophy and bone development. *Developmental cell* **12**: 377-389.

Arnold SJ. 1992. Constraints on phenotypic evolution. *Am Nat* **140 Suppl 1**: S85-107.

Arratia G, Schultze HP, Casciotta J. 2001. Vertebral column and associated elements in dipnoans and comparison with other fishes: development and homology. *J Morphol* **250**: 101-172.

Askary A, Mork L, Paul S, He X, Izuhara AK, Gopalakrishnan S, Ichida JK, McMahon AP, Dabizljevic S, Dale R et al. 2015. Iroquois Proteins Promote Skeletal Joint Formation by Maintaining Chondrocytes in an Immature State. *Developmental cell* **35**: 358-365.

Aslam F, McCabe L, Frenkel B, van Wijnen AJ, Stein GS, Lian JB, Stein JL. 1999. AP-1 and vitamin D receptor (VDR) signaling pathways converge at the rat osteocalcin VDR element: requirement for the internal activating protein-1 site for vitamin D-mediated trans-activation. *Endocrinology* **140**: 63-70.

Atake OJ, Cooper DML, Eames BF. 2019. Bone-like features in skate suggest a novel elasmobranch synapomorphy and deep homology of trabecular mineralization patterns. *Acta biomaterialia* **84**: 424-436.

Aumailley M, Gayraud B. 1998. Structure and biological activity of the extracellular matrix. *J Mol Med (Berl)* **76**: 253-265.

Azpeitia E, Benitez M, Vega I, Villarreal C, Alvarez-Buylla ER. 2010. Single-cell and coupled GRN models of cell patterning in the Arabidopsis thaliana root stem cell niche. *BMC Syst Biol* **4**: 134.

Bagheri-Fam S, Barrionuevo F, Dohrmann U, Gunther T, Schule R, Kemler R, Mallo M, Kanzler B, Scherer G. 2006. Long-range upstream and downstream enhancers control distinct subsets of the complex spatiotemporal Sox9 expression pattern. *Developmental biology* **291**: 382-397.

Bairati A, Comazzi M, Gioria M, Rigo C. 1998. The ultrastructure of chondrocytes in the cartilage of Sepia officinalis and Octopus vulgaris (Mollusca, Cephalopoda). *Tissue Cell* **30**: 340-351.

Baker CV, Bronner-Fraser M. 1997. The origins of the neural crest. Part II: an evolutionary perspective. *Mech Dev* **69**: 13-29.

Bardack D, Zangerl R. 1968. First fossil lamprey: a record from the Pennsylvanian of Illinois. *Science* **162**: 1265-1267.

Barrionuevo F, Taketo MM, Scherer G, Kispert A. 2006. Sox9 is required for notochord maintenance in mice. *Developmental biology* **295**: 128-140.

Bauer O, Sharir A, Kimura A, Hantisteanu S, Takeda S, Groner Y. 2015. Loss of Osteoblast Runx3 Produces Severe Congenital Osteopenia. *Molecular and Cellular Biology* **35**: 1097-1109.

Bayliss MT, Osborne D, Woodhouse S, Davidson C. 1999. Sulfation of chondroitin sulfate in human articular cartilage. The effect of age, topographical position, and zone of cartilage on tissue composition. *J Biol Chem* **274**: 15892-15900.

Bell DM, Leung KK, Wheatley SC, Ng LJ, Zhou S, Ling KW, Sham MH, Koopman P, Tam PP, Cheah KS. 1997. SOX9 directly regulates the type-II collagen gene. *Nat Genet* **16**: 174-178.

Belluoccio D, Bernardo BC, Rowley L, Bateman JF. 2008. A microarray approach for comparative expression profiling of the discrete maturation zones of mouse growth plate cartilage. *Biochim Biophys Acta* **1779**: 330-340.

Belogortseva N, Krezalek M, Guyton K, Labno C, Poroyko V, Zaborina O, Alverdy JC. 2017. Media from macrophages co-incubated with Enterococcus faecalis induces epithelial cell monolayer reassembly and altered cell morphology. *PLOS ONE* **12**: e0182825.

Benest AV, Harper SJ, Herttuala SY, Alitalo K, Bates DO. 2007. VEGF-C induced angiogenesis preferentially occurs at a distance from lymphangiogenesis. *Cardiovascular Research* **78**: 315-323.

Benjamin M. 1989. Hyaline-cell cartilage (chondroid) in the heads of teleosts. *Anat Embryol (Berl)* **179**: 285-303.
Benjamin M. 1990. The cranial cartilages of teleosts and their classification. *J Anat* **169**: 153-172.
Benjamin M, Evans EJ. 1990. Fibrocartilage. *J Anat* **171**: 1-15.
Benjamin M, Ralphs JR. 1991. Extracellular matrix of connective tissues in the heads of teleosts. *J Anat* **179**: 137-148.
Benjamin M, Ralphs JR, Eberewariye OS. 1992. Cartilage and related tissues in the trunk and fins of teleosts. *J Anat* **181 ( Pt 1)**: 113-118.
Beresford WA. 1981. *Chondroid bone, secondary cartilage, and metaplasia*. Urban & Schwarzenberg, Baltimore, Md.
Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. 2000. The Protein Data Bank. *Nucleic Acids Research* **28**: 235-242.
Bertin A, Hanna P, Otarola G, Fritz A, Henriquez JP, Marcellini S. 2015. Cellular and molecular characterization of a novel primary osteoblast culture from the vertebrate model organism Xenopus tropicalis. *Histochemistry and Cell Biology* **143**: 431-442.
Bertrand S, Fuentealba J, Aze A, Hudson C, Yasuo H, Torrejon M, Escriva H, Marcellini S. 2013. A dynamic history of gene duplications and losses characterizes the evolution of the SPARC family in eumetazoans. *Proceedings Biological sciences* **280**: 20122963.
Beverdam A, Brouwer A, Reijnen M, Korving J, Meijlink F. 2001. Severe nasal clefting and abnormal embryonic apoptosis in Alx3/Alx4 double mutant mice. *Development* **128**: 3975-3986.
Bhatt S, Diaz R, Trainor PA. 2013. Signals and switches in Mammalian neural crest cell differentiation. *Cold Spring Harbor perspectives in biology* **5**: a008326.
Bhattacherjee V, Mukhopadhyay P, Singh S, Johnson C, Philipose JT, Warner CP, Greene RM, Pisano MM. 2007. Neural crest and mesoderm lineage-dependent gene expression in orofacial development. *Differentiation* **75**: 463-477.
Bi W, Deng JM, Zhang Z, Behringer RR, de Crombrugghe B. 1999a. Sox9 is required for cartilage formation. *Nature genetics* **22**: 85-89.
Bi W, Deng JM, Zhang Z, Behringer RR, de Crombrugghe B. 1999b. Sox9 is required for cartilage formation. *Nat Genet* **22**: 85-89.
Bialek P, Kern B, Yang X, Schrock M, Sosic D, Hong N, Wu H, Yu K, Ornitz DM, Olson EN et al. 2004. A twist code determines the onset of osteoblast differentiation. *Dev Cell* **6**: 423-435.
Blumer MJ, Longato S, Richter E, Perez MT, Konakci KZ, Fritsch H. 2005. The role of cartilage canals in endochondral and perichondral bone formation: are there similarities between these two processes? *J Anat* **206**: 359-372.
Blumer MJF, Longato S, Fritsch H. 2004. Cartilage canals in the chicken embryo are involved in the process of endochondral bone formation within the epiphyseal growth plate. *The Anatomical Record Part A: Discoveries in Molecular, Cellular, and Evolutionary Biology* **279A**: 692-700.
Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114-2120.
Bonczek O, Balcar VJ, Sery O. 2017. PAX9 gene mutations and tooth agenesis: A review. *Clin Genet* **92**: 467-476.
Bond SR, Lau A, Penuela S, Sampaio AV, Underhill TM, Laird DW, Naus CC. 2011. Pannexin 3 is a novel target for Runx2, expressed by osteoblasts and mature growth plate chondrocytes. *J Bone Miner Res* **26**: 2911-2922.

Bonilla-Claudio M, Wang J, Bai Y, Klysik E, Selever J, Martin JF. 2012. Bmp signaling regulates a dose-dependent transcriptional program to control facial skeletal development. *Development* **139**: 709-719.

Bonucci E, Silvestrini G, Bianco P. 1992. Extracellular alkaline phosphatase activity in mineralizing matrices of cartilage and bone: ultrastructural localization using a cerium-based method. *Histochemistry* **97**: 323-327.

Bonyadi Rad E, Musumeci G, Pichler K, Heidary M, Szychlinska MA, Castrogiovanni P, Marth E, Böhm C, Srinivasaiah S, Krönke G et al. 2017. Runx2 mediated Induction of Novel Targets ST2 and Runx3 Leads to Cooperative Regulation of Hypertrophic Differentiation in ATDC5 Chondrocytes. *Scientific Reports* **7**: 17947.

Bordat C. 1987. Ultrastructural study of the vertebrae of the selachian *Scyliorhinus canicula*. *Canadian Journal of Zoology* **65**: 1435-1444.

Bosch Grau M, Gonzalez Curto G, Rocha C, Magiera MM, Marques Sousa P, Giordano T, Spassky N, Janke C. 2013. Tubulin glycylases and glutamylases have distinct functions in stabilization and motility of ependymal cilia. *J Cell Biol* **202**: 441-451.

Botelho JF, Smith-Paredes D, Soto-Acuña S, O'Connor J, Palma V, Vargas AO. 2016. Molecular development of fibular reduction in birds and its evolution from dinosaurs. *Evolution* **70**: 543-554.

Boumah CE, Lee M, Selvamurugan N, Shimizu E, Partridge NC. 2009. Runx2 recruits p300 to mediate parathyroid hormone's effects on histone acetylation and transcriptional activation of the matrix metalloproteinase-13 gene. *Mol Endocrinol* **23**: 1255-1263.

Boyan BD, Olivares-Navarrete R, Berger MB, Hyzy SL, Schwartz Z. 2018. Role of Wnt11 during Osteogenic Differentiation of Human Mesenchymal Stem Cells on Microstructured Titanium Surfaces. *Scientific Reports* **8**: 8588.

Braasch I, Gehrke AR, Smith JJ, Kawasaki K, Manousaki T, Pasquier J, Amores A, Desvignes T, Batzel P, Catchen J et al. 2016. The spotted gar genome illuminates vertebrate evolution and facilitates human-teleost comparisons. *Nat Genet* **48**: 427-437.

Brawand D, Soumillon M, Necsulea A, Julien P, Csardi G, Harrigan P, Weier M, Liechti A, Aximu-Petri A, Kircher M et al. 2011. The evolution of gene expression levels in mammalian organs. *Nature* **478**: 343-348.

Brekke TD, Henry LA, Good JM. 2016. Genomic imprinting, disrupted placental expression, and speciation. *Evolution* **70**: 2690-2703.

Bridgewater LC, Lefebvre V, de Crombrugghe B. 1998. Chondrocyte-specific enhancer elements in the Col11a2 gene resemble the Col2a1 tissue-specific enhancer. *J Biol Chem* **273**: 14998-15006.

Bridgewater LC, Walker MD, Miller GC, Ellison TA, Holsinger LD, Potter JL, Jackson TL, Chen RK, Winkel VL, Zhang Z et al. 2003. Adjacent DNA sequences modulate Sox9 transcriptional activation at paired Sox sites in three chondrocyte-specific enhancer elements. *Nucleic Acids Res* **31**: 1541-1553.

Britton CS, Sorrells TR, Johnson AD. 2020. Protein-coding changes preceded cis-regulatory gains in a newly evolved transcription circuit. *Science* **367**: 96-100.

Brown RM, Davis MG, Hayashi K, MacLean JA. 2013. Regulated expression of Rhox8 in the mouse ovary: evidence for the role of progesterone and RHOX5 in granulosa cells. *Biol Reprod* **88**: 126.

Bruyninx M, Hennuy B, Cornet A, Houssa P, Daukandt M, Reiter E, Poncin J, Closset J, Hennen G. 1999. A novel gene overexpressed in the prostate of castrated rats: hormonal regulation,

relationship to apoptosis and to acquired prostatic cell androgen independence. *Endocrinology* **140**: 4789-4799.

Buchholz DR. 2015. More similar than you think: Frog metamorphosis as a model of human perinatal endocrinology. *Developmental Biology* **408**: 188-195.

Buckwalter JA, Rosenberg LC, Ungar R. 1987. Changes in proteoglycan aggregates during cartilage mineralization. *Calcif Tissue Int* **41**: 228-236.

Buckwalter JA, Mower D, Ungar R, Schaeffer J, Ginsberg B. 1986. Morphometric analysis of chondrocyte hypertrophy. *The Journal of bone and joint surgery American volume* **68**: 243-255.

Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. 2015. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol* **109**: 21.29.21-21.29.29.

Burdan F. 2005. [Development and congenital malformations of the skeleton]. *Polski merkuriusz lekarski : organ Polskiego Towarzystwa Lekarskiego* **19**: 94-97.

Cain CJ, Gaborit N, Lwin W, Barruet E, Ho S, Bonnard C, Hamamy H, Shboul M, Reversade B, Kayserili H et al. 2016. Loss of Iroquois homeobox transcription factors 3 and 5 in osteoblasts disrupts cranial mineralization. *Bone Rep* **5**: 86-95.

Calabrese G, Bennett BJ, Orozco L, Kang HM, Eskin E, Dombret C, De Backer O, Lusis AJ, Farber CR. 2012. Systems Genetic Analysis of Osteoblast-Lineage Cells. *PLOS Genetics* **8**: e1003150.

Cameron SL, Lo N, Bourguignon T, Svenson GJ, Evans TA. 2012. A mitochondrial genome phylogeny of termites (Blattodea: Termitoidae): Robust support for interfamilial relationships and molecular synapomorphies define major clades. *Molecular Phylogenetics and Evolution* **65**: 163-173.

Camilleri S, McDonald F. 2006. Runx2 and dental development. *Eur J Oral Sci* **114**: 361-373.

Campo RD, Romano JE. 1986. Changes in cartilage proteoglycans associated with calcification. *Calcif Tissue Int* **39**: 175-184.

Capa M, Nogueira JM, Rossi MC. 2011. Comparative internal structure of dorsal lips and radiolar appendages in Sabellidae (Polychaeta) and phylogenetic implications. *J Morphol* **272**: 302-319.

Carlsen S, Nandakumar KS, Holmdahl R. 2006. Type IX collagen deficiency enhances the binding of cartilage-specific antibodies and arthritis severity. *Arthritis Research & Therapy* **8**: R102.

Cattell M, Lai S, Cerny R, Medeiros DM. 2011. A new mechanistic scenario for the origin and evolution of vertebrate cartilage. *PLoS One* **6**: e22474.

Cervantes-Diaz F, Contreras P, Marcellini S. 2017. Evolutionary origin of endochondral ossification: the transdifferentiation hypothesis. *Dev Genes Evol* **227**: 121-127.

Chai Y, Jiang X, Ito Y, Bringas P, Jr., Han J, Rowitch DH, Soriano P, McMahon AP, Sucov HM. 2000. Fate of the mammalian cranial neural crest during tooth and mandibular morphogenesis. *Development* **127**: 1671-1679.

Chan ET, Quon GT, Chua G, Babak T, Trochesset M, Zirngibl RA, Aubin J, Ratcliffe MJ, Wilde A, Brudno M et al. 2009. Conservation of core gene expression in vertebrate tissues. *J Biol* **8**: 33.

Chang SF, Chang TK, Peng HH, Yeh YT, Lee DY, Yeh CR, Zhou J, Cheng CK, Chang CA, Chiu JJ. 2009. BMP-4 induction of arrest and differentiation of osteoblast-like cells via p21 CIP1 and p27 KIP1 regulation. *Mol Endocrinol* **23**: 1827-1838.

Charoenlarp P, Rajendran AK, Iseki S. 2017. Role of fibroblast growth factors in bone regeneration. *Inflamm Regen* **37**: 10-10.

Chau M, Lui JC, Landman EB, Spath SS, Vortkamp A, Baron J, Nilsson O. 2014. Gene expression profiling reveals similarities between the spatial architectures of postnatal articular and growth plate cartilage. *PLoS One* **9**: e103061.

Chen J-Y, Huang D-Y, Li C-W. 1999. An early Cambrian craniate-like chordate. *Nature* **402**: 518-522.

Chen J, Long F. 2013. beta-catenin promotes bone formation and suppresses bone resorption in postnatal growing mice. *J Bone Miner Res* **28**: 1160-1169.

Chen J, Zhang Q, McCulloch CA, Sodek J. 1991. Immunohistochemical localization of bone sialoprotein in foetal porcine bone tissues: comparisons with secreted phosphoprotein 1 (SPP-1, osteopontin) and SPARC (osteonectin). *Histochem J* **23**: 281-289.

Chen S, Santos L, Wu Y, Vuong R, Gay I, Schulze J, Chuang HH, MacDougall M. 2005. Altered gene expression in human cleidocranial dysplasia dental pulp cells. *Arch Oral Biol* **50**: 227-236.

Cheng A, Genever PG. 2010. SOX9 determines RUNX2 transactivity by directing intracellular degradation. *J Bone Miner Res* **25**: 2680-2689.

Choesmel V, Foucault F, Thiery JP, Blin N. 2004. Design of a real time quantitative PCR assay to assess global mRNA amplification of small size specimens for microarray hybridisation. *Journal of Clinical Pathology* **57**: 1278-1287.

Choudhuri S. 2014. Chapter 2 - Fundamentals of Molecular Evolution**The opinions expressed in this chapter are the author's own and they do not necessarily reflect the opinions of the FDA, the DHHS, or the Federal Government. In *Bioinformatics for Beginners*, doi:https://doi.org/10.1016/B978-0-12-410471-6.00002-5 (ed. S Choudhuri), pp. 27-53. Academic Press, Oxford.

Chu Q, Gao Y, Gao X, Dong Z, Song W, Xu Z, Xiang L, Wang Y, Zhang L, Li M et al. 2018. Ablation of Runx2 in Ameloblasts Suppresses Enamel Maturation in Tooth Development. *Sci Rep* **8**: 9594.

Chuang LSH, Lai SK, Murata-Hori M, Yamada A, Li H-Y, Gunaratne J, Ito Y. 2012. RUNX3 interactome reveals novel centrosomal targeting of RUNX family of transcription factors. *Cell Cycle* **11**: 1938-1947.

Clark CD, Zhang B, Lee B, Evans SI, Lassar AB, Lee KH. 2013. Evolutionary conservation of Nkx2.5 autoregulation in the second heart field. *Dev Biol* **374**: 198-209.

Clark CT, Smith KK. 1993. Cranial osteogenesis in Monodelphis domestica (Didelphidae) and Macropus eugenii (Macropodidae). *J Morphol* **215**: 119-149.

Coates MI, Finarelli JA, Sansom IJ, Andreev PS, Criswell KE, Tietjen K, Rivers ML, La Riviere PJ. 2018. An early chondrichthyan and the evolutionary assembly of a shark body plan. *Proc Biol Sci* **285**.

Coates MI, Sequeira SEK, Sansom IJ, Smith MM. 1998. Spines and tissues of ancient sharks. *Nature* **396**: 729-730.

Coble DW. 1966. Alkaline Phosphatase in Fish Scales. *Journal of the Fisheries Research Board of Canada* **23**: 149-152.

Colden M, Dar AA, Saini S, Dahiya PV, Shahryari V, Yamamura S, Tanaka Y, Stein G, Dahiya R, Majid S. 2017. MicroRNA-466 inhibits tumor growth and bone metastasis in prostate cancer by direct regulation of osteogenic transcription factor RUNX2. *Cell Death Dis* **8**: e2572.

Cole AG. 2011. A review of diversity in the evolution and development of cartilage: the search for the origin of the chondrocyte. *Eur Cell Mater* **21**: 122-129.

Cole AG, Hall BK. 2004a. Cartilage is a metazoan tissue; integrating data from nonvertebrate sources. *Acta Zoologica* **85**: 69-80.

Cole AG, Hall BK. 2004b. The nature and significance of invertebrate cartilages revisited: distribution and histology of cartilage and cartilage-like tissues within the Metazoa. *Zoology (Jena)* **107**: 261-273.

Conen KL, Nishimori S, Provot S, Kronenberg HM. 2009. The transcriptional cofactor Lbh regulates angiogenesis and endochondral bone formation during fetal bone development. *Developmental Biology* **333**: 348-358.

Conner JR, Hornick JL. 2013. SATB2 is a novel marker of osteoblastic differentiation in bone and soft tissue tumours. *Histopathology* **63**: 36-49.

Cooper KL, Oh S, Sung Y, Dasari RR, Kirschner MW, Tabin CJ. 2013. Multiple phases of chondrocyte enlargement underlie differences in skeletal proportions. *Nature* **495**: 375-378.

Corallo D, Trapani V, Bonaldo P. 2015. The notochord: structure and functions. *Cell Mol Life Sci* **72**: 2989-3008.

Couly G, Creuzet S, Bennaceur S, Vincent C, Le Douarin NM. 2002. Interactions between Hox-negative cephalic neural crest cells and the foregut endoderm in patterning the facial skeleton in the vertebrate head. *Development* **129**: 1061-1073.

Couly GF, Coltey PM, Douarin NML. 1992. The developmental fate of the cephalic mesoderm in quail-chick chimeras. *Development* **114**: 1-15.

Couly GF, Coltey PM, Le Douarin NM. 1993. The triple origin of skull in higher vertebrates: a study in quail-chick chimeras. *Development* **117**: 409-429.

Criswell KE, Coates MI, Gillis JA. 2017a. Embryonic development of the axial column in the little skate, Leucoraja erinacea. *J Morphol* **278**: 300-320.

Criswell KE, Coates MI, Gillis JA. 2017b. Embryonic origin of the gnathostome vertebral skeleton. *Proc Biol Sci* **284**.

Cunliffe VT. 2003. Zebrafish: A Practical Approach. Edited by C. NÜSSLEIN-VOLHARD and R. DAHM. Oxford University Press. 2002. 322 pages. ISBN 0 19 963808 X. Price £40.00 (paperback). ISBN 0 19 963809 8. Price £80.00 (hardback). *Genetical Research* **82**: 79-79.

D'Souza RN, Aberg T, Gaikwad J, Cavender A, Owen M, Karsenty G, Thesleff I. 1999. Cbfa1 is required for epithelial-mesenchymal interactions regulating tooth development in mice. *Development* **126**: 2911-2920.

Dale RM, Topczewski J. 2011a. Identification of an evolutionarily conserved regulatory element of the zebrafish col2a1a gene. *Developmental biology* **357**: 518-531.

Dale RM, Topczewski J. 2011b. Identification of an evolutionarily conserved regulatory element of the zebrafish col2a1a gene. *Developmental biology* **357**: 518-531.

Daniel JF. 1926. *The lateral blood supply of primitive elasmobranch fishes*. University of California Press.

Daniel JF. 1934. *The Elasmobranch Fishes*. University of California Press, Berkeley.

Datta HK, Ng WF, Walker JA, Tuck SP, Varanasi SS. 2008. The cell biology of bone metabolism. *J Clin Pathol* **61**: 577-587.

Davidson EH, Erwin DH. 2006. Gene regulatory networks and the evolution of animal body plans. *Science* **311**: 796-800.

Davidson EH, Levine MS. 2008. Properties of developmental gene regulatory networks. *Proc Natl Acad Sci U S A* **105**: 20063-20066.

Davidson EH, McClay DR, Hood L. 2003. Regulatory gene networks and the properties of the developmental process. *Proc Natl Acad Sci U S A* **100**: 1475-1480.

Day TF, Guo X, Garrett-Beal L, Yang Y. 2005. Wnt/beta-catenin signaling in mesenchymal progenitors controls osteoblast and chondrocyte differentiation during vertebrate skeletogenesis. *Developmental cell* **8**: 739-750.

de Crombrugghe B, Lefebvre V, Behringer RR, Bi W, Murakami S, Huang W. 2000. Transcriptional mechanisms of chondrocyte differentiation. *Matrix biology : journal of the International Society for Matrix Biology* **19**: 389-394.

Dean MN, Summers AP. 2006. Mineralized cartilage in the skeleton of chondrichthyan fishes. *Zoology (Jena)* **109**: 164-168.

Decker RS, Koyama E, Pacifici M. 2015. Articular Cartilage: Structural and Developmental Intricacies and Questions. *Current osteoporosis reports* **13**: 407-414.

Delany AM, Amling M, Priemel M, Howe C, Baron R, Canalis E. 2000. Osteopenia and decreased bone formation in osteonectin-deficient mice. *J Clin Invest* **105**: 915-923.

Delsuc F, Brinkmann H, Chourrout D, Philippe H. 2006. Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature* **439**: 965-968.

Dobreva G, Chahrour M, Dautzenberg M, Chirivella L, Kanzler B, Fariñas I, Karsenty G, Grosschedl R. 2006. SATB2 Is a Multifunctional Determinant of Craniofacial Patterning and Osteoblast Differentiation. *Cell* **125**: 971-986.

Dodig M, Tadic T, Kronenberg MS, Dacic S, Liu YH, Maxson R, Rowe DW, Lichtler AC. 1999. Ectopic Msx2 overexpression inhibits and Msx2 antisense stimulates calvarial osteoblast differentiation. *Developmental biology* **209**: 298-307.

Donoghue PC. 1998. Growth and patterning in the conodont skeleton. *Philos Trans R Soc Lond* **363**: 633-666.

Donoghue PC, Forey PL, Aldridge RJ. 2000. Conodont affinity and chordate phylogeny. *Biol Rev Camb Philos Soc* **75**: 191-251.

Donoghue PC, Graham A, Kelsh RN. 2008. The origin and evolution of the neural crest. *Bioessays* **30**: 530-541.

Donoghue PC, Rucklin M. 2016. The ins and outs of the evolutionary origin of teeth. *Evol Dev* **18**: 19-30.

Donoghue PC, Sansom IJ. 2002. Origin and early evolution of vertebrate skeletonization. *Microsc Res Tech* **59**: 352-372.

Donoghue PC, Sansom IJ, Downs JP. 2006. Early evolution of vertebrate skeletal tissues and cellular interactions, and the canalization of skeletal development. *J Exp Zool B Mol Dev Evol* **306**: 278-294.

Drabek K, van de Peppel J, Eijken M, van Leeuwen JP. 2011. GPM6B regulates osteoblast function and induction of mineralization by controlling cytoskeleton and matrix vesicle release. *J Bone Miner Res* **26**: 2045-2051.

Drissi H, Luc Q, Shakoori R, Chuva De Sousa Lopes S, Choi JY, Terry A, Hu M, Jones S, Neil JC, Lian JB et al. 2000. Transcriptional autoregulation of the bone related CBFA1/RUNX2 gene. *J Cell Physiol* **184**: 341-350.

Ducy P, Desbois C, Boyce B, Pinero G, Story B, Dunstan C, Smith E, Bonadio J, Goldstein S, Gundberg C et al. 1996. Increased bone formation in osteocalcin-deficient mice. *Nature* **382**: 448-452.

Ducy P, Zhang R, Geoffroy V, Ridall AL, Karsenty G. 1997. Osf2/Cbfa1: a transcriptional activator of osteoblast differentiation. *Cell* **89**: 747-754.

Dudek M, Gossan N, Yang N, Im HJ, Ruckshanthi JP, Yoshitane H, Li X, Jin D, Wang P, Boudiffa M et al. 2016. The chondrocyte clock gene Bmal1 controls cartilage homeostasis and integrity. *J Clin Invest* **126**: 365-376.

Dujon B. 1996. The yeast genome project: what did we learn? *Trends Genet* **12**: 263-270.

Dunlop LL, Hall BK. 1995. Relationships between cellular condensation, preosteoblast formation and epithelial-mesenchymal interactions in initiation of osteogenesis. *Int J Dev Biol* **39**: 357-371.

Dunn CW, Giribet G, Edgecombe GD, Hejnol A. 2014. Animal Phylogeny and Its Evolutionary Implications. **45**: 371-395.

Dy P, Wang W, Bhattaram P, Wang Q, Wang L, Ballock RT, Lefebvre V. 2012. Sox9 directs hypertrophic maturation and blocks osteoblast differentiation of growth plate chondrocytes. *Developmental cell* **22**: 597-609.

Eames BF, Allen N, Young J, Kaplan A, Helms JA, Schneider RA. 2007. Skeletogenesis in the swell shark Cephaloscyllium ventriosum. *J Anat* **210**: 542-554.

Eames BF, Amores A, Yan YL, Postlethwait JH. 2012. Evolution of the osteoblast: skeletogenesis in gar and zebrafish. *BMC evolutionary biology* **12**: 27.

Eames BF, de la Fuente L, Helms JA. 2003. Molecular ontogeny of the skeleton. *Birth Defects Res C Embryo Today* **69**: 93-101.

Eames BF, Helms JA. 2004. Conserved molecular program regulating cranial and appendicular skeletogenesis. *Dev Dyn* **231**: 4-13.

Eames BF, Sharpe PT, Helms JA. 2004. Hierarchy revealed in the specification of three skeletal fates by Sox9 and Runx2. *Dev Biol* **274**: 188-200.

Eames BF, Yan YL, Swartz ME, Levic DS, Knapik EW, Postlethwait JH, Kimmel CB. 2011. Mutations in fam20b and xylt1 reveal that cartilage matrix controls timing of endochondral ossification by inhibiting chondrocyte maturation. *PLoS Genet* **7**: e1002246.

Edgar R, Domrachev M, Lash AE. 2002. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* **30**: 207-210.

Enault S, Munoz DN, Silva WT, Borday-Birraux V, Bonade M, Oulion S, Venteo S, Marcellini S, Debiais-Thibaud M. 2015b. Molecular footprinting of skeletal tissues in the catshark Scyliorhinus canicula and the clawed frog Xenopus tropicalis identifies conserved and derived features of vertebrate calcification. *Front Genet* **6**: 283.

Enomoto H, Enomoto-Iwamoto M, Iwamoto M, Nomura S, Himeno M, Kitamura Y, Kishimoto T, Komori T. 2000. Cbfa1 is a positive regulatory factor in chondrocyte maturation. *J Biol Chem* **275**: 8695-8702.

Erwin DH. 2019. Prospects for a General Theory of Evolutionary Novelty. *Journal of Computational Biology* **26**: 735-744.

Espinoza J, Sanchez M, Sanchez A, Hanna P, Torrejon M, Buisine N, Sachs L, Marcellini S. 2010. Two families of Xenopus tropicalis skeletal genes display well-conserved expression patterns with mammals in spite of their highly divergent regulatory regions. *Evolution & development* **12**: 541-551.

Fan X, Loebel D, Bildsoe H, Wilkie E, Qin J, Wang J, Tam P. 2016. Tissue interactions, cell signaling and transcriptional control in the cranial mesoderm during craniofacial development. *AIMS Genetics* **3**: 74-98.

Fang J, Hall BK. 1997. Chondrogenic cell differentiation from membrane bone periostea. *Anat Embryol (Berl)* **196**: 349-362.

Farnum CE, Lee R, O'Hara K, Urban JP. 2002. Volume increase in growth plate chondrocytes during hypertrophy: the contribution of organic osmolytes. *Bone* **30**: 574-581.
Farquharson C, Whitehead CC, Loveridge N. 1994. Alterations in glycosaminoglycan concentration and sulfation during chondrocyte maturation. *Calcified Tissue International* **54**: 296-303.
Feldman AL, Costouros NG, Wang E, Qian M, Marincola FM, Alexander HR, Libutti SK. 2002. Advantages of mRNA amplification for microarray analysis. *BioTechniques* **33**: 906-912, 914.
Ferguson CM, Miclau T, Hu D, Alpern E, Helms JA. 1998. Common molecular pathways in skeletal morphogenesis and repair. *Ann N Y Acad Sci* **857**: 33-42.
Ferrari D, Kosher RA. 2002. Dlx5 is a positive regulator of chondrocyte differentiation during endochondral ossification. *Developmental biology* **252**: 257-270.
Fisher S, Franz-Odendaal T. 2012. Evolution of the bone gene regulatory network. *Curr Opin Genet Dev* **22**: 390-397.
Fonseca BF, Couly G, Dupin E. 2017. Respective contribution of the cephalic neural crest and mesoderm to SIX1-expressing head territories in the avian embryo. *BMC developmental biology* **17**: 13.
Frey JL, Kim SP, Li Z, Wolfgang MJ, Riddle RC. 2017. β-Catenin Directs Long-Chain Fatty Acid Catabolism in the Osteoblasts of Male Mice. *Endocrinology* **159**: 272-284.
Fung Ling Chau J, Fook Leong W, Li B. 2009. Signaling pathways governing osteoblast proliferation, differentiation and function. *Histology and histopathology*.
Fyhrie DP, Christiansen BA. 2015. Bone Material Properties and Skeletal Fragility. *Calcified tissue international* **97**: 213-228.
Gaikwad JS, Cavender A, D'Souza RN. 2001. Identification of tooth-specific downstream targets of Runx2. *Gene* **279**: 91-97.
Galli GG, Honnens de Lichtenberg K, Carrara M, Hans W, Wuelling M, Mentz B, Multhaupt HA, Fog CK, Jensen KT, Rappsilber J et al. 2012. Prdm5 regulates collagen gene transcription by association with RNA polymerase II in developing bone. *PLoS Genet* **8**: e1002711.
Gans C, Northcutt R. 1983. Neural Crest and the Origin of Vertebrates: A New Head. *Science* **220**: 268-274.
Gentili C, Cancedda R. 2009. Cartilage and bone extracellular matrix. *Curr Pharm Des* **15**: 1334-1348.
Gerhart J, Kirschner M. 2007. The theory of facilitated variation. *Proc Natl Acad Sci U S A* **104 Suppl 1**: 8582-8589.
Gillis JA, Alsema EC, Criswell KE. 2017. Trunk neural crest origin of dermal denticles in a cartilaginous fish. *Proc Natl Acad Sci U S A* **114**: 13200-13205.
Gilmour DT, Lyon GJ, Carlton MB, Sanes JR, Cunningham JM, Anderson JR, Hogan BL, Evans MJ, Colledge WH. 1998. Mice deficient for the secreted glycoprotein SPARC/osteonectin/BM40 develop normally but show severe age-onset cataract formation and disruption of the lens. *Embo j* **17**: 1860-1870.
Giraud G, Terrone S, Bourgeois CF. 2018. Functions of DEAD box RNA helicases DDX5 and DDX17 in chromatin organization and transcriptional regulation. *BMB reports* **51**: 613-622.
Goldberg M, Kulkarni AB, Young M, Boskey A. 2011. Dentin: structure, composition and mineralization. *Front Biosci (Elite Ed)* **3**: 711-735.

Golding RE, Ponder WF, Byrne M. 2009. Three-dimensional reconstruction of the odontophoral cartilages of Caenogastropoda (Mollusca: Gastropoda) using micro-CT: Morphology and phylogenetic significance. *J Morphol* **270**: 558-587.
Goldschmidt R. 1982. *The material basis of evolution*. Yale University Press.
Göllner H, Dani C, Phillips B, Philipsen S, Suske G. 2001. Impaired ossification in mice lacking the transcription factor Sp3. *Mechanisms of Development* **106**: 77-83.
Golub E, Boesze-Battaglia K. 2007. The role of alkaline phosphatase in mineralization. *Current Opinion in Orthopaedics* **18**: 444-448.
Gomez-Picos P, Eames BF. 2015. On the evolutionary relationship between chondrocytes and osteoblasts. *Front Genet* **6**: 297.
Goodwin ML, Jin H, Straessler K, Smith-Fry K, Zhu JF, Monument MJ, Grossmann A, Randall RL, Capecchi MR, Jones KB. 2014. Modeling alveolar soft part sarcomagenesis in the mouse: a role for lactate in the tumor microenvironment. *Cancer Cell* **26**: 851-862.
Gordon CT, Rodda FA, Farlie PG. 2009. The RCAS retroviral expression system in the study of skeletal development. *Dev Dyn* **238**: 797-811.
Goto N, Fujimoto K, Fujii S, Ida-Yonemochi H, Ohshima H, Kawamoto T, Noshiro M, Shukunami C, Kozai K, Kato Y. 2016. Role of MSX1 in Osteogenic Differentiation of Human Dental Pulp Stem Cells. *Stem Cells International* **2016**: 8035759.
Goudemand N, Orchard MJ, Urdy S, Bucher H, Tafforeau P. 2011. Synchrotron-aided reconstruction of the conodont feeding apparatus and implications for the mouth of the first vertebrates. *Proc Natl Acad Sci U S A* **108**: 8720-8724.
Gould SJ, Lewontin RC, Smith JM, Holliday R. 1979. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proceedings of the Royal Society of London Series B Biological Sciences* **205**: 581-598.
Gould SJ. 2002a. *The structure of evolutionary theory*. Harvard University Press.
Gould SJ. 2002b. *The structure of evolutionary theory*. Belknap Press of Harvard University Press, Cambridge, Mass.
Granda JL, Posner AS. 1971. Distribution of four hydrolases in the epiphyseal plate. *Clin Orthop Relat Res* **74**: 269-272.
Gray H. 2001. Gray s Anatomy of the Human Body.
Gray H, Williams PL. 1989. *Gray's anatomy*. C. Livingstone, Edinburgh ; New York.
Guralnick R, Smith K. 1999. Historical and biomechanical analysis of integration and dissociation in molluscan feeding, with special emphasis on the true limpets (Patellogastropoda: Gastropoda). *J Morphol* **241**: 175-195.
Guweidhi A, Kleeff J, Adwan H, Giese NA, Wente MN, Giese T, Büchler MW, Berger MR, Friess H. 2005. Osteonectin influences growth and invasion of pancreatic cancer cells. *Ann Surg* **242**: 224-234.
H. NW, M. ED, J. A, I. H, H. WH, A. BR, J. SR. 2019. Current concepts in osteogenesis imperfecta: bone structure, biomechanics and medical management. *Journal of Children's Orthopaedics* **13**: 1-11.
Haeckel E. 1866. *Generelle Morphologie der Organismen. Allgemeine Grundzüge der organischen Formen-Wissenschaft, mechanisch begründet durch die von C. Darwin reformirte Descendenz-Theorie, etc*.
Halfon MS. 2017. Perspectives on Gene Regulatory Network Evolution. *Trends Genet* **33**: 436-447.
Hall BK. 1975. Evolutionary consequences of skeletal Differentiation. *American Zoologist* **15**: 329-350.

Hall BK. 1978. Initiation of osteogenesis by mandibular mesenchyme of the embryonic chick in response to mandibular and non-mandibular epithelia. *Archives of Oral Biology* **23**: 1157-1161.
Hall BK. 1986. The role of movement and tissue interactions in the development and growth of bone and secondary cartilage in the clavicle of the embryonic chick. *J Embryol Exp Morphol* **93**: 133-152.
Hall BK. 2005a. Bones and cartilage : developmental and evolutionary skeletal biology. pp. xxviii, 760 p. Elsevier Academic Press,, San Diego, Calif. ; London.
Hall BK. 2005b. Chapter 3 - Cartilage. In *Bones and Cartilage*, doi:https://doi.org/10.1016/B978-012319060-4/50005-1, pp. 33-47. Academic Press, San Diego.
Hall BK. 2015. *Bones and cartilage : developmental and evolutionary skeletal biology*. Elsevier/AP, Academic Press is an imprint of Elsevier, Amsterdam.
Hall BK, Gillis JA. 2013a. Incremental evolution of the neural crest, neural crest cells and neural crest-derived skeletal tissues. *Journal of Anatomy* **222**: 19-31.
Hall BK, Gillis JA. 2013b. Incremental evolution of the neural crest, neural crest cells and neural crest-derived skeletal tissues. *J Anat* **222**: 19-31.
Hall BK, Miyake T. 1992. The membranous skeleton: the role of cell condensations in vertebrate skeletogenesis. *Anat Embryol (Berl)* **186**: 107-124.
Hall BK, Miyake T. 1995. Divide, accumulate, differentiate: cell condensation in skeletal development revisited. *Int J Dev Biol* **39**: 881-893.
Hall BK, Miyake T. 2000. All for one and one for all: condensations and the initiation of skeletal development. *Bioessays* **22**: 138-147.
Ham AW, Cormack DH. 1987. *Ham's histology*. Lippincott, Philadelphia.
Hamburger V, Hamilton HL. 1951. A series of normal stages in the development of the chick embryo. *Journal of Morphology* **88**: 49-92.
Hammond CL, Schulte-Merker S. 2009. Two populations of endochondral osteoblasts with differential sensitivity to Hedgehog signalling. *Development* **136**: 3991-4000.
Hanken J, Thomson K. 1990. Morphogenesis and Evolution. *Copeia* **1990**: 247.
Hankenson KD, Dishowitz M, Gray C, Schenker M. 2011. Angiogenesis in bone regeneration. *Injury* **42**: 556-561.
Harada H, Tagashira S, Fujiwara M, Ogawa S, Katsumata T, Yamaguchi A, Komori T, Nakatsuka M. 1999. Cbfa1 isoforms exert functional differences in osteoblast differentiation. *The Journal of biological chemistry* **274**: 6972-6978.
Hardingham T. 1981. Proteoglycans: their structure, interactions and molecular organization in cartilage. *Biochemical Society transactions* **9**: 489-497.
Hatori M, Klatte KJ, Teixeira CC, Shapiro IM. 1995. End labeling studies of fragmented DNA in the avian growth plate: evidence of apoptosis in terminally differentiated chondrocytes. *J Bone Miner Res* **10**: 1960-1968.
Hattori T, Muller C, Gebhard S, Bauer E, Pausch F, Schlund B, Bosl MR, Hess A, Surmann-Schmitt C, von der Mark H et al. 2010. SOX9 is a major negative regulator of cartilage vascularization, bone marrow formation and endochondral ossification. *Development* **137**: 901-911.
He K, Wu G, Li WX, Guan D, Lv W, Gong M, Ye S, Lu A. 2017. A transcriptomic study of myogenic differentiation under the overexpression of PPARgamma by RNA-Seq. *Sci Rep* **7**: 15308.

Hecht J, Stricker S, Wiecha U, Stiege A, Panopoulou G, Podsiadlowski L, Poustka AJ, Dieterich C, Ehrich S, Suvorova J et al. 2008. Evolution of a core gene network for skeletogenesis in chordates. *PLoS Genet* **4**: e1000025.

Heinen TJ, Staubach F, Haming D, Tautz D. 2009. Emergence of a new gene from an intergenic region. *Curr Biol* **19**: 1527-1531.

Herbst EC, Doube M, Smithson TR, Clack JA, Hutchinson JR. 2019. Bony lesions in early tetrapods and the evolution of mineralized tissue repair. *Paleobiology* **45**: 676-697.

Herlofsen SR, Hoiby T, Cacchiarelli D, Zhang X, Mikkelsen TS, Brinchmann JE. 2014. Brief report: importance of SOX8 for in vitro chondrogenic differentiation of human mesenchymal stromal cells. *Stem Cells* **32**: 1629-1635.

Higashihori N, Song Y, Richman JM. 2008. Expression and regulation of the decoy bone morphogenetic protein receptor BAMBI in the developing avian face. *Dev Dyn* **237**: 1500-1508.

Hill CR, Jacobs BH, Brown CB, Barnett JV, Goudy SL. 2015. Type III transforming growth factor beta receptor regulates vascular and osteoblast development during palatogenesis. *Developmental Dynamics* **244**: 122-133.

Hill TP, Spater D, Taketo MM, Birchmeier W, Hartmann C. 2005. Canonical Wnt/beta-catenin signaling prevents osteoblasts from differentiating into chondrocytes. *Developmental cell* **8**: 727-738.

Hilton MJ, Tu X, Long F. 2007. Tamoxifen-inducible gene deletion reveals a distinct cell type associated with trabecular bone, and direct regulation of PTHrP expression and chondrocyte morphology by Ihh in growth region cartilage. *Developmental Biology* **308**: 93-105.

Hinsbergh VWMv, Rabelink TJ. 2005. FGFR1 and the Bloodline of the Vasculature. *Arteriosclerosis, Thrombosis, and Vascular Biology* **25**: 883-886.

Hinton RJ, Jing Y, Jing J, Feng JQ. 2017. Roles of Chondrocytes in Endochondral Bone Formation and Fracture Repair. *J Dent Res* **96**: 23-30.

Hirasawa T, Kuratani S. 2015. Evolution of the vertebrate skeleton: morphology, embryology, and development. *Zoological Lett* **1**: 2.

Hirasawa T, Oisi Y, Kuratani S. 2016. Palaeospondylus as a primitive hagfish. *Zoological Lett* **2**: 20.

Hirschman A, Dziewiatkowski DD. 1966. Protein-polysaccharide loss during endochondral ossification: immunochemical evidence. *Science* **154**: 393-395.

His W. 1868. *Untersuchungen über die erste Anlage des Wirbelthierleibes: die erste Entwickelung des Hühnchens im Ei*. FCW Vogel.

Hoeppner LH, Secreto F, Jensen ED, Li X, Kahler RA, Westendorf JJ. 2009. Runx2 and bone morphogenic protein 2 regulate the expression of an alternative Lef1 transcript during osteoblast maturation. *Journal of Cellular Physiology* **221**: 480-489.

Hoffmann HM, Beumer TL, Rahman S, McCabe LR, Banerjee C, Aslam F, Tiro JA, van Wijnen AJ, Stein JL, Stein GS et al. 1996. Bone tissue-specific transcription of the osteocalcin gene: Role of an activator osteoblast-specific complex and suppressor hox proteins that bind the OC box. *Journal of Cellular Biochemistry* **61**: 310-324.

Hogg DA. 1982. Ossification of the laryngeal, tracheal and syringeal cartilages in the domestic fowl. *J Anat* **134**: 57-71.

Hojo H, Ohba S. 2019. Insights into Gene Regulatory Networks in Chondrocytes. *International journal of molecular sciences* **20**.

Holleville N, Matéos S, Bontoux M, Bollerot K, Monsoro-Burq AH. 2007. Dlx5 drives Runx2 expression and osteogenic differentiation in developing cranial suture mesenchyme. *Developmental Biology* **304**: 860-874.

Holmberg Olausson K, Maire CL, Haidar S, Ling J, Learner E, Nister M, Ligon KL. 2014. Prominin-1 (CD133) defines both stem and non-stem cell populations in CNS development and gliomas. *PLoS One* **9**: e106694.

Homberger DG, Meyers RA. 1989. Morphology of the lingual apparatus of the domestic chicken, Gallus gallus, with special attention to the structure of the fasciae. *Am J Anat* **186**: 217-257.

Horigome N, Myojin M, Ueki T, Hirano S, Aizawa S, Kuratani S. 1999. Development of cephalic neural crest cells in embryos of Lampetra japonica, with special reference to the evolution of the jaw. *Developmental biology* **207**: 287-308.

Hoshi K, Komori T, Ozawa H. 1999. Morphological characterization of skeletal cells in Cbfa1-deficient mice. *Bone* **25**: 639-651.

Hu G, Codina M, Fisher S. 2012. Multiple enhancers associated with ACAN suggest highly redundant transcriptional regulation in cartilage. *Matrix biology : journal of the International Society for Matrix Biology* **31**: 328-337.

Hu G, Hovav R, Grover CE, Faigenboim-Doron A, Kadmon N, Page JT, Udall JA, Wendel JF. 2016. Evolutionary Conservation and Divergence of Gene Coexpression Networks in Gossypium (Cotton) Seeds. *Genome Biol Evol* **8**: 3765-3783.

Hu G, Lee H, Price SM, Shen MM, Abate-Shen C. 2001. Msx homeobox genes inhibit differentiation through upregulation of cyclin D1. *Development* **128**: 2373-2384.

Hu K, Olsen BR. 2016. The roles of vascular endothelial growth factor in bone repair and regeneration. *Bone* **91**: 30-38.

Hubbard TJ, Aken BL, Ayling S, Ballester B, Beal K, Bragin E, Brent S, Chen Y, Clapham P, Clarke L et al. 2009. Ensembl 2009. *Nucleic Acids Res* **37**: D690-697.

Hutchinson SA, Cheesman SE, Hale LA, Boone JQ, Eisen JS. 2007. Nkx6 proteins specify one zebrafish primary motoneuron subtype by regulating late islet1 expression. *Development* **134**: 1671-1677.

Huycke TR, Eames BF, Kimmel CB. 2012. Hedgehog-dependent proliferation drives modular growth during morphogenesis of a dermal bone. *Development* **139**: 2371-2380.

Huysseune A, Sire JY. 1998. Evolution of patterns and processes in teeth and tooth-related tissues in non-mammalian vertebrates. *Eur J Oral Sci* **106 Suppl 1**: 437-481.

Hwang SP, Tsou MF, Lin YC, Liu CH. 1997. The zebrafish BMP4 gene: sequence analysis and expression pattern during embryonic development. *DNA and cell biology* **16**: 1003-1011.

Ikegami D, Akiyama H, Suzuki A, Nakamura T, Nakano T, Yoshikawa H, Tsumaki N. 2011. Sox9 sustains chondrocyte survival and hypertrophy in part through Pik3ca-Akt pathways. *Development* **138**: 1507-1519.

Inada M, Wang Y, Byrne MH, Rahman MU, Miyaura C, López-Otín C, Krane SM. 2004. Critical roles for collagenase-3 (Mmp13) in development of growth plate cartilage and in endochondral ossification. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 17192-17197.

Inada M, Yasui T, Nomura S, Miyake S, Deguchi K, Himeno M, Sato M, Yamagiwa H, Kimura T, Yasui N et al. 1999. Maturational disturbance of chondrocytes in Cbfa1-deficient mice. *Dev Dyn* **214**: 279-290.

Indjeian Vahan B, Kingman Garrett A, Jones Felicity C, Guenther Catherine A, Grimwood J, Schmutz J, Myers Richard M, Kingsley David M. 2016. Evolving New Skeletal Traits by cis-Regulatory Changes in Bone Morphogenetic Proteins. *Cell* **164**: 45-56.

Irisarri I, Baurain D, Brinkmann H, Delsuc F, Sire JY, Kupfer A, Petersen J, Jarek M, Meyer A, Vences M et al. 2017. Phylotranscriptomic consolidation of the jawed vertebrate timetree. *Nat Ecol Evol* **1**: 1370-1378.

Ishida A, Nakayama A, Koide H, Tatsuno I, Tanaka T, Yokote K. 2019. SAT-520 The Effects Of AKAP13 On Osteogenesis By Modulation Of RhoA And Wnt Signaling. *Journal of the Endocrine Society* **3**.

Itoh T, Ando M, Tsukamasa Y, Akao Y. 2012. Expression of BMP-2 and Ets1 in BMP-2-stimulated mouse pre-osteoblast differentiation is regulated by microRNA-370. *FEBS letters* **586**: 1693-1701.

Izadifar Z, Chang T, Kulyk W, Chen X, Eames BF. 2016. Analyzing Biological Performance of 3D-Printed, Cell-Impregnated Hybrid Constructs for Cartilage Tissue Engineering. *Tissue engineering Part C, Methods* **22**: 173-188.

Jacob AL, Smith C, Partanen J, Ornitz DM. 2006. Fibroblast growth factor receptor 1 signaling in the osteo-chondrogenic cell lineage regulates sequential steps of osteoblast maturation. *Developmental Biology* **296**: 315-328.

James MJ, Järvinen E, Wang X-P, Thesleff I. 2006. Different Roles of Runx2 During Early Neural Crest–Derived Bone and Tooth Development. *Journal of Bone and Mineral Research* **21**: 1034-1044.

Jandzik D, Garnett AT, Square TA, Cattell MV, Yu JK, Medeiros DM. 2015. Evolution of the new vertebrate head by co-option of an ancient chordate skeletal tissue. *Nature* **518**: 534-537.

Janis CM, Devlin K, Warren DE, Witzmann F. 2012. Dermal bone in early tetrapods: a palaeophysiological hypothesis of adaptation for terrestrial acidosis. *Proceedings Biological sciences* **279**: 3035-3040.

Janvier P. 1996a. *Early vertebrates*.

Janvier P. 1996b. *Early vertebrates*. Oxford University Press, Oxford.

Janvier P. 1997. Pteraspidomorphi The Tree of Life Web Project.

Janvier P, Arsenault M. 2002. Palaeobiology: calcification of early vertebrate cartilage. *Nature* **417**: 609.

Jensen ED, Gopalakrishnan R, Westendorf JJ. 2009. Bone morphogenic protein 2 activates protein kinase D to regulate histone deacetylase 7 localization and repression of Runx2. *J Biol Chem* **284**: 2225-2234.

Jensen ED, Niu L, Caretti G, Nicol SM, Teplyuk N, Stein GS, Sartorelli V, van Wijnen AJ, Fuller-Pace FV, Westendorf JJ. 2008. p68 (Ddx5) interacts with Runx2 and regulates osteoblast differentiation. *J Cell Biochem* **103**: 1438-1451.

Jing Y, Zhou X, Han X, Jing J, von der Mark K, Wang J, de Crombrugghe B, Hinton RJ, Feng JQ. 2015. Chondrocytes Directly Transform into Bone Cells in Mandibular Condyle Growth. *J Dent Res* **94**: 1668-1675.

Johanson Z, Kearsley A, den Blaauwen J, Newman M, Smith MM. 2010. No bones about it: an enigmatic Devonian fossil reveals a new skeletal framework--a potential role of loss of gene regulation. *Semin Cell Dev Biol* **21**: 414-423.

Johnson DR. 1980. Formation of marrow cavity and ossification in mouse limb buds grown in vitro. *J Embryol Exp Morphol* **56**: 301-307.

Johnston C, Horney B, Deluca S, MacKenzie A, Eales J, Angus R. 1994. Changes in alkaline phosphatase isoenzyme activity in tissues and plasma of Atlantic salmon ( Salmo salar ) before and during smoltification and gonadal maturation. *Fish Physiology and Biochemistry* **12**: 485-497.

Jung Y-K, Park H-R, Cho H-J, Jang J-A, Lee E-J, Han M-S, Kim G-W, Han S. 2019. Degrading products of chondroitin sulfate can induce hypertrophy-like changes and MMP-13/ADAMTS5 production in chondrocytes. *Scientific Reports* **9**: 15846.

Kague E, Roy P, Asselin G, Hu G, Simonet J, Stanley A, Albertson C, Fisher S. 2016. Osterix/Sp7 limits cranial bone initiation sites and is required for formation of sutures. *Dev Biol* **413**: 160-172.

Kahler RA, Westendorf JJ. 2003. Lymphoid enhancer factor-1 and β-catenin inhibit Runx2-dependent transcriptional activation of the osteocalcin promoter. *Journal of Biological Chemistry* **278**: 11937-11944.

Kahler RA, Yingst SMC, Hoeppner LH, Jensen ED, Krawczak D, Oxford JT, Westendorf JJ. 2008. Collagen 11a1 is indirectly activated by lymphocyte enhancer-binding factor 1 (Lef1) and negatively regulates osteoblast maturation. *Matrix biology : journal of the International Society for Matrix Biology* **27**: 330-338.

Kaneto S, Wada H. 2011. Regeneration of amphioxus oral cirri and its skeletal rods: implications for the origin of the vertebrate skeleton. *J Exp Zool B Mol Dev Evol* **316**: 409-417.

Kang X, Sun Y, Zhang Z. 2019. Identification of key transcription factors - gene regulatory network related with osteogenic differentiation of human mesenchymal stem cells based on transcription factor prognosis system. *Exp Ther Med* **17**: 2113-2122.

Kapoor BG, Khanna B. 2004. *Ichthyology Handbook*. Springer-Verlag Berlin Heidelberg.

Karsenty G, Kronenberg HM, Settembre C. 2009. Genetic control of bone formation. *Annu Rev Cell Dev Biol* **25**: 629-648.

Kawane T, Komori H, Liu W, Moriishi T, Miyazaki T, Mori M, Matsuo Y, Takada Y, Izumi S, Jiang Q et al. 2014. Dlx5 and Mef2 Regulate a Novel Runx2 Enhancer for Osteoblast-Specific Expression. *Journal of Bone and Mineral Research* **29**: 1960-1969.

Kawane T, Qin X, Jiang Q, Miyazaki T, Komori H, Yoshida CA, Matsuura-Kawata VKdS, Sakane C, Matsuo Y, Nagai K et al. 2018. Runx2 is required for the proliferation of osteoblast progenitors and induces proliferation by regulating Fgfr2 and Fgfr3. *Scientific Reports* **8**: 13551.

Kawasaki K, Suzuki T, Weiss KM. 2004. Genetic basis for the evolution of vertebrate mineralized tissue. *Proc Natl Acad Sci U S A* **101**: 11356-11361.

Kawasaki K, Weiss KM. 2006. Evolutionary genetics of vertebrate tissue mineralization: the origin and evolution of the secretory calcium-binding phosphoprotein family. *J Exp Zool B Mol Dev Evol* **306**: 295-316.

Keating JN, Marquart CL, Donoghue PCJ. 2015. Histology of the heterostracan dermal skeleton: Insight into the origin of the vertebrate mineralised skeleton. *Journal of Morphology* **276**: 657-680.

Kemp NE, Westrin SK. 1979. Ultrastructure of calcified cartilage in the endoskeletal tesserae of sharks. *J Morphol* **160**: 75-109.

Kerkhofs J, Roberts SJ, Luyten FP, Van Oosterwyck H, Geris L. 2012. Relating the chondrocyte gene network to growth plate morphology: from genes to phenotype. *PLoS One* **7**: e34729.

Kern B, Shen J, Starbuck M, Karsenty G. 2001. Cbfa1 contributes to the osteoblast-specific expression of type I collagen genes. *The Journal of biological chemistry* **276**: 7101-7107.

Kerney R, Gross JB, Hanken J. 2007. Runx2 is essential for larval hyobranchial cartilage formation in Xenopus laevis. *Dev Dyn* **236**: 1650-1662.

Khosravi P, Gazestani VH, Pirhaji L, Law B, Sadeghi M, Goliaei B, Bader GD. 2015. Inferring interaction type in gene regulatory networks using co-expression data. *Algorithms Mol Biol* **10**: 23.

Kiani C, Chen L, Wu YJ, Yee AJ, Yang BB. 2002. Structure and function of aggrecan. *Cell research* **12**: 19-32.

Kida J, Hata K, Nakamura E, Yagi H, Takahata Y, Murakami T, Maeda Y, Nishimura R. 2018. Interaction of LEF1 with TAZ is necessary for the osteoblastogenic activity of Wnt3a. *Scientific Reports* **8**: 10375.

Kielty CM, Kwan AP, Holmes DF, Schor SL, Grant ME. 1985. Type X collagen, a product of hypertrophic chondrocytes. *Biochem J* **227**: 545-554.

Kieslinger M, Folberth S, Dobreva G, Dorn T, Croci L, Erben R, Consalez GG, Grosschedl R. 2005. EBF2 regulates osteoblast-dependent differentiation of osteoclasts. *Developmental cell* **9**: 757-767.

Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* **14**: R36.

Kim IS, Otto F, Zabel B, Mundlos S. 1999. Regulation of chondrocyte differentiation by Cbfa1. *Mech Dev* **80**: 159-170.

Kim Y-J, Lee M-H, Wozney JM, Cho J-Y, Ryoo H-M. 2004. Bone morphogenetic protein-2-induced alkaline phosphatase expression is stimulated by Dlx5 and repressed by Msx2. *Journal of Biological Chemistry* **279**: 50773-50780.

Kimmel CB, Miller CT, Kruze G, Ullmann B, BreMiller RA, Larison KD, Snyder HC. 1998. The shaping of pharyngeal cartilages during early development of the zebrafish. *Developmental biology* **203**: 245-263.

Kimura H, Kwan KM, Zhang Z, Deng JM, Darnay BG, Behringer RR, Nakamura T, de Crombrugghe B, Akiyama H. 2008. Cthrc1 Is a Positive Regulator of Osteoblastic Bone Formation. *PLOS ONE* **3**: e3174.

Kimura H, Miyashita H, Suzuki Y, Kobayashi M, Watanabe K, Sonoda H, Ohta H, Fujiwara T, Shimosegawa T, Sato Y. 2009. Distinctive localization and opposed roles of vasohibin-1 and vasohibin-2 in the regulation of angiogenesis. *Blood* **113**: 4810-4818.

Kinoshita K, Suzuki T, Koike M, Nishida C, Koike A, Nunome M, Uemura T, Ichiyanagi K, Matsuda Y. 2020. Combined deletions of IHH and NHEJ1 cause chondrodystrophy and embryonic lethality in the Creeper chicken. *Communications Biology* **3**: 144.

Kirsch T, Nah HD, Shapiro IM, Pacifici M. 1997. Regulated production of mineralization-competent matrix vesicles in hypertrophic chondrocytes. *J Cell Biol* **137**: 1149-1160.

Kirschner M, Gerhart J. 1998. Evolvability. *Proc Natl Acad Sci U S A* **95**: 8420-8427.

Kishiya M, Sawada T, Kanemaru K, Kudo H, Numasawa T, Yokoyama T, Tanaka S, Motomura S, Ueyama K, Harata S et al. 2008. A functional RNAi screen for Runx2-regulated genes associated with ectopic bone formation in human spinal ligaments. *J Pharmacol Sci* **106**: 404-414.

Knight RD, Schilling TF. 2006. Cranial neural crest and development of the head skeleton. *Adv Exp Med Biol* **589**: 120-133.

Knudson CB, Knudson W. 2001. Cartilage proteoglycans. *Semin Cell Dev Biol* **12**: 69-78.

Koehler A, Desser S, Chang B, MacDonald J, Tepass U, Ringuette M. 2009. Molecular evolution of SPARC: absence of the acidic module and expression in the endoderm of the starlet sea anemone, Nematostella vectensis. *Dev Genes Evol* **219**: 509-521.

Kokabu S, Gamer L, Cox K, Lowery J, Tsuji K, Raz R, Economides A, Katagiri T, Rosen V. 2012. BMP3 Suppresses Osteoblast Differentiation of Bone Marrow Stromal Cells via Interaction with Acvr2b. *Molecular Endocrinology* **26**: 87-94.

Komori T. 2010. Regulation of bone development and extracellular matrix protein genes by RUNX2. *Cell Tissue Res* **339**: 189-195.

Komori T. 2015. The functions of Runx family transcription factors and Cbfb in skeletal development. *Oral Science International* **12**: 1-4.

Komori T. 2019. Regulation of Proliferation, Differentiation and Functions of Osteoblasts by Runx2. *International journal of molecular sciences* **20**.

Komori T, Yagi H, Nomura S, Yamaguchi A, Sasaki K, Deguchi K, Shimizu Y, Bronson RT, Gao YH, Inada M et al. 1997. Targeted disruption of Cbfa1 results in a complete lack of bone formation owing to maturational arrest of osteoblasts. *Cell* **89**: 755-764.

Komori T. 2006. Regulation of osteoblast differentiation by transcription factors. *Journal of Cellular Biochemistry* **99**: 1233-1239.

Komori T. 2009. Regulation of bone development and extracellular matrix protein genes by RUNX2. *Cell and Tissue Research* **339**: 189.

Komori T. 2015. The functions of Runx family transcription factors and Cbfb in skeletal development. *Oral Science International* **12**: 1-4.

Komori T. 2018. Runx2, an inducer of osteoblast and chondrocyte differentiation. *Histochem Cell Biol* **149**: 313-323.

Komori T, Yagi H, Nomura S, Yamaguchi A, Sasaki K, Deguchi K, Shimizu Y, Bronson RT, Gao YH, Inada M et al. 1997. Targeted disruption of Cbfa1 results in a complete lack of bone formation owing to maturational arrest of osteoblasts. *Cell* **89**: 755-764.

Kream BE, Harrison JR, Krebsbach PH, Bogdanovic Z, Bedalov A, Pavlin D, Woody CO, Clark SH, Rowe D, Lichtler AC. 1995. Regulation of type I collagen gene expression in bone. *Connect Tissue Res* **31**: 261-264.

Kronenberg HM. 2003. Developmental regulation of the growth plate. *Nature* **423**: 332-336.

Kumar S, Stecher G, Suleski M, Hedges SB. 2017. TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. *Molecular Biology and Evolution* **34**: 1812-1819.

Kundu M, Javed A, Jeon JP, Horner A, Shum L, Eckhaus M, Muenke M, Lian JB, Yang Y, Nuckolls GH et al. 2002. Cbfbeta interacts with Runx2 and has a critical role in bone development. *Nat Genet* **32**: 639-644.

Kurakazu I, Akasaki Y, Hayashida M, Tsushima H, Goto N, Sueishi T, Toya M, Kuwahara M, Okazaki K, Duffy T et al. 2019. FOXO1 transcription factor regulates chondrogenic differentiation through transforming growth factor β1 signaling. doi:10.1074/jbc.RA119.009409.

Langille RM, Hall BK. 1993a. Calcification of Cartilage from the Lamprey Petromyzon marinus (L.) in vitro. *Acta Zoologica* **74**: 31-41.

Langille RM, Hall BK. 1993b. Pattern Formation and the Neural Crest. In *The Skull*, Vol 1: Development (ed. J Hanken, BK Hall), pp. 77-111. the University of Chicago Press, Chicago.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357-359.

Laue K, Janicke M, Plaster N, Sonntag C, Hammerschmidt M. 2008. Restriction of retinoic acid activity by Cyp26b1 is required for proper timing and patterning of osteogenesis during zebrafish development. *Development* **135**: 3775-3787.

Leboy PS, Shapiro IM, Uschmann BD, Oshima O, Lin D. 1988. Gene expression in mineralizing chick epiphyseal cartilage. *J Biol Chem* **263**: 8515-8520.

Lee HK, Lee DS, Ryoo HM, Park JT, Park SJ, Bae HS, Cho MI, Park JC. 2010. The odontogenic ameloblast-associated protein (ODAM) cooperates with RUNX2 and modulates enamel mineralization via regulation of MMP-20. *J Cell Biochem* **111**: 755-767.

Lee KS, Kim HJ, Li QL, Chi XZ, Ueta C, Komori T, Wozney JM, Kim EG, Choi JY, Ryoo HM et al. 2000. Runx2 is a common target of transforming growth factor beta1 and bone morphogenetic protein 2, and cooperation between Runx2 and Smad5 induces osteoblast-specific gene expression in the pluripotent mesenchymal precursor cell line C2C12. *Mol Cell Biol* **20**: 8783-8792.

Lee MH, Kim YJ, Kim HJ, Park HD, Kang AR, Kyung HM, Sung JH, Wozney JM, Kim HJ, Ryoo HM. 2003. BMP-2-induced Runx2 expression is mediated by Dlx5, and TGF-beta 1 opposes the BMP-2-induced osteoblast differentiation by suppression of Dlx5 expression. *J Biol Chem* **278**: 34387-34394.

Lee MH, Kim YJ, Yoon WJ, Kim JI, Kim BG, Hwang YS, Wozney JM, Chi XZ, Bae SC, Choi KY et al. 2005. Dlx5 specifically regulates Runx2 type II expression by binding to homeodomain-response elements in the Runx2 distal promoter. *J Biol Chem* **280**: 35579-35587.

Lefebvre V, Behringer RR, de Crombrugghe B. 2001. L-Sox5, Sox6 and Sox9 control essential steps of the chondrocyte differentiation pathway. *Osteoarthritis Cartilage* **9 Suppl A**: S69-75.

Lefebvre V, de Crombrugghe B. 1998. Toward understanding SOX9 function in chondrocyte differentiation. *Matrix biology : journal of the International Society for Matrix Biology* **16**: 529-540.

Lefebvre V, Dvir-Ginzberg M. 2017. SOX9 and the many facets of its regulation in the chondrocyte lineage. *Connect Tissue Res* **58**: 2-14.

Lefebvre V, Garofalo S, de Crombrugghe B. 1995. Type X collagen gene expression in mouse chondrocytes immortalized by a temperature-sensitive simian virus 40 large tumor antigen. *J Cell Biol* **128**: 239-245.

Lefebvre V, Huang W, Harley VR, Goodfellow PN, de Crombrugghe B. 1997. SOX9 is a potent activator of the chondrocyte-specific enhancer of the pro alpha1(II) collagen gene. *Mol Cell Biol* **17**: 2336-2346.

Leprevost A, AzaIs T, Trichet M, Sire JY. 2017. Vertebral Development and Ossification in the Siberian Sturgeon (Acipenser Baerii), with New Insights on Bone Histology and Ultrastructure of Vertebral Elements and Scutes. *Anat Rec (Hoboken)* **300**: 437-449.

Leung VY, Gao B, Leung KK, Melhado IG, Wynn SL, Au TY, Dung NW, Lau JY, Mak AC, Chan D et al. 2011. SOX9 governs differentiation stage-specific gene expression in growth plate chondrocytes via direct concomitant transactivation and repression. *PLoS Genet* **7**: e1002356.

Levin M, Hashimshony T, Wagner F, Yanai I. 2012. Developmental milestones punctuate gene expression in the Caenorhabditis embryo. *Developmental cell* **22**: 1101-1108.

Levine M, Davidson EH. 2005. Gene regulatory networks for development. *Proc Natl Acad Sci U S A* **102**: 4936-4942.

Li B, Balasubramanian K, Krakow D, Cohn DH. 2017. Genes uniquely expressed in human growth plate chondrocytes uncover a distinct regulatory network. *BMC Genomics* **18**: 983.

Li CY, Zhang Y, Wang Z, Zhang Y, Cao C, Zhang PW, Lu SJ, Li XM, Yu Q, Zheng X et al. 2010. A human-specific de novo protein-coding gene associated with human brain functions. *PLoS Comput Biol* **6**: e1000734.

Li J, Jin D, Fu S, Mei G, Zhou J, Lei L, Yu B, Wang G. 2013. Insulin-like growth factor binding protein-3 modulates osteoblast differentiation via interaction with vitamin D receptor. *Biochem Biophys Res Commun* **436**: 632-637.

Li Z, Xu Z, Duan C, Liu W, Sun J, Han B. 2018. Role of TCF/LEF Transcription Factors in Bone Development and Osteogenesis. *Int J Med Sci* **15**: 1415-1422.

Lin K-L, Chou C-H, Hsieh S-C, Hwa S-Y, Lee M-T, Wang F-F. 2010. Transcriptional upregulation of DDR2 by ATF4 facilitates osteoblastic differentiation through p38 MAPK-mediated Runx2 activation. *Journal of Bone and Mineral Research* **25**: 2489-2503.

Li H, Marijanovic I, Kronenberg MS, Erceg I, Stover ML, Velonis D, Mina M, Heinrich JG, Harris SE, Upholt WB et al. 2008. Expression and function of Dlx genes in the osteoblast lineage. *Dev Biol* **316**: 458-470.

Li J, Dong S. 2016. The Signaling Pathways Involved in Chondrocyte Differentiation and Hypertrophic Differentiation. *Stem Cells International* **2016**: 2470351.

Li N, Wang Q, Zhu T, Qiao L, Zhang F, Mi R, Wang B, Chen L, Gu J, Lu Y et al. 2016. In vitro functional characterization of prostaglandin-endoperoxide synthase 2 during chondrocyte hypertrophic differentiation. *Oncotarget* **7**.

Li S, Kong H, Yao N, Yu Q, Wang P, Lin Y, Wang J, Kuang R, Zhao X, Xu J et al. 2011. The role of runt-related transcription factor 2 (Runx2) in the late stage of odontoblast differentiation and dentin formation. *Biochem Biophys Res Commun* **410**: 698-704.

Li SW, Arita M, Kopen GC, Phinney DG, Prockop DJ. 1998. A 1,064 bp fragment from the promoter region of the Col11a2 gene drives lacZ expression not only in cartilage but also in osteoblasts adjacent to regions undergoing both endochondral and intramembranous ossification in mouse embryos. *Matrix biology : journal of the International Society for Matrix Biology* **17**: 213-221.

Li Y, Lacerda DA, Warman ML, Beier DR, Yoshioka H, Ninomiya Y, Oxford JT, Morris NP, Andrikopoulos K, Ramirez F et al. 1995. A fibrillar collagen gene, Col11a1, is essential for skeletal morphogenesis. *Cell* **80**: 423-430.

Li Y, Varala K, Coruzzi GM. 2015. From milliseconds to lifetimes: tracking the dynamic behavior of transcription factors in gene networks. *Trends Genet* **31**: 509-515.

Li Z, Xu Z, Duan C, Liu W, Sun J, Han B. 2018. Role of TCF/LEF Transcription Factors in Bone Development and Osteogenesis. *Int J Med Sci* **15**: 1415-1422.

Lian JB, Stein GS. 2003. Runx2/Cbfa1: a multifunctional regulator of bone formation. *Curr Pharm Des* **9**: 2677-2685.

Liang H, Xu J, Xue M, Jackson C. 2016. Matrix metalloproteinases in bone development and pathology: current knowledge and potential clinical utility. *Metalloproteinases In Medicine* **Volume 3**: 93-102.

Liu C-F, Lefebvre V. 2015a. The transcription factors SOX9 and SOX5/SOX6 cooperate genome-wide through super-enhancers to drive chondrogenesis. *Nucleic Acids Research* **43**: 8183-8203.

Liu CF, Lefebvre V. 2015b. The transcription factors SOX9 and SOX5/SOX6 cooperate genome-wide through super-enhancers to drive chondrogenesis. *Nucleic Acids Res* **43**: 8183-8203.

Liu C-F, Angelozzi M, Haseeb A, Lefebvre V. 2018. SOX9 is dispensable for the initiation of epigenetic remodeling and the activation of marker genes at the onset of chondrogenesis. *Development* **145**: dev164459.

Liu X, Wang Y, Zhang L, Xu Z, Chu Q, Xu C, Sun Y, Gao Y. 2018. Combination of Runx2 and Cbfbeta upregulates Amelotin gene expression in ameloblasts by directly interacting with cisenhancers during amelogenesis. *Mol Med Rep* **17**: 6068-6076.

Liu Y, Li H, Tanaka K, Tsumaki N, Yamada Y. 2000. Identification of an enhancer sequence within the first intron required for cartilage-specific transcription of the alpha2(XI) collagen gene. *J Biol Chem* **275**: 12712-12718.

Livingston BT, Killian CE, Wilt F, Cameron A, Landrum MJ, Ermolaeva O, Sapojnikov V, Maglott DR, Buchanan AM, Ettensohn CA. 2006. A genome-wide analysis of biomineralization-related proteins in the sea urchin Strongylocentrotus purpuratus. *Developmental biology* **300**: 335-348.

Lohmander S, Hjerpe A. 1975. Proteoglycans of mineralizing rib and epiphyseal cartilage. *Biochim Biophys Acta* **404**: 93-109.

Long JA, Burrow CJ, Ginter M, Maisey JG, Trinajstic KM, Coates MI, Young GC, Senden TJ. 2015. First shark from the Late Devonian (Frasnian) Gogo Formation, Western Australia sheds new light on the development of tessellated calcified cartilage. *PloS one* **10**: e0126066.

Longabaugh WJ, Davidson EH, Bolouri H. 2005. Computational representation of developmental genetic regulatory networks. *Dev Biol* **283**: 1-16.

Longabaugh WJR, Davidson EH, Bolouri H. 2009. Visualization, documentation, analysis, and communication of large-scale gene regulatory networks. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1789**: 363-374.

Lopes CT, Franz M, Kazi F, Donaldson SL, Morris Q, Bader GD. 2010. Cytoscape Web: an interactive web-based network browser. *Bioinformatics* **26**: 2347-2348.

Losos JB. 2011. Convergence, adaptation, and constraint. *Evolution* **65**: 1827-1840.

Lowe CJ, Clarke DN, Medeiros DM, Rokhsar DS, Gerhart J. 2015. The deuterostome context of chordate origins. *Nature* **520**: 456-465.

Lu C, Miclau T, Hu D, Hansen E, Tsui K, Puttlitz C, Marcucio RS. 2005. Cellular basis for age-related changes in fracture repair. *J Orthop Res* **23**: 1300-1307.

Lu Y, Qiao L, Lei G, Mira RR, Gu J, Zheng Q. 2014. Col10a1 gene expression and chondrocyte hypertrophy during skeletal development and disease. *Frontiers in Biology* **9**: 195-204.

Luderer HF, Bai S, Longmore GD. 2008. The LIM protein LIMD1 influences osteoblast differentiation and function. *Experimental cell research* **314**: 2884-2894.

Lui JC, Nilsson O, Baron J. 2014. Recent research on the growth plate: Recent insights into the regulation of the growth plate. *Journal of molecular endocrinology* **53**: T1-9.

Luo G, Ducy P, McKee MD, Pinero GJ, Loyer E, Behringer RR, Karsenty G. 1997. Spontaneous calcification of arteries and cartilage in mice lacking matrix GLA protein. *Nature* **386**: 78-81.

Luu HH, Song WX, Luo X, Manning D, Luo J, Deng ZL, Sharff KA, Montag AG, Haydon RC, He TC. 2007. Distinct roles of bone morphogenetic proteins in osteogenic differentiation of mesenchymal stem cells. *J Orthop Res* **25**: 665-677.

Maes C, Carmeliet P, Moermans K, Stockmans I, Smets N, Collen D, Bouillon R, Carmeliet G. 2002. Impaired angiogenesis and endochondral bone formation in mice lacking the vascular endothelial growth factor isoforms VEGF164 and VEGF188. *Mech Dev* **111**: 61-73.

Mahanta P, Ahmed HA, Bhattacharyya DK, Kalita JK. 2012. An effective method for network module extraction from microarray data. *BMC Bioinformatics* **13 Suppl 13**: S4.

Mailhot G, Yang M, Mason-Savas A, Mackay CA, Leav I, Odgren PR. 2008. BMP-5 expression increases during chondrocyte differentiation in vivo and in vitro and promotes proliferation and cartilage matrix synthesis in primary chondrocyte cultures. *Journal of cellular physiology* **214**: 56-64.

Mak AB, Pehar M, Nixon AM, Williams RA, Uetrecht AC, Puglielli L, Moffat J. 2014. Post-translational regulation of CD133 by ATase1/ATase2-mediated lysine acetylation. *J Mol Biol* **426**: 2175-2182.

Mak KK, Chen MH, Day TF, Chuang PT, Yang Y. 2006. Wnt/{beta}-catenin signaling interacts differentially with Ihh signaling in controlling endochondral bone and synovial joint formation. *Development* **133**: 3695-3707.

Mak KK, Kronenberg HM, Chuang PT, Mackem S, Yang Y. 2008. Indian hedgehog signals independently of PTHrP to promote chondrocyte hypertrophy. *Development* **135**: 1947-1956.

Mallatt J, Chen JY. 2003. Fossil sister group of craniates: predicted and found. *J Morphol* **258**: 1-31.

Marletaz F, Peijnenburg K, Goto T, Satoh N, Rokhsar DS. 2019. A New Spiralian Phylogeny Places the Enigmatic Arrow Worms among Gnathiferans. *Current biology : CB* **29**: 312-318 e313.

Martinek N, Shahab J, Sodek J, Ringuette M. 2007. Is SPARC an evolutionarily conserved collagen chaperone? *J Dent Res* **86**: 296-305.

Matsushiro A, Miyashita T. 2004. Evolution of hard-tissue mineralization: comparison of the inner skeletal system and the outer shell system. *J Bone Miner Metab* **22**: 163-169.

Mavrogiannis LA, Antonopoulou I, Baxova A, Kutilek S, Kim CA, Sugayama SM, Salamanca A, Wall SA, Morriss-Kay GM, Wilkie AO. 2001. Haploinsufficiency of the human homeobox gene ALX4 causes skull ossification defects. *Nature genetics* **27**: 17-18.

McCall MN. 2013. Estimation of Gene Regulatory Networks. *Postdoc J* **1**: 60-69.

McCauley DW, Bronner-Fraser M. 2003. Neural crest contributions to the lamprey head. *Development* **130**: 2317-2327.

McCauley DW, Bronner-Fraser M. 2006. Importance of SoxE in neural crest development and the evolution of the pharynx. *Nature* **441**: 750-752.

McCune AR, Schimenti JC. 2012. Using genetic networks and homology to understand the evolution of phenotypic traits. *Curr Genomics* **13**: 74-84.

McKee MD, Glimcher MJ, Nanci A. 1992. High-resolution immunolocalization of osteopontin and osteocalcin in bone and cartilage during endochondral ossification in the chicken tibia. *Anat Rec* **234**: 479-492.

McKenzie SK, Oxley PR, Kronauer DJ. 2014. Comparative genomics and transcriptomics in ants provide new insights into the evolution and function of odorant binding and chemosensory proteins. *BMC Genomics* **15**: 718.

McKitrick MC. 1993. Phylogenetic Constraint in Evolutionary Theory: Has It Any Explanatory Power? *Annual Review of Ecology and Systematics* **24**: 307-330.

McNamara KJ. 2012. Heterochrony: the Evolution of Development. *Evolution: Education and Outreach* **5**: 203-218.

Mead TJ, Wang Q, Bhattaram P, Dy P, Afelik S, Jensen J, Lefebvre V. 2013. A far-upstream (-70 kb) enhancer mediates Sox9 auto-regulation in somatic tissues during development and adult regeneration. *Nucleic Acids Res* **41**: 4459-4469.

Meech R, Edelman DB, Jones FS, Makarenkova HP. 2005. The homeobox transcription factor Barx2 regulates chondrogenesis during limb development. *Development* **132**: 2135-2146.

Meulemans D, Bronner-Fraser M. 2005. Central role of gene cooption in neural crest evolution. *J Exp Zool B Mol Dev Evol* **304**: 298-303.

Meulemans D, Bronner-Fraser M. 2007. Insights from amphioxus into the evolution of vertebrate cartilage. *PLoS One* **2**: e787.

Miclea RL, Karperien M, Bosch CAJ, van der Horst G, van der Valk MA, Kobayashi T, Kronenberg HM, Rawadi G, Akçakaya P, Löwik CWGM et al. 2009. Adenomatous polyposis coli-mediated control of β-catenin is essential for both chondrogenic and osteogenic differentiation of skeletal precursors. *BMC developmental biology* **9**: 26.

Miclea RL, van der Horst G, Robanus-Maandag EC, Löwik CW, Oostdijk W, Wit JM, Karperien M. 2011. Apc bridges Wnt/β-catenin and BMP signaling during osteoblast differentiation of KS483 cells. *Exp Cell Res* **317**: 1411-1421.

Miller J, Horner A, Stacy T, Lowrey C, Lian JB, Stein G, Nuckolls GH, Speck NA. 2002. The core-binding factor beta subunit is required for bone formation and hematopoietic maturation. *Nat Genet* **32**: 645-649.

Milz S, Boszczyk A, Putz R. 2002. [Development and functional structure of the epiphyseal plate]. *Der Orthopade* **31**: 835-840.

Minarik M, Stundl J, Fabian P, Jandzik D, Metscher BD, Psenicka M, Gela D, Osorio-Perez A, Arias-Rodriguez L, Horacek I et al. 2017. Pre-oral gut contributes to facial structures in non-teleost fishes. *Nature* **547**: 209-212.

Minina E, Wenzel HM, Kreschel C, Karp S, Gaffield W, McMahon AP, Vortkamp A. 2001. BMP and Ihh/PTHrP signaling interact to coordinate chondrocyte proliferation and differentiation. *Development* **128**: 4523-4534.

Mishina Y, Snider TN. 2014. Neural crest cell signaling pathways critical to cranial bone development and pathology. *Experimental cell research* **325**: 138-147.

Miura S, Hanaoka K, Togashi S. 2008. Skeletogenesis in Xenopus tropicalis: characteristic bone development in an anuran amphibian. *Bone* **43**: 901-909.

Miyazaki T, Kanatani N, Rokutanda S, Yoshida C, Toyosawa S, Nakamura R, Takada S, Komori T. 2008. Inhibition of the terminal differentiation of odontoblasts and their transdifferentiation into osteoblasts in Runx2 transgenic mice. *Archives of histology and cytology* **71**: 131-146.

Mizoguchi I, Takahashi I, Sasano Y, Kagayama M, Kuboki Y, Mitani H. 1997. Localization of types I, II and X collagen and osteocalcin in intramembranous, endochondral and chondroid bone of rats. *Anat Embryol (Berl)* **196**: 291-297.

Mochida Y, Parisuthiman D, Pornprasertsuk-Damrongsri S, Atsawasuwan P, Sricholpech M, Boskey AL, Yamauchi M. 2009. Decorin modulates collagen matrix assembly and mineralization. *Matrix biology : journal of the International Society for Matrix Biology* **28**: 44-52.

Moczek AP. 2008. On the origins of novelty in development and evolution. *Bioessays* **30**: 432-447.

Moradian-Oldak J. 2012. Protein-mediated enamel mineralization. *Front Biosci (Landmark Ed)* **17**: 1996-2023.

Mori-Akiyama Y, Akiyama H, Rowitch DH, de Crombrugghe B. 2003. Sox9 is required for determination of the chondrogenic cell lineage in the cranial neural crest. *Proc Natl Acad Sci U S A* **100**: 9360-9365.

Moriishi T, Shibata Y, Tsukazaki T, Yamaguchi A. 2005. Expression profile of Xenopus banded hedgehog, a homolog of mouse Indian hedgehog, is related to the late development of endochondral ossification in Xenopus laevis. *Biochem Biophys Res Commun* **328**: 867-873.

Morris SC, Caron JB. 2012. Pikaia gracilens Walcott, a stem-group chordate from the Middle Cambrian of British Columbia. *Biol Rev Camb Philos Soc* **87**: 480-512.

Morris SC, Caron JB. 2014. A primitive fish from the Cambrian of North America. *Nature* **512**: 419-422.

Mundlos S, Olsen BR. 1997. Heritable diseases of the skeleton. Part I: Molecular insights into skeletal development-transcription factors and signaling pathways. *Faseb Journal* **11**: 125-132.

Mundlos S, Schwahn B, Reichert T, Zabel B. 1992. Distribution of osteonectin mRNA and protein during human embryonic and fetal development. *J Histochem Cytochem* **40**: 283-291.

Murray PG, Hanson D, Coulson T, Stevens A, Whatmore A, Poole RL, Mackay DJ, Black GC, Clayton PE. 2013. 3-M syndrome: a growth disorder associated with IGF2 silencing. *Endocrine connections* **2**: 225-235.

Nah GS, Tay BH, Brenner S, Osato M, Venkatesh B. 2014a. Characterization of the Runx gene family in a jawless vertebrate, the Japanese lamprey (Lethenteron japonicum). *PLoS One* **9**: e113445.

Nah GSS, Lim ZW, Tay B-H, Osato M, Venkatesh B. 2014b. Runx family genes in a cartilaginous fish, the elephant shark (Callorhinchus milii). *PloS one* **9**: e93816-e93816.

Nakase T, Takaoka K, Hirakawa K, Hirota S, Takemura T, Onoue H, Takebayashi K, Kitamura Y, Nomura S. 1994. Alterations in the expression of osteonectin, osteopontin and osteocalcin mRNAs during the development of skeletal tissues in vivo. *Bone Miner* **26**: 109-122.

Nakashima K, de Crombrugghe B. 2003. Transcriptional mechanisms in osteoblast differentiation and bone formation. *Trends in Genetics* **19**: 458-466.

Nakashima K, Zhou X, Kunkel G, Zhang Z, Deng JM, Behringer RR, de Crombrugghe B. 2002. The novel zinc finger-containing transcription factor osterix is required for osteoblast differentiation and bone formation. *Cell* **108**: 17-29.

Nakatani T, Partridge NC. 2017. MEF2C Interacts With c-FOS in PTH-Stimulated Mmp13 Gene Expression in Osteoblastic Cells. *Endocrinology* **158**: 3778-3791.

Nassif A, Senussi I, Meary F, Loiodice S, Hotton D, Robert B, Bensidhoum M, Berdal A, Babajko S. 2014. Msx1 role in craniofacial bone morphogenesis. *Bone* **66**: 96-104.

Naumann A, Dennis JE, Awadallah A, Carrino DA, Mansour JM, Kastenbauer E, Caplan AI. 2002. Immunochemical and mechanical characterization of cartilage subtypes in rabbit. *J Histochem Cytochem* **50**: 1049-1058.

Necsulea A, Kaessmann H. 2014. Evolutionary dynamics of coding and non-coding transcriptomes. *Nat Rev Genet* **15**: 734-748.

Neufeld SJ, Wang F, Cobb J. 2014. Genetic interactions between Shox2 and Hox genes during the regional growth and development of the mouse limb. *Genetics* **198**: 1117-1126.

Neuhold LA, Killar L, Zhao W, Sung ML, Warner L, Kulik J, Turner J, Wu W, Billinghurst C, Meijers T et al. 2001. Postnatal expression in hyaline cartilage of constitutively active human collagenase-3 (MMP-13) induces osteoarthritis in mice. *J Clin Invest* **107**: 35-44.

Newton PT, Staines KA, Spevak L, Boskey AL, Teixeira CC, Macrae VE, Canfield AE, Farquharson C. 2012. Chondrogenic ATDC5 cells: an optimised model for rapid and physiological matrix mineralisation. *International journal of molecular medicine* **30**: 1187-1193.

Ng JK, Kawakami Y, Büscher D, Raya Á, Itoh T, Koth CM, Esteban CR, Rodríguez-León J, Garrity DM, Fishman MC et al. 2002. The limb identity gene Tbx5 promotes limb initiation by interacting with Wnt2b and Fgf10. *Development* **129**: 5161-5170.

Ng LJ, Wheatley S, Muscat GE, Conway-Campbell J, Bowles J, Wright E, Bell DM, Tam PP, Cheah KS, Koopman P. 1997. SOX9 binds DNA, activates transcription, and coexpresses with type II collagen during chondrogenesis in the mouse. *Dev Biol* **183**: 108-121.

Nguyen JKB, Eames BF. 2020. Evolutionary repression of chondrogenic genes in the vertebrate osteoblast. *The FEBS journal* doi:10.1111/febs.15228.

Nicolae C, Ko YP, Miosge N, Niehoff A, Studer D, Enggist L, Hunziker EB, Paulsson M, Wagener R, Aszodi A. 2007. Abnormal collagen fibrils in cartilage of matrilin-1/matrilin-3-deficient mice. *J Biol Chem* **282**: 22163-22175.

Nie X, Luukko K, Kettunen P. 2006. BMP signalling in craniofacial development. *Int J Dev Biol* **50**: 511-521.

Nieuwkoop P, Faber J. 1956. Normal Table of Xenopus laevis (Daudin) North Holland Publ. *Co, Amsterdam*.

Nikitina N, Bronner-Fraser M, Sauka-Spengler T. 2009. DiI cell labeling in lamprey embryos. *Cold Spring Harb Protoc* **2009**: pdb prot5124.

Nilsson O, Parker EA, Hegde A, Chau M, Barnes KM, Baron J. 2007. Gradients in bone morphogenetic protein-related gene expression across the growth plate. *J Endocrinol* **193**: 75-84.

Nishimura R, Wakabayashi M, Hata K, Matsubara T, Honma S, Wakisaka S, Kiyonari H, Shioi G, Yamaguchi A, Tsumaki N et al. 2012. Osterix regulates calcification and degradation of chondrogenic matrices through matrix metalloproteinase 13 (MMP13) expression in association with transcription factor Runx2 during endochondral ossification. *J Biol Chem* **287**: 33179-33190.

Nishimura R, Hata K, Matsubara T, Wakabayashi M, Yoneda T. 2012. Regulation of bone and cartilage development by network between BMP signalling and transcription factors. *The Journal of Biochemistry* **151**: 247-254.

Niu P, Zhong Z, Wang M, Huang G, Xu S, Hou Y, Yan Y, Wang H. 2017. Zinc finger transcription factor Sp7/Osterix acts on bone formation and regulates col10a1a expression in zebrafish. *Science Bulletin* **62**: 174-184.

Noden DM. 1988. Interactions and fates of avian craniofacial mesenchyme. *Development* **103 Suppl**: 121-140.

Northcutt RG, Gans C. 1983. The genesis of neural crest and epidermal placodes: a reinterpretation of vertebrate origins. *Q Rev Biol* **58**: 1-28.

Nyman JS, Lynch CC, Perrien DS, Thiolloy S, O'Quinn EC, Patil CA, Bi X, Pharr GM, Mahadevan-Jansen A, Mundy GR. 2011. Differential effects between the loss of MMP-2 and MMP-9 on structural and tissue-level properties of bone. *J Bone Miner Res* **26**: 1252-1260.

Oh CD, Lu Y, Liang S, Mori-Akiyama Y, Chen D, de Crombrugghe B, Yasuda H. 2014. SOX9 regulates multiple genes in chondrocytes, including genes encoding ECM proteins, ECM modification enzymes, receptors, and transporters. *PLoS One* **9**: e107577.

Ohazama A, Sharpe PT. 2004. TNF signalling in tooth development. *Curr Opin Genet Dev* **14**: 513-519.

Ohba S, He X, Hojo H, McMahon AP. 2015. Distinct Transcriptional Programs Underlie Sox9 Regulation of the Mammalian Chondrocyte. *Cell Rep* **12**: 229-243.

Ohtani K, Yao T, Kobayashi M, Kusakabe R, Kuratani S, Wada H. 2008a. Expression of Sox and fibrillar collagen genes in lamprey larval chondrogenesis with implications for the evolution of vertebrate cartilage. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* **310B**: 596-607.

Ohtani K, Yao T, Kobayashi M, Kusakabe R, Kuratani S, Wada H. 2008b. Expression of Sox and fibrillar collagen genes in lamprey larval chondrogenesis with implications for the evolution of vertebrate cartilage. *J Exp Zool B Mol Dev Evol* **310**: 596-607.

Ohyama Y, Tanaka T, Shimizu T, Matsui H, Sato H, Koitabashi N, Doi H, Iso T, Arai M, Kurabayashi M. 2012. Runx2/Smad3 complex negatively regulates TGF-β-induced connective tissue growth factor gene expression in vascular smooth muscle cells. *Journal of atherosclerosis and thrombosis* **19**: 23-35.

Okuma T, Hirata M, Yano F, Mori D, Kawaguchi H, Chung UI, Tanaka S, Saito T. 2015. Regulation of mouse chondrocyte differentiation by CCAAT/enhancer-binding proteins. *Biomed Res* **36**: 21-29.

Oldknow KJ, MacRae VE, Farquharson C. 2015. Endocrine role of bone: recent and emerging perspectives beyond osteocalcin. *The Journal of endocrinology* **225**: R1-19.

Ortega N, Behonick DJ, Werb Z. 2004. Matrix remodeling during endochondral ossification. *Trends Cell Biol* **14**: 86-93.

Orth MW, Luchene LJ, Schmid TM. 1996. Type X collagen isolated from the hypertrophic cartilage of embryonic chick tibiae contains both hydroxylysyl- and lysylpyridinoline cross-links. *Biochem Biophys Res Commun* **219**: 301-305.

Ota KG, Fujimoto S, Oisi Y, Kuratani S. 2011. Identification of vertebra-like elements and their possible differentiation from sclerotomes in the hagfish. *Nat Commun* **2**: 373.

Ota KG, Kuraku S, Kuratani S. 2007. Hagfish embryology with reference to the evolution of the neural crest. *Nature* **446**: 672-675.

Otto F, Thornell AP, Crompton T, Denzel A, Gilmour KC, Rosewell IR, Stamp GW, Beddington RS, Mundlos S, Olsen BR et al. 1997. Cbfa1, a candidate gene for cleidocranial dysplasia syndrome, is essential for osteoblast differentiation and bone development [see comments]. *Cell* **89**: 765-771.

Ovchinnikov D. 2009. Alcian Blue/Alizarin Red Staining of Cartilage and Bone in Mouse. *Cold Spring Harbor protocols* **2009**: pdb.prot5170.

Owen TA, Bortell R, Yocum SA, Smock SL, Zhang M, Abate C, Shalhoub V, Aronin N, Wright KL, van Wijnen AJ et al. 1990. Coordinate occupancy of AP-1 sites in the vitamin D-responsive and CCAAT box elements by Fos-Jun in the osteocalcin gene: model for phenotype suppression of transcription. *Proc Natl Acad Sci U S A* **87**: 9990-9994.

Pacifici M, Oshima O, Fisher LW, Young MF, Shapiro IM, Leboy PS. 1990. Changes in osteonectin distribution and levels are associated with mineralization of the chicken tibial growth cartilage. *Calcif Tissue Int* **47**: 51-61.

Panganiban G, Rubenstein JL. 2002. Developmental functions of the Distal-less/Dlx homeobox genes. *Development* **129**: 4371-4386.

Pantalacci S, Semon M. 2015. Transcriptomics of developing embryos and organs: A raising tool for evo-devo. *J Exp Zool B Mol Dev Evol* **324**: 363-371.

Park J, Gebhardt M, Golovchenko S, Perez-Branguli F, Hattori T, Hartmann C, Zhou X, deCrombrugghe B, Stock M, Schneider H et al. 2015. Dual pathways to endochondral osteoblasts: a novel chondrocyte-derived osteoprogenitor cell identified in hypertrophic cartilage. *Biol Open* **4**: 608-621.

Park JY, Jang H, Curry TE, Sakamoto A, Jo M. 2013. Prostate androgen-regulated mucin-like protein 1: a novel regulator of progesterone metabolism. *Mol Endocrinol* **27**: 1871-1886.

Park SH, Lee J, Kang MA, Moon YJ, Wang SI, Kim KM, Park BH, Jang KY, Kim JR. 2016. Potential of l-thyroxine to differentiate osteoblast-like cells via Angiopoietin1. *Biochem Biophys Res Commun* **478**: 1409-1415.

Pataquiva-Mateus AY, Wu HC, Lucchesi C, Ferraz MP, Monteiro FJ, Spector M. 2012. Supplementation of collagen scaffolds with SPARC to facilitate mineralization. *Journal of biomedical materials research Part B, Applied biomaterials* **100**: 862-870.

Pawaputanon Na Mahasarakham C, Ezura Y, Kawasaki M, Aryal Ac S, Moriya S, Yamada T, Izu Y, Nishimori K, Izumi Y, Noda M. 2015. BMP-2 Enhances Lgr4 Gene Expression in Osteoblastic Cells. *Journal of Cellular Physiology* **231**.

Pawlowska E, Wojcik KA, Synowiec E, Szczepanska J, Blasiak J. 2015. Expression of RUNX2 and its signaling partners TCF7, FGFR1/2 in cleidocranial dysplasia. *Acta Biochim Pol* **62**: 123-126.

Peacock JD, Huk DJ, Ediriweera HN, Lincoln J. 2011. Sox9 transcriptionally represses Spp1 to prevent matrix mineralization in maturing heart valves and chondrocytes. *PLoS One* **6**: e26769.

Peichel CL, Nereng KS, Ohgi KA, Cole BL, Colosimo PF, Buerkle CA, Schluter D, Kingsley DM. 2001. The genetic architecture of divergence between threespine stickleback species. *Nature* **414**: 901-905.

Peignoux-Deville J, Lallier F, Vidal B. 1982. Evidence for the presence of osseous tissue in dogfish vertebrae. *Cell Tissue Res* **222**: 605-614.

Person P, Philpott DE. 1969. The nature and significance of invertebrate cartilages. *Biol Rev Camb Philos Soc* **44**: 1-16.

Peruzzi B, Cappariello A, Del Fattore A, Rucci N, De Benedetti F, Teti A. 2012. c-Src and IL-6 inhibit osteoblast differentiation and integrate IGFBP5 signalling. *Nature Communications* **3**: 630.

Peter IS, Davidson EH. 2011. A gene regulatory network controlling the embryonic specification of endoderm. *Nature* **474**: 635-639.

Poole AR, Pidoux I, Rosenberg L. 1982. Role of proteoglycans in endochondral ossification: immunofluorescent localization of link protein and proteoglycan monomer in bovine fetal epiphyseal growth plate. *J Cell Biol* **92**: 249-260.

Poole AR, Laverty S, Mwale F. 2000. Endochondral bone formation and development in the axial and appendicular skeleton. In *The Osteoporosis Primer*, doi:DOI: 10.1017/CBO9780511545795.002 (ed. D Goltzman, JE Henderson), pp. 3-17. Cambridge University Press, Cambridge.

Pratap J, Javed A, Languino LR, van Wijnen AJ, Stein JL, Stein GS, Lian JB. 2005. The Runx2 Osteogenic Transcription Factor Regulates Matrix Metalloproteinase 9 in Bone Metastatic Cancer Cells and Controls Cell Invasion. *Molecular and Cellular Biology* **25**: 8581-8591.

Provot S, Kempf H, Murtaugh LC, Chung U-i, Kim D-W, Chyung J, Kronenberg HM, Lassar AB. 2006. Nkx3.2/Bapx1 acts as a negative regulator of chondrocyte maturation. *Development* **133**: 651-662.

Prummel KD, Nieuwenhuize S, Mosimann C. 2020. The lateral plate mesoderm. *Development* **147**: dev175059.

Pullig O, Weseloh G, Gauer S, Swoboda B. 2000. Osteopontin is expressed by adult human osteoarthritic chondrocytes: protein and mRNA analysis of normal and osteoarthritic cartilage. *Matrix Biology* **19**: 245-255.

Purnell BA. 2020. Generating a new transcriptional network. *Science* **367**: 36-37.

Qin X, Jiang Q, Miyazaki T, Komori T. 2018. Runx2 regulates cranial suture closure by inducing hedgehog, Fgf, Wnt and Pthlh signaling pathway gene expressions in suture mesenchymal cells. *Human Molecular Genetics* **28**: 896-911.

Ramaesh T, Bard JBL. 2003. The growth and morphogenesis of the early mouse mandible: a quantitative analysis. *Journal of anatomy* **203**: 213-222.

Rees SG, Hughes Wassell DT, Waddington RJ, Embery G. 2001. Interaction of bone proteoglycans and proteoglycan components with hydroxyapatite. *Biochim Biophys Acta* **1568**: 118-128.

Reif W-E. 1982. Evolution of Dermal Skeleton and Dentition in Vertebrates. In *Evolutionary Biology: Volume 15*, doi:10.1007/978-1-4615-6968-8_7 (ed. MK Hecht, et al.), pp. 287-368. Springer US, Boston, MA.

Renn J, Winkler C. 2014. Osterix/Sp7 regulates biomineralization of otoliths and bone in medaka (Oryzias latipes). *Matrix biology : journal of the International Society for Matrix Biology* **34**: 193-204.

Roach HI. 1994. Why does bone matrix contain non-collagenous proteins? The possible roles of osteocalcin, osteonectin, osteopontin and bone sialoprotein in bone mineralisation and resorption. *Cell biology international* **18**: 617-628.

Roach HI. 1997. New aspects of endochondral ossification in the chick: chondrocyte apoptosis, bone formation by former chondrocytes, and acid phosphatase activity in the endochondral bone matrix. *J Bone Miner Res* **12**: 795-805.

Roach HI. 1999. Association of matrix acid and alkaline phosphatases with mineralization of cartilage and endochondral bone. *Histochem J* **31**: 53-61.

Robert AW, Angulski ABB, Spangenberg L, Shigunov P, Pereira IT, Bettes PSL, Naya H, Correa A, Dallagiovanna B, Stimamiglio MA. 2018. Gene expression analysis of human adipose tissue-derived stem cells during the initial steps of in vitro osteogenesis. *Scientific Reports* **8**: 4739.

Robertson G, Hirst M, Bainbridge M, Bilenky M, Zhao Y, Zeng T, Euskirchen G, Bernier B, Varhol R, Delaney A et al. 2007. Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat Methods* **4**: 651-657.

Robledo RF, Rajan L, Li X, Lufkin T. 2002. The Dlx5 and Dlx6 homeobox genes are essential for craniofacial, axial, and appendicular skeletal development. *Genes & development* **16**: 1089-1101.

Robson P, Wright GM, Sitarz E, Maiti A, Rawat M, Youson JH, Keeley FW. 1993. Characterization of lamprin, an unusual matrix protein from lamprey cartilage. Implications for evolution, structure, and assembly of elastin and other fibrillar proteins. *The Journal of biological chemistry* **268**: 1440-1447.

Rocha C, Papon L, Cacheux W, Marques Sousa P, Lascano V, Tort O, Giordano T, Vacher S, Lemmers B, Mariani P et al. 2014. Tubulin glycylases are required for primary cilia, control of cell proliferation and tumor development in colon. *EMBO J* **33**: 2247-2260.

Rosset EM, Bradshaw AD. 2016. SPARC/osteonectin in mineralized tissue. *Matrix biology : journal of the International Society for Matrix Biology* **52-54**: 78-87.

Rotllant J, Liu D, Yan YL, Postlethwait JH, Westerfield M, Du SJ. 2008. Sparc (Osteonectin) functions in morphogenesis of the pharyngeal skeleton and inner ear. *Matrix biology : journal of the International Society for Matrix Biology* **27**: 561-572.

Rotwein P. 2018a. The insulin-like growth factor 2 gene and locus in nonmammalian vertebrates: Organizational simplicity with duplication but limited divergence in fish. *The Journal of biological chemistry* **293**: 15912-15932.

Rotwein P. 2018b. Insulinlike Growth Factor 1 Gene Variation in Vertebrates. *Endocrinology* **159**: 2288-2305.

Roughley PJ. 2006. The structure and function of cartilage proteoglycans. *Eur Cell Mater* **12**: 92-101.

Rychel AL, Smith SE, Shimamoto HT, Swalla BJ. 2006. Evolution and development of the chordates: collagen and pharyngeal cartilage. *Mol Biol Evol* **23**: 541-549.

Rychel AL, Swalla BJ. 2007. Development and evolution of chordate cartilage. *J Exp Zool B Mol Dev Evol* **308**: 325-335.

Salie R, Kneissel M, Vukevic M, Zamurovic N, Kramer I, Evans G, Gerwin N, Mueller M, Kinzel B, Susa M. 2010. Ubiquitous overexpression of Hey1 transcription factor leads to osteopenia and chondrocyte hypertrophy in bone. *Bone* **46**: 680-694.

Samee N, Geoffroy V, Marty C, Schiltz C, Vieux-Rochas M, Levi G, de Vernejoul MC. 2008. Dlx5, a positive regulator of osteoblastogenesis, is essential for osteoblast-osteoclast coupling. *The American journal of pathology* **173**: 773-780.

Sanderson MJ, Doyle JJ. 1992. Reconstruction of Organismal and Gene Phylogenies from Data on Multigene Families: Concerted Evolution, Homoplasy, and Confidence. *Systematic Biology* **41**: 4-17.

Sansom IJ, Smith MP, Armstrong HA, Smith MM. 1992. Presence of the earliest vertebrate hard tissue in conodonts. *Science* **256**: 1308-1311.

Sasaki T, Amizuka N, Irie K, Ejiri S, Ozawa H. 2000. Localization of alkaline phosphatase and osteopontin during matrix mineralization in the developing cartilage of coccygeal vertebrae. *Archives of histology and cytology* **63**: 271-284.

Sato M, Morii E, Komori T, Kawahata H, Sugimoto M, Terai K, Shimizu H, Yasui T, Ogihara H, Yasui N et al. 1998. Transcriptional regulation of osteopontin gene in vivo by PEBP2alphaA/CBFA1 and ETS1 in the skeletal tissues. *Oncogene* **17**: 1517-1525.

Sato Y. 2012. The vasohibin family: a novel family for angiogenesis regulation. *The Journal of Biochemistry* **153**: 5-11.

Satoh N, Tagawa K, Lowe CJ, Yu JK, Kawashima T, Takahashi H, Ogasawara M, Kirschner M, Hisata K, Su YH et al. 2014. On a possible evolutionary link of the stomochord of hemichordates to pharyngeal organs of chordates. *Genesis* **52**: 925-934.

Satokata I, Ma L, Ohshima H, Bei M, Woo I, Nishizawa K, Maeda T, Takano Y, Uchiyama M, Heaney S et al. 2000. Msx2 deficiency in mice causes pleiotropic defects in bone growth and ectodermal organ formation. *Nat Genet* **24**: 391-395.

Satokata I, Maas R. 1994. Msx1 deficient mice exhibit cleft palate and abnormalities of craniofacial and tooth development. *Nat Genet* **6**: 348-356.

Satokata I, Ma L, Ohshima H, Bei M, Woo I, Nishizawa K, Maeda T, Takano Y, Uchiyama M, Heaney S et al. 2000. Msx2 deficiency in mice causes pleiotropic defects in bone growth and ectodermal organ formation. *Nat Genet* **24**: 391-395.

Schilling TF, Le Pabic P. 2014. Chapter 7 - Neural Crest Cells in Craniofacial Skeletal Development. In *Neural Crest Cells*, doi:https://doi.org/10.1016/B978-0-12-401730-6.00008-9 (ed. PA Trainor), pp. 127-151. Academic Press, Boston.

Schneider RA, Hu D, Helms JA. 1999. From head to toe: conservation of molecular signals regulating limb and craniofacial morphogenesis. *Cell Tissue Res* **296**: 103-109.

Seidel R, Blumer M, Zaslansky P, Knotel D, Huber DR, Weaver JC, Fratzl P, Omelon S, Bertinetti L, Dean MN. 2017. Ultrastructural, material and crystallographic description of endophytic masses - A possible damage response in shark and ray tessellated calcified cartilage. *J Struct Biol* **198**: 5-18.

Seidel R, Lyons K, Blumer M, Zaslansky P, Fratzl P, Weaver JC, Dean MN. 2016. Ultrastructural and developmental features of the tessellated endoskeleton of elasmobranchs (sharks and rays). *J Anat* **229**: 681-702.

Sekiya I, Tsuji K, Koopman P, Watanabe H, Yamada Y, Shinomiya K, Nifuji A, Noda M. 2000. SOX9 enhances aggrecan gene promoter/enhancer activity and is up-regulated by retinoic acid in a cartilage-derived cell line, TC6. *J Biol Chem* **275**: 10738-10744.

Senarath-Yapa K, Li S, Meyer N, Longaker M, Quarto N. 2013. Integration of Multiple Signaling Pathways Determines Differences in the Osteogenic Potential and Tissue Regeneration of Neural Crest-Derived and Mesoderm-Derived Calvarial Bone. *International journal of molecular sciences* **14**: 5978-5997.

Seto T, Yamamoto T, Shimojima K, Shintaku H. 2017. A novel COL1A1 mutation in a family with osteogenesis imperfecta associated with phenotypic variabilities. *Human Genome Variation* **4**: 17007.

Shadwick JDL, Ruiz-Trillo I. 2012. A genomic survey shows that the haloarchaeal type tyrosyl tRNA synthetase is not a synapomorphy of opisthokonts. *European Journal of Protistology* **48**: 89-93.

Shao J, Zhou Y, Xiao Y. 2018. The regulatory roles of Notch in osteocyte differentiation via the crosstalk with canonical Wnt pathways during the transition of osteoblasts to osteocytes. *Bone* **108**: 165-178.

Sharff KA, Song W-X, Luo X, Tang N, Luo J, Chen J, Bi Y, He B-C, Huang J, Li X. 2009. Hey1 basic helix-loop-helix protein plays an important role in mediating BMP9-induced osteogenic differentiation of mesenchymal progenitor cells. *Journal of Biological Chemistry* **284**: 649-659.

Shubin N, Tabin C, Carroll S. 2009. Deep homology and the origins of evolutionary novelty. *Nature* **457**: 818-823.

Shubin NH. 2002. Origin of evolutionary novelty: examples from limbs. *J Morphol* **252**: 15-28.

Si Y, Liu P, Li P, Brutnell TP. 2014. Model-based clustering for RNA-seq data. *Bioinformatics* **30**: 197-205.

Siqueira MF, Flowers S, Bhattacharya R, Faibish D, Behl Y, Kotton DN, Gerstenfeld L, Moran E, Graves DT. 2011. FOXO1 modulates osteoblast differentiation. *Bone* **48**: 1043-1051.

Sire JY, Davit-Beal T, Delgado S, Gu X. 2007. The origin and evolution of enamel mineralization genes. *Cells Tissues Organs* **186**: 25-48.

Sire JY, Donoghue PC, Vickaryous MK. 2009. Origin and evolution of the integumentary skeleton in non-tetrapod vertebrates. *J Anat* **214**: 409-440.

Smith MM, Hall BK. 1990a. Development and evolutionary origins of vertebrate skeletogenic and odontogenic tissues. *Biol Rev Camb Philos Soc* **65**: 277-373.

Smith MM, Hall BK. 1990b. Development and evolutionary origins of vertebrate skeletogenic and odontogenic tissues. *Biological Reviews of the Cambridge Philosophical Society* **65**: 277-373.

Smith MM, Hall BK. 1993. A Developmental Model for Evolution of the Vertebrate Exoskeleton and Teeth. In *Evolutionary Biology*, Vol 27 (ed. MK Hecht). Plenum Press, New York.

Smith SE, Douglas R, Silva KBd, Swalla BJ. 2003. Morphological and molecular identification of Saccoglossus species (Hemichordata: Harrimaniidae) in the Pacific Northwest. *Canadian journal of zoology* **81**: 133-141.

Smits P, Dy P, Mitra S, Lefebvre Vr. 2004. Sox5 and Sox6 are needed to develop and maintain source, columnar, and hypertrophic chondrocytes in the cartilage growth plate. *Journal of Cell Biology* **164**: 747-758.

Smits P, Li P, Mandel J, Zhang Z, Deng JM, Behringer RR, de Crombrugghe B, Lefebvre V. 2001a. The Transcription Factors L-Sox5 and Sox6 Are Essential for Cartilage Formation. *Developmental cell* **1**: 277-290.

Smits P, Li P, Mandel J, Zhang Z, Deng JM, Behringer RR, de Croumbrugghe B, Lefebvre V. 2001b. The transcription factors L-Sox5 and Sox6 are essential for cartilage formation. *Dev Cell* **1**: 277-290.

Sodek J, Li JJ, Kim RH, Ogata Y, Yamauchi M. 1996. Characterization of the bone sialoprotein (BSP) gene promoter. *Connect Tissue Res* **35**: 23-31.

Sohaskey ML, Jiang Y, Zhao JJ, Mohr A, Roemer F, Harland RM. 2010. Osteopotentia regulates osteoblast maturation, bone formation, and skeletal integrity in mice. *J Cell Biol* **189**: 511-525.

Someren EPv, Vaes BLT, Steegenga WT, Sijbers AM, Dechering KJ, Reinders MJT. 2005. Unravelling the murine osteoblast differentiation pathway by network structure analysis using time-series microarray data. In *2005 IEEE Computational Systems Bioinformatics Conference - Workshops (CSBW'05)*, doi:10.1109/CSBW.2005.138, pp. 172-173.

Sophia Fox AJ, Bedi A, Rodeo SA. 2009. The basic science of articular cartilage: structure, composition, and function. *Sports health* **1**: 461-468.

Spater D, Hill TP, O'Sullivan R J, Gruber M, Conner DA, Hartmann C. 2006. Wnt9a signaling is required for joint integrity and regulation of Ihh during chondrogenesis. *Development* **133**: 3039-3049.

St-Jacques B, Hammerschmidt M, McMahon AP. 1999. Indian hedgehog signaling regulates proliferation and differentiation of chondrocytes and is essential for bone formation. *Genes Dev* **13**: 2072-2086.

Stamper BD, Park SS, Beyer RP, Bammler TK, Cunningham ML. 2012. Unique sex-based approach identifies transcriptomic biomarkers associated with non-syndromic craniosynostosis. *Gene Regul Syst Bio* **6**: 81-92.

Stayton CT. 2008. Is convergence surprising? An examination of the frequency of convergence in simulated datasets. *J Theor Biol* **252**: 1-14.

Stemple DL. 2005. Structure and function of the notochord: an essential organ for chordate development. *Development* **132**: 2503-2512.

Stock M, Schäfer H, Fliegauf M, Otto F. 2004. Identification of Novel Target Genes of the Bone-Specific Transcription Factor Runx2. *Journal of Bone and Mineral Research* **19**: 959-972.

Streit A, Tambalo M, Chen J, Grocott T, Anwar M, Sosinsky A, Stern CD. 2013. Experimental approaches for gene regulatory network construction: the chick as a model system. *Genesis* **51**: 296-310.

Stricker S, Fundele R, Vortkamp A, Mundlos S. 2002. Role of Runx genes in chondrocyte differentiation. *Dev Biol* **245**: 95-108.

Stuart JM, Segal E, Koller D, Kim SK. 2003. A gene-coexpression network for global discovery of conserved genetic modules. *Science* **302**: 249-255.

Su N, Jin M, Chen L. 2014. Role of FGF/FGFR signaling in skeletal development and homeostasis: learning from mouse models. *Bone Research* **2**: 14003.

Sun MGF, Sikora M, Costanzo M, Boone C, Kim PM. 2012. Network evolution: rewiring and signatures of conservation in signaling. *PLoS Comput Biol* **8**: e1002411-e1002411.

Tadic T, Dodig M, Erceg I, Marijanovic I, Mina M, Kalajzic Z, Velonis D, Kronenberg MS, Kosher RA, Ferrari D et al. 2002. Overexpression of Dlx5 in chicken calvarial cells accelerates osteoblastic differentiation. *J Bone Miner Res* **17**: 1008-1014.

Takeda S, Bonnamy JP, Owen MJ, Ducy P, Karsenty G. 2001. Continuous expression of Cbfa1 in nonhypertrophic chondrocytes uncovers its ability to induce hypertrophic chondrocyte differentiation and partially rescues Cbfa1-deficient mice. *Genes Dev* **15**: 467-481.

Talwar RM, Wong BS, Svoboda K, Harper RP. 2006. Effects of estrogen on chondrocyte proliferation and collagen synthesis in skeletally mature articular cartilage. *Journal of oral and maxillofacial surgery : official journal of the American Association of Oral and Maxillofacial Surgeons* **64**: 600-609.

Tamamura Y, Katsube K, Mera H, Itokazu M, Wakitani S. 2017. Irx3 and Bmp2 regulate mouse mesenchymal cell chondrogenic differentiation in both a Sox9-dependent and -independent manner. *J Cell Physiol* **232**: 3317-3336.

Tamamura Y, Otani T, Kanatani N, Koyama E, Kitagaki J, Komori T, Yamada Y, Costantini F, Wakisaka S, Pacifici M et al. 2005. Developmental regulation of Wnt/beta-catenin signals is required for growth plate assembly, cartilage integrity, and endochondral ossification. *J Biol Chem* **280**: 19185-19195.

Tanoue H, Morinaga J, Yoshizawa T, Yugami M, Itoh H, Nakamura T, Uehara Y, Masuda T, Odagiri H, Sugizaki T et al. 2018. Angiopoietin-like protein 2 promotes chondrogenic differentiation during bone growth as a cartilage matrix factor. *Osteoarthritis and Cartilage* **26**: 108-117.

Tarazona OA, Slota LA, Lopez DH, Zhang G, Cohn MJ. 2016. The genetic program for cartilage development has deep homology within Bilateria. *Nature* **533**: 86-89.

Tarkkonen K, Hieta R, Kytölä V, Nykter M, Kiviranta R. 2017. Comparative analysis of osteoblast gene expression profiles and Runx2 genomic occupancy of mouse and human osteoblasts in vitro. *Gene* **626**: 119-131.

Tautz D, Domazet-Loso T. 2011. The evolutionary origin of orphan genes. *Nat Rev Genet* **12**: 692-702.

Teixeira CC, Liu Y, Thant LM, Pang J, Palmer G, Alikhani M. 2010. Foxo1, a Novel Regulator of Osteoblast Differentiation and Skeletogenesis. *Journal of Biological Chemistry* **285**: 31055-31065.

Termine JD, Belcourt AB, Conn KM, Kleinman HK. 1981a. Mineral and collagen-binding proteins of fetal calf bone. *The Journal of biological chemistry* **256**: 10403-10408.

Termine JD, Kleinman HK, Whitson SW, Conn KM, McGarvey ML, Martin GR. 1981b. Osteonectin, a bone-specific protein linking mineral to collagen. *Cell* **26**: 99-105.

Thompson AC, Capellini TD, Guenther CA, Chan YF, Infante CR, Menke DB, Kingsley DM. 2018. A novel enhancer near the Pitx1 gene influences development and evolution of pelvic appendages in vertebrates. *Elife* **7**.

Thorogood P. 1988. The developmental specification of the vertebrate skull. *Development* **103 Suppl**: 141-153.

Tickle C. 2004. The contribution of chicken embryology to the understanding of vertebrate limb development. *Mechanisms of Development* **121**: 1019-1029.

Tickle C. 2015. How the embryo makes a limb: determination, polarity and identity. *Journal of Anatomy* **227**: 418-430.

Toll-Riera M, Castelo R, Bellora N, Alba MM. 2009. Evolution of primate orphan proteins. *Biochemical Society transactions* **37**: 778-782.

Tou L, Quibria N, Alexander JM. 2003. Transcriptional regulation of the human Runx2/Cbfa1 gene promoter by bone morphogenetic protein-7. *Mol Cell Endocrinol* **205**: 121-129.

Trinajstic K, Boisvert C, Long J, Maksimenko A, Johanson Z. 2015. Pelvic and reproductive structures in placoderms (stem gnathostomes). *Biol Rev Camb Philos Soc* **90**: 467-501.

Tucker AS, Al Khamis A, Ferguson CA, Bach I, Rosenfeld MG, Sharpe PT. 1999. Conserved regulation of mesenchymal gene expression by Fgf-8 in face and limb development. *Development* **126**: 221-228.

Uchimura T, Hollander JM, Nakamura DS, Liu Z, Rosen CJ, Georgakoudi I, Zeng L. 2017. An essential role for IGF2 in cartilage development and glucose metabolism during postnatal long bone growth. *Development* **144**: 3533-3546.

Ueta C, Iwamoto M, Kanatani N, Yoshida C, Liu Y, Enomoto-Iwamoto M, Ohmori T, Enomoto H, Nakata K, Takada K et al. 2001. Skeletal malformations caused by overexpression of Cbfa1 or its dominant negative form in chondrocytes. *J Cell Biol* **153**: 87-100.

Uy BR, Simoes-Costa M, Sauka-Spengler T, Bronner ME. 2012. Expression of Sox family genes in early lamprey development. *Int J Dev Biol* **56**: 377-383.

Vadlamudi U, Espinoza HM, Ganga M, Martin DM, Liu X, Engelhardt JF, Amendt BA. 2005. PITX2, β-catenin and LEF-1 interact to synergistically regulate the LEF-1 promoter. *Journal of cell science* **118**: 1129-1137.

van Someren EP, Vaes BLT, Steegenga WT, Sijbers AM, Dechering KJ, Reinders MJT. 2005. Least absolute regression network analysis of the murine osteoblast differentiation network. *Bioinformatics* **22**: 477-484.

Vasan NS. 1987. Somite chondrogenesis: the role of the microenvironment. *Cell Differ* **21**: 147-159.

Veis A, Barss J, Dahl T, Rahima M, Stock S. 2002. Mineral-related proteins of sea urchin teeth: Lytechinus variegatus. *Microscopy research and technique* **59**: 342-351.

Venkatesh B, Erdmann MV, Brenner S. 2001. Molecular synapomorphies resolve evolutionary relationships of extant jawed vertebrates. *Proceedings of the National Academy of Sciences of the United States of America* **98**: 11382-11387.

Venkatesh B, Lee AP, Ravi V, Maurya AK, Lian MM, Swann JB, Ohta Y, Flajnik MF, Sutoh Y, Kasahara M et al. 2014. Elephant shark genome provides unique insights into gnathostome evolution. *Nature* **505**: 174-179.

Vieira FA, Thorne MA, Stueber K, Darias M, Reinhardt R, Clark MS, Gisbert E, Power DM. 2013. Comparative analysis of a teleost skeleton transcriptome provides insight into its regulation. *Gen Comp Endocrinol* **191**: 45-58.

Vieux-Rochas M, Bouhali K, Mantero S, Garaffo G, Provero P, Astigiano S, Barbieri O, Caratozzolo MF, Tullo A, Guerrini L et al. 2013. BMP-Mediated Functional Cooperation between Dlx5;Dlx6 and Msx1;Msx2 during Mammalian Limb Development. *PLOS ONE* **8**: e51700.

Villanueva AR, Kujawa M, Mathews CH, Parfitt AM. 1983. Identification of the mineralization front: comparison of a modified toluidine blue stain with tetracycline fluorescence. *Metab Bone Dis Relat Res* **5**: 41-45.

Voldoire E, Brunet F, Naville M, Volff JN, Galiana D. 2017. Expansion by whole genome duplication and evolution of the sox gene family in teleost fish. *PLoS One* **12**: e0180936.

Volkmann D, Baluska F. 2006. Gravity: one of the driving forces for evolution. *Protoplasma* **229**: 143-148.

Vortkamp A. 2001. Interaction of growth factors regulating chondrocyte differentiation in the developing embryo. *Osteoarthritis Cartilage* **9 Suppl A**: S109-117.

Vortkamp A, Lee K, Lanske B, Segre GV, Kronenberg HM, Tabin CJ. 1996a. Regulation of rate of cartilage differentiation by Indian hedgehog and PTH-related protein. *Science* **273**: 613-622.

Vortkamp A, Lee K, Lanske B, Segre GV, Kronenberg HM, Tabin CJ. 1996b. Regulation of rate of cartilage differentiation by Indian hedgehog and PTH-related protein [see comments]. *Science* **273**: 613-622.

Wachsmuth L, Soder S, Fan Z, Finger F, Aigner T. 2006. Immunolocalization of matrix proteins in different human cartilage subtypes. *Histol Histopathol* **21**: 477-485.

Wada H. 2010. Origin and genetic evolution of the vertebrate skeleton. *Zoological science* **27**: 119-123.

Wagner DO, Aspenberg P. 2011. Where did bone come from? *Acta Orthop* **82**: 393-398.

Wagner GP, Lynch VJ. 2010. Evolutionary novelties. *Curr Biol* **20**: R48-52.

Wahba GM, Hostikka SL, Carpenter EM. 2001. The Paralogous Hox Genes Hoxa10 and Hoxd10 Interact to Pattern the Mouse Hindlimb Peripheral Nervous System and Skeleton. *Developmental Biology* **231**: 87-102.

Wallace A. 1997. The Origin of Animal Body Plans: A Study in Evolutionary Developmental Biology. Cambridge: Cambridge University Press.

Wang D, Gilbert JR, Zhang X, Zhao B, Ker DFE, Cooper GM. 2019. Calvarial Versus Long Bone: Implications for Tailoring Skeletal Tissue Engineering. *Tissue Engineering Part B: Reviews* **26**: 46-63.

Wang J, Zhou J, Bondy CA. 1999. Igf1 promotes longitudinal bone growth by insulin-like actions augmenting chondrocyte hypertrophy. *Faseb j* **13**: 1985-1990.

Wang L, Jie Q, Yang L. 2017. Chondrocytes-osteoblast transition in endochondral ossification. *Annals of Joint* **2**.

Wang W, Wang Y-G, Reginato AM, Glotzer DJ, Fukai N, Plotkina S, Karsenty G, Olsen BR. 2004. Groucho homologue Grg5 interacts with the transcription factor Runx2–Cbfa1 and modulates its activity during postnatal growth in mice. *Developmental Biology* **270**: 364-381.

Wang X, Manner PA, Horner A, Shum L, Tuan RS, Nuckolls GH. 2004a. Regulation of MMP-13 expression by RUNX2 and FGF2 in osteoarthritic cartilage. *Osteoarthritis Cartilage* **12**: 963-973.

Wang Y, Middleton F, Horton JA, Reichel L, Farnum CE, Damron TA. 2004b. Microarray analysis of proliferative and hypertrophic growth plate zones identifies differentiation markers and signal pathways. *Bone* **35**: 1273-1293.

Wang Z, Gerstein M, Snyder M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* **10**: 57-63.

Watanabe H, Yamada Y, Kimata K. 1998. Roles of aggrecan, a large chondroitin sulfate proteoglycan, in cartilage structure and function. *J Biochem* **124**: 687-693.

Weir EC, Philbrick WM, Amling M, Neff LA, Baron R, Broadus AE. 1996. Targeted overexpression of parathyroid hormone-related peptide in chondrocytes causes chondrodysplasia and delayed endochondral bone formation. *Proc Natl Acad Sci U S A* **93**: 10240-10245.

Welsch U, Erlinger R, Potter IC. 1991. Proteoglycans in the notochord sheath of lampreys. *Acta Histochem* **91**: 59-65.

Weng JJ, Su Y. 2013. Nuclear matrix-targeting of the osteogenic factor Runx2 is essential for its recognition and activation of the alkaline phosphatase gene. *Biochim Biophys Acta* **1830**: 2839-2852.

Whelan NV, Kocot KM, Moroz TP, Mukherjee K, Williams P, Paulay G, Moroz LL, Halanych KM. 2017. Ctenophore relationships and their placement as the sister group to all other animals. *Nat Ecol Evol* **1**: 1737-1746.

Will AJ, Cova G, Osterwalder M, Chan WL, Wittler L, Brieske N, Heinrich V, de Villartay JP, Vingron M, Klopocki E et al. 2017. Composition and dosage of a multipartite enhancer cluster control developmental expression of Ihh (Indian hedgehog). *Nat Genet* **49**: 1539-1545.

Wilt FH, Killian CE, Livingston BT. 2003. Development of calcareous skeletal elements in invertebrates. *Differentiation* **71**: 237-250.

Winterbach W, Van Mieghem P, Reinders M, Wang H, de Ridder D. 2013. Topology of molecular interaction networks. *BMC systems biology* **7**: 90-90.

Witten PE, Huysseune A, Hall BK. 2010. A practical approach for the identification of the many cartilaginous tissues in teleost fish. *Journal of Applied Ichthyology* **26**: 257-262.

Wittler L, Hilger A, Proske J, Pennimpede T, Draaken M, Ebert AK, Rosch W, Stein R, Nothen MM, Reutter H et al. 2012. Murine expression and mutation analyses of the prostate androgen-regulated mucin-like protein 1 (Parm1) gene, a candidate for human epispadias. *Gene* **506**: 392-395.

Wood TWP, Nakamura T. 2018. Problems in Fish-to-Tetrapod Transition: Genetic Expeditions Into Old Specimens. *Frontiers in Cell and Developmental Biology* **6**.

Wright GM, Keeley FW, Robson P. 2001. The unusual cartilaginous tissues of jawless craniates, cephalochordates and invertebrates. *Cell Tissue Res* **304**: 165-174.

Wright GM, Youson JH. 1982. Ultrastructure of mucocartilage in the larval anadromous sea lamprey, Petromyzon marinus L. *Am J Anat* **165**: 39-51.

Wu D, Mandal S, Choi A, Anderson A, Prochazkova M, Perry H, Gil-Da-Silva-Lopes VL, Lao R, Wan E, Tang PL et al. 2015. DLX4 is associated with orofacial clefting and abnormal jaw development. *Hum Mol Genet* **24**: 4340-4352.

Wu H, Whitfield TW, Gordon JA, Dobson JR, Tai PW, van Wijnen AJ, Stein JL, Stein GS, Lian JB. 2014. Genomic occupancy of Runx2 with global expression profiling identifies a novel dimension to control of osteoblastogenesis. *Genome Biol* **15**: R52.

Wu M, Hesse E, Morvan F, Zhang JP, Correa D, Rowe GC, Kiviranta R, Neff L, Philbrick WM, Horne WC et al. 2009. Zfp521 antagonizes Runx2, delays osteoblast differentiation in vitro, and promotes bone formation in vivo. *Bone* **44**: 528-536.

Xiao G, Jiang D, Ge C, Zhao Z, Lai Y, Boules H, Phimphilai M, Yang X, Karsenty G, Franceschi RT. 2005. Cooperative interactions between activating transcription factor 4 and Runx2/Cbfa1 stimulate osteoblast-specific osteocalcin gene expression. *Journal of Biological Chemistry* **280**: 30689-30696.

Xiong Q, Liu Y, Xue Y, Liu S, Wang J, Li P, Wu C, Yang Y, Xiao H. 2018. A novel de novo mutation in COL2A1 leading to spondyloepiphyseal dysplasia congenita in a Chinese family. *Human Genome Variation* **5**: 17059.

Xu L, Zhang G, Zhou Y, Chen Y, Xu W, Wu S, Zhang X. 2011. Stimulation of B7-H3 (CD276) directs the differentiation of human marrow stromal cells to osteoblasts. *Immunobiology* **216**: 1311-1317.

Yamamoto Y, Yoshizawa T, Fukuda T, Shirode-Fukuda Y, Yu T, Sekine K, Sato T, Kawano H, Aihara K, Nakamichi Y et al. 2013. Vitamin D receptor in osteoblasts is a negative regulator of bone mass control. *Endocrinology* **154**: 1008-1020.

Yamashita S, Andoh M, Ueno-Kudoh H, Sato T, Miyaki S, Asahara H. 2009. Sox9 directly promotes Bapx1 gene expression to repress Runx2 in chondrocytes. *Exp Cell Res* **315**: 2231-2240.

Yamashita S, Kataoka K, Yamamoto H, Kato T, Hara S, Yamaguchi K, Renard-Guillet C, Katou Y, Shirahige K, Ochi H et al. 2019. Comparative analysis demonstrates cell type-specific conservation of SOX9 targets between mouse and chicken. *Scientific Reports* **9**: 12560.

Yanagishita M. 1993. Function of proteoglycans in the extracellular matrix. *Acta Pathol Jpn* **43**: 283-293.

Yang J-Y, Cho SW, An JH, Jung JY, Kim SW, Kim SY, Kim JE, Shin CS. 2013. Osteoblast-Targeted Overexpression of TAZ Increases Bone Mass In Vivo. *PLOS ONE* **8**: e56585.

Yang L, Tsang KY, Tang HC, Chan D, Cheah KS. 2014. Hypertrophic chondrocytes can become osteoblasts and osteocytes in endochondral bone formation. *Proc Natl Acad Sci U S A* **111**: 12097-12102.

Yang X, Matsuda K, Bialek P, Jacquot S, Masuoka HC, Schinke T, Li L, Brancorsini S, Sassone-Corsi P, Townes TM et al. 2004a. ATF4 Is a Substrate of RSK2 and an Essential Regulator of Osteoblast Biology: Implication for Coffin-Lowry Syndrome. *Cell* **117**: 387-398.

Yang X, Matsuda K, Bialek P, Jacquot S, Masuoka HC, Schinke T, Li L, Brancorsini S, Sassone-Corsi P, Townes TM et al. 2004b. ATF4 is a substrate of RSK2 and an essential regulator of osteoblast biology; implication for Coffin-Lowry Syndrome. *Cell* **117**: 387-398.

Yang Y, Topol L, Lee H, Wu J. 2003. Wnt5a and Wnt5b exhibit distinct activities in coordinating chondrocyte proliferation and differentiation. *Development* **130**: 1003-1015.

Yano H, Hamanaka R, Nakamura-Ota M, Adachi S, Zhang JJ, Matsuo N, Yoshioka H. 2014. Sp7/Osterix induces the mouse pro-α2(I) collagen gene (Col1a2) expression via the proximal promoter in osteoblastic cells. *Biochemical and Biophysical Research Communications* **452**: 531-536.

Yao Y, Wang Y. 2013. ATDC5: an excellent in vitro model cell line for skeletal development. *J Cell Biochem* **114**: 1223-1229.

Yasui N, Ono K, Konomi H, Nagai Y. 1984. Transitions in collagen types during endochondral ossification in human growth cartilage. *Clin Orthop Relat Res*: 215-218.

Yong LW, Yu JK. 2016. Tracing the evolutionary origin of vertebrate skeletal tissues: insights from cephalochordate amphioxus. *Curr Opin Genet Dev* **39**: 55-62.

Yoon BS, Lyons KM. 2004. Multiple functions of BMPs in chondrogenesis. *Journal of cellular biochemistry* **93**: 93-103.

Yoshida CA, Yamamoto H, Fujita T, Furuichi T, Ito K, Inoue K, Yamana K, Zanma A, Takada K, Ito Y et al. 2004. Runx2 and Runx3 are essential for chondrocyte maturation, and Runx2 regulates limb growth through induction of Indian hedgehog. *Genes Dev* **18**: 952-963.

Yoshida CA, Furuichi T, Fujita T, Fukuyama R, Kanatani N, Kobayashi S, Satake M, Takada K, Komori T. 2002. Core-binding factor beta interacts with Runx2 and is required for skeletal development. *Nat Genet* **32**: 633-638.

Young B, Minugh-Purvis N, Shimo T, St-Jacques B, Iwamoto M, Enomoto-Iwamoto M, Koyama E, Pacifici M. 2006. Indian and sonic hedgehogs regulate synchondrosis growth plate and cranial base development and function. *Developmental Biology* **299**: 272-282.

Ytrehus B, Carlson CS, Lundeheim N, Mathisen L, Reinholt FP, Teige J, Ekman S. 2004. Vascularisation and osteochondrosis of the epiphyseal growth cartilage of the distal femur in pigs--development with age, growth rate, weight and joint shape. *Bone* **34**: 454-465.

Yu JK, Meulemans D, McKeown SJ, Bronner-Fraser M. 2008. Insights from the amphioxus genome on the origin of vertebrate neural crest. *Genome Res* **18**: 1127-1132.

Yu LY, Pei Y, Xia WB, Xing XP, Meng XW, Zhou XY. 2007. Effect of fibroblast growth factor 9 on Runx2 gene promoter activity in MC3T3-E1 and C2C12 cells. *Chinese medical journal* **120**: 491-495.

Yu S, Franceschi RT, Luo M, Fan J, Jiang D, Cao H, Kwon T-G, Lai Y, Zhang J, Patrene K et al. 2009. Critical Role of Activating Transcription Factor 4 in the Anabolic Actions of Parathyroid Hormone in Bone. *PLOS ONE* **4**: e7583.

Yu S, Franceschi RT, Luo M, Fan J, Jiang D, Cao H, Kwon T-G, Lai Y, Zhang J, Patrene K et al. 2009. Critical Role of Activating Transcription Factor 4 in the Anabolic Actions of Parathyroid Hormone in Bone. *PLOS ONE* **4**: e7583.

Yu T, Graf M, Renn J, Schartl M, Larionova D, Huysseune A, Witten PE, Winkler C. 2017. A vertebrate-specific and essential role for osterix in osteogenesis revealed by gene knockout in the teleost medaka. *Development* **144**: 265-271.

Yuan B, Takaiwa M, Clemens TL, Feng JQ, Kumar R, Rowe PS, Xie Y, Drezner MK. 2008. Aberrant Phex function in osteoblasts and osteocytes alone underlies murine X-linked hypophosphatemia. *J Clin Invest* **118**: 722-734.

Zakany J, Duboule D. 2007. The role of HOX genes during vertebrate limb development. *Current opinion in genetics & development* **17**: 359-366.

Zakin L, De Robertis EM. 2004. Inactivation of mouse Twisted gastrulation reveals its role in promoting Bmp4 activity during forebrain development. *Development* **131**: 413-424.

Zangerl R. 1966. A new shark in the family Edestidae, *Ornithoprion hertwigi* from the Pennsylvania Mecca and Logan Quarry Shales of Indiana. *Fieldiana Geol* **16**: 1-43.

Zaragoza C, Lopez-Rivera E, Garcia-Rama C, Saura M, Martinez-Ruiz A, Lizarbe TR, Martin-de-Lara F, Lamas S. 2006. Cbfa-1 mediates nitric oxide regulation of MMP-13 in osteoblasts. *Journal of cell science* **119**: 1896-1902.

Zelzer E, Glotzer DJ, Hartmann C, Thomas D, Fukai N, Soker S, Olsen BR. 2001. Tissue specific regulation of VEGF expression during bone development requires Cbfa1/Runx2. *Mech Dev* **106**: 97-106.

Zelzer E, Mamluk R, Ferrara N, Johnson RS, Schipani E, Olsen BR. 2004. VEGFA is necessary for chondrocyte survival during bone development. *Development* **131**: 2161-2171.

Zeng L, Kempf H, Murtaugh LC, Sato ME, Lassar AB. 2002. Shh establishes an Nkx3.2/Sox9 autoregulatory loop that is maintained by BMP signals to induce somitic chondrogenesis. *Genes & development* **16**: 1990-2005.

Zhang F, Luo K, Rong Z, Wang Z, Luo F, Zhang Z, Sun D, Dong S, Xu J, Dai F. 2017. Periostin Upregulates Wnt/β-Catenin Signaling to Promote the Osteogenesis of CTLA4-Modified Human Bone Marrow-Mesenchymal Stem Cells. *Scientific Reports* **7**: 41634.

Zhang G, Cohn MJ. 2006a. Hagfish and lancelet fibrillar collagens reveal that type II collagen-based cartilage evolved in stem vertebrates. *Proc Natl Acad Sci U S A* **103**: 16829-16833.

Zhang G, Cohn MJ. 2006b. Hagfish and lancelet fibrillar collagens reveal that type II collagen-based cartilage evolved in stem vertebrates. *Proceedings of the National Academy of Sciences* **103**: 16829-16833.

Zhang G, Eames BF, Cohn MJ. 2009. Chapter 2. Evolution of vertebrate cartilage development. *Curr Top Dev Biol* **86**: 15-42.

Zhang G, Miyamoto MM, Cohn MJ. 2006a. Lamprey type II collagen and Sox9 reveal an ancient origin of the vertebrate collagenous skeleton. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 3180-3185.

Zhang G, Miyamoto MM, Cohn MJ. 2006b. Lamprey type II collagen and Sox9 reveal an ancient origin of the vertebrate collagenous skeleton. *Proc Natl Acad Sci U S A* **103**: 3180-3185.

Zhang P, Jimenez SA, Stokes DG. 2003. Regulation of human COL9A1 gene expression. Activation of the proximal promoter region by SOX9. *J Biol Chem* **278**: 117-123.

Zhang Y, Hassan MQ, Xie R-L, Hawse JR, Spelsberg TC, Montecino M, Stein JL, Lian JB, Van Wijnen AJ, Stein GS. 2009. Co-stimulation of the bone-related Runx2 P1 promoter in mesenchymal cells by SP1 and ETS transcription factors at polymorphic purine-rich DNA sequences (Y-repeats). *Journal of Biological Chemistry* **284**: 3125-3135.

Zhao Q, Eberspaecher H, Lefebvre V, De Crombrugghe B. 1997. Parallel expression of Sox9 and Col2a1 in cells undergoing chondrogenesis. *Dev Dyn* **209**: 377-386.

Zheng Q, Zhou G, Morello R, Chen Y, Garcia-Rojas X, Lee B. 2003. Type X collagen gene regulation by Runx2 contributes directly to its hypertrophic chondrocyte-specific expression in vivo. *J Cell Biol* **162**: 833-842.

Zhou G, Zheng Q, Engin F, Munivez E, Chen Y, Sebald E, Krakow D, Lee B. 2006. Dominance of SOX9 function over RUNX2 during skeletogenesis. *Proc Natl Acad Sci U S A* **103**: 19004-19009.

Zhou X, von der Mark K, Henry S, Norton W, Adams H, de Crombrugghe B. 2014. Chondrocytes transdifferentiate into osteoblasts in endochondral bone during development, postnatal growth and fracture healing in mice. *PLoS Genet* **10**: e1004820.

Zhu C, Zheng XF, Yang YH, Li B, Wang YR, Jiang SD, Jiang LS. 2016. LGR4 acts as a key receptor for R-spondin 2 to promote osteogenesis through Wnt signaling pathway. *Cell Signal* **28**: 989-1000.

Zhu Y, Wu Y, Cheng J, Wang Q, Li Z, Wang Y, Wang D, Wang H, Zhang W, Ye J et al. 2018. Pharmacological activation of TAZ enhances osteogenic differentiation and bone formation of adipose-derived stem cells. *Stem Cell Research & Therapy* **9**: 53.