ARYLSULFATASE-I: A NOVEL SULFATASE REGULATING PROTEOGLYCAN SULFATION DURING ENDOCHONDRAL OSSIFICATION

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

Rafaela Grecco Machado

PERMISSION TO USE

In presenting this thesis 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 in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis 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.

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 Road, University of Saskatchewan

Saskatoon, Saskatchewan

S7N 5E5 Canada

OR

Dean of the College of Graduate and Postdoctoral Studies

University of Saskatchewan

116 Thorvaldson Building, 110 Science Place

Saskatoon, Saskatchewan

S7N 5C9 Canada

ABSTRACT

Endochondral ossification consists of bone formation from a previous cartilaginous template. This process starts during embryonic development when mesenchymal cells condense in presumptive skeletal regions and differentiate into chondrocytes, which undergo a progressive maturation process. Chondrocytes are arranged in morphologically distinct zones within the cartilaginous template, reflecting their maturation states. While mature chondrocytes are in the medial region, immature chondrocytes are in the distal areas of the template. The chondrocytes pronounced cellular changes during maturation are accompanied by prominent extracellular matrix (ECM) remodelling and reorganization. Sulfated proteoglycans (PGs) are one of the main components of cartilage ECM, and their sulfation levels are critical for proper endochondral ossification. PG sulfation is regulated by sulfotransferases and sulfatases - that add and remove, respectively, sulfate esters from PGs. Deficiencies in sulfatases lead to diseases, including mucopolysaccharidosis and multiple sulfatase deficiency. Disease phenotypes include various skeletal defects seemingly related to chondrocyte maturation, including impaired chondrocyte proliferation and hypertrophy, severe growth retardation, and delayed primary and secondary ossification center formation. The mechanisms of how sulfation leads to those phenotypes and how sulfur levels are modulated across endochondral ossification are still to be unveiled. Here the hypothesis that a novel sulfatase - Arylsulfatase I (ARSI) - decreases PG sulfation during cartilage maturation thus promoting endochondral ossification, was tested. Synchrotron X-ray fluorescence (XRF) imaging confirmed that sulfate esters decreased significantly in mature cartilage during endochondral ossification. ARSB and GALNS- the two known chondroitin sulfate PG sulfatases- were not specifically expressed in mature cartilage. Utilizing laser-capture microdissection and RNAseq data previously obtained in the lab, we were able to identify a novel sulfatase - Arsi - to be upregulated in mature cartilage. Increased ARSI protein expression in mature cartilage was further demonstrated *in vivo* in chick and mouse sections and during the maturation of ATDC5 chondrocytes *in vitro*. Biochemical assays overexpressing ARSI in HeLa cells confirmed that ARSI is a novel

PG sulfatase. Taken together, biochemical and expression results strongly indicate that ARSI controls the decrease of sulfate esters in mature cartilage. To test ARSI function during endochondral ossification, four putative transgenic zebrafish (Tg) lines for this gene were generated by Gateway-Tol2 transgenesis. All generated Gateway constructs had the correct sequence and were integrated into the zebrafish genome, but possibly due to posttranscriptional mechanisms, protein overexpression was not confirmed in putative Tg fish. Our hypothesis was partially addressed, and more studies are needed to investigate the potential ARSI roles in endochondral ossification. In this thesis, we showed for the first time *ARSI* mRNA expression in chick, ARSI protein expression in chick and mouse tissues, and we presented the first biochemical characterization of ARSI as a PG sulfatase. Our results supplement the scarce literature about ARSI and potentiate future work uncovering its role in skeletal development and disease.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Brian Eames, for allowing me to pursue my Ph.D. studies in his lab. Thank you for trusting that I could learn and work in an area very different from my previous background, for allowing me to try and learn from my mistakes, and for all the guidance over the years. Furthermore, I would like to thank my committee members, Dr. Valerie Verge, Dr. Julia Boughner, Dr. Troy Harkness, and Dr. Ingrid Pickering, for their valuable feedback, which pushed me to improve this project. Special thanks to Dr. Mark Hackett, who taught me much about XRF through e-mails, zoom calls, and in-person meetings, and to the Canadian Light Source (CLS), acknowledged ij chapters 3 and 4.

I would also like to thank my lab mates for keeping up with me listening to music at the lab and constantly complaining about life: Oghenevwogaga Atake, Hamed Alizadeh, Elham Koosha, Ayda Hassanzadeh, Neethi Satish, Michelle Whalen, Monica Salud, J-S Gauthier, and Albert Mark. Special thanks to Devin Brown, who introduced me to the long shifts at the synchrotron and whose results were fundamental for the development of this thesis; to Jason Nguyen, that made me laugh so much with his bad jokes and lifted my spirits more times than he can imagine; and to my great *amiga* Patsy Gomez Picos who taught me a whole lot about science and friendship and was my family away from home. Thanks also to the summer students I co-supervised. Your interest in learning reminds me of why I'm doing science.

Thanks to the Government of Saskatchewan, the Eunice Bilokreli family, the APP department, and the Canadian Federation of University Women, for the financial support.

Thanks to my parents for prioritizing my education over their own needs and for letting me follow my dreams (even though they still don't understand what I do).

Lastly, I would like to thank my life partner, José Alvim, for being my greatest supporter and friend, and my daughter Nina, for always welcoming me with the sweetest smile that gives me a reason to be better every day.

DEDICATION

To the pioneering women in science, whose bravery allowed me to stand here today.

To current women in science, who are fighting invisible battles, yet still remain supportive of each other.

To future women in science, may your challenges be more about the science itself than about surviving in it. May you break stereotypes and pave the way for future women in STEM.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 2: BACKGROUND

CHAPTER 3: ARYLSULFATASE I, EXPRESSED DURING CARTILAGE MATURATION, REDUCES CHONDROITIN SULFATE LEVELS

LIST OF FIGURES

CHAPTER 2: BACKGROUND

CHAPTER 3: ARYLSULFATASE I, EXPRESSED DURING CARTILAGE MATURATION, REDUCES CHONDROITIN SULFATE LEVELS

CHAPTER 4: GENERATING TRANSGENIC ZEBRAFISH TO ANALYZE ARSI BIOLOGICAL FUNCTION

CHAPTER 5: SUMMARY, LIMITATIONS, AND FUTURE DIRECTIONS

LIST OF ABBREVIATIONS

°C	Celsius
2D	Two dimensional
3'UTR	Three prime untranslated region
3D	Three dimensional
4-MUS	4-methylumbelliferyl sulphate
ACAN	Aggrecan
AMP	Ampicillin
ARPE-19	Human spontaneously arising retinal pigment epithelia cell line
ARSA	Arylsulfatase-A
ARSB	Arylsulfatase-B
ARSC	Arylsulfatase-C
ARSD	Arylsulfatase-D
ARSE	Arylsulfatase-E
ARSF	Arylsulfatase-F
ARSG	Arylsulfatase-G
ARSH	Arylsulfatase-H
ARSI	Arylsulfatase-I
arsia	*arylsulfatase family, member Ia*
arsib	*arylsulfatase family, member Ib*
ARSJ	Arylsulfatase-J

ARSK	Arylsulfatase-K
AUP	Animal use protocol
B3GAT3	Beta-1,3-Glucuronyltransferase 3
bp	Base pair
C	Carbon
C4st1	Chondroitin-4-sulfotransferase 1
Ca	Calcium
cDNA	Complementary DNA
CDS	Coding sequence
ch	Ceratohyal
CHST11	Carbohydrate Sulfotransferase 11
CHST12	Carbohydrate Sulfotransferase 12
CHST3	Carbohydrate Sulfotransferase 3
Cl	Chlorine
CLS	Canadian Light Source
cm	Centimeters
cmd	Cartilage-matrix-deficient
cmlc2	*Myosin light chain 2*
COL1	Collagen type I protein
COL11	Collagen type XI protein
COL2	Collagen type II protein
Col2a1	Collagen Type II Alpha 1 Chain

col2a1a	*Collagen type II alpha 1a*
COL9	Collagen type IX protein
CPM	Counts per million
CRE	Cis-regulatory element
CS	Chondroitin sulfate
CSPGs	Chondroitin sulfate PGs
Cu	Copper
DAPI	4′,6-diamidino-2-phenylindole
DEPC	Diethyl pyrocarbonate
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dpf	Days post fertilization
DS	Dermatan sulfate
DSBs	Double-strand breaks
E.coli	Escherichia coli
E12.5	Embryonic day 12.5
E14.5	Embryonic day 14.5
E16.5	Embryonic day 16.5
ECL	Enhanced chemiluminescence
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid

EFTUD2	Elongation Factor Tu GTP Binding Domain Containing 2
EGFP	Enhanced green fluorescent protein
ER	Endoplasmic reticulum
eV	Electron volts
F0	Injected zebrafish
F1	First generation
FBS	Fetal bovine serum
FC	Fold change
Fe	Iron
FGF	Fibroblast Growth Factor
FGFR2	Fibroblast Growth Factor Receptor 2
FGly	Formylglycine
FPA	Focal plane array
FTIR	Fourier transform infrared
GAGs	Glycosaminoglycans
GalNAc	N-acetylgalactosamine
GALNS	N-acetylgalactosamine 6-sulfatase
GFP	Green fluorescent protein
GlcA	Glucuronic acid
GlcNAc	D-glucosamine N-acetylglucosamine
GlcNS	Heparan-N-sulfatase
gRNAs	Guide RNAs

GS	Goat serum
GWAS	Genome-wide association studies
h	Hours
hARSI	Human Arylsulfatase-I
HCl	Hydrochloric acid
HH32	Hamburger–Hamilton stage 32
HH36	Hamburger–Hamilton stage 36
hm	Hyomandibula
hpf	Hours post fertilization
HS	Heparan sulfate
HSPGs	Heparan sulfate PGs
HT-1080	Human fibrosarcoma cell line
IdoA	Iduronic acid
IDS	Iduronate-2-Sulfatase
IHC	Immunohistochemistry
Ihh	Indian hedgehog
IMM	Immature cartilage
IR	Infrared
ISH	*In situ* hybridization
kDa	Kilodalton
keV	Kiloelectron volt
KS	Keratan sulfate
LASU	Lab Animal Services Unit

LB	Lysogeny broth
LCM	Laser Capture Microdissection
MABT	Maleic acid buffer containing Tween 20
MAT	Mature cartilage
MCT	Mercury-cadmium-telluride
mg	Milligram
min	Minutes
ml	Milliliters
MLD	Metachromatic leukodystrophy
mm	Millimeters
mM	Millimolar
MMP13	Matrix metalloproteinase 13
MMP9	Matrix metalloproteinase 9
MOs	Morpholino oligomers
MPS	Mucopolysaccharidosis
mRNA	Messenger RNA
MSD	Multiple sulfatase deficiency
Mut	Mutant
N-CAM	Neural cell adhesion molecule
NaOH	Sodium hydroxide
NCBI	National Center for Biotechnology Information
nm	Nanomelic

NT2/D1	Human embryonal carcinoma cell line NTERA-2 cl.D1
O	oxygen
OA	Osteoarthritis
OCT	Optimal cutting temperature
OI	Osteogenesis imperfecta
OMIM	Online Mendelian Inheritance in Man
ON	Overnight
OPT	Optical projection tomography
P	Phosphorus
P2A	2A self-cleaving peptide
Papss2	3'-Phosphoadenosine 5'-Phosphosulfate Synthase 2
PBS	Phosphate-buffered saline
PBST	1X PBS with 0.5% Triton X-100
PC	Perichondrium
PCR	Polymerase chain reaction
PDGFRA	Platelet Derived Growth Factor Receptor Alpha
PFA	Paraformaldehyde
PGs	Proteoglycans
pME	Middle entry clone
pNTS	p-nitrophenol sulphate
PSMD12	Proteasome 26S Subunit, Non-ATPase 12
PTHrP	Parathyroid hormone-related protein

RNA	Ribonucleic acid
RNAseq	RNA sequencing
rpm	Revolutions per minute
RT	Room temperature
RT-PCR	Reverse transcription polymerase chain reaction
SDS-PAGE	Sodium dodecyl-sulfate polyacrylamide gel electrophoresis
SEM	Standard error of the mean
Slc26a2	Solute Carrier Family 26 Member 2
SMAK	Sam's microprobe analysis toolkit
SMCHD1	Structural Maintenance of Chromosomes Flexible Hinge Domain Containing 1
SPP1	Secreted Phosphoprotein 1
SS	Sheep serum
SULF1	Sulfatase 1
SULF2	Sulfatase 2
SUMF1	Sulfatase modifying factor 1
SXRMB	Soft X-ray Microcharacterization Beamline
TAE	Tris-acetate-EDTA
TALENs	Transcription activator-like effector nucleases
Tg	Transgenic
TGFβ	Transforming Growth Factor Beta
uxs1	UDP-Glucuronate Decarboxylase 1
UACC AREB	University Animal Care Committee Animal Research Ethics Board

VESPERS	Very sensitive Elemental and Structural Probe Employing Radiation from a Synchrotron
WES	Whole exome sequencing
WGS	Whole genome sequencing
WT	Wild type
XANES	X-ray absorption near-edge spectroscopy
XRF	X-Ray Fluorescence
Xyt1	Xylosyltransferase 1
ZFNs	Zinc finger nucleases
Zn	Zinc
µm	Micrometers

CHAPTER 1: THESIS OVERVIEW

This chapter provides a general overview of my thesis. Here I describe how the chapters are organized, the main objectives and hypothesis, and what is already published and what is not. I also highlight contributions made by collaborators in each chapter. This thesis is comprised of five chapters, and it is written in a manuscript-based format. Chapter 2 is a background chapter that explains important concepts needed to understand this thesis. Chapters 3 and 4 contain experiments investigating our hypothesis that Arylsulfatase I (ARSI) decreases proteoglycan sulfation during cartilage maturation thus promoting endochondral ossification. Chapter 5 is the final chapter and brings an overall discussion, limitations, and future directions for this project.

Chapter 2 is comprised of two portions. The first portion includes a literature review about endochondral ossification, cartilage extracellular matrix (ECM), sulfatases, and ARSI. More specifically, I emphasize how ECM and sulfatases might influence endochondral ossification and what is known about the relatively understudied ARSI. The second portion of this chapter consists of an invited review paper I co-authored with Dr. Eames entitled “Using Zebrafish to Test the Genetic Basis of Human Craniofacial Diseases”. This review paper was published in the *Journal of Dental Research* in the special issue Critical Reviews in Oral Biology & Medicine. I wrote all the sections with input from Dr. Eames and made figures 2 and 4 of the paper. Dr. Eames prepared figures 1 and 3. This review was added in this introductory chapter because it describes how the zebrafish model can be used to study gene function and disease. This paper highlights anatomic, embryonic, and genetic features of craniofacial development that are conserved among mammals and zebrafish and allow comparative developmental studies to be performed. A summary and a historical perspective of the genetic manipulation methods used in zebrafish to interrogate gene function, focusing on the

Gateway Tol2 system, are also presented. The topics discussed in the review serve as the theoretical base for Chapter 4, where genetic manipulation techniques are used to study the function of ARSI during endochondral ossification.

Chapter 3 is comprised of a manuscript in preparation named “Arylsulfatase I, expressed during cartilage maturation, reduces chondroitin sulfate levels”. In this manuscript, we show that ARSI is a novel cartilage PG sulfatase that is expressed in mature chondrocytes during endochondral ossification. The manuscript contains data from myself and collaborators (Dr. Amir Ashique, Dr. Katie Ovens, Mr. Devin Brown, Dr. Yuko Naito-Matsui, Dr. Hiroshi Kitagawa, and Dr. Brian Eames). Dr. Amir Ashique and Dr. Katie Ovens generated and analyzed the laser capture and RNA sequencing results; Mr. Devin Brown analyzed sulfation levels of chick humerus at the synchrotron; and Dr. Yuko Naito-Matsui and Dr. Hiroshi Kitagawa showed that ARSI is a PG sulfatase using Hela cells. Together with Dr. Brian Eames, I analyzed the gene list generated by RNA sequencing results and looked for potential sulfatases involved in cartilage maturation. I conducted histological and *in situ* hybridization analyses to better understand the developing chick growth plate and re-quantified sulfur levels in synchrotron images of chick humeri. I made sequencing comparison analyses of ARSI in different species, showed that ARSI is specifically expressed *in vivo* in mouse and chick mature cartilage, and showed that ARSI expression increases with time *in vitro* in ATDC5 cells. I wrote most of the manuscript and reworked most of the figures, graphs, and statistical analyses based on data provided by our collaborators.

Chapter 4 focuses on the generation of transgenic (Tg) zebrafish lines overexpressing human ARS,I tagged with GFP or not, to investigate the biological function of ARSI during endochondral ossification. The absence of available Tg or mutant animal models for ARSI underscores the significance of this study. The chapter details the steps and outcomes of producing four putative Tg zebrafish using the Gateway-Tol2 system. All construct sequences were confirmed by enzymatic digestions and sequencing results. Nonetheless, expected green fluorescent protein (GFP) expression was not detected in the cartilages of the putative Tg zebrafish. Subsequently, human ARSI (hARSI) integration into the fish genome and the RNA

transcription in the putative Tg fish are demonstrated. Overexpression of ARSI protein was not proven, probably due to post-translational mechanisms. The final part of this chapter outlines experiments conducted to assess skeletal development and sulfur levels in the putative Tg fish, carried out in parallel with zebrafish genotyping. Putative Tg fish had increased potassium and sulfur levels in their ceratohyal and an increased perichondral bone in their hyomandibula. As it could not be demonstrated that the putative Tg fish overexpressed ARSI, the hypothesis and objectives for this chapter were partially addressed. More studies are needed to assess ARSI's biological function. Most of the experiments in this chapter were performed by me, including designing primers to check for the genomic integration of the constructs. Two summer students I co-supervised - Asawari Albal and Maya Berscheid - performed PCRs to amplify the *col2a1a/hARSI* and *hARSI/EGFP* regions. I designed, ordered, and prepared the reagents for the *SUMF1* experiments. Maya Berscheid performed the *SUMF1* injections and skeletal scoring of *SUMF1*-injected fish.

Chapter 5 concludes the thesis with a general summary of the previous chapters, limitations of experiments, and future directions for this research.

CHAPTER 2: BACKGROUND

This chapter includes a literature review written by me (pages 4 -21) and a review paper entitled: *"Using zebrafish to test the genetic basis of human craniofacial diseases"* which I wrote with Dr. Brian Eames (pages 22-38). The review paper was published in Critical Reviews in Oral Biology & Medicine in the *Journal of Dental Research* in 2017. I wrote all the sections of the review paper with input from Dr. Eames and made Figures 2 and 4. Figures 1 and 3 were prepared by Dr. Eames

2.2 Literature Review

2.2.1 Endochondral ossification

In vertebrates, osteogenesis, or bone formation, is governed by two main processes: intramembranous and endochondral ossification. While both processes start with precursor mesenchymal cells, how those cells give rise to bone differs (Eames et al., 2003). During intramembranous ossification, mesenchymal cells are directly converted into bone and form intramembranous (dermal) bones, such as the flat bones of the skull, clavicles, and fin rays. On the other hand, during endochondral ossification, mesenchymal cells first differentiate into chondrocytes that generate a cartilaginous template for bone deposition (Long & Ornitz, 2013). Endochondral ossification starts during embryonic development and is responsible for the formation of endochondral (chondral) bones of most of the axial (ribs and vertebrae) and appendicular skeleton (e.g. upper and lower limbs) (Gilbert, 2000). Additionally, endochondral ossification is responsible for postnatal longitudinal chondral bone growth, and it can also be witnessed during pathological events like advanced cases of osteoarthritis and bone fracture healing (Aghajanian & Mohan, 2018; Kronenberg, 2003; Staines et al., 2013)

The multistep endochondral ossification process starts when undifferentiated mesenchymal cells migrate to presumptive skeletal regions in the embryos. Mesenchymal cells then change their adhesiveness to the ECM and aggregate with each other forming clusters. At the condensation stage, there is a transitory up-regulation of the PGs versican and syndecan and molecules like N-cadherin, hyaluronan, and Neural Cell Adhesion Molecule (N-CAM), which are important to recruit cells, and maintain and stabilize the mesenchymal condensation (Figure 2.1A) (Eames et al., 2003; Hall & Miyake, 2000). Following mesenchymal condensation, mesenchymal cells differentiate into chondrocytes, which secrete an ECM that is rich in collagens (mainly Collagen Type II (COL2), but also Collagen type IX (COL9), and Collagen type XI (COL11)) and proteoglycans (mainly Aggrecan, ACAN) (Behonick & Werb, 2003; Mackie et al., 2011). Cells at the periphery of the condensation form an epithelium layer called perichondrium, which expresses collagen type I (COL1) and separates the cartilage template from the surrounding mesenchyme (Long & Ornitz, 2013) (Figure 2.1B). At the same time that chondrogenesis is happening, the cartilage primordia starts to take the shape of the skeletal element for which it will serve as a template.

After chondrogenesis, chondrocytes start a progressive maturation process inside the cartilage template, which proceeds outward from the center (diaphysis) to the ends (epiphyses) (Behonick & Werb, 2003) (Figure 2.1C). Inside the cartilage primordia, chondrocytes are organized in a differentiation continuum where immature cartilage (composed of resting and proliferative chondrocytes) is at the epiphysis, and mature cartilage (consisting of pre-hypertrophic and hypertrophic chondrocytes) is at the diaphysis. Depending on which stage of differentiation the chondrocytes are at, they express different genes and secrete distinctive molecules, including ECM molecules like collagens and PGs. For example, resting and proliferative chondrocytes express *Col2* and *Acan*, while hypertrophic chondrocytes exhibit high levels of *Col10* but lower levels of *Col2* and *Acan*, indicating their distinct differentiation state. (Behonick & Werb, 2003; Gomez-Picos & Eames, 2015; Karsenty & Wagner, 2002) (Figure 2.1C).

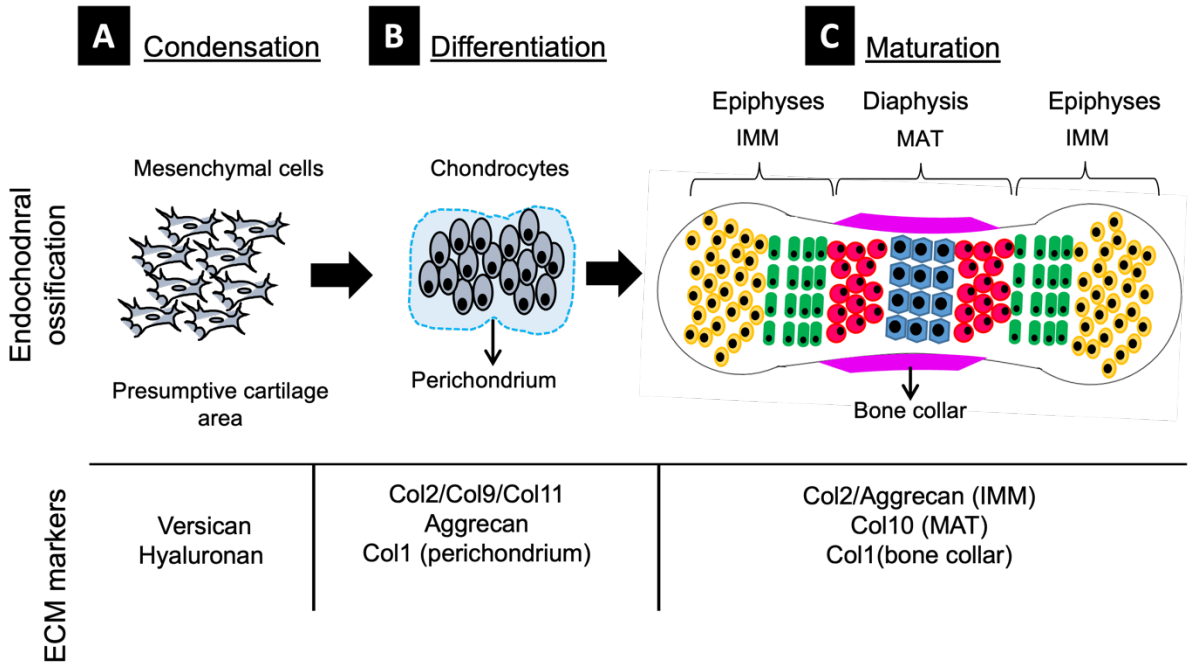


Figure 2.1 Overview of endochondral ossification. During embryonic development, mesenchymal cells migrate to presumptive cartilage areas. Changes in gene expression make those cells aggregate and form condensations where versican and hyaluronan are upregulated. (B) Mesenchymal cells differentiate into chondrocytes, which secrete an ECM rich in collagens and PGs and form a cartilaginous template. Mesenchymal cells from the border of the template form a perichondral layer which separates the cartilaginous template from the surrounding mesenchyme. (C) chondrocytes start a progressive maturation process from the epiphysis to the diaphysis. Resting (yellow) and proliferative (green) chondrocytes form the immature cartilage (IMM) region and prehypertrophic (red) and hypertrophic (blue) chondrocytes constitute the mature cartilage (MAT) region. The perichondral cells surrounding the MAT receive signals from hypertrophic chondrocytes stimulating their differentiation into osteoblasts, which deposit perichondral bone (bone collar) in the diaphysis.

The following steps in the progression of endochondral ossification involve blood vessel invasion, cartilage degradation, and bone deposition, which initiate in the mature cartilage region area of the cartilaginous template influenced by hypertrophic chondrocytes. Hypertrophic chondrocytes increase their volume dramatically while secreting a specific ECM rich in COL10 that ultimately becomes mineralized (Gilbert, 2000). They also have an important function in ECM remodelling by secreting enzymes such as matrix metalloproteinase 13 (MMP13), which cleave ECM in the mature cartilage and facilitate vascular invasion (Inada et al., 2004; Stickens et al., 2004). Most

hypertrophic chondrocytes go through cell death. At the same time, blood vessels penetrate this apoptotic cartilage bringing in red blood cells, chondroclasts, osteoclasts, and osteoblasts (Kronenberg, 2003; Long & Ornitz, 2013; Mackie et al., 2008; Mackie et al., 2011). Chondroclasts and osteoclasts degrade cartilage ECM and remove leftover chondrocytes in the site of vascular invasion. At the same time, osteoblasts start depositing trabecular bone in this same region which is the primary ossification centre (Ortega et al., 2004). More recent studies show that hypertrophic chondrocytes can also transdifferentiate into osteoblasts and participate in the formation of trabecular bone during development and disease (Aghajanian & Mohan, 2018). Similarly, secondary ossification centers are formed postnatally in the epiphyses of the chondral bones, where the trabecular bone is also deposited after the vascular invasion (Mackie et al., 2008).

Perichondral bone is an important contributor to the formation of the mature chondral bone. The literature often describes perichondral ossification as an intramembranous process since mesenchymal cells of the perichondrium layer differentiate directly into osteoblasts (Egawa et al., 2014). However, there is a need for crosstalk between cartilage and the mesenchymal cells for perichondral ossification to happen, which is not true for other intramembranous bones (Eames et al., 2003). Perichondral ossification starts in the bone diaphysis, where mature chondrocytes stimulate overlying mesenchymal cells to differentiate into osteoblasts by secreting molecules like *Ihh* (Egawa et al., 2014; Long et al., 2004; St-Jacques et al., 1999). Osteoblasts then deposit bone matrix, forming a bone collar surrounding the mature cartilage region (Figure 2.1C). Cartilage maturation, endochondral, and perichondral ossification proceed in a continuous and orchestrated way outwards from the diaphysis. Chondrocyte proliferation and consequent bone elongation cease after puberty leading to the fusion of the primary and secondary ossification centers and the formation of a complete mature chondral bone with just a thin cartilage layer left at its articular surfaces (Mackie et al., 2011; White & Wallis, 2001). Altogether, endochondral ossification is a complex process involving different cell types, transcription factors, growth factors, and distinct ECM microenvironments.

2.2.2 Cartilage extracellular matrix

The cartilage ECM is mainly composed of collagens and PGs, and it starts to be deposited while cells differentiate into chondrocytes (Heinegard & Oldberg, 1989). The long-established view about cartilage ECM has been that PG and collagens work together to maintain cartilage structure and stability. While highly negatively charged PGs attract water to the tissue, the collagen fibres restrain the cartilage swelling, avoiding significant changes in its conformation (White & Wallis, 2001). More recent studies, however, uncovered that the ECM performs much more than a structural function. ECM is involved in cell migration, proliferation, differentiation, and death. Specifically, there is growing evidence that ECM has essential participation in the progression of events leading to endochondral bone formation, such as chondrocyte hypertrophy and matrix calcification (Behonick & Werb, 2003; Deutsch et al., 1995; Ortega et al., 2004).

Essentially PGs are composed of a core protein and attached sulfated glycosaminoglycans (GAGs) (Figure 2.2 A); however, they present tremendous structural variation. Depending on protein core variations and on which GAGs are attached to them, PGs are classified into different types (Schwartz & Domowicz, 2002). Examples of GAG classes are chondroitin sulfate (CS), composed of repeating disaccharide units of N-acetylgalactosamine (GalNAc) and glucuronic acid (GlcA) (Figure 2.2 B) and heparan sulfate (HS) composed of either Iduronic acid (IdoA) or GlcA and a D-glucosamine N-acetylglucosamine (GlcNAc) (Rabenstein, 2002; Wang et al., 2016). CS and HS chains are linked to specific serine residues in the protein core of the PGs via a common tetrasaccharide link region composed of xylose, galactose, galactose, and glucuronic acid sugars (Prydz & Dalen, 2000). PGs may exhibit variations in the number and length of the attached GAG chains. Also, different sulfation positions increase the diversity of PG composition and function and will be further discussed later in this chapter (Habuchi et al., 2004). In cartilage ECM, the most abundant class of GAG is CS, which, attached to ACAN, forms chondroitin sulfate PGs (CSPGs) (Roughley & Lee, 1994).

PG synthesis is a complex multi-step process involving enzymes like glycosyltransferases, epimerases, and sulfotransferases. Briefly, it starts when PG core proteins are synthesized and translocated to the cell endoplasmic reticulum (ER), where the first monosaccharide of the tetrasaccharide link (xylose) is added to a specific serine residue in the core protein by a xylosyltransferase (Briggs & Hohenester, 2018; Brown & Eames, 2016). After that, the nascent PG is transferred to the *cis* Golgi network, where the other 3 sugars (galactose-galactose-glucuronic acid) are added by two galactosyltransferases and a glucuronyltransferase, respectively. (Kearns et al., 1993; Vertel et al., 1993). The addition of the repeating disaccharides by specific glycosyltransferases and their subsequent sulfation by sulfotransferases also occurs in the Golgi apparatus before PGs are targeted to the extracellular space, cell membrane or other organelles (Dick et al., 2012; Sugumaran et al., 1992).

Mutations in PGs core proteins, GAGs elongation, PG secretion, or enzymes controlling post-translational modifications like sulfation can cause various diseases (Schwartz & Domowicz, 2002). Likely due to the abundance of PGs in skeletal tissues, mutations in PG synthesis genes are known to cause at least 20 different human skeletal diseases, including osteoarthritis, spondyloepimetaphyseal dysplasia, and diastrophic dysplasia (Brown & Eames, 2016; Schwartz & Domowicz, 2002). Animal models have also been used to understand further the extension of PG function during skeletal development, such as the cartilage-matrix-deficient (*cmd*), chondroitin-4-sulfotransferase 1 (*C4st1)* deficient, and Xylosyltransferase 1 (*Xylt1)* deficient mice; nanomelic (*nm*) chicken; and f*am20b* and *xylt1* zebrafish mutants. Interestingly, reduced levels of cartilage PGs in these animals lead to accelerated hypertrophy of chondrocytes, and early bone formation. Their bones were frequently decreased in length, and their growth plates presented diverse abnormalities, highlighting the importance of the PGs for a proper endochondral ossification (Domowicz et al., 2009; Eames et al., 2011; Kluppel et al., 2005; Mis et al., 2014; Orkin et al., 1977; Schwartz & Domowicz, 2002; Watanabe et al., 1994). The underlying mechanisms responsible for most observed phenotypes in animal models and human diseases have not been elucidated.

Once in the extracellular space, spatial immobilization of growth factors may be one of the most critical functions of PGs. PGs can serve as a sink, binding growth factors and making them unavailable to their receptors while protecting them from degradation; they can also bind growth factors enabling their diffusion toward target cells, or they can still act as coreceptors facilitating the binding of growth factors to their receptors. (Cortes et al., 2009; Lu et al., 2011; Schwartz & Domowicz, 2002). In this sense, the impact of PGs on growth factor signaling can vary, either positively or negatively, based on the growth factor, receptor, PG, and tissue being examined. The ability of different PGs to bind growth factors allows a spatial- and time-controlled presentation of these molecules to target cells, controlling the activation of signalling pathways and consequent changes in gene expression (da Costa et al., 2017; Esko & Selleck, 2002; Mizumoto et al., 2015).

2.2.3 Proteoglycan sulfation and endochondral ossification

PG sulfation may affect cartilage biomechanics and its biological function. Because most PGs are highly sulfated molecules, their negative charges attract water to the cartilage, increasing its compressive-resistant strength (Chahine et al., 2005). At the same time, the positions and levels of sulfation influence the binding of growth factors and PGs, leading to potential changes in gene expression (Cortes et al., 2009; da Costa et al., 2017; Otsuki et al., 2010; Rosen & Lemjabbar-Alaoui, 2010). Considering the high variability in PGs structure and the possibility of different sulfation positions, not all interactions between PGs and growth factors are well described. Examples of how sulfation levels affect the binding of GFs and PGs and gene expression come from animal models with perturbed PG sulfation (Cortes et al., 2009; da Costa et al., 2017; Otsuki et al., 2010; Rosen & Lemjabbar-Alaoui, 2010). For example, mice with undersulfated CSPGs have been found to have reduced binding of Indian hedgehog (Ihh) to non-sulfated CS, resulting in abnormal Ihh protein distribution in the extracellular matrix of their growth plates and diminished Ihh signaling. These mice also exhibited reduced expression of downstream targets of Ihh (*Ptch1* and *Gli1*)

and *Fgfr3*. These findings suggest that sulfation reduction in CS may lead to changes in gene expression by altering Ihh binding and diffusion (Cortes et al., 2009). *Ihh* plays an important role in chondrocyte proliferation, and its reduced diffusion would lead to reduced chondrocyte proliferation and changes in chondrocyte maturation (Ohba, 2016).

Two classes of intracellular enzymes control PG sulfation: sulfotransferases and sulfatases, which catalyze the incorporation or removal of sulfate esters from GAGs, respectively (Figure 2.2 C and D)(Bulow & Hobert, 2006; Diez-Roux & Ballabio, 2005; Mougous et al., 2002). Sulfate moieties are attached to growing PG chains in the form of O – linked sulfate esters through the action of sulfotransferases. CSPGs, specifically, may be sulfated on the 4- (C4S) and 6- (C6S) carbons of the GalNAc sugar or the 2 (C2S) -carbon of GlcA, but those sulfation positions vary in different GAGs (Kusche-Gullberg & Kjellen, 2003; Uyama et al., 2007) (Figure 2.2 B-D).

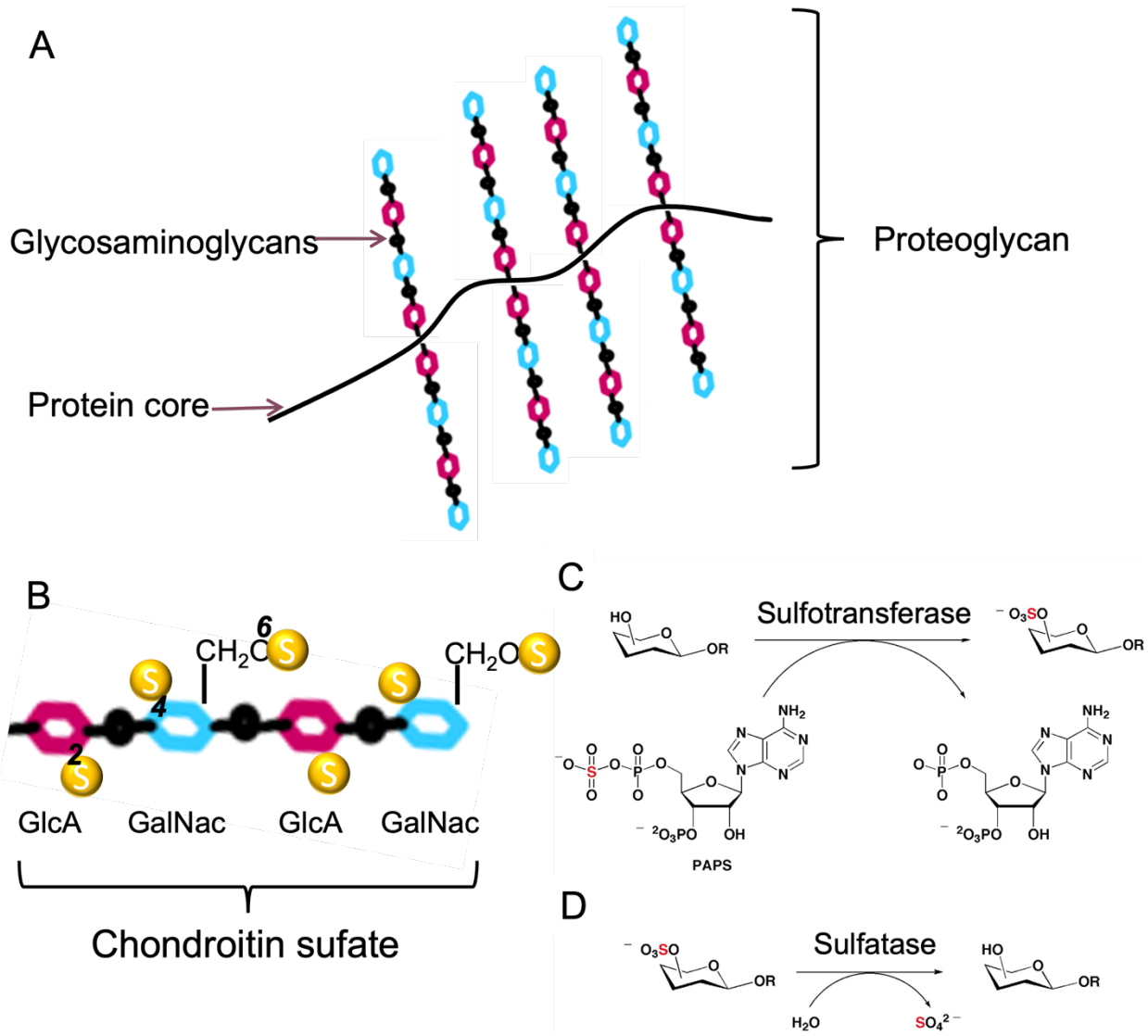


Figure 2.2 Chondroitin sulfate proteoglycans, sulfation positions, sulfotransferases, and sulfatases.Proteoglycans are formed by a core protein with attached glycosaminoglycans at specific sites. Glycosaminoglycans are repeating disaccharides (represented in pink and blue). (B) Chondroitin sulfate glycosaminoglycans are formed by repetitions of GlcA and GalNac. These GAGs can be sulfated in different positions (carbons, 2, 4, and 6. Sulfation represented by yellow circles). (C) Sulfotransferases catalyze the transfer of sulfur from an activated donor, 3′-phosphoadenosine-5′-phosphosulfate (PAPS), to molecules such as sugars. (D) Sulfatases catalyze the hydrolysis of sulfate-esters from molecules such as GAGs. (Figures C and D Adapted from Mougous et al., 2022)

The regulation of sulfation levels in cartilage PGs has been shown to be critical for proper endochondral bone formation. Mice with mutations in genes such as *Papss2* (catalyze the synthesis of 3'-phosphoadenosine 5'-phosphosulfate - PAPs - sulfate donor) and *Slc26a2* (sulfate transporter) showed that inadequate sulfation of PGs leads to a variety of skeletal defects, including reduced chondrocyte proliferation at growth plates, severe skeletal growth retardation, and delayed secondary ossification center formation (Cortes et al., 2009; Gualeni et al., 2013). Even though these are indications that PG sulfation regulates chondrocyte differentiation, how sulfur levels are modulated across endochondral ossification and the mechanisms by which sulfur dysregulation causes skeletal phenotypes are still to be revealed.

Two examples that may help to elucidate the mechanism by which ECM sulfation, particularly CSPG sulfation, regulates endochondral ossification are seen in the *C4st1*-mutant mouse and the bm mouse (Kluppel et al., 2005). *C4st1* stands for chondroitin-4- sulfotransferase 1 gene, which encodes an enzyme that adds sulfate groups to the 4-O-position in chondroitin. When this gene is mutated, mice have lower C4S levels, severe chondrodysplasia characterized by a disorganized cartilage growth plate, accelerated chondrocyte maturation, and dwarfism. Analysis of signaling pathways revealed upregulation of TGFβ signaling, downregulation of BMP signaling, and unaffected Ihh signaling in the *C4st1*-mutant mouse (Kluppel et al., 2005). Similarly, the bm mouse carries a mutation in the *Papss2* gene, which catalyzes the synthesis of the universal sulfate donor PAPS in mammals (Stelzer et al., 2007). These mice have a severe undersulfation of CSPGs, significant reduction in chondrocyte proliferation, dome-shaped skull, and shortened limbs and tail (Cortes et al., 2009; Orkin et al., 1977; Schwartz & Domowicz, 2002). Using the bm mouse model, it was demonstrated that IHH binds to sulfated CSPGs, specifically ACAN, suggesting a mechanism in which CSPGs, together with Heparan sulfate PGs (HSPGs), modulate IHH signalling by controlling its distribution across the growth plate (Cortes et al., 2009).

2.2.4 Sulfatases

Sulfatases are a family of enzymes that hydrolyze sulfate esters from various substrates, such as glycolipids, steroids, and GAGs (Sardiello et al., 2005). This enzyme class is highly conserved regarding their sequence, catalytic site, and protein structure in prokaryotes and eukaryotes, suggesting that the sulfatase gene family shares a common ancestor. In prokaryotes, sulfatases mainly act as sulfur scavengers, removing sulfate groups from substrates to make them available for their hosts. Still, they also work in processes of osmoprotection and bacterial pathogenesis (Hanson et al., 2004). In eukaryotes, sulfatases participate in the turnover and degradation of sulfated molecules such as GAGs, the production of steroid hormones, and the modulation of cell signalling pathways (Obaya, 2006). The knowledge that sulfatases are implicated in diverse cellular processes has placed these enzymes as potential targets for developing therapies for diseases like cancers, osteoarthritis, and mucopolysaccharidosis (MPS) (Lai et al., 2008; Otsuki et al., 2010; Tomatsu et al., 2010).

The first mammalian sulfatases were identified by studying patients affected by MPS and metachromatic leukodystrophy (MLD). Those patients had an accumulation of GAGs or sulfolipids in their cells. It was observed that mutations in specific sulfatase genes lead to a block in catabolic pathways and consequent accumulation of those molecules in cells and tissues (Diez-Roux & Ballabio, 2005). After the completion of the human genome sequence, novel sulfatase genes were identified based on sequence similarities and incorporated into the sulfatase gene family. Today there are 17 reported sulfatase sequences in the human genome. The subcellular localization, physiological substrate, metabolic function, and disease involvement of many of these sulfatases remain to be discovered (Schlotawa et al., 2020). A summary of the 17 sulfatases is compiled in Table 2.1.

Table 2.1 Subcellular localization, substrates, and diseases related to human sulfatases

Sulfatase	Gene	Subcelular localization	Artificial substrate	Physiological substrate	Related diseases	Main disease manifestation
Arylsulfatase A	ARSA	Lysosomal	4MUS, pNCS	sulfatide	Metachromatic leukodystrophy (MLD)	Demyelinization of central and peripheral nervous system
Arylsulfatase B	ARSB	Lysosomal	4MUS	GalN4S - (CS/DS)	Maroteaux-Lamy syndrome (MPSVI)	Dysostosis multiplex Corneal clouding
Arylsulfatase C	ARSC	ER	4MUS, pNPS	steroid sulfates	X-linked Ichthyosis (XLI)	Dark scaly skin, Mild corneal opacities
Arylsulfatase D	ARSD	ER	4MUS	**unknown**	unknown	unknown
Arylsulfatase E	ARSE	Golgi	4MUS	**unknown**	Chondrodysplasia Punctata 1 (CDPX1)	Aberrant bone mineralization, Nasal hypoplasia, Short stature Distal phalangeal hypoplasia
Arylsulfatase F	ARSF	ER	4MUS	**unknown**	unknown	unknown
Arylsulfatase G	ARSG	Lysosomal	4MUS	GlcNS3S (HS)	Usher syndrome type IV	Late onset of retinitis pigmentosa and progressive sensorineural hearing loss
Arylsulfatase H	ARSH	**unknown**	**unknown**	**unknown**	**unknown**	**unknown**
Arylsulfatase I	ARSI	**unknown**	**unknown**	**unknown**	**unknown**	**unknown**
Arylsulfatase J	ARSJ	**unknown**	**unknown**	**unknown**	**unknown**	**unknown**
Arylsulfatase K	ARSK (TSULF)	Lysosomal	pNCS, pNPS	2-O-sulfate groups from D-glucuronate (HS/DS/CS)	MPSX	Short stature, coarse facial features and dysostosis multiplex
Galactosamine-6- sulfatase	GalN6S (GALNS/GAL6S)	Lysosomal	4MUS	GalNAc6S (CS), Gal6s (KS)	Morquio A (MPS IVa)	Severe skeletal abnormalities Corneal clouding
Heparan-N-sulfatase	GlcNS (SGSH)	Lysosomal	4MUS	GalNs (HS)	Sanfilippo A syndrome (MPSIIIA)	Hyperactivity, mild somatic features, Severe neurological manifestations
Glucosamine-6-sulfatase	GlcN6S (GNS)	Lysosomal	4MUS	GlcNAc6S (HS), GlcNAc6S (KS)	Sanfilippo D syndrome (MPSIIID)	
Iduronate-2-Sulfatase	IdoA2S (IDS)	Lysosomal	4MUS	IdoA2s (HS, DS)	Hunter syndrome	Dysostosis multiplex, Organomegaly, Mild to severe mental retardation
Sulfatase-1	Sulf-1	Cell surface	4MUS	GlcNS(6S) (HS)	**unknown**	**unknown**
Sulfatase-2	Sulf2	Cell surface	4MUS	GlcNS(6S) (HS)	**unknown**	**unknown**

Initially, sulfatases were classified into two subgroups- arylsulfatases and non-arylsulfatases - depending on their activity against small aromatic sulfated compounds such as 4-methylumbelliferyl sulfate (4-MUS) or p-nitrophenol sulfate (pNTS). Sulfatases that acted on those aromatic compounds were classified as arylsulfatases (ARS) (e.g., Arylsulfatase A, Arylsulfatase B, Arylsulfatase C) (Oshikawa et al., 2009). However, this nomenclature system needs to be revised since some of the newly discovered sulfatases were classified as arylsulfatases based on sequence homology even if their ability to cleave those aromatic compounds was either absent or not tested (e.g., Arylsulfatase H, Arylsulfatase I, Arylsulfatase J). Also, more recently, described sulfatases are being named "Sulf" (E.g. SULF1 and SULF2) even though they present arylsulfatase activity (Diez-Roux & Ballabio, 2005) (Table 2.1).

Many sulfatases act against small aromatic sulfated compounds *in vitro.* Still, they seem very specific for their physiological substrates *in vivo* and have no or poor redundancy (Hanson et al., 2004). PG sulfatases are thought to act on specific sugars and sulfation positions within a GAG chain. For example, IDS catalyzes the hydrolyses of the C2 sulfate ester bonds of IdoA (2S) residues in dermatan sulfate and heparan sulfate; ARSB catalyzes the hydrolyses of the C4 sulfate ester bonds from GalNAc4S present in CS and DS; and GALNS removes sulfate groups from C6 from GalNAc6S (CS) and Gal6s (KS) (Bhattacharyya et al., 2014; Demydchuk et al., 2017; Rivera-Colon et al., 2012) (Table 2.1).

In eukaryotes, sulfatases are targeted to the secretory pathway and are glycosylated before being transported to the cell surface or different subcellular compartments, including the ER, Golgi complex, and lysosomes. Different cell localization indicates sulfatases have different substrates and functions (Brown & Eames, 2016) (Table 2.1). Enzymes that are present in lysosomes (e.g., ARSB, GALNS) participate in the molecular degradation pathway and have better enzymatic activity in acidic environments, while other sulfatases localized in the ER (e.g. ARSC, ARSD, ARSF) or secreted (e.g. Sulf1 and Sulf2) work better in a neutral pH (Hanson et al., 2004). There are some conflicting reports about the subcellular localization of some sulfatases (e.g. ARSI has been suggested to be secreted in the ECM and present in the

lysosomes (Obaya, 2006; Oshikawa et al., 2009), and it is still reported to be in Golgi at the Human Protein Atlas website (Ponten et al., 2008). ARSB has been reported as a lysosomal protein but has been localized on the cell surface of rat hepatocytes (Mitsunaga-Nakatsubo et al., 2009) (Table 2.1). It is unclear why these variations happen, but sulfatase subcellular localization might vary depending on the cell type analyzed (Ratzka et al., 2010).

Sulfatases have similar sizes (500-800 amino acids), are extensively glycosylated, and have a highly similar active site which undergoes a unique post-translational modification by the sulfatase modifying factor 1 (SUMF1). In eukaryotes, a specific cysteine residue in the sulfatase catalytic site is converted into formylglycine (FGly) by SUMF1 in the endoplasmic reticulum before sulfatases are sorted into different cell compartments. This modification is essential and limiting for the sulfatase catalytic activity (Cosma et al., 2003). Mutations in *Sumf1* lead to a reduction of all sulfatase activities and a rare genetic disorder called multiple sulfatase deficiency (MSD). Patients affected by this condition present a combination of symptoms of single sulfatases deficiencies and a variety of phenotypes, including neurologic impairments, cardiac defects, and bone and cartilage abnormalities (Schlotawa et al., 2020; Settembre et al., 2008). The importance of sulfatases to human metabolism is further emphasized by at least eight human monogenic diseases caused by the deficiency of individual sulfatase activities (Hopwood & Ballabio, 2019) (Table 2.1). Most of these diseases are MPSs, where mutations in lysosomal sulfatases such as GALNS, ARSB, SGSH, G6S, and IDS block GAG degradation and lead to GAG accumulation inside the cell, loss of cellular functions, tissue damage and organ dysfunction (Diez-Roux & Ballabio, 2005; Neufeld & Muenzer, 2019).

Due to their different compartmentalization at the subcellular level and ability to act on different substrates, sulfatases are involved in many cellular processes, including endochondral ossification and cartilage homeostasis (Jiang et al., 2020; Otsuki et al., 2010; Settembre et al., 2008). Sulfatases that act on GAGs are particularly interesting for endochondral ossification since cartilage ECM contains large amounts of highly sulfated GAGs. Seven sulfatases are known to be involved in hydrolyzing sulfate esters

from GAGs. Among those, five are known to act on HS disaccharides and just two – ARSB and GALNS- on CS (Table 2.1). Sulfation levels of cartilage ECM have been shown to influence the binding and diffusion of signalling molecules and downstream gene expression (Brown & Eames, 2016; Cortes et al., 2009). By removing sulfate groups, sulfatases may influence those processes while allowing PG degradation and turnover (Settembre et al., 2008). Sulfate groups are attached to specific positions on the sugar chains of PGs, contributing to their structure and function. When sulfatases remove these sulfate groups, it causes structural changes in PGs, increasing their vulnerability to degradation by other enzymes. By exposing specific sites on the PG molecule, the removal of sulfate groups enables enzymes like proteases to efficiently break down the PG. This process facilitates the turnover of PGs, as they are continuously synthesized and degraded as part of regular physiological processes.

An example of how sulfatase activity regulates chondrogenesis comes from a study that inhibited all sulfatase activity by knocking out the *Sumf1* gene in mice. *Sumf1* mice mutants presented growth retardation, reduced length of long bones, spinal deformities, short and coarse skulls, and joint deformities (Settembre et al., 2007). Those skeletal phenotypes were caused by increased Fibroblast Growth Factor (FGF) signalling, decreased chondrocyte autophagy, proliferation, and differentiation (Settembre et al., 2008). Failures in endochondral ossification are also observed in patients and animal models with different kinds of MPS, which have mutations in specific lysosomal sulfatase genes (Table 2.1). The cellular and molecular mechanisms of MPS are not completely understood, but skeletal cells in MPS present elevated GAG storage and enlarged lysosomes. The GAG accumulation contributes to abnormal growth factor activity, disrupted cell cycle, and impaired autophagy, leading to endochondral ossification failures. These failures include delayed primary and secondary ossification center formation, impaired chondrocyte proliferation and hypertrophy, impaired longitudinal bone growth, and reduced calcification, among others (Jiang et al., 2020).

Recently discovered extracellular sulfatases (SULF-1 and SULF-2) have also been described to control cartilage homeostasis. These enzymes localize to the cell

membrane and ECM and were shown to be overexpressed in the osteoarthritic cartilage (Otsuki et al., 2008). Follow-up studies showed that Sulfs work regulating the balance between BMP (promotes chondrogenesis and cartilage repair) and FGF (stimulates catabolic responses in chondrocytes) signalling to maintain cartilage homeostasis (Otsuki et al., 2010). Both BMP antagonists (noggin) and FGF ligands and receptors were shown to rely on HS to perform their functions in chondrocytes, so desulfation of HS by Sulfs leads to a release of noggin and a decreased FGF-HS complex formation, resulting in an enhancement of BMP and inhibition of FGF signalling in chondrocytes (Viviano et al., 2004; Wang et al., 2004). Considering the importance of different sulfatases for development and disease, the investigation of other sulfatases with still unknown functions is necessary (Table 2.1).

2.2.5 Arylsulfatase-I (ARSI)

The complete set of 17 human sulfatases genes was reported in 2005, where four sulfatases - *Arylsulfatase H (ARSH), Arylsulfatase J (ARSJ), Arylsulfatase K (ARSK), and Arylsulfatase I (ARSI)* - were added to the previous human sulfatases list. These novel sulfatases were identified through a sequencing comparison against known mammalian sulfatases and their catalytic sites; still, much of their biological and biochemical functions remain unknown (e.g., which are their physiological substrates and subcellular localization? Where/when are those sulfatases expressed during development? Which signalling pathways/diseases do they impact?) (Sardiello et al., 2005). Specifically, this thesis focused on studying *ARSI* and its involvement in endochondral ossification. This choice was made mainly because this sulfatase has been shown to be differentially expressed in mature cartilage and will be better explained throughout this section. Sequence similarities indicate that *ARSI* is closely related to *ARSJ* and the well-known *ARSB.* These three sulfatase genes are on human chromosome 5, have around 54% of sequence identities, and the proteins have conserved active sites. On the other hand, *ARSI* and *ARSJ* have a unique feature compared to the other sulfatases: they have just two coding exons, while other

sulfatases have 8 or more exons (Obaya, 2006). Phylogenetic analyses based on the protein sequences suggest that *ARSI*, *ARSJ*, and *ARSB* are derived from an initial gene duplication event of a primordial invertebrate *Sul-3* gene, which generated two subfamilies. The first contains genes with 8 exons (*ARSB*), and the second contains genes with just 2 exons (*ARSI/ARSJ*). The *ARSI/ARSJ* sub-family underwent a second gene duplication event and sequence divergence, generating separate *ARSI* and *ARSJ* genes (Holmes, 2016). Besides containing fewer exons, ARSI and ARSJ genes seem shorter than ARSB. According to the database GeneCards (https://www.genecards.org) the *ARSI* gene size is 42,965 bases, while ARSB is 208, 879 bases long, and *ARSJ is* 79,444 bases long. Interestingly, even though ARSI and ARSJ have fewer exons, the size of ARSI and ARSJ protein is very similar to other sulfases (between 500-600 aas).

Even though similarities between *ARSI* and *ARSB* genes lead to speculations regarding the protein ARSI being lysosomal and acting similarly to the protein ARSB, studies focusing on the ARSI biochemical and biological roles are scarce and sometimes contradictory. The full-length cDNA of *ARSI* was first cloned in 2006 and used to transfect HeLa cells. *ARSI*-transfected cells showed no activity against 4-MUS (a known arylsulfatase substrate). Further western blot analyses of *ARSI*-transfected cells showed that ARSI was mainly in the cellular pellet, suggesting that this protein is bound to cellular organelles or membranes. Co-localization immunofluorescence was performed to check for ARSI presence in lysosomes, but the results were not clear (Obaya, 2006).

The only other study regarding *ARSI* function overexpressed *ARSI-FLAG* in human spontaneously arising retinal pigment epithelia cell line (ARPE-19 cells). Interestingly, this study shows that, like other sulfatases, ARSI is activated by SUMF1 through modification of Cys93 into FGly. ARSI arylsulfatase activity against 4-MUS was just observed in the conditioned media of cells co-expressing *ARSI-FLAG* and *SUMF1*. Conditioned media refers to the culture medium that has been exposed to living cells for a specific period, during which the cells release various secreted factors, such as growth factors, cytokines, hormones, and other signalling molecules. Western blot analyses of the soluble cellular extract, the insoluble cellular fraction, and the

concentrated cell medium showed ARSI-FLAG just in the insoluble cell fraction. However, when *SUMF1* was co-expressed with *ARSI-FLAG*, the Western blot showed the presence of ARSI in cellular extracts and in the conditioned medium (after precipitation with Trichloroacetic acid (TCA), suggesting this enzyme is being secreted. *SUMF1* might be needed for *ARSI* overexpression since this is a necessary and limiting factor for sulfatase activity. Immunofluorescence and confocal microscopy were used to localize ARSI, which was observed in intracellular granular structures and the ER but not in the lysosomes. The distribution pattern of ARSI-FLAG was similar regardless of the co-expression of SUMF1 data (Oshikawa et al., 2009).

Little is known about ARSI expression in different cell types, tissues, organs, and developmental stages. Some existing expression data come from Northern blot analyses where ARSI expression was assessed in more than 15 adult tissues, 4 fetal tissues, and 7 cell lines, all from human. ARSI expression was observed in the placenta, fetal lung and kidney, and a cell line derived from a lung carcinoma (A549) (Obaya, 2006). Another study examined *ARSI* expression through the UniGene expressed sequence tag count database. Thirteen tissues expressing *ARSI* were found. The highest *ARSI* counts were in the placenta, followed by retinoic acid-treated embryonic stem cells, retinoic acid-treated NT2 embryonic carcinoma cell line, colon tumour, fetal eyes, and lens. RT-PCR was also performed to analyze *ARSI* expression in the following cell lines: ARPE-19); human retinoblastoma Y79; human embryonal carcinoma cell line NTERA-2 cl.D1 (NT2/D1); human fibrosarcoma cell line (HT-1080); and human adenocarcinoma cell line HeLa S3. *ARSI* was mostly present in ARPE-19 and weakly expressed in NT2/D1 and HT-1080 (Oshikawa et al., 2009). Lastly, expression of *ARSI* in embryonic tissues, such as the eye lens, stomach, esophagus, and cartilage, was also observed by *in situ* hybridization (ISH) in mouse embryos at different developmental stages (Ratzka et al., 2010). The fact that ARSI is primarily present in fetal tissues and cancer cell lines may indicate that this enzyme participates in events of tissue remodelling which are very prominent during development and tumour growth.

The participation of ARSI in cartilage and bone development is not currently known. One indication that ARSI is present in developing bones comes from a study by Ratzka et al. in 2010. In this study, the mRNA expression pattern of 9 sulfatases, including *Arsi* (mouse transcripts of *ARSI)*, was investigated in mouse embryos at different developmental stages (E12.5, E14.5 and E16.5) by ISH. *Arsi* expression was observed in Meckel's cartilage, the optic capsule, the parachordal plate (future basioccipital and basisphenoid bone), in the tips of the digits, and in the ulna prehypertrophic and hypertrophic chondrocytes partially overlapping *Ihh* domains. Interestingly, *Arsi* expression was not observed in bone or immature cartilage in chondral bones (Ratzka et al., 2010). Recent RNA sequencing data from our research group – further discussed in Chapter 3 – also showed that *Arsi* was differentially expressed in mature mouse cartilage compared to immature cartilage during endochondral bone development, making it an interesting candidate gene to regulate cartilage maturation.

One significant way to test for gene function is by creating animals that overexpress or do not express the gene of interest, observing their phenotypes, and checking for altered molecular pathways (Simmons, 2008). Animal models from various vertebrate classes, including mammals, birds, and fish, have been utilized to study genes involved in bone development and disease due to the conservation of gene function and homology of skeletal elements across species (Domowicz et al., 2009; Elefteriou & Yang, 2011; Kwon et al., 2019; Marí-Beffa et al., 2021). The use of zebrafish to interrogate gene function has been growing over the years due to many factors such as rapid and external development, quicker generational time, and large availability of embryos. The following review paper addresses the reasoning behind using zebrafish as an animal model to study candidate genes related to skeletal development, including a description of different molecular biology techniques to generate transgenic and mutant animals.

2.3 Review paper: Using zebrafish to test the genetic basis of human craniofacial diseases

R. Grecco Machado and B. Frank Eames. Critical Reviews in Oral Biology & Medicine. Journal of Dental Research 2017, Vol. 96(11) 1192 –1199. https://doi.org/10.1177/0022034517722776

Keywords: zebrafish genetics, GWA, transgenesis, CRISPR, non-coding mutations, adult phenotype

2.3.1 Abstract

Genome-wide association studies (GWAS) opened an innovative and productive avenue to investigate the molecular basis of human craniofacial disease. However, GWAS only identify candidate genes, but do not prove that any particular one is the functional villain underlying disease or just an unlucky genomic bystander. Genetic manipulation of animal models is the best approach to reveal which genetic loci identified from human GWAS are related functionally to specific diseases. The purpose of this review is to discuss the potential of zebrafish to resolve which candidate genetic loci are mechanistic drivers of craniofacial diseases. Many anatomic, embryonic, and genetic features of craniofacial development are conserved among zebrafish and mammals, making zebrafish a good model of craniofacial diseases. Also, the ability to manipulate gene function in zebrafish was greatly expanded over the past 20 years, enabling systems like Gateway Tol2 and CRISPR-Cas9 to test both gain- and loss-of-function alleles identified from human GWAS in both coding and non-coding regions of DNA. With the optimization of genetic editting methods, large numbers of candidate genes can be interrogated efficiently. Finding the functional villains that underlie diseases will permit new treatments and prevention strategies, and also will increase understanding of how gene pathways operate during normal development.

2.3.2 Resolving the cliffhanger of human GWAS

Genome-wide association studies (GWAS) opened an innovative avenue to investigate the molecular basis of human health and disease. These studies identify specific genomic loci that correlate with particular diseases (Dermitzakis & Clark, 2009). GWAS use either single nucleotide polymorphism microarrays, whole genome sequencing (WGS), or whole exome sequencing (WES) (Li et al., 2008; Rabbani et al., 2014). WES scans only the coding regions of the genome, while the other two methods interrogate also non-coding regions. Currently, there are 7,723 GWAS in the GWAS Catalog, describing 31,231 unique SNP-trait associations (MacArthur et al., 2017), and these numbers will continue to increase dramatically in the future.

In some ways, however, GWAS are classic cliffhangers, referring to old movies that end with a train hanging on the edge of a cliff, so that the customer will pay to see the next one. Unfortunately, resolution of a GWA cliffhanger often never comes. GWAS only identify candidate “mutations”, but do not prove that any particular one is the functional villain underlying disease or just an unlucky genomic bystander (Ioannidis et al., 2009). Statistical and computational analyses are necessary to prioritize which GWA candidates should be analyzed, but really to resolve the GWA cliffhanger, functional analyses of candidate mutations are necessary (Edwards et al., 2013; Leslie et al., 2015).

Genetic manipulation of animal models is the best approach to reveal which genetic loci identified from human GWAS are related functionally to specific diseases (Ioannidis et al., 2009). Given this technical reliance, suitable animal models to resolve the GWA cliffhanger need to have extensively studied genomes. The mouse is most often employed, since it is evolutionarily closer to humans and thought to have the most “human-like” physiological processes, embryonic morphogenesis, and gene expression and function (Lieschke & Currie, 2007). However, high costs of maintenance, low fecundity, and *in utero* embryonic development lead to practical difficulties using mice to test GWA candidates, especially in large scale (Lieschke & Currie, 2007; Van Otterloo et al., 2016). Compared to mice, zebrafish have lower maintenance costs, higher

fecundity rates, external fertilization and development (*in ovo*), quick maturation, and more transparent embryos. Due to an increasing appreciation for conservation of gene function and cell and tissue processes among vertebrates (O'Brien et al., 2004), zebrafish (*Danio rerio*) has become increasingly popular to model human skeletal and craniofacial diseases (Van Otterloo et al., 2016).

The purpose of this review is to discuss the potential of zebrafish to resolve which candidates from human GWAS are mechanistic drivers of craniofacial diseases. We highlight anatomic, embryonic, and genetic features of craniofacial development that are conserved among mammals and zebrafish, and current molecular genetic techniques to manipulate gene function in zebrafish.

2.3.3 Zebrafish as a genetic model of human craniofacial disease

The once fanciful idea that gene function and even complex developmental processes are heavily conserved among disparate animal clades is now accepted scientific dogma. Indeed, craniofacial development employs almost the same embryological processes and molecular pathways in most vertebrates. Neural crest and mesodermal progenitor cells migrate to homologous regions of the head, and, through conserved molecular and cellular interactions, homologous elements of the vertebrate craniofacial skeleton take shape (Santagati & Rijli, 2003; Yelick & Schilling, 2002). Therefore, vertebrate models can provide insights into the genetic causes of human craniofacial defects.

How is a zebrafish skull related to that of mammals? By some measures, any comparison is immediately called into question. The adult zebrafish skull has 73 bones, whereas its mammalian counterpart only has approximately 22 (Mork & Crump, 2015). These vastly different numbers are misleading, however, since most mammalian head bones have homolog(s) in fish (Figure 2.3). For example, in both zebrafish and mammalian skulls, the cerebrum is protected by frontal and parietal bones, the dentary is a lower jaw bone surrounding Meckel's cartilage, and an occipital complex protects

the brain stem (Cubbage & Mabee, 1996; Richardson et al., 2014). And yes, zebrafish have teeth (Huysseune et al., 1998). Of course, there are also clade-specific adaptations that can mask underlying homologies. Zebrafish teeth are not strictly homologous, since they do not form in the mouth, as in mammals, but are located along the most posterior pharyngeal skeletal element (Huysseune et al., 1998). In perhaps the best-known example, mammalian middle ear ossicles are homologous to bones of the zebrafish jaw joint (Anthwal et al., 2013). Other craniofacial discrepancies in mammals compared to fish are the loss of posterior pharyngeal skeletal elements and the addition of a nasal airway separate from the mouth, massively transforming the palate.

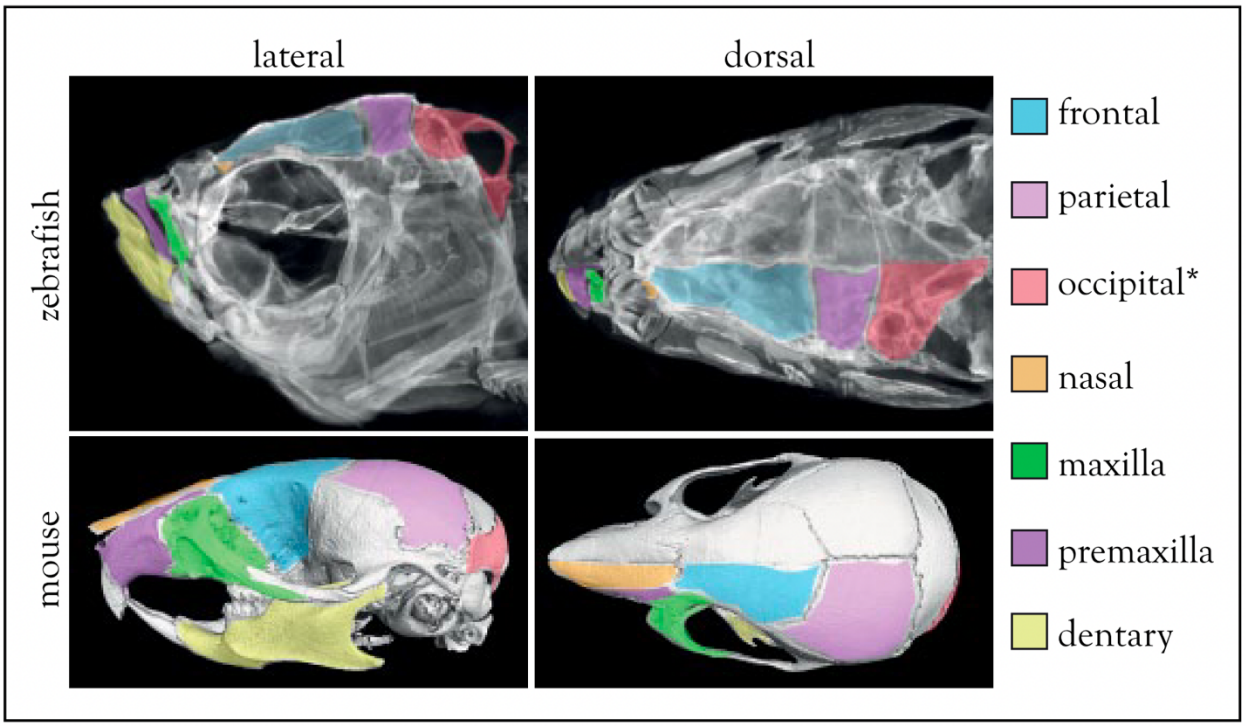


Figure 2.3 - Conservation of skeletal anatomy between zebrafish and mouse skulls. Optical projection tomography (OPT) projection images of an adult zebrafish skull and microCT 3D renderings of an adult mouse skull are color-coded for easily visible homologous skeletal elements that retain the same name (only left side of midline elements are colored in dorsal views). Additional homologous bones are not depicted, because they are either not visible easily or have different names. *Although the exact relationships can be unclear, the mammalian occipital might be homologous to a fusion of several zebrafish bones, including the epioccipital, exoccipital, supraoccipital, and basioccipital (all colored red). Many other examples of putative fusions, or alternatively, simple evolutionary losses, might explain the decreased number of skeletal elements in the mouse skull compared to that of the zebrafish. (Mouse images courtesy of Nathan Young)

Despite some anatomical differences in the craniofacial skeleton between zebrafish and mammals, homologous regions share an incredible amount of cellular and molecular features (Mackay et al., 2013; Medeiros & Crump, 2012). For example, even though zebrafish form teeth in their throats, the molecular pathways mediating tooth formation are conserved among zebrafish and mammals (Harris et al., 2008). While later stages of mammalian palate formation do not occur in zebrafish, such as morphogenesis of the palatal shelves, cellular and molecular control of early events are conserved (Ghassibe-Sabbagh et al., 2011; Swartz et al., 2011). For instance, *pdgfra* mutant zebrafish demonstrated neural crest migration defects in palatal precursors (Eberhart et al., 2008), and subsequent human studies confirmed that *PDGFRA* mutations are associated with cleft palate (Rattanasopha et al., 2012). Finally, molecular pathways of skeletal cell differentiation are also highly conserved between zebrafish and mammals. Due to such findings, studying zebrafish craniofacial development can shed light on the etiology of human craniofacial diseases. Indeed, many human craniofacial diseases have been modelled in zebrafish (Mork & Crump, 2015).

Before discussing molecular genetic tools available to probe candidate craniofacial disease-causing genes, a few notes are necessary on the similarities and differences between zebrafish and human genomes. Completion of the zebrafish genome project allowed a direct comparison with the human genome (K. Howe et al., 2013). 71% of human genes have orthologs in zebrafish, and 82% of the human genes bearing morbidity descriptions at the Online Mendelian Inheritance in Man (OMIM) have zebrafish orthologs. Interestingly, 76% of human genes identified in GWAS have at least one zebrafish ortholog (K. Howe et al., 2013). This high degree of gene similarity supports the use of zebrafish to model human diseases.

Far before the zebrafish genome was sequenced, however, it was argued that the ancestor of teleost fish underwent a whole genome duplication event approximately 340 million years ago, in principle giving that fish two copies of most genes that are found in mammals today (Postlethwait et al., 2000). Although this might present a major

problem using zebrafish as a genetic model for human diseases, the case is not that dire. Duplicated genes might be lost by attrition (nonfunctionalization), divide the original functions of the ancestral gene among the duplicates (subfunctionalization), evolve a new function (neofunctionalization), or simply function redundantly (Glasauer & Neuhauss, 2014). Nonfunctionalization appears to be the most common fate for genes duplicated in the ancestral teleost, a fact that helps greatly using zebrafish as a genetic model of human disease. Estimates suggest that only 20-30% of mammalian genes have two co-orthologs in zebrafish today (Postlethwait et al., 2000; Woods et al., 2005). Subfunctionalization among some of these remaining duplicates can favor genetic studies in zebrafish, because it permits analyses of traits that might be masked by pleiotropy when the single mammalian copy is targeted (Postlethwait et al., 2004). So really, only neofunctionalization and redundancy complicate the use of zebrafish to model genetics of human disease. In the case of redundancy, phenotypes might only emerge if both co-orthologs are targeted, but new techniques make such an approach feasible.

2.3.4 The ever-expanding toolkit to interrogate gene function in zebrafish

Over the past two decades, the ability to manipulate gene function in zebrafish has grown dramatically (Figure 2.4). From mutagenesis and transgenesis protocols to morpholinos and site-specific gene editting, zebrafish has now surpassed the utility of mouse by many accounts. Historically, zebrafish emerged as a genetic model in the 80's (Streisinger et al., 1981), but this role was enhanced exponentially by large forward genetic screens in the 90's (Driever et al., 1996; Haffter et al., 1996). While these studies isolated 48 mutations related to craniofacial development, the random feature of forward genetic approaches make them less useful to test promising candidate genes from GWAS. Reverse genetic approaches in zebrafish, however, interrogate the function of a candidate gene in a more efficient manner.

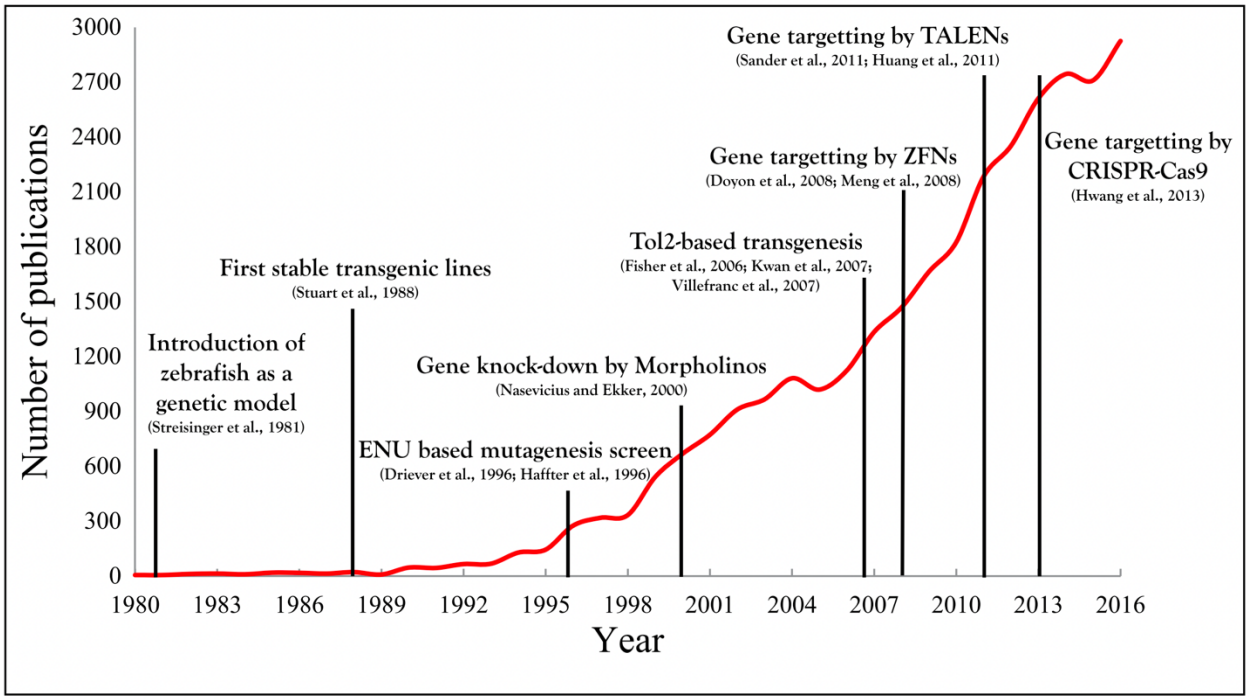


Figure 2.4 - Landmarks in zebrafish genetic research have increased the number of published studies. Data from SCOPUS shows how publications with the term “zebrafish” (red line) have increased over the past three decades. This increase is accompanied by hallmarks in genetics that facilitated the manipulation of gene function in zebrafish.

2.3.4.1 Gain-of-function approaches

Transgenic techniques permit evaluation of the normal function of a given gene, and in the case of diseases that demonstrate dominant inheritance (i.e., gain-of-function alleles), can be used to reveal causative genes from human GWAS. The most common transgenic technique is the Gateway Tol2 system, which flanks the gene of interest with two Tol2 recombination *cis* sequences (Kawakami, 2007). The Tol2 construct is mixed with Tol2 *transposase* mRNA and microinjected into early zebrafish embryos, where it is then inserted randomly into the zebrafish genome (Kawakami, 2007). Transgenic approaches are gain-of-function approaches, allowing elucidation of the normal function of a candidate gene or the dominant function of a particular allele identified from GWAS.

By contrast, transgenic approaches are not useful if the candidate allele identified from GWAS is a loss-of-function allele.

To elucidate which cell lineage might drive human disease, the Gateway Tol2 system can target specific cell lineages using cell-specific gene enhancers (Figure 2.5). Visualizing those cells by confocal microscopy or even time-lapse movies in transgenic or mutant backgrounds reveals the effects of a given gene on the location and/or behavior of those cells (Le Pabic et al., 2014; McGurk et al., 2014). Diverse transgenic reporter lines allow the visualization of skeletal lineages at different stages of differentiation in zebrafish embryos/larvae, including: *sox10:EGFP* (Wada et al., 2005) and *fli1a:EGFP* (Lawson & Weinstein, 2002) for neural crest cells; *sox9a:EGFP* (Eames et al., 2013) and *col2a1a:EGFP* (Dale & Topczewski, 2011) for chondrocytes; and *sp7:EGFP* (DeLaurier et al., 2010) and *RUNX2:EGFP* (Knopf et al., 2011) for osteoblasts, among many others (Hammond & Moro, 2012). Reporter lines also have cells that glow in response to growth factor signalling or manipulate activity of signalling pathways known to be involved in craniofacial development and disease (Moro et al., 2013). Consequently, the ability to observe signalling and specific cell lineages helps reveal the molecular and cellular basis of craniofacial morphogenesis and disease (McGurk et al., 2014).

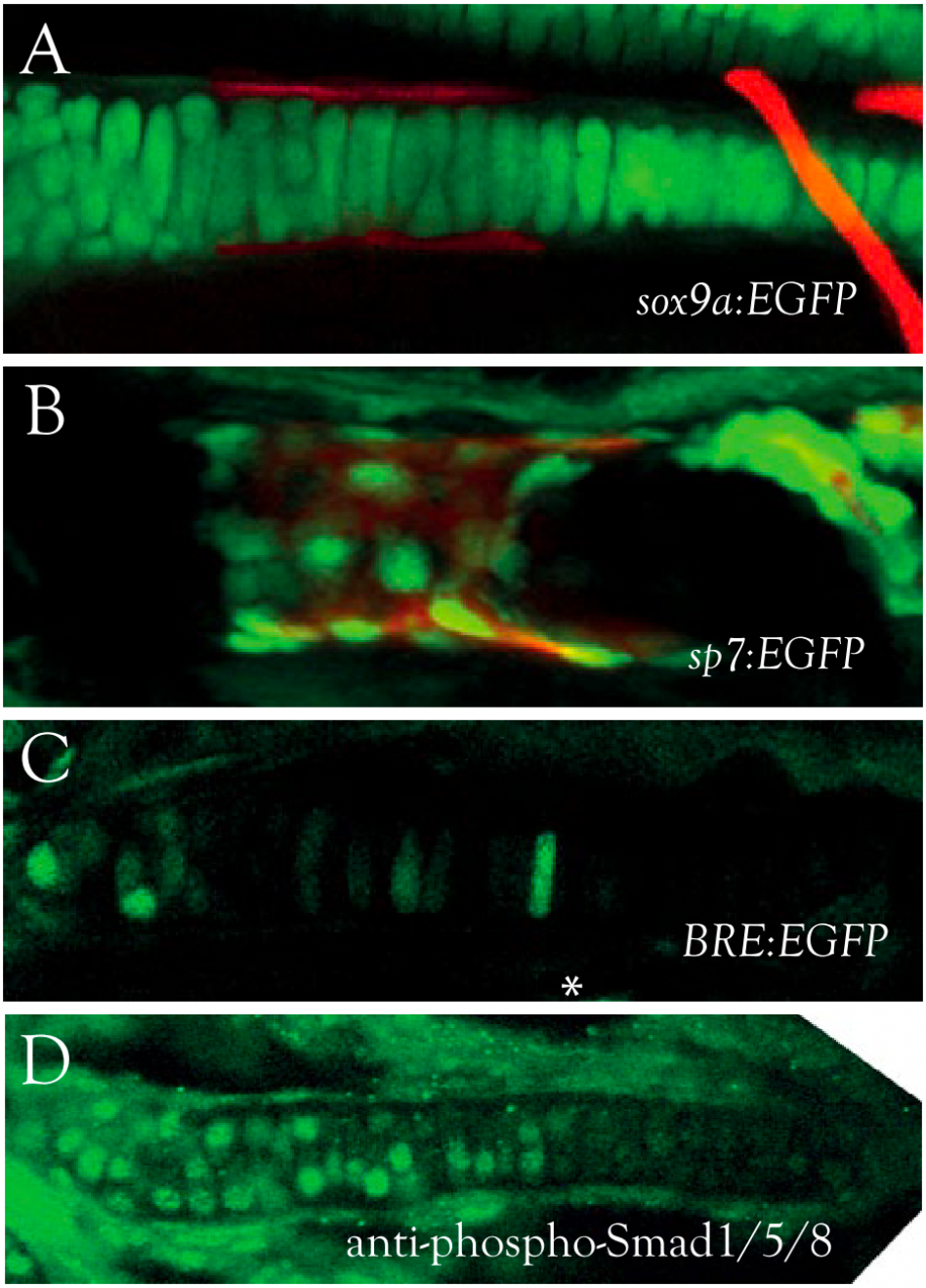


Figure 2.5 - Transgenic zebrafish highlight specific cell lineages and/or growth factor signaling activity. Confocal projections of the zebrafish ceratohyal demonstrate cells of developing cartilage (A, *sox9a:EGFP*) or bone (B, *sp7:EGFP*) in living animals. Other transgenic zebrafish lines reveal which cells are responding to growth factor signalling, such as BMP signalling (C, *BRE:EGFP*), producing similar results to more traditional immunocytological techniques (D, anti-phospho-Smad1/5/8), except that the former approach can be done in living animals.

2.3.4.2 Loss-of-function approaches

Morpholino oligomers (MOs), or modified antisense oligonucleotides that affect mRNA translation or splicing, were one of the first reverse genetic tools to knock down specific gene function (Figure 2.4) (Nasevicius & Ekker, 2000). Allowing rapid, albeit transient, knockdown upon injection into early embryos, MOs can help to identify causative genes identified from GWAS (Bielczyk-Maczynska et al., 2014; Liu et al., 2013).

Results from MOs should be carefully controlled for the effectiveness of gene knockdown, possible off-target effects, and non-specific toxicity (Eisen & Smith, 2008). The need for so many controls to validate MO knockdowns might make them less attractive, since one of their advantages is the rapid knockdown and assessment of gene function. Probably because of these drawbacks, up to 80% of gene mutations made via other approaches failed to recapitulate the morphant phenotypes (Kok et al., 2015; Law & Sargent, 2014).

Without a doubt, the development of sequence-specific nucleases for genome editting was a great breakthrough for loss-of-function studies in zebrafish (Gaj et al., 2013). The customization of these systems to recognize and bind to any desired position in the genome and create double-strand breaks (DSBs) finally brought the study of gene function in zebrafish up to levels achieved in mouse (Bedell & Ekker, 2015). DSB repair is error-prone, resulting in insertions and deletions (indels) that typically cause null alleles if the 5' end of a gene is targetted (Gaj et al., 2013). By introducing a custom-made template during DSB repair, specific DNA sequences of interest can be introduced at the target site (Gaj et al., 2013). Template-driven repair can 1) replace a critical exon with a reporter protein (simultaneously creating a loss-of-function in your gene of interest and a lineage-specific marker for your cell of interest); 2) "knock-in" a point mutation identified from human GWAS into the coding sequence; or 3) introduce *loxP* sites around your gene of interest to make a conditional loss-of-

function allele. However, the ease of the non-template repair approach makes it more common currently.

Zinc finger nucleases (ZFNs) were the first tools for site-specific genome modification in zebrafish (Figure 2.4)(Doyon et al., 2008; Meng et al., 2008), followed closely by Transcription activator-like effector nucleases (TALENs)(Cade et al., 2012). Both ZFNs and TALENs bind specific, nearby genomic sequences on opposite DNA strands, producing a DSB in the targetted genomic region (Li et al., 2016). Although TALENs generate higher rates of deletions than insertions compared to ZFNs (Kim et al., 2013), the functions of ZFNs and TALENs are similar. TALENs present some advantages over ZFNs, because they are easier to design and present a higher targetting efficiency and lower off-target effects (Gaj et al., 2013).

For a relatively short time, TALENs led gene function studies in zebrafish (Figure 2.6) (Bedell et al., 2012), including validation of human GWA candidate genes (Mackay et al., 2013). For example, WES from a patient with syndromic colobomatous microphthalmia, with ocular and craniofacial phenotypes, suggested *EFTUD2* as a gene of interest (Deml et al., 2015). The zebrafish homolog *eftud2* was mutated using TALENs, and mutants presented reduced head size, small eyes, and massive cell death. Similarities between the phenotypes of the patient and mutant zebrafish suggested that *eftud2* has an important conserved role during vertebrate craniofacial development (Deml et al., 2015). Analyses of zebrafish *eftud2* mutants could reveal mechanisms and potential treatments of this human disease.

And just as TALENs became the poster child for genome editting in zebrafish, along came the CRISPR-Cas9 system, which uses guide RNAs (gRNAs) to direct the Cas9 protein to generate a DSB at a specific genomic sequence (Ran et al., 2013). Compared to ZFNs and TALENS, the CRISPR-Cas9 system is cheaper to use, is easier to design and assemble, reduces off-target effects, and generates mutations six times more efficiently (Varshney et al., 2015). CRISPR-Cas9 was both efficient and robust in zebrafish, generating indels at 8 of 10 targetted genes using 11 different gRNAs (Hwang et al., 2013). Besides creating indels that often produce null alleles, the

CRISPR system can generate knock-ins (Hruscha et al., 2013), multiple knockouts (Ota et al., 2014), and conditional knockouts (Yin et al., 2015). The future is bright for expanding utilization of the CRISPR system, such as high-throughput mutagenesis screens (Varshney et al., 2015).

Data from ZFIN, the Zebrafish Model Organism Database (http://zfin.org), illuminate not only the history of techniques that alter gene function in zebrafish, but also future trends. ZFIN is an extensive online resource that assembles zebrafish genetic, genomic, phenotypic, and developmental data (D. G. Howe et al., 2013) . Maybe because of the advantages of the CRISPR system compared to TALENs, ZFIN already reports 1982 CRISPR-mediated vs. 372 TALEN-mediated mutations. MOs still lead the way with 8935 search results, given their historical precedence. Considering the versatility of the CRISPR-Cas9 system and the suitability of zebrafish as an animal model to study gene function and development, we note that the number of publications of mutations in zebrafish genes using CRISPR-Cas9 is increasing exponentially (Figure 2.6)

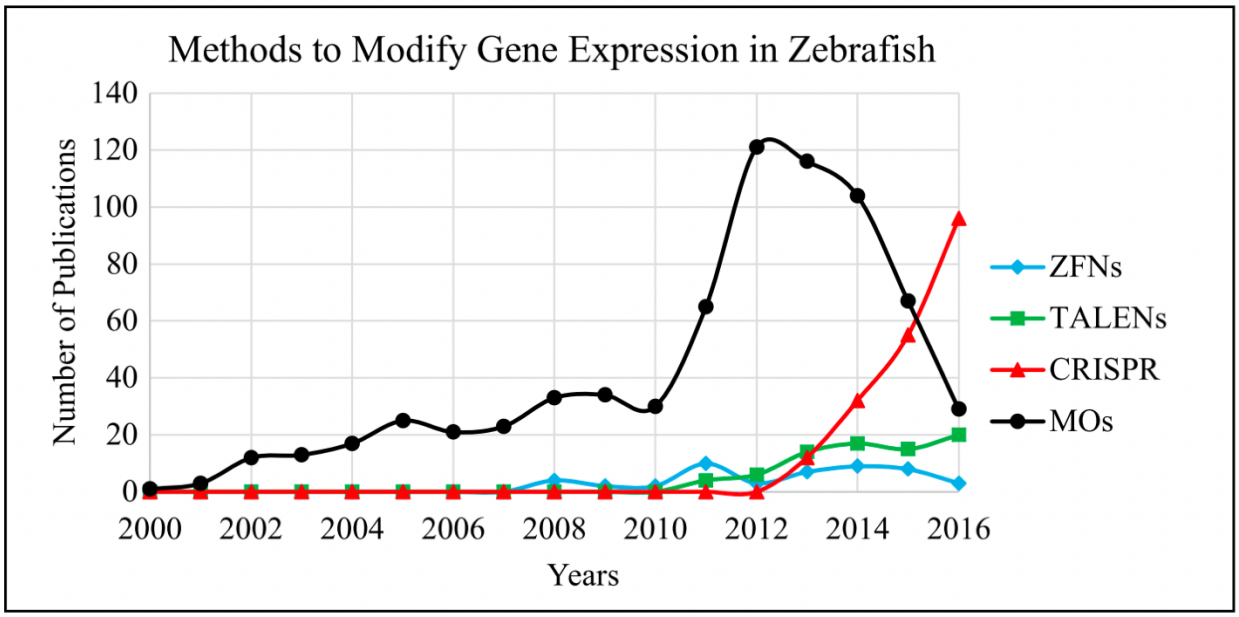


Figure 2.6. CRISPR has eclipsed the dominance of morpholino oligomers (MOs) in zebrafish research. Number of scientific papers employing specific methods to modify gene expression in zebrafish over the years was tracked searching SCOPUS for the terms “zebrafish” and “ZFN” (zinc finger nuclease), “TALEN” (transcription activator-like effector nuclease), “CRISPR” (clustered regularly interspaced short palindromic repeats), or “morpholino”. Publications with ZFNs and TALENs did not increase much since their use in zebrafish in 2008 and 2011, respectively. MOs, first used in zebrafish in 2000, had the lead in the publication numbers until 2015, when it was surpassed by CRISPR, which appears to be increasing exponentially since its development in 2013.

Modulating zebrafish genes with CRISPR-Cas9 already led to understanding how human GWA candidate genes might cause syndromes with craniofacial phenotypes (Kury et al., 2017; Shaw et al., 2017). A WES of individuals presenting a variety of craniofacial and neurological defects identified deletions and point mutations in *PSMD12* (Kury et al., 2017). To demonstrate a causative role for this gene, the orthologous zebrafish gene *psmd12* was disrupted by CRISPR-Cas9. Zebrafish *psmd12* mutant larvae presented a variety of craniofacial defects, including microcephaly and altered skeletal element patterning. Accordingly, Psmd12 is important for proper craniofacial development, and abnormalities in *psmd12* loss of function zebrafish resembled those seen in human patients with *PSMD12* mutations (Kury et al., 2017). Through WES and WGS studies of 40 patients and their families, mutations in *SMCHD1* were candidates for arhinia (absence of external nose) and associated craniofacial defects (Shaw et al., 2017). *smchd1* function was evaluated in zebrafish by MOs and CRISPR-Cas9. Morphants and mutants had similar craniofacial phenotypes, including microphthalmia and altered skeletal patterning. Interestingly, morphant phenotypes could be rescued by WT *SMCHD1* mRNA, but not three recurrent arhinia-associated human mRNA variants (Shaw et al., 2017). This last result demonstrates the ability to use zebrafish to analyze whether mutant alleles identified from human GWAS might be null or hypomorphic alleles. Once a null mutation in zebrafish is established using CRISPR-Cas9, if a good assay exists to assess gene function (i.e., biochemical or phenotypic), then rescue experiments can be performed by injecting wild-type or mutant alleles into mutant embryos. Hypomorphic mutant alleles should rescue only partially the mutant phenotype, while null alleles should not rescue at all.

Regardless of the chosen method to analyze gene function, dosage of MOs, enzymes, and RNAs injected into zebrafish embryos need to be controlled carefully. Concentrations below the optimal values will not generate phenotypes. On the other hand, excessive concentrations can lead to embryonic lethality and off-target effects, causing phenotypes that are not really related to the target gene (Lawson, 2016). In the case of MOs that are coinjected with mRNA to rescue the morphant phenotypes, for example, two dosages must be considered (Eisen & Smith, 2008). For ZFNs, effective doses reported in the literature vary by more than 1000-fold, ranging from 5ng (Doyon et al., 2008) down to 5pg (Meng et al., 2008). Cas9 and gRNA concentrations need optimization, and when targeting multiple genes at once, each individual gRNA ideally should be optimized (Shah et al., 2015). Dose-response curves for each of these reagents should reveal a reasonable balance between phenotypic efficiency and toxicity in the zebrafish embryos.

2.3.5 Mutations in non-coding genomic regions

Many GWA candidate mutations actually are located in non-coding genomic regions, raising the prospect that *cis*-regulatory element (CRE) mutations are causing the disease (Bhatia et al., 2015). Although non-coding regions once were considered "junk DNA", they can control where, when, and at what levels genes are expressed (Bhatia & Kleinjan, 2014). Mutations in CREs already have been implicated in a wide range of craniofacial diseases (Attanasio et al., 2013; Leslie et al., 2015). One way that CRE mutations can cause disease is by altering essential transcription factor binding sites (Bhatia et al., 2015). However, CREs often are located very far from their target genes (sometimes megabases away), hidden in introns or gene deserts, making it hard to correlate them with their specific target genes. Perhaps because of this challenge, most functional studies of GWA candidate mutations focus on coding regions (Attanasio et al., 2013).

Accepting this challenge, zebrafish are a good model to study candidate human non-coding sequences for many reasons. First, important CREs are highly conserved

among humans and zebrafish (Woolfe et al., 2005). Second, even if nucleotide sequence varies among zebrafish and human CREs, they often retain transcription factor binding activity and overall function (Taher et al., 2011). Third, specific CREs can be targetted using the genomic editting techniques detailed above. Finally, transgenic assays to test potential CRE functions are simply easier and quicker in zebrafish than in mice (Bhatia et al., 2015; Woolfe et al., 2005), and CRISPR-Cas9 approaches even permit high-throughput dissection of non-coding elements (Wright & Sanjana, 2016).

Mutations in non-coding genomic regions identified from GWAS of human craniofacial disease have been studied in zebrafish. 13 genomic regions identified from published GWAS on nonsyndromic cleft lip with or without cleft palate underwent targetted sequencing in over 4000 patients, revealing 66 candidate mutations (Leslie et al., 2015). Of these, 63 were in non-coding regions (including 11 in predicted CREs). Functional analysis of a non-coding mutation at the *FGFR2* locus was performed in zebrafish, since this gene is known to participate in craniofacial development and disease (Stanier & Pauws, 2012). The wild-type allele of this putative human CRE (located 254.6Kb downstream of the *FGFR2* transcription start site) was cloned into a plasmid containing a basal promoter driving GFP expression and injected into zebrafish embryos. Indeed, this human CRE drove expression in the zebrafish neural keel, brain, and neural crest, sites of endogenous *fgfr2* expression. On the other hand, mutations of this putative human CRE showed reduced enhancer activity in zebrafish compared to the wild-type sequence, suggesting that the identified mutation suppresses a neural crest enhancer that regulates *FGFR2* expression (Leslie et al., 2015).

2.3.6 Zebrafish can reveal gene functions in late development

Increasingly, adult zebrafish phenotypes are being examined, since some physiologic processes, such as bone turnover, repair, and remodeling, must be studied later in life (Mariotti et al., 2015). Gene knockdown by MOs is transient, precluding study of juvenile and adult phenotypes. Methods that produce heritable mutations, such as ZFNs, TALENs, and CRISPR-Cas9, enable analyses of all developmental stages. While

some mutations are incompatible with life, conditional mutant strategies allow the study of these mutants in later periods of development (Maddison et al., 2014). Also, viable zebrafish mutants have shed light on nonlethal human birth defects. For example, the zebrafish *Chihuahua* mutant exhibitted normal cartilage formation, but had a general defect in bone growth and mineralization that led to bone fragility, resembling osteogenesis imperfecta (OI) in humans (Fisher et al., 2003). Indeed, *Chihuahua* zebrafish had a mutation in *col1a1a*, encoding one of the co-orthologues of human *COLLAGEN TYPE 1, ALPHA 1*, which is one of two causative genes that account for 90% of OI (Benusiene & Kucinskas, 2003). In another example, mutations in the Ectodysplasin (Eda) pathway in zebrafish cause ectodermal defects that resemble those in humans with hypohidrotic ectodermal dysplasia, caused by Eda pathway mutations (Harris et al., 2008; Lind et al., 2006).

Viable craniofacial mutants can also be used to investigate the process of endochondral ossification during development and disease. Zebrafish larvae with mutations in genes encoding proteoglycan (PG) synthesis enzymes (*xylt1* and *fam20b*) exhibitted a reduction in cartilage PGs and an accelerated endochondral ossification in their craniofacial skeletons (Eames et al., 2011). Interestingly, these mutants grow into viable adults with more dramatic phenotypes, exhibiting severe reductions to the size of the craniofacial skeleton, foreshortened upper and lower jaws, hypoplastic midface, and bulging eyes. Humans with Raine syndrome have mutations in *FAM20C* and display similar facial dysmorphies as seen in *fam20b* mutant zebrafish (Simpson et al., 2009), suggesting that these zebrafish can be a model to help understand the etiology of Raine syndrome.

2.3.7 Conclusion

Advances in genome sequencing technologies allowed GWA studies to identify many human genes that are candidates to generate craniofacial disease when mutated. Even though it is exciting to have so many candidate genes, efficiently translating these findings into clinical applications depends upon functional verification that the mutation

causes the disease phenotype. While the general population has been bombarded for many years with promises that genome sequencing will lead to disease cures, scientists are also eager to show that GWAS will be a major step in translational research. Among the available animal models, zebrafish can be used to test rapidly and efficiently a large number of candidate genes for craniofacial diseases. Specifically, the Tol2 and CRISPR-Cas9 systems allow for customized genetic manipulation of zebrafish, enabling tests of both gain- and loss-of-function alleles identified from human GWAS. Considering zebrafish as the star of the next chapters of this story, resolution of the GWA cliffhanger seems to be closer every day. Validation of the functional villains of diseases is critical to understand gene pathways, purpose new treatments, and develop prevention strategies for diseases. We expect that, with the optimization of zebrafish genome editting techniques, this validation can happen in a faster and more reliable way in the coming years.

ACKNOWLEDGMENTS

This work was supported by Saskatchewan Health Research Foundation and Canadian Institutes of Health Research. Authors have no conflicts of interest.

2.4 Hypothesis and specific aims

To identify novel candidate genes controlling cartilage maturation during endochondral ossification, I searched for genes differentially expressed in developing mature cartilage compared to immature cartilage. *Arsi* appeared to be a promising candidate gene since its mRNA was shown to be highly expressed in mature cartilage by Ratzka et al. and our recent RNA sequencing results confirm this gene is differentially expressed in mouse mature cartilage (Ratzka et al., 2010). This led to our hypothesis that ARSI decreases sulfate ester levels in mature cartilage thus promoting endochondral ossification

To investigate our hypothesis five specific aims were considered:

1. To analyze the protein and mRNA expression pattern of ARSI, and two other known chondroitin sulfate sulfatases -GALNS, and ARSB- on mature and immature cartilage in the chicken humerus (Chapter 3).
2. To assess ARSI protein expression in sections of mouse humerus and during ATDC5 chondrogenic differentiation (Chapter 3).
3. To generate transgenic zebrafish lines overexpressing ARSI (Chapter 4)
4. To assess if ARSI controls sulfur levels during cartilage maturation using transgenic zebrafish (Chapter 4)
5. To analyze if *ARSI* promotes endochondral ossification *in vivo* in transgenic zebrafish (Chapter 4)

CHAPTER 3: ARYLSULFATASE I, EXPRESSED DURING CARTILAGE MATURATION, REDUCES CHONDROITIN SULFATE LEVELS

Rafaela Grecco-Machado[1], Devin S. Brown[1], Mark J. Hackett[2], Yuko Naito-Matsui[3], Amir Ashique[1], Katie Ovens[4], Ian McQuillan[4], Hiroshi Kitagawa[3], Ingrid J. Pickering[5], Graham N. George[5], B. Frank Eames[1*]

This chapter is comprises a manuscript in preparation containing data from myself and collaborators (Dr. Amir Ashique, Dr. Katie Oven, MSc. Devin Brown, Dr. Yuko Naito-Matsui, Dr. Hiroshi Kitagawa, and Dr. Brian Eames). I wrote the manuscript, organized most of the figures, graphs, and statistics, and analyzed the results. More specifically, I did all the experiments and analyses for figure 3.1 and for figures 3.7 to 3.14. MSc. Devin Brown generated the data for figures 3.2 and 3.4. Using his specimens, I quantified sulfur, sulfate esters, and taurine levels on chick samples and made the graphs and statistical analyses. MSc. Devin Brown also generated the data and figures 3.3 and 3.5 which I formatted. Dr. Amir Ashique and Dr. Katie Oven generated and analyzed the laser capture and RNA sequencing results. Using their data, I looked for ARSI expression and made figure 3.6 and tables 3.1 and 3.2. Dr. Yuko Naito-Matsui and Dr. Hiroshi Kitagawa provided us with data from ARSI biochemical function, and with their data, I made the graphs and statistics in figures 3.6 and 3.7.

3.1 Abstract

During endochondral ossification, cartilage undergoes a process called maturation, during which chondrocytes biochemically modify their extracellular matrix. For example, mature chondrocytes express new collagens, such as type X collagen, and enzymes, such as matrix metalloproteinase 13, that degrade collagens. Biochemical modifications to proteoglycans (PGs), which are predominantly chondroitin

sulfate, have not been characterized fully during cartilage maturation. PG sulfation - reflected specifically by sulfate esters - has been shown to influence cartilage biomechanics and signalling pathways that may change gene expression. Here we tested the hypothesis that the orphan sulfatase Arylsulfatase I (ARSI) decreases PG sulfation during cartilage maturation. Synchrotron x-ray fluorescence (XRF) imaging demonstrated that sulfate esters decreased significantly in developing mature cartilage**.** While expression of the two known chondroitin sulfate PG sulfatases (ARSB and GALNS) was not specific to mature cartilage, laser-capture microdissection and RNAseq revealed up-regulation of the sulfatase *Arylsulfatase I* (*Arsi*) in mature cartilage. Increased ARSI protein expression was further demonstrated in mature cartilage *in vivo* and during maturation of ATDC5 chondrocytes *in vitro*. Biochemical assays overexpressing human *ARSI* in HeLa cells confirmed that ARSI is a novel PG sulfatase. Disruption of PG sulfation is associated with a variety of endochondral ossification defects; however, there is limited understanding of the underlying mechanisms. This first biochemical characterization of ARSI as a chondroitin sulfate sulfatase in cell lines *in vitro*, and its apparent role in promoting endochondral ossification, potentiates future work uncovering the role of ARSI during skeletal development and disease.

3.2 Introduction

During endochondral ossification, mesenchymal cells differentiate into chondrocytes and secrete an extracellular matrix (ECM) of collagens and proteoglycans (PGs) that serves as a cartilage template for subsequent bone formation (Long & Ornitz, 2013). Typically, in the middle region of this template, some chondrocytes undergo a maturation process, resulting in prehypertrophic and hypertrophic chondrocytes (Provot & Schipani, 2005). Mature chondrocytes secrete molecules that induce the formation of an overlying perichondral bone. In addition to acquiring such traits as cellular hypertrophy, chondrocytes biochemically remodel their ECM during maturation. For example, while the initial cartilage template contains abundant collagen type II (Col2), mature chondrocytes downregulate *Col2a1* expression and upregulate collagen type X

(Col10) (Behonick & Werb, 2003). Degradation of ECM molecules also occurs during maturation, and mature chondrocytes express enzymes, such as matrix metalloproteinase 13 that can degrade collagens (Melrose et al 2016, Ortega 2004).

Biochemical modifications of PGs during cartilage maturation remain to be characterized fully but understanding such PG modifications could provide insight into the mechanisms of human disease. Mutations in PG synthesis genes underlie at least 20 different human skeletal diseases, including osteoarthritis, spondyloepimetaphyseal dysplasia, and diastrophic dysplasia (Brown & Eames, 2016). These mutations may affect the core proteins of PGs or enzymes that control the addition of glycosaminoglycan (GAG) side chains to those core proteins (Schwartz and Domowicz 2002). Another class of these mutations affects GAG sulfation, a biochemical modification of PGs that can influence cartilage function and even chondrocyte maturation (Dzobo et al., 2016; Gama et al., 2006; Gao et al., 2014; van der Kraan et al., 2002). Higher PG sulfation adds negative charges, which attract more water, increasing the compressive-resistant strength of cartilage (Chahine et al., 2005). Moreover, PG sulfation influences the binding affinities and signalling activities of growth factors involved in cartilage maturation and bone development (e.g., members of the transforming growth factor-beta family, Indian hedgehog, and fibroblast growth factors) (Cortes et al., 2009; da Costa et al., 2017; Otsuki et al., 2010; Rosen & Lemjabbar-Alaoui, 2010).

During PG sulfation, sulfate groups are attached to sugars of the repeating disaccharide GAG side chains in the form of O–linked sulfate esters (Hanson et al., 2004). Sulfated GAGs are categorized into different classes depending on the disaccharide composition and sulfation pattern. Chondroitin sulfate (CS), the most abundant class of GAG in cartilage, is composed of glucuronic acid (GlcA) and N-acetyl galactosamine (GalNAc) disaccharide repeats (Knudson & Knudson, 2001). Chondroitin sulfate proteoglycans (CSPGs) may be sulfated on the 2-carbon of GlcA and the 4- and 6-carbons of the GalNAc sugar (Kusche-Gullberg & Kjellen, 2003; Uyama et al., 2007). O-linked sulfate esters are attached to growing CS chains by specific sulfotransferases and removed by specific sulfatases (Sardiello et al., 2005).

Among 17 sulfatases that have been identified in the human genome from gene sequence analyses (Sardiello et al., 2005), only two have been shown so far to specifically regulate CS. ARSB acts on GalNAc-4-S, and GALNS acts on GalNAc-6-S (Bhattacharyya et al., 2015; Tomatsu et al., 2005). In human patients, mutations to *ARSB* or *GALNS* can cause mucopolysaccharidosis (MPS) types IV or VI, respectively (Muenzer, 2011). Both conditions cause a range of stunting during skeletal development depending on the severity of the mutations, but how this happens is unknown (Montano et al., 2007; Valayannopoulos et al., 2010). Other known mutations in sulfatases also cause human diseases, such as ARSA (Metachromatic leukodystrophy), IDS (Hunter syndrome- MPSII), ARSE (Chondrodysplasia Punctuata), and STS/ARSC (X-linked ichthyosis) (Diez-Roux & Ballabio, 2005). Given the relevance of sulfatase mutations to human disease, the biological substrates of the six remaining orphan sulfatases need to be revealed.

Regulation of cartilage PG sulfation levels is critical for proper endochondral ossification. Mice with mutations in either *Papss2* (regulates the sulfate donor within the cell), *Slc26a2* (transports sulfate within the cell), or *Sumf1* (activates sulfatases) demonstrate that inadequate sulfation of PGs can cause a variety of skeletal defects (Cortes et al., 2009; Gualeni et al., 2013; Settembre et al., 2007; Settembre et al., 2008). For example, these mutant mice can have reduced chondrocyte proliferation, severe growth retardation, and delayed secondary ossification center formation. How PG sulfation levels might be modulated during endochondral ossification is still to be unveiled, but previous studies suggest that PG sulfation might decrease during cartilage maturation (Farquharson et al., 1994; Wuthier, 1969). If true, then sulfatases that are specifically expressed during cartilage maturation should be identified.

In this study, we hypothesized that Arylsulfatase I (ARSI), an sulfatase with a previously unknown biochemical function, decreases PG sulfation during cartilage maturation. Collectively, our findings showed that expression of ARSI, but not ARSB nor GALNS, increases as sulfate esters decrease in developing mature cartilage. *ARSI* overexpression *in vitro* revealed that ARSI is a novel PG sulfatase, opening the pathway

for understanding how biochemical modification of PGs affects endochondral ossification.

3.3 Methods

3.3.1 Embryo collection and tissue processing

All animal experiments were approved by the University of Saskatchewan's Animal Research Ethics Board and adhered to the Canadian Council on Animal Care guidelines for humane animal use. White leghorn chicken eggs were incubated in a humidified, rocking incubator at 37 °C until they reached Hamburger-Hamilton stage 36. Chick humeri were harvested, washed in 1x phosphate-buffered saline (PBS), embedded in optimal cutting temperature (OCT) compound (Tissue Tek, Torrance, CA, United States), and immediately flash-frozen in liquid nitrogen-cooled isopentane for all the synchrotron-related analyses and laser capture microdissection. For all other analyses, humeri were fixed in 4% paraformaldehyde (PFA) overnight before embedding in OCT. Mouse embryos were harvested at stage E14.5 and their humeri were harvested and prepared as described for chick samples.

3.3.2 Histology

Safranin O/Fast Green staining was performed on 7 μm thick chick humerus frozen sections. Cryosectioned tissues were dehydrated in 70% EtOH for 2 min, placed in Wiegert’s hematoxylin for 5 min, washed with tap water for 3 mins, and stained with 0.02% fast green (CI 42053) for 30 secs. The tissues were then de-stained in 1 % acetic acid for 30 sec. Sections were then stained with 0.1 % safranin O stain for 40 min, then dehydrated with 95% EtOH and 100% EtOH for 12 dips each, and xylene for 5 mins before mounting the slide for imaging.

A modified Safranin O protocol, that allowed the visualization of less sulfated PGs in mature cartilage, was utilized for the chick sample in Figure 3.2. In this case, cryosectioned tissues were re-hydrated in water for 1 min before staining. Sections were placed in Wiegert's hematoxylin for 5 min, rinsed in running tap water for 1 min, and then counter-stained in fast green FCF (0.05 %) for 1 min. The tissues were then de-stained in 1 % acetic acid for 10 dips. Sections were then left in 0.1 % safranin O stain for 20 min and then dehydrated and mounted for imaging.

Trichrome staining was performed on 7 µm thick frozen sections of HH36 chick humeri as described elsewhere, except for the deparaffinization steps (Ashique et al., 2022). Whole chick humeri were subjected to Alizarin red and Alcian blue whole-mount staining, performed as previously described (Eames et al., 2011).

3.3.3 Slides immunohistochemistry (IHC)

While slides from unfixed animals were dried at 60°C for 30 minutes and fixed in 4% PFA overnight at 4°C, fixed sections were air-dried for 10 min and post-fixed in 4%PFA at room temperature (RT) for 20 min. After that, all samples were treated equally. Slides were rinsed twice in PBST (PBS/0.5% triton-100), and sections were digested with trypsin (1mg/mL in EDTA/PBS) for 30 min, rinsed twice in PBST and digested with hyaluronidase (5mg/mL in PBST) for 30 min. Sections were then incubated with a blocking solution (4% sheep serum (SS) and 2% goat serum (GS) in PBST) for at least 1 hour at room temperature, followed by incubation (overnight at 4°C) with primary antibodies: anti-COL10 (DSHB - 1:100), anti-COL2 (DSHB - 1:100), anti-ARSI (Sigma - 1:300), anti-GALNS (Abcam - 1:100) and anti-ARSB (Abcam - 1:300). The slides were then washed twice with PBST for 5 minutes and incubated (overnight at 4°C) with specific anti-mouse or anti-rabbit (1:1000) secondary antibodies. Slides were rinsed twice in PBST and the cell nuclei were stained with Hoechst 33258 for 10 minutes. Slides were stored in PBST at 4˚C until the time images were taken.

3.3.4 *Synchrotron analyses.*

All the synchrotron analyses described below were performed on 10 µm-thick chick humerus sections. Sections were melted onto thermanox plastic for x-ray fluorescence (XRF), XANES (X-Ray Absorption Near-Edge Spectroscopy), and chemically-specific XRF imaging or into an Infrared-grade CaF_2 window (Crystran) for Fourier transform infrared (FTIR) analysis. Adjacent sections were collected onto Superfrost Plus™ glass slides (Fischer Scientific) for histological and IHC analyses.

XRF data collection at the Canadian Light Source (CLS) VESPERS beamline.

XRF imaging of chick humeri was performed as previously described, but with a larger step-size (Hackett et al., 2016). Briefly, ten chick humeri were raster-scanned at VESPERS beamline with a 15 µm step size using a of 4 × 6 µm (horizontal × vertical) sized pink beam (15 keV) with a 1-second dwell time per pixel, to generate XRF maps.

XANES at the CLS SXRMB Beamline.

Sulfur K-edge XANES was carried out at the Soft X-ray Microcharacterization Beamline (SXRMB) using chick samples, similarly to previously described for zebrafish samples (Hackett et al., 2016). XAS spectra were calibrated against the spectrum of an anhydrous $Na_2S_2O_3$ powder solid standard, with the lowest energy peak thought to be 2469.2 eV. Spectra were collected across the energy range of 2450–2515 eV, with a total collection time of approximately 10 min, using the Aquaman software. An average spectrum was calculated for each sample from three replicate measurements. Spot scans were collected for at least 3 spots in the immature cartilage region and at least 3 spots in the mature region. Spectra of calibration standards and model compounds were recorded at room temperature. Powdered solid standard compounds used for fitting routines represented disulfides (oxidized glutathione), thiols (reduced glutathione), thio-ethers (methionine), sulfoxides (methionine sulfoxide), sulfinic acids (hypotaurine), sulfonic acids (taurine), O-linked sulfate esters (*N*-acetyl D-galactosamine 6-sulfate), N-linked sulfate esters (D-galactosamine 2-sulfate), and inorganic sulfates (Na_2SO_4). Spectra were collected from four tissue sections from different animals. XANES spectra

were processed using the EXAFSPAK suite of programs (http://www-ssrl.slac.stanford.edu/exafspak.html). Using the DATFIT program, spectra from tissue sections were fitted with a linear combination of normalized reference spectra. Standards were excluded from the refinement algorithm if they contributed to <1% of the total spectra.

Chemically-specific XRF imaging at the CLS SXRMB Beamline.

Chemical-specific XRF imaging was performed to assess the amount and distribution of distinctive sulfur chemical forms similar to previously described (Hackett et al., 2016). SXRMB relies on a bending magnet source and operates in the 1-10 keV range with an energy resolution $\Delta E/E$ of 1.0×10^{-4}. Thermanox™ slides with humerus sections were mounted at a 45° angle relative to the incoming in a vacuum chamber and raster scanned. Spectra were collected at room temperature across the energy range of 2450−2515 eV. Each humerus was raster-scanned with a step size between 15-20 μm with a 1-second dwell time per pixel. XRF maps at different energies in the sulfur K-edge region and one map of total sulfur were created and processed as previously described (Hackett et al., 2016). Briefly, 4 maps from each sample were collected: a background map to account for both non-sulfur and lower-excitation energy sulfur (thiols, disulfides, thio-ethers, and sulfoxides), one map at the sulfonic acid peak, one map at the sulfate ester peak, and one total sulfur map collected above the sulfur edge. The exact energies at which the 4 sulfur maps were collected were calibrated to the lowest energy peak of anhydrous sodium thiosulfate, thought to be 2469.2 eV.

XRF data processing.

Maps collected at VESPERS and SXRMB were opened in Sam's microprobe analysis toolkit (SMAK) version 1.40 (https://www.sams-xrays.com/smak) and normalized to the pre-sample ion chamber channel. All maps were then imported into the same file and the background channel was subtracted from the sulfate esters and sulfonic acids. The XANES fitting sub-program was used to create the sulfate esters and sulfonic acids maps using the peak ratio parameters determined by XANES measurements of pure standard compounds (Hackett et al., 2016). All XRF files were

then converted into a gray-scale text image using the program smak2txt created by Graham George (University of Saskatchewan), allowing them to be opened in ImageJ (https://imagej.nih.gov/ij/). The chemically specific maps for total sulfur and sulfonic acids and their adjacent sections stained with COL10 were opened in Image J and saved as 32-bit images. ROIs were drawn surrounding the COL10 expression domain in different samples and were pasted on the respective chemically specific maps. Average pixel intensity, total pixel intensity, and standard deviation were measured in the immature and mature cartilage regions and collected for statistical analysis.

Fourier transform infrared (FTIR) spectroscopic imaging at the Mid-IR beamline at the CLS.

Infrared data were collected with a Hyperion 3000 FTIR imaging system (Bruker Optics, Ettlingen, Germany) fitted with a mercury-cadmium-telluride (MCT) 64×64 focal plane array (FPA) detector with an upper objective of 15x magnification and a numerical aperture of 0.4, combined with a lower condenser of 15x magnification and 0.4 numerical aperture. This arrangement yielded a pixel size of 21.4 μm^2, using 8x8 binning. All scans were performed in transmission mode. Samples were mounted on IR-grade CaF_2 windows (Crystran Ltd.). A focal plane array detect was illuminated with a Globar source (aperture 3.5 mm, part of the Hyperion 3000 FTIR system used). FTIR-spectroscopic images were collected across the spectral region 3600 – 900 cm^{-1}, on a liquid nitrogen-cooled FPA detector with a spectral resolution of 4 cm^{-1} with the co-addition of 128 scans, with a background image similarly collected from a blank substrate using 128 co-added scans. The background was collected immediately before each sample. In total, 8 humerus samples were measured at this beamline. All data processing and image generation was performed using Cytospec software (Cytospec, Version 1.2.04) and Opus software (Version 6.5, Bruker, Ettlingen, Germany).

3.3.5 Laser capture icrodissection (LCM)

LCM was performed in mouse humeri samples at stage E14.5 as previously described for chick embryos (Gomez-Picos et al., 2022). Immature and mature chondrocytes were captured from unfixed 10 μm frozen sections. The capture of each tissue was done in triplicate from three distinct embryos.

3.3.6 RNA isolation and amplification.

RNA from immature cartilage (*n* = 3) and mature cartilage (*n* = 3) was extracted in triplicate from three unique tissue samples that were obtained from three different embryos, and thus a total of 6 cDNA libraries were constructed. RNA isolation, amplification and library preparation were performed as previously described (Gomez-Picos et al., 2022).

3.3.7 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 from 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 log2 expression across three replicates of each tissue produced three bimodal distributions, which were used to set the count thresholds to 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 log2 fold change greater than 2 ($p<0.01$).

3.3.7 ARSI sequence alignments.

The full-length ARSI protein sequences were obtained from the NCBI website for the following species: *Gallus gallus* (XP_004945003.1), *Homo sapiens* (NP_001012301.1), *Mus musculus* (NP_001033588.1), and *Danio rerio* arsia (XP_002664306.1) and arsib (XP_692237.2). Using the "Identify Conserved Domains" tool on the NCBI website, the sequences of the sulfatase domain of the ARSI protein from those same species were obtained. Full-length ARSI protein sequences and sulfatase domain sequences were aligned using the Clustal Omega (Multiple sequence alignment tool) and comparison matrices were generated.

3.3.8 RNA in situ hybridization (ISH).

Probe making and preparation

ISH was performed on chick HH36 humeri 7 µm frozen sections. Some RNA probes were obtained from collaborators: *COL2A1 and SPP1* from Richard Schneider, and *COL10A1* and *IHH* from Ralph Marcucio. *ARSB* and *GALNS* probes were previously designed in Eames lab by Devin Brown. Genes were cloned from PCR of chick cDNA. Three different *ARSI* probes were generated targeting just the coding sequence (CDS) region, just the three prime untranslated region (3'UTR), and the end of CDS plus the 3'UTR region of the *ARSI* gene. Specific primers were designed to amplify each region, PCRs were performed, and *ARSI* DNA fragments were purified and cloned into a pCR™4-TOPO® vector (Thermofisher). Cloned DNA fragments were transformed into E.coli cells (OneShot, Invitrogen), selected for with ampicillin and cultured overnight in LB liquid medium. The plasmid DNA was extracted and purified to produce minipreps. These minipreps were then sequenced and compared to RNA

sequences using the Basic Local Alignment Search Tool (BLAST) on NCBI. This comparison was conducted to verify that the cloned DNA sequences were accurate. Plasmids were linearized accordingly to produce the proper sense and anti-sense probes, transcribed using T7 or T3 RNA polymerases, and labelled with digoxigenin. In addition, an *ARSI* anti-sense probe was also hydrolyzed following a Cold Spring Harbor Protocol to improve probe access through the cellular membrane (Ferrandiz & Sessions, 2008). Hydrolysis was done by incubating the probe with twice the volume of RNase-free hydrolysis buffer (40 mM NaHCO3/60 nM Na2CO3) at 58 °C for about 10 mins, followed by re-precipitation and dissolution with 0.1% DEPC.

In-situ hybridization on frozen sections

ISH was performed in 7 µm frozen section of chick humeri. On day 1 the slides were dried at 58°C for 40 min, fixed in 4% PFA/1X PBS/0.1% DEPC for 20min, rinsed in 1xPBS/0.1% DEPC for 5min, incubated in 0.2 N HCl/0.1% DEPC at RT for 10 mins, rinsed twice in 1X PBS/0.1% DEPC for 5 mins each, permeabilized with proteinase K/1X PBS/0.1% DEPC (3 ug/ml) at 37 °C for 15 mins, rinsed in 1X PBS/0.1% DEPC for 5 mins, post-fixed in 4% PFA/0.1% DEPC for 15 mins, rinsed in 1X PBS/0.1% DEPC for 2 mins, and pre-hybridized in hybridization buffer for at least 3 hrs at 58 °C. Sections were then incubated with different RNA probes ON at 58°C. On day 2 slides were washed at 58°C with washing solution one time for 15 min and 2 times for 30 min, followed by 2 washes with 1XMABT for 30 min each. A blocking solution (1XMABT /2%blocking reagent/20% heat inactivated sheep serum) was used on the slides for 2-3 hours at RT to reduce non-specific antibody binding. Slides were then incubated ON at 4 °C with blocking solution (containing active sheep serum) with a 1:1000 dilution of anti-DIG alkaline phosphatase antibody (Roche, Sigma-Aldrich). On day 3 slides were washed with 1XMABT 4 times of 30 min at RT, washed with alkaline phosphatase staining buffer 2 times of 10 min at RT, and put in a humidified chamber protected from light with BM Purple (Roche, Sigma-Aldrich) covering the tissues until the signal was

visible. Slides were then washed with 1X PBS, three times for 5 mins, dehydrated, and mounted for imaging.

3.3.9 ATDC5 Cell culture.

Cell experiments were performed using the chondrogenic mouse cell line ATDC5 (Yao & Wang, 2013). ATDC5 cells were cultured in a humidified atmosphere (37℃, 5% CO_2) in DMEM/F12 with 5% FBS until they reached 80% confluency. Cells were then trypsinized, counted in a hemocytometer, and spun down (1000 rpm for 5 mins). One 10 µl dot containing $2.5x10^5$ cells was plated in the center of each well in 24-well plates. Plates were incubated for 90 min to allow for cells to attach to the wells. Finally, 1 mL of differentiation medium (DMEM/F12, 5% FBS, 1% insulin-transferrin-selenium, 50ug/ml ascorbate-2-phosphate, and 10mM β-Glycerophosphate) was gently added to each well, and it was changed every other day until micromasses were harvested.

3.3.10 Alcian blue staining on micromass cultures.

Micromasses were stained with Alcian Blue at 7, 14, and 21 days to examine the secretion of sulfated proteoglycans as previously described (Izadifar et al., 2016) and were imaged using light microscopy.

3.3.11 Micromasses IHC.

ATDC5 micromasses were fixed with 4% PFA rocking overnight at 4˚C and later washed 3 times (5 min) in PBS with gentle rocking. Samples were then digested with trypsin (1mg/mL in EDTA/PBS) for 45 min (7 days), 1 hour (14 days), or 1.5 hours (21 days), washed 3 times (5 min) with PBST and digested with hyaluronidase (5mg/mL in PBST) following the same times used for trypsin digestion. Micromasses were then

washed 3 times (5 min) with PBST and incubated with a blocking solution (1XPBS, 0.5% TritonX-100, 4% GS 2% SS; 1:100) overnight at 4°C. Samples were covered with a primary COL10 (1:100) or ARSI antibody solution (1:300) and incubated overnight at 4˚C with shaking. Micromasses were washed 6-8X in blocking solution over 1-2 hours and incubated with an anti-rabbit secondary antibody (1:1000) overnight. Samples were then washed with PBST 6 times over 1-2 hours and stored in PSBT at 4˚C until the time images were taken.

3.3.12 Micromasses Western Blot.

For western blot analysis, cells were washed three times with PBS and then dissociated in laemmli buffer after the addition of proteinase inhibitor (Thermo Fisher). Lysates were centrifuged at 12,000 rpm for 10 min at 4 °C and the supernatant was used for the experiments. Protein extracts were boiled for 5 minutes at 95^{o}C before being loaded onto 10% SDS-PAGE gels and transferred onto nitrocellulose membranes. After blocking with 5% skim milk in PBST (PBS/0.05%Tween-20), the membranes were incubated with anti-ARSI antibody (Sigma) at 4°C with gentle shaking overnight. Next, the membranes were washed 3 times with 5% skim milk in PBST while shaking at room temperature and incubated with a secondary antibody for 1 h. Membranes were then washed 3 times with 5% skim milk in PBST, 1 time with PBST, and 1 time with PBS. Finally, the chemiluminescent signals on the membranes were detected using an ECL reagent (Thermo Fisher) and a ChemiDoc MP Imaging System (Biorad).

3.3.13 Cloning of the human ARSI cDNA.

ARSI cDNA was obtained by reverse transcription using human placenta cDNA as a template and cloned into pSP72-EF1-IRES-Blast vector, which biocistronically encodes a blasticidin-resistant gene under the control of an internal ribosomal entry site.

3.3.14 Establishment of ARSI-expressing cells.

The pSP72-EF1-hARSI-IRES-Blast plasmid containing hARSI under control of Human elongation factor-1 alpha (EF1a) promoter, was linearized with ClaI and then transfected into HeLa cells using FuGENE6 (Promega). Transfected cells were selected with blasticidin and multiple clones stably expressing *ARSI* were established. To obtain negative control cells, pSP72-EF1-IRES-Blast empty vector was transfected into HeLa cells.

Disclaimer about HeLa cells: The HeLa cells were taken from Henrietta Lacks, an African-American woman, without her knowledge or consent in 1951 and were used in scientific research without her or her family's consent. This was legal at the time, but it is now recognized that the use of human tissue for research requires informed consent. The lack of proper consent can raise ethical and legal concerns such as privacy and control over one's own body. The Lacks family did not receive any compensation for the use of her cells until many years later, despite the significant contributions that HeLa cells have made to scientific advancement.

3.3.15 Reverse transcription real-time PCR for Hela cells

Total RNA was extracted using High Pure RNA Isolation Kit (Roche) and reverse transcribed with MMLV reverse transcriptase. Real-time PCR was performed using FastStart Essential DNA Green Master (Roche) and LightCycler 96 System (Roche). Primers used for real-time PCR were as follows: hARSI-qPCR-1-S, 5'-CAAGGGGGTCAAGTTGGAGA-3'; hARSI-qPCR-1-AS, 5'-CAGCTTCTGTGGCAGTGTCA-3' (*ARSI*); hC4ST1-S1, 5'-AAACGCCAGCGGAAGAA-3'; hC4ST1-A1, 5'-GGGATGGCAGAGTGAGTAGA-3' (C4ST-1); hC4ST2-S1, 5'-GAGGGAAAGTTCTTTGTTTAAGTG-3'; hC4ST2-A1, 5'-CGGCCTTAACAGCCATAAT-3' (*C4ST-2*); hC6ST1-S1, 5'-TCTGCCATTGGCTTGAAC-3'; hC6ST1-A1, 5'-CATGCAGACATGAAATAGCAAAC-3' (*C6ST-1*); hGlcAT-I-RT-S2, 5'-

CCTGCCTACTATCTATGTTGTTAC; hGlcAT-I-RT-A2, 5'-ACCACCAGGTGTGTGAA-3' (*GlcAT-I*); hGAPDH-S2, 5'-ATGGGTGTGAACCATGAGAAGTA-3'; hGAPDH-A2, 5'-GGCAGTGATGGCATGGAC-3' (*GAPDH*).

3.3.16 Analysis of chondroitin sulfate and heparan sulfate disaccharide composition.

HeLa cells were washed with PBS, collected using a cell scraper and homogenized in cold acetone. The dried powder was digested with actinase E (Kaken Pharmaceutical) for 2 days at 55˚C. Trichloroacetic acid was added to the digested samples (final 5%) and the soluble fraction was obtained. After extraction with diethyl ether, the aqueous phase was subjected to ethanol precipitation. The precipitate was dissolved in a pyridine-acetate buffer and gel-filtrated on a PD-MiniTrap G-25 column (GE Healthcare). The flow-through fraction was collected, dried up and dissolved in MilliQ water. An aliquot of the samples was digested with chondroitinase (ChABC and AC-II, Seikagaku) for chondroitin sulfate analyses and another aliquot was digested with heparinase and heparitinase for heparan sulfate analyses. Digestions proceeded for 2 h at 37˚C and samples were derivatized with 2-aminobenzamide for 2 h at 65˚C. The derivatized samples were analyzed by high-performance liquid chromatography (CBM-20A, Shimadzu) on a YMC-Pack PA-G column (YMC). Identification and quantification of the disaccharides were achieved by comparison with chondroitin sulfate and heparan sulfate disaccharide standards as described previously (Kitagawa et al., 1995).

3.3.17 Statistical analysis.

Different statistical methods were used depending on the experiments. One sample T-test were used to compare data from immature and mature cartilage of chick humeri obtained by XRF, chemically specific XRF, and FTIR. Unpaired T-test were used to compare CS and HS levels in *ARSI* transfected and non-transfected HeLa cells. Lastly, two-way ANOVA was used to compare the amount of different disaccharides in

ARSI transfected and non-transfected HeLa cells. The results are reported as mean ± SEM. $P < 0.05$ was considered statistically significant. P values and statistical tests used are indicated in each figure legend.

3.4 Results

3.4.1 Embryonic chick growth plate characterization

Embryonic chick humeri at HH 36 - approximately 10.5 days of development - were used to study the relationship between the maturation state of cartilage and sulfate levels in the ECM (Hamburger & Hamilton, 1951) (Figure 3.1 A, B). To characterize the embryonic chick growth plate at this stage Safranin O and trichrome staining were performed. Safranin O, which binds to sulfated PGs, was observed in the whole humerus and allowed for the visualization of chondrocytes in different stages of maturation (Figure 3.1C). Resting chondrocytes were small and rounded and were localized closer to the humerus epiphyses (Figure 3.1D). Those were followed by proliferative chondrocytes, which were flatter and organized in a more columnar manner (Figure 3.1E). Prehyperthophic chondrocytes, bigger in size, were observed towards the bone diaphysis and were followed by larger hypertrophic chondrocytes in the center of the developing bone (Figure 3.1F, G). Trichrome staining showed that, as expected, perichondral bone (in Aniline blue) was located surrounding the mature cartilage region and ceased in the proliferative chondrocyte region (Figure 3.1H and I). Higher magnification of perichondral bone showed tightly wound collagen fibres into the bony extracellular matrix (blue perichondral bone) and trapped osteocytes (Figure 3.1J). Based on these histological features, HH36 was a suitable stage to study cartilage maturation because immature, mature cartilage, and perichondral bone regions were readily observed, while there was still no vascular invasion of the cartilaginous template. Vascular invasion brings different cell types to the developing bone and leads to cartilage degradation and bone deposition. Therefore, it is preferable to study cartilage maturation before vascular invasion occurs.

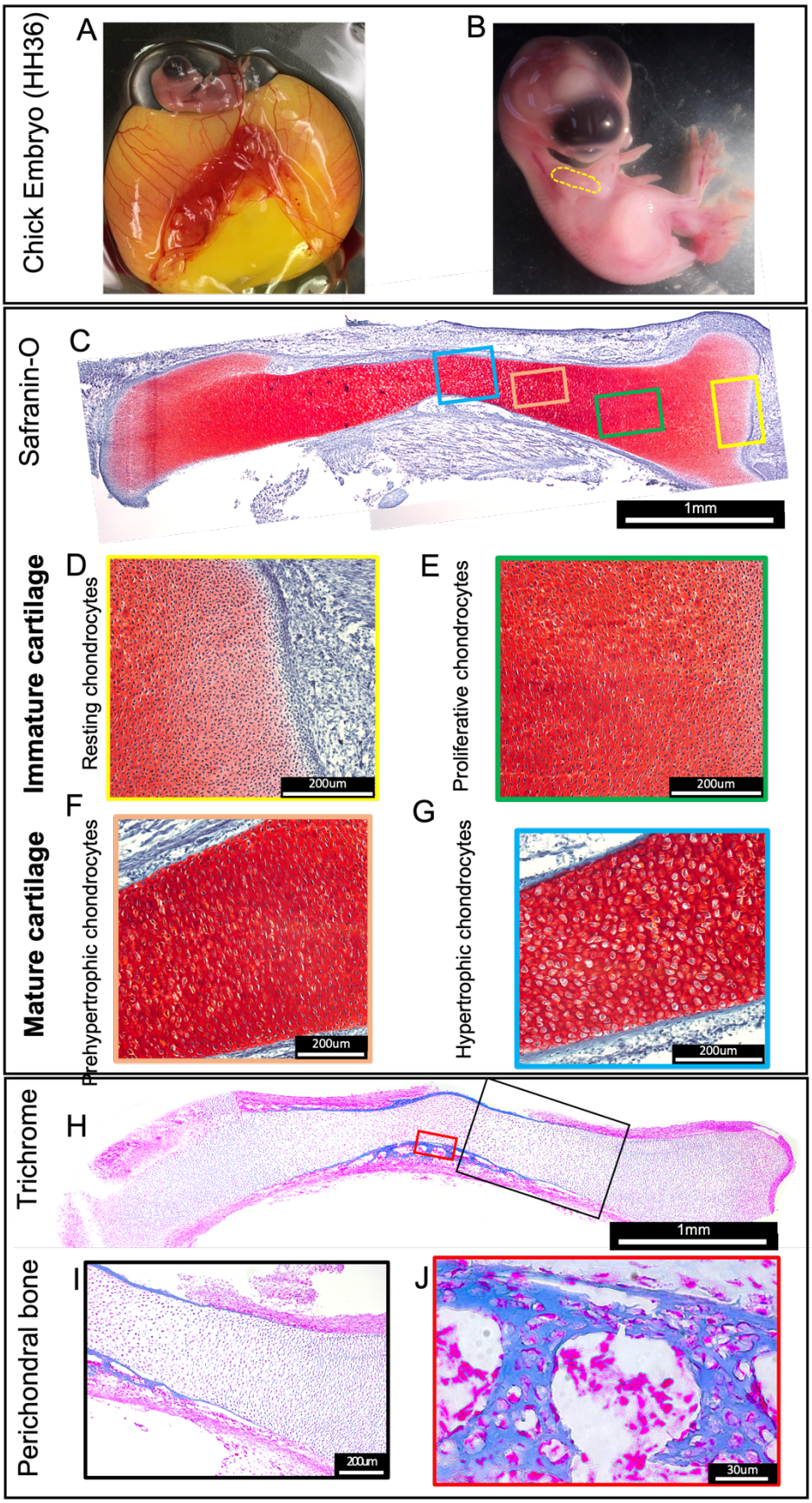

Chick Embryo (HH36)
A
B
Safranin-O
C
1mm
Immature cartilage
D
Resting chondrocytes
200um
E
Proliferative chondrocytes
200um
Mature cartilage
F
Prehypertrophic chondrocytes
200um
G
Hypertrophic chondrocytes
200um
Trichrome
H
1mm
Perichondral bone
I
200um
J
30um

Figure 3.1 - Histological characterization of the embryonic chick growth plate (HH36). (A) After 10.5 days at 37ºC chick embryos were removed from the eggs. (B) The chicken humeri (yellow circle) were dissected and embedded in oct. (C) 7 µm cryosections stained with Safranin-o show that at HH36 the chick humerus is comprised by MAT and IMM cartilage regions, and no vascular invasion occurred yet. (C-G) Higher magnification images show resting chondrocytes (in a yellow box) closer to the epiphyses, proliferative chondrocytes (green box), prehyperthophic chondrocytes (orange box), and hypertrophic chondrocytes (blue box). Resting and proliferative chondrocytes comprise the immature cartilage region, while prehyperthophic and hypertrophic chondrocytes comprise the mature cartilage region. (H) Trichrome staining shows perichondral bone in blue surrounding the mature cartilage region. (I) Higher magnification confirms that perichondral bone ceased in the proliferative chondrocyte region. (J) perichondral bone is formed by tightly wound collagen fibers (blue perichondral bone) and trapped osteocytes.

3.4.2 PG sulfation decreases during cartilage maturation

Previous studies suggested that PG sulfation might decrease during cartilage maturation (Farquharson et al., 1994; Hackett et al., 2016; Wuthier, 1969), so sulfur levels were quantitated in the developing chick humerus using synchrotron-based X-ray fluorescence (XRF) imaging. Histological and molecular analyses were used to identify mature cartilage regions. Alizarin red, which binds mineralized tissue, stained perichondral bone in the mid-diaphysis of the HH36 chick humerus (Figure 3.2A). Staining levels of Alcian blue, which binds sulfated PGs, appeared reduced in the cartilage that is directly underlying the perichondral bone (Figure 3.2B). Staining levels of Safranin O, which also binds sulfated PGs, on histological sections of the HH36 chick humerus also were decreased in this cartilage region (after 20 min staining), which was confirmed as mature cartilage by the presence of hypertrophic chondrocytes and expression of COL10 (Figure 3.2C,D) (Kong et al., 1993).

Using COL10 IHC on adjacent sections to define mature and nearby immature cartilage regions (dotted lines in Figure 3.2D), XRF imaging of the HH36 chick humerus revealed each region's relative sulfur levels (N=9; Figure 3.2E). Qualitatively, COL10 immunostained regions mostly fit well within a low sulfur region (Figure 3.2E). The

decrease of sulfur in mature cartilage was visible in all samples. These qualitative differences served as an important indicator of the biologically significance of sulfur decrease in mature cartilage. Quantitation of these images verified a significant decrease (4.5 ± 1.2%) in the average total sulfur content in mature cartilage compared to immature cartilage (Figure 3.2F). The proliferative zone in the epiphyses of developing cartilage also appeared to have lower levels of total sulfur, although this was not quantitated.

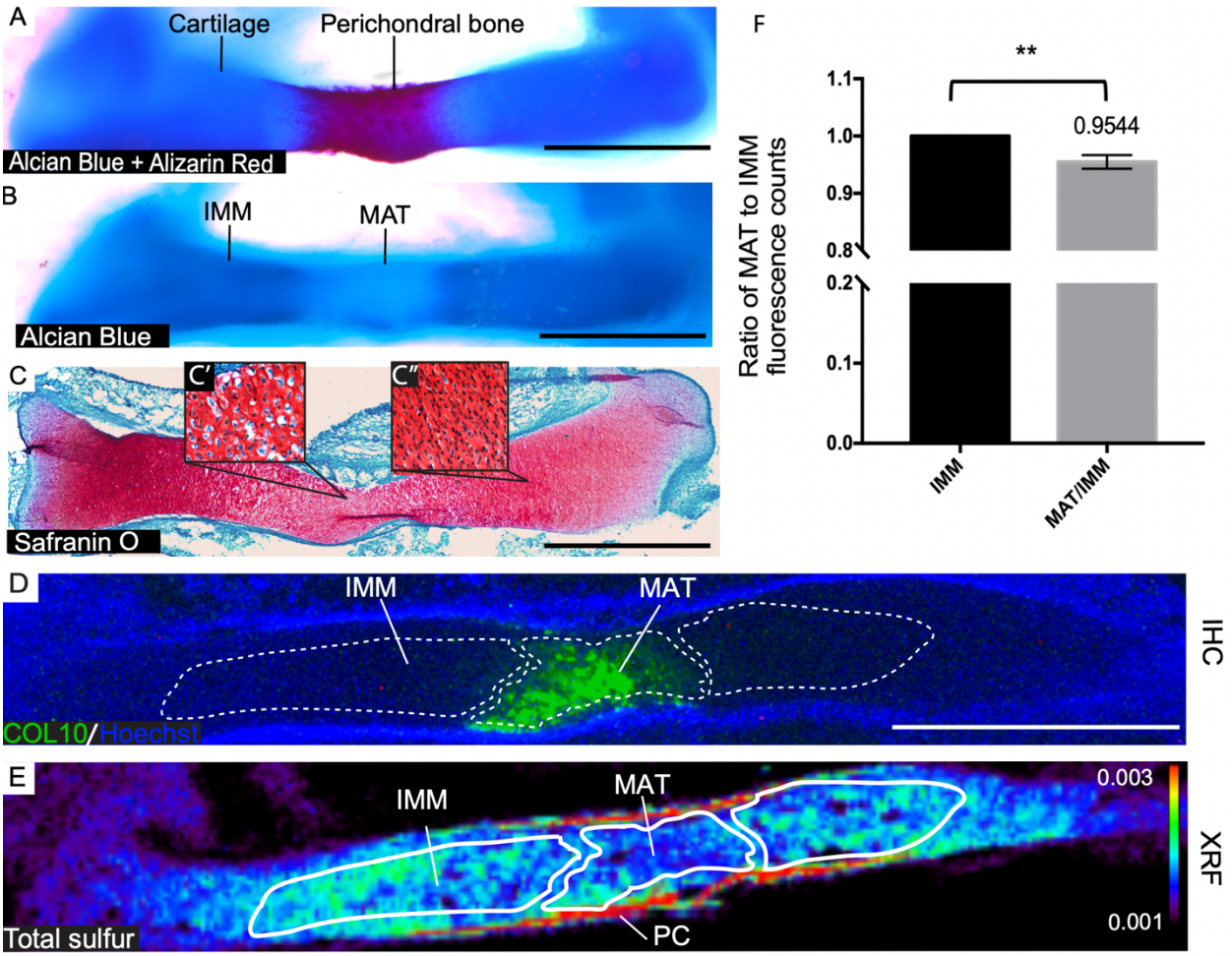
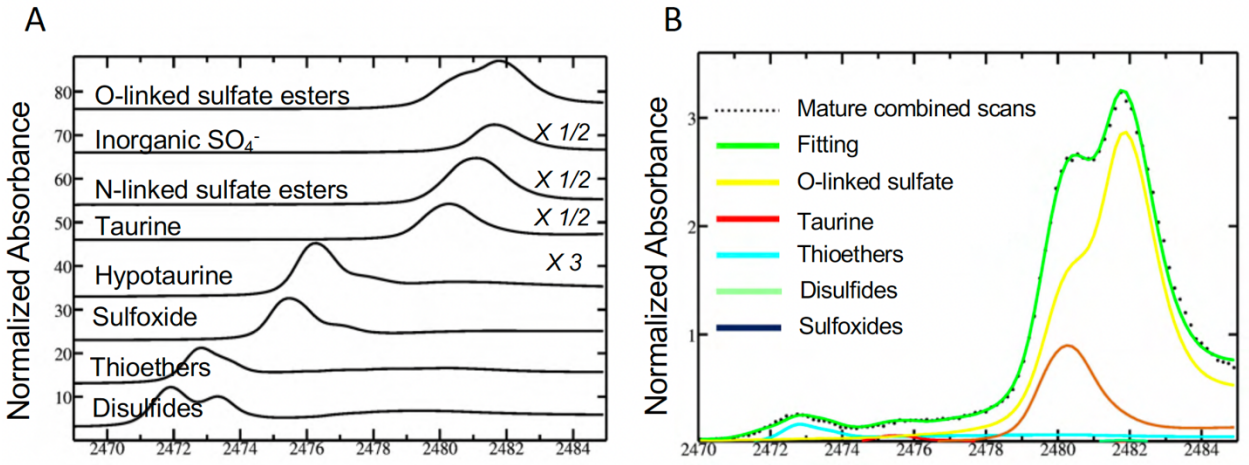


Figure 3.2 PG sulfation decreases during cartilage maturation in chick humerus (HH36). (A) Alcian blue and Alizarin red staining indicate cartilage and perichondral bone regions, respectively. (B) Alcian blue, which binds to sulphated PGs, is decreased in the region of MAT underlying the perichondral bone. (C) Safranin O staining shows hypertrophic chondrocytes (C') in the MAT, and non-hypertrophic chondrocytes (C'') in the IMM. (D) COL10 expression was used to define MAT and IMM regions for XRF image analysis (white dashed line). (E) XRF imaging of adjacent sections depicts total sulfur levels in MAT and IMM regions. (F) Quantitation of XRF images shows a decrease of 4.5 ± 1.2% in the average total sulfur in MAT relative to immature cartilage (N = 9). Error bars represent the standard error of the mean. Statistics were calculated using one sample T-test, ** p=0.005. IMM: immature cartilage; MAT: mature cartilage; PC: perichondrium; XRF: X-ray fluorescence; IHC: immunohistochemistry. Scale bar: 1mm.

Different chemical forms of sulfur are present in biological tissues like cartilage, but PG sulfation is reflected specifically by sulfate esters (Hackett et al., 2016; Mikami & Kitagawa, 2013). To investigate relative levels of specific chemical forms of sulfur in developing immature and mature cartilage, X-ray absorption near edge structure

(XANES) analysis was performed on HH36 chick humerus sections (Figure 3.3). O-linked sulfate esters were the most abundant (77.3% ± 2.1% of sulfur in immature cartilage and 69.5% ± 10.7% in mature cartilage), followed by sulfonic acids, thioethers, disulfides, and sulfoxides (Figure 3.3C). The high O-linked sulfate esters levels support the idea that the majority of sulfur signal in developing chick cartilage was from proteoglycans. N-linked sulfate esters were practically negligible in developing cartilage and were not included in the final XANES fitting, as reported previously (Hackett et al., 2016).

Figure 3.3 Analysis of XANES spectra fittings with standard sulfur compounds has revealed the five



C

	Immature	Mature
O-linked sulfate esters	77.3% ± 2.1%	69.5% ± 10.7%
Sulfonic acids (taurine)	13.6% ± 2.0%	22.0% ± 10.5%
Thioethers	7.1% ±1.0%	5.5% ± 1.5%
Disulfides	1.1% ±0.1%	2.0% ± 0.9%
Sulfoxides	0.9% ±0.2%	1.0% ± 0.4%

major forms of sulfur present in chick cartilage (HH36). (A) Normalized standard absorption curves for the different forms of sulfate in biology. (B) Standard curve fittings for combined averaged XANES spot scans of the five more abundant sulfur chemical forms in mature cartilage. (C) Contributions of the five major forms of sulfur present in mature and immature cartilage are listed in the table. The relative abundances of the five forms of sulfur given in table C are very similar between mature and immature cartilage (N = 4). Abundances were reported in percentage +/- standard error of the mean. Values were calculated from an averaged spectrum of 3-4 spots per region per sample. The number of spots analyzed was not sufficient for a statistical comparison to be made between mature and immature cartilage regions.

To visualize *in situ* the distribution of different chemical forms of sulfur in developing cartilage, chemically specific XRF maps were generated (N=5; Figure 3.4). Again, COL10 IHC on adjacent sections was used to define the mature and nearby immature cartilage regions (dotted lines in Figure 3.4A). Qualitatively, the relative intensities of total sulfur and sulfate esters had similar distributions in the HH36 chick humerus, with lower levels apparent in the mature cartilage region (Figure 3.4B, C). The similarity in results between total sulfur and sulfate esters also supported that PG sulfation accounted for most of the sulfur in cartilage. By contrast, maps of sulfonic acid and lower oxidation states of sulfur did not clearly follow this distribution pattern across developing cartilage (Figure 3.4D, E). Quantitation of chemically specific XRF maps revealed that sulfate esters decreased significantly (13.4 ± 4.1%) in mature cartilage, compared to immature cartilage (Figure 3.4F). Total sulfur also tended to decrease (8.8 ± 4.2%) in mature cartilage, but not significantly in these samples (Figure 3.4F). Like the XANES analyses (Figure 3.3), sulfonic acids tended to increase slightly but not significantly in mature cartilage and showed a large variation in patterns between samples (25.4 ± 35%) (Figure 3.4F).

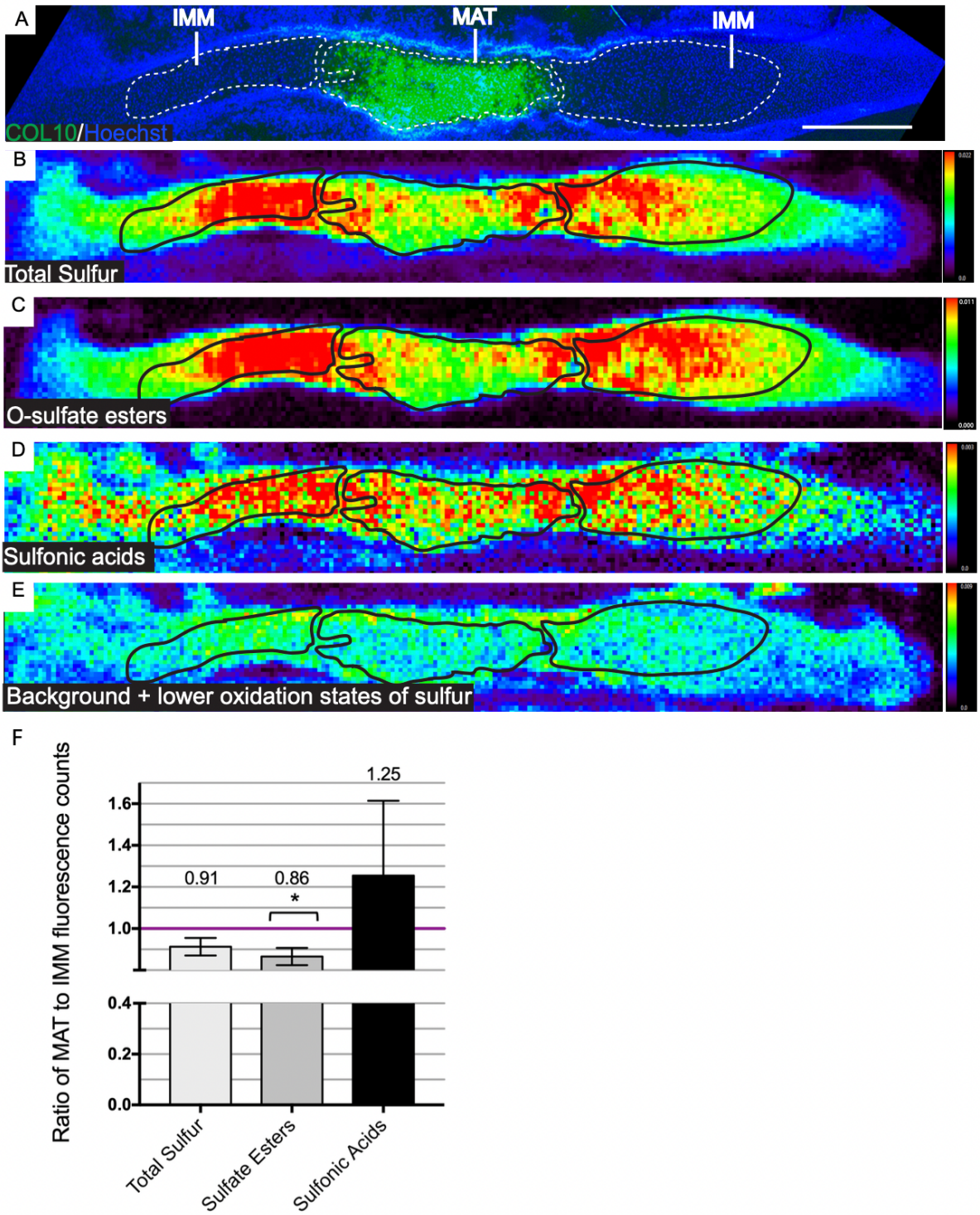


Figure 3.4. Chemically specific XRF maps show the *in situ* distribution of different chemical forms of sulfur in developing chick humerus (HH36). (A) COL10 IHC was used to define regions of interest for MAT versus IMM (dotted lines). The drawn regions of interest were analyzed in each map for quantification. (B-C) Total sulfur and O-sulfate esters have similar distributions in the chick humerus, with lower levels in the MAT. (D-E) Sulfonic acids and lower oxidation states of sulfur did not follow a clear distribution pattern across developing cartilage. (F) Quantitation of the chemically specific imaging data (N=5) shows that sulfate esters decreased 13.4 ± 4.1 % in MAT compared to IMM. Total sulfur and sulfonic acids did not change significantly between the cartilage regions. Error bars represent the standard error of the mean. Statistics were calculated using one sample T-test, * p = 0.003. IMM: immature cartilage; MAT: mature cartilage. Scale bar: 1mm.

Fourier transform infrared (FTIR) imaging of the HH36 chick humerus showed a trend towards PG and GAG decrease (4 ± 4 % and 3 ± 3 %, respectively) in mature cartilage, but not significantly (Figure 3.5), indicating that at least some of the sulfate esters in the mature matrix are diminishing independently of overall PG levels. Reduction in PG sulfation levels may be the first step driving PG degradation. In summary, O-linked sulfate esters, the main form of sulfur in developing cartilage, decrease during cartilage maturation. Given that sulfate ester levels reflect PG sulfation, PG sulfation decreases during cartilage maturation.

Three different infrared (IR) regions of interest were used to determine the relative abundance of total proteins and CS proteoglycans in mature versus immature cartilage on chick humerus. Protein was estimated using the amide I band (1590 – 1720 cm^{-1}) arising from C=O stretching of peptide bonds (Kim et al., 2005, Boskey and Pleshko Camacho, 2007). Two IR regions, 985 – 1140 cm^{-1} and the 1374 cm^{-1} second derivative band were chosen to give independent estimates of CS content based on previously validated studies (Kim et al.,2 005, Boskey and Pleshko Camacho, 2007, Saarakkala et al., 2010, Rieppo et al., 2012). Normalization of spectra was made by using total protein content to account for variations in thickness as described previously (Baker et al., 2014).

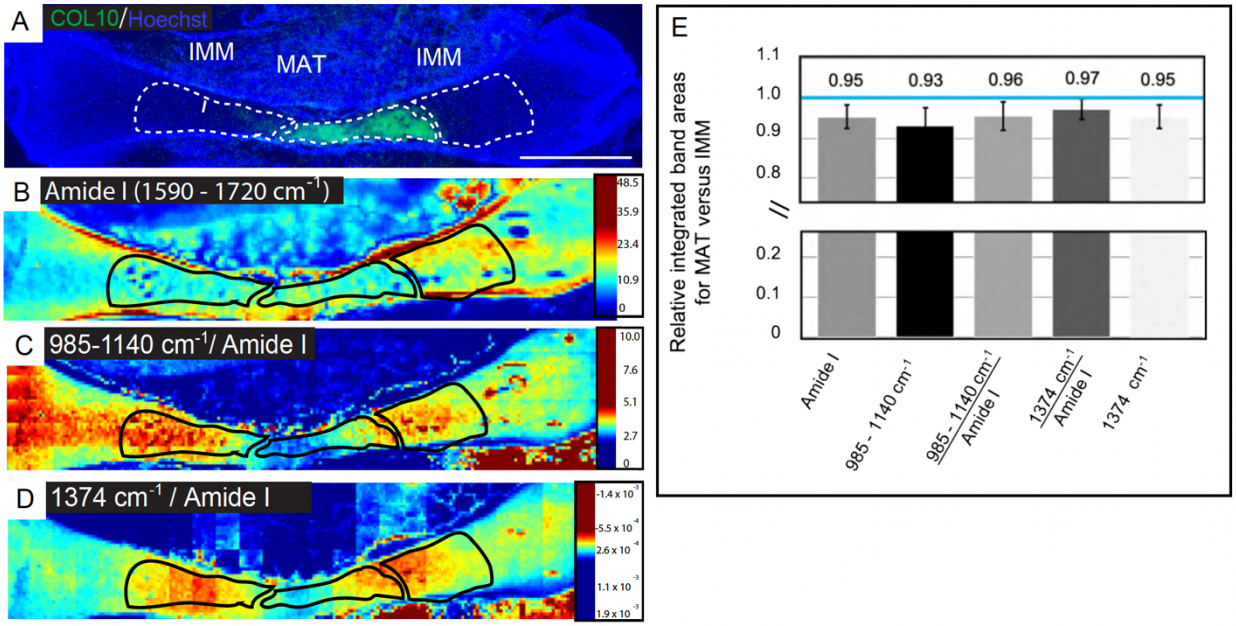


Figure 3.5 FTIR imaging shows no significant decrease in proteins or PGs in mature cartilage (A) IMM and MAT cartilage regions were assigned based on COL10 IHC of adjacent sections. (B) The integrated Amide I peak (1590 – 1720 cm^{-1}) in the FTIR map was used to approximate total protein levels. (C-D) Maps of PG representative band (985 – 1140 cm^{-1}) and GAG second derivative peak (1374 cm^{-1}), normalized to the Amide I band. (E) No bands showed significant differences between MAT and IMM regions, although all PG bands and the amide I band showed slight decreases in mature cartilage. N = 7. Error bars represent the standard error of the mean. Statistics were calculated using one sample T-test. IMM: immature cartilage; MAT: mature cartilage. Scale bar: 1mm.

3.4.3 ARSI, but not GALNS or ARSB, is expressed specifically in mature cartilage

To find a PG sulfatase that was expressed specifically in mature cartilage, laser capture microdissection of mature and immature cartilage in the developing E14.5 mouse humerus was combined with RNAseq (N=3 for each tissue; Figure 3.6). Over 9000 genes were expressed above threshold in each mature and immature cartilage. *Galns* and *Arsb* encode the only known animal sulfatases that specifically remove sulfate esters from chondroitin sulfate PGs, the main PG in cartilage matrix (Bhattacharyya et al., 2015; Tomatsu et al., 2005). However, neither *Galns* nor *Arsb*

transcripts were enriched in mature cartilage (Table 3.1). Genes encoding known heparan sulfate PG sulfatases, such as *Gns, Ids, Arsg, Sulf1,* and *Sulf2 (*Table 2.1), also were not enriched in mature cartilage (Table 3.1). Many sulfatases have been identified only from sequence analyses, including *Arsd, Arse, Arsf, Arsh, Arsi,* and *Arsj*, and since they have no biochemical function ascribed to them, they are orphan sulfatases (Hanson et al., 2004; Sardiello et al., 2005). Of these, *Arsi* was enriched 2.59 log_2FC in mature over immature cartilage (Table 3.2), supporting a previous report (Ratzka et al., 2010). The enrichment of *Col10a1* transcripts by 11.94 log_2FC in mature cartilage over immature cartilage verified the accuracy of these expression data (Table 3.2). In summary, comparative transcriptomic analyses between mature and immature cartilage identified *Arsi* as an orphan sulfatase whose expression might explain the decreased PG sulfation during cartilage maturation.

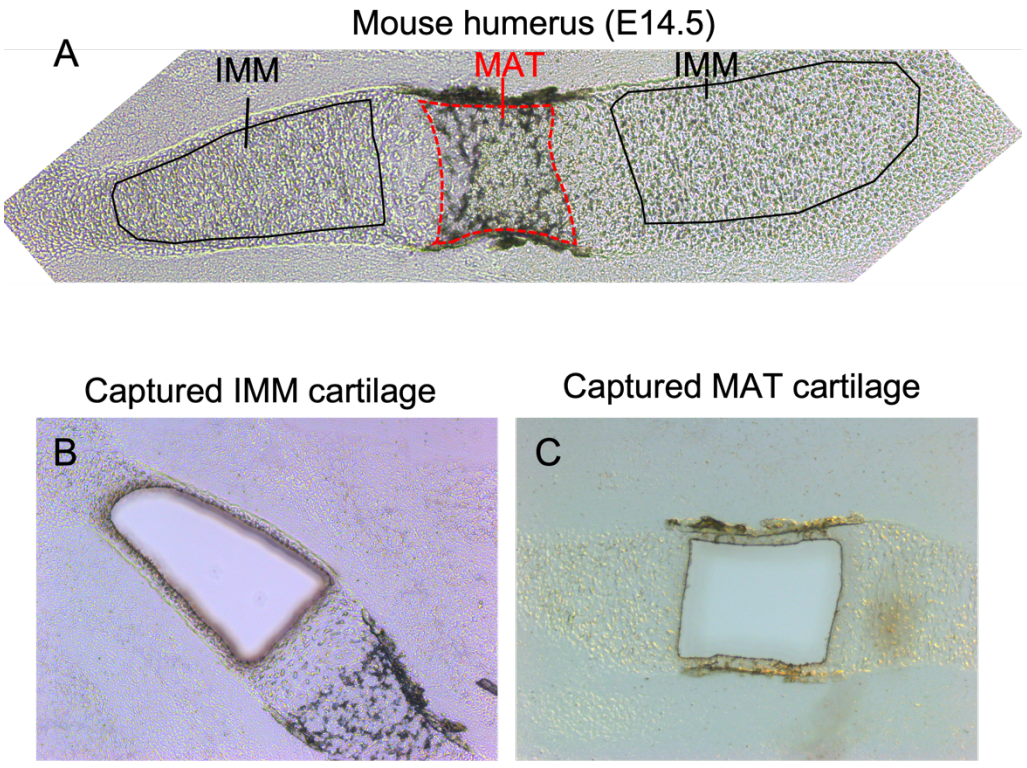


Figure 3.6 Laser-capture microdissection cartilage in the developing mouse humerus.

(A) Unstained mouse humerus (stage E14.5) before laser capture microdissection. (B) Mouse humerus after Laser-capture microdissection of immature cartilage. (C) Mouse humerus after Laser-capture microdissection of mature cartilage. Cells were captured and later submitted to RNA sequencing. (N=3 for each tissue)

Table 3.1 Gene counts in immature and mature cartilage regions of developing mouse humerus (E14.5)

	Immature Cartilage			Mature Cartilage		
	Sample 1	**Sample 2**	**Sample 3**	**Sample 1**	**Sample 2**	**Sample 3**
Galns	102	104	102	92	167	96
Arsb	5	45	26	9	7	12
Gns	366	505	348	552	496	242
Arsg	10	22	33	29	26	22
Ids	280	372	63	376	155	99
Sulf1	1286	2596	210	1280	381	988
Sulf2	1865	2789	1796	674	353	595
Arsi*	245	293	576	1008	1222	1255
Col10a1*	9	9	8	17573	13763	11186

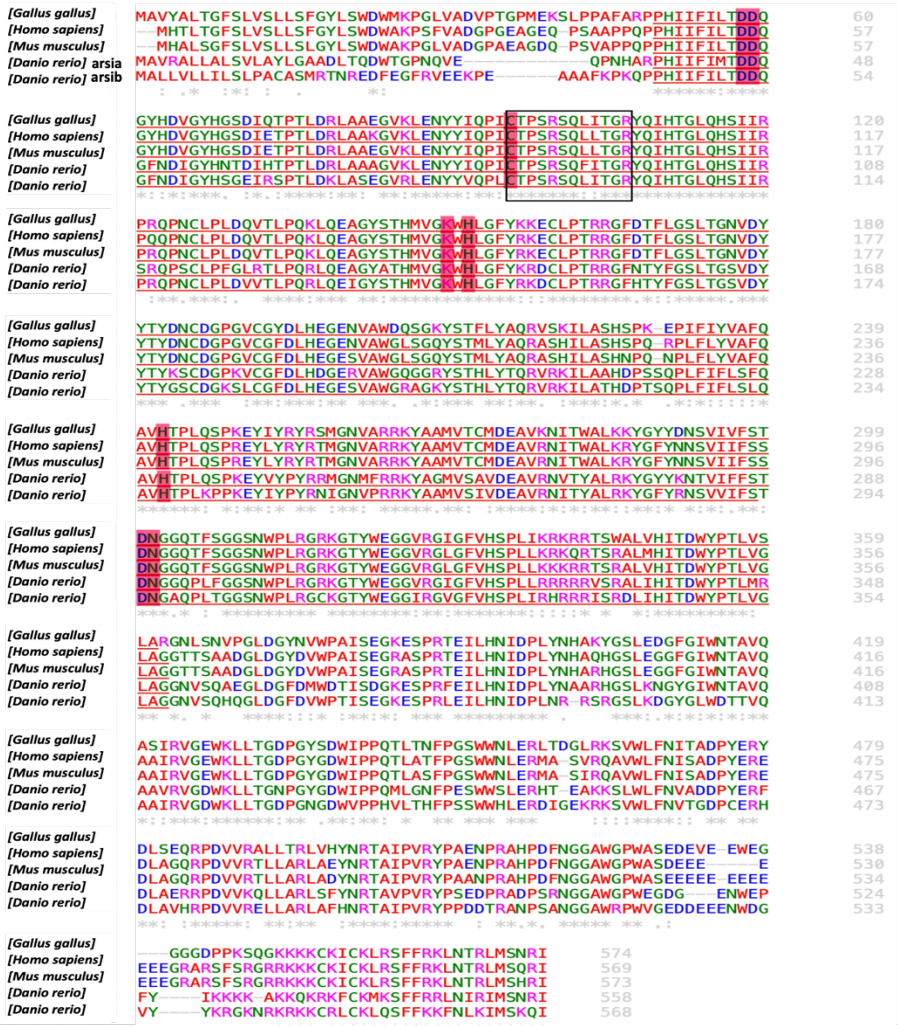
* Genes that are differentially expressed between immature and mature cartilage

Table 3.2 Differential expression counts for ARSI and Col10a1 in immature vs. mature cartilage in the developing mouse humerus (E14.5).

Gene	log2FC	log2CPM	PValue
Arsi	2.588097	5.791335	1.08E-05
Col10a1	11.93918	9.459523	7.98E-35

Sulfatases have substantial sequence conservation among different vertebrate species (Figure 3.7). ARSI amino acid sequences obtained from NCBI and aligned with Clustal Omega showed that 83.51% of the full-length ARSI protein sequence was conserved between *Gallus gallus* and *Mus musculus* (Figure 3.7B). This conservation is even more evident in the sulfatase domain, which contains the enzyme putative active sites and is 90.38% similar between *Mus musculus* and *Gallus gallus* (Figure 3.7C). These sequence similarities suggest that ARSI function might be conserved across vertebrates.

A CLUSTAL 0(1.2.4) multiple sequence alignment



B Percentage Identity Matrix - full length ARSI protein

	[Gallus gallus]	[Homo sapiens]	[Mus musculus]	[Danio rerio] arsia	[Danio rerio] arsib
[*Gallus gallus*] XP_004945003.2	100	83.39	83.51	70.58	66.67
[*Homo sapiens*] NP_001012301.1	83.39	100	95.96	68.24	65.77
[*Mus musculus*] NP_001033588.1	83.51	95.96	100	68.23	66.37
[*Danio rerio*] arsia XP_002664306.1	70.58	68.24	68.23	100	71.27
[*Danio rerio*] arsib XP_692237.2	66.67	65.77	66.37	71.27	100

C Percentage Identity Matrix - Sulfatase domain of ARSI protein

	[Gallus gallus]	[Homo sapiens]	[Mus musculus]	[Danio rerio] arsia	[Danio rerio] arsib
[*Gallus gallus*] XP_004945003.2	100	90.03	90.38	77.56	75.96
[*Homo sapiens*] NP_001012301.1	90.03	100	97.44	75.32	74.68
[*Mus musculus*] NP_001033588.1	90.38	97.44	100	75.72	76.04
[*Danio rerio*] arsia XP_002664306.1	77.56	75.32	75.72	100	80.25
[*Danio rerio*] arsib XP_692237.2	75.96	74.68	76.04	80.25	100

Figure 3.7 ARSI is highly conserved across vertebrates.

(A) Alignments of ARSI sequences from human, mouse, chicken, and zebrafish show abundant amino acid similarities. Residues highlighted in pink are putative catalytic sites, which are the same for all compared species. The catalytic core of sulfatases (boxed in black) is conserved throughout the entire enzyme class. The conserved sulfatase domain is underlined in red. Asterisks indicate positions which have a single, fully conserved residue, colon indicates conservation between groups of strongly similar properties, and period indicates conservation between groups of weakly similar properties. (B) The percentage identity matrix shows the percentage of similarities of amino acids among human, mouse, and chick and their respective accession numbers. (C) When just the conserved sulfatase domain is considered, the amino acid correspondences across species are even more evident.

ISH was performed in adjacent sections of HH36 chick humerus to assess the expression of ARSI, ARSB and GALNS and typical markers of cartilage maturation. While *COL2A1* was highly expressed in immature cartilage and part of the mature cartilage, *COL10A1* and was expressed specifically in mature cartilage (Figure 3.8A-B'). Both *COL2A1* and *COL10A1* were not expressed in the center of the humerus, where *SPP1* - a marker for hypertrophy and early mineralization - had higher levels of expression (Figure 3.8C and C'). *IHH*, a molecule known to regulate the onset of hypertrophy, was observed in prehypertrophic chondrocytes as expected (Figure 3.8D and D'). While GALNS did not seem to be highly expressed in any region of the humerus (Figure 3.8E and E'), ARSB had a more general and unspecific expression all over the element (Figure 3.8F and F'). *ARSI* expression was observed in prehypertrophic chondrocytes, partially overlapping the *IHH* domains. Interestingly, *ARSI* expression varied with the depth of the tissue, where deeper regions presented a smaller domain of expression and peripheral sections had a broader domain of expression (Figure 3.8G-G'), suggesting a ring-like shape surrounding the prehypertrophic chondrocytes (Figure 3.8H).

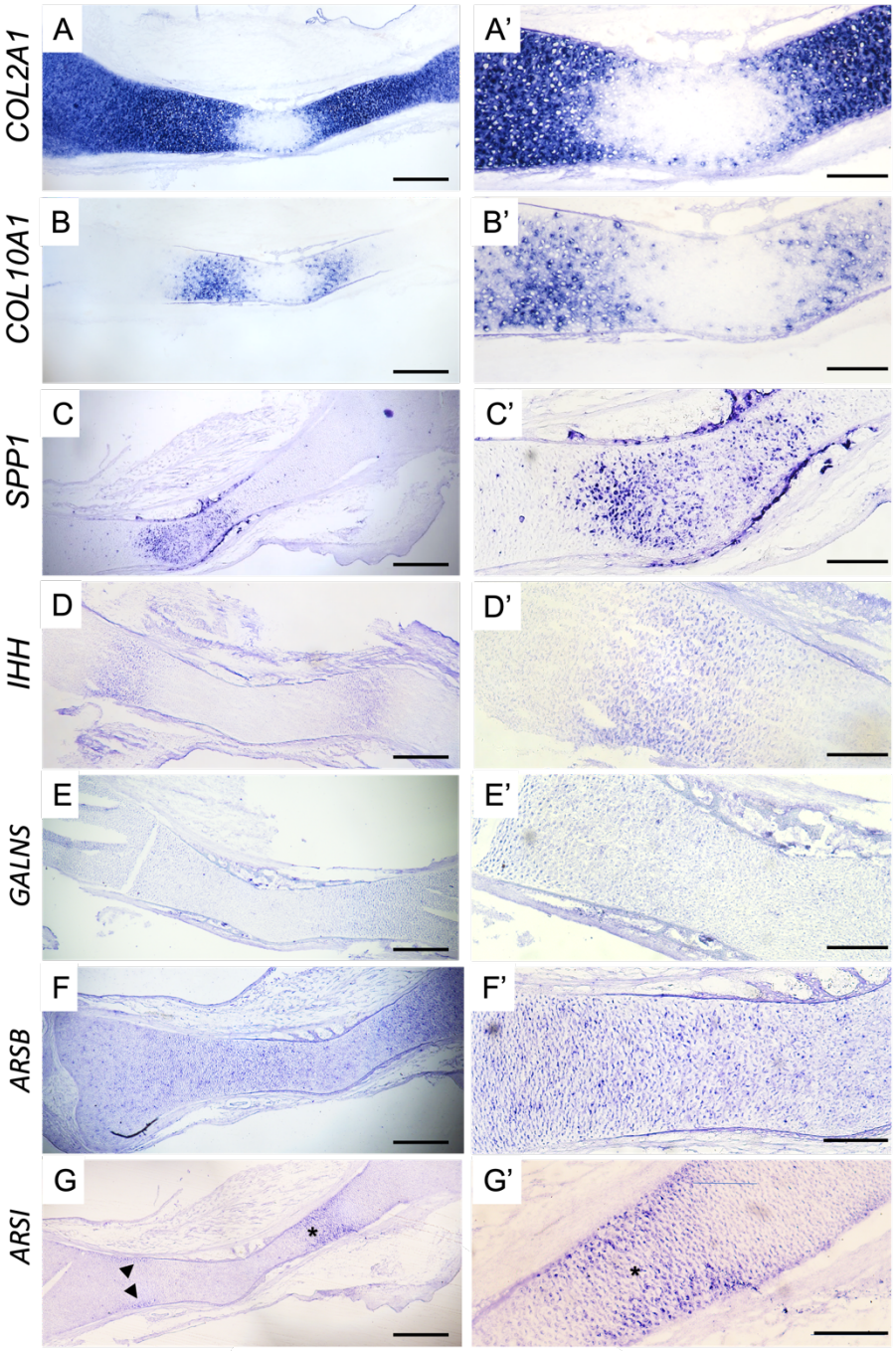

COL2A1
A
A'
COL10A1
B
B'
SPP1
C
C'
IHH
D
D'
GALNS
E
E'
ARSB
F
F'
ARSI
G
G'
H
PHC
HC
PHC
IMM
MAT
IMM

Figure 3.8 ISH of endochondral ossification markers and sulfatases in developing chick humerus (HH36). (A-A') *COL2* was highly expressed in the whole humerus except the mostly central region. (B-B') *COL10* was specifically expressed in MAT. (C-C') Hypertrophic chondrocytes expressed *SPP1*, a marker for hypertrophy and early mineralization. (D-D') *IHH* was expressed just in prehypertrophic chondrocytes. (E-E') The expression of *GALNS* was not observed above background levels in any region of the humerus. (F-F') The expression of *ARSB* seemed diffused throughout the humerus and didn't differ between IMM and MAT regions. (G'-G') *ARSI* was expressed in prehypertrophic chondrocytes, partially overlapping *IHH* domain. *ARSI* expression varied depending on the depth of the sections. Deeper regions presented a smaller domain of expression (arrows), while peripheric sections had a broader domain of expression (asterisk). (H) Hypothetical representation of ARSI domains of expression in a ring-like shape (purple) around prehypertrophic chondrocytes in the chick humerus. Scale bars: A-G 500um; A'-G' 200um. HC: hypertrophic chondrocytes; IMM: Immature cartilage; MAT: Mature cartilage; PHC: prehypertrophic chondrocytes

Since no previous work reported ARSI expression in chicken, more ISH analyses were performed to confirm the observed expression pattern. The *ARSI* ring-like expression was further confirmed by ISH of serial adjacent sections from the same humerus, where it was possible to observe the domains of expression increasing and decreasing in both humerus heads in different tissue depths (Figure 3.9). To make our results more robust, 3 different probes were designed to target 3 different regions of the *ARSI* mRNA: (1) just the coding sequence region (CDS), (2) the end of the CDS plus part of the 3'untranslated region (3'UTR), and (3) just the 3'UTR. ISH results showed the same pattern of ARSI expression using the 3 different probes (Figure 3.10A-C). Probe hydrolyses did not change the observed *ARSI* expression (Figure 3.10D), and sense probes showed no specific binding in the chick humerus. (Figure 3.10E, F)

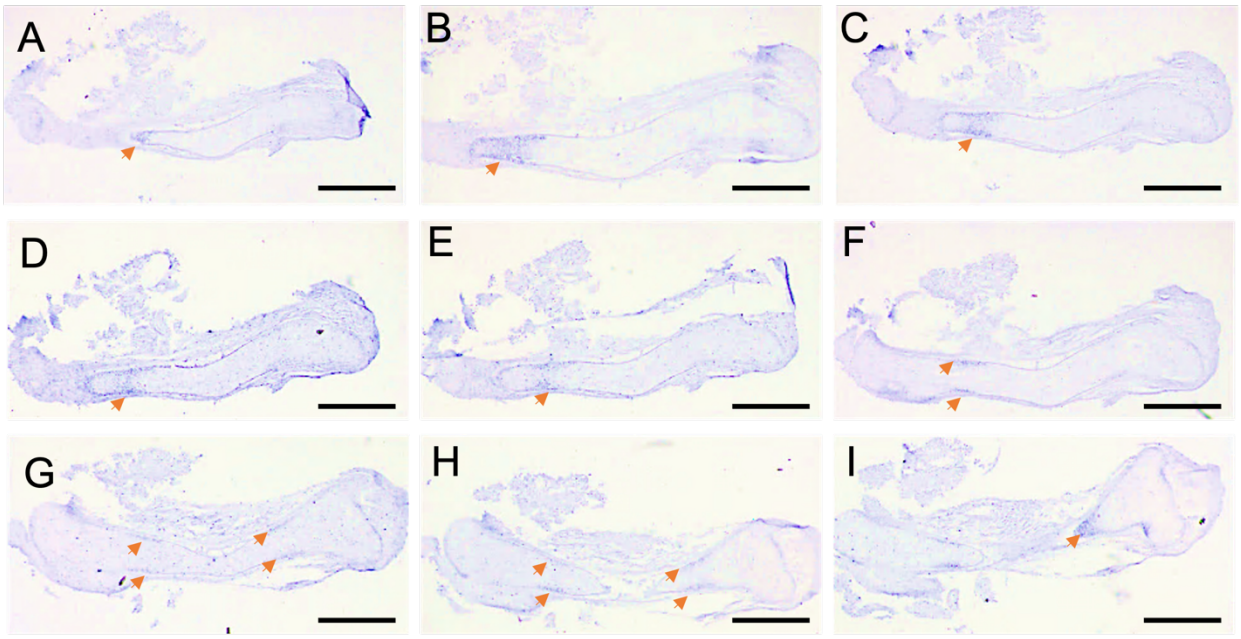


Figure 3.9 ISH in adjacent sections indicates a "ring-like" expression pattern of *ARSI* surrounding prehypertrophic chondrocytes in chick humerus. (A) Arrowhead indicating the *ARSI* expression domain starting to be exposed on the left humerus head. (B-C). The domain of *ARSI* expression was wider because more of the humerus head is being exposed. (D-E) *ARSI* expression starts to decrease when sections get too deep into the tissue. (F-H) In deeper sections, the expression is restricted to smaller domains in the periphery of the humerus. (I) Closer to the end of the tissue domain of expression starts to increase again. Arrowheads show *ARSI* expression domains. Scale bar: 1mm.

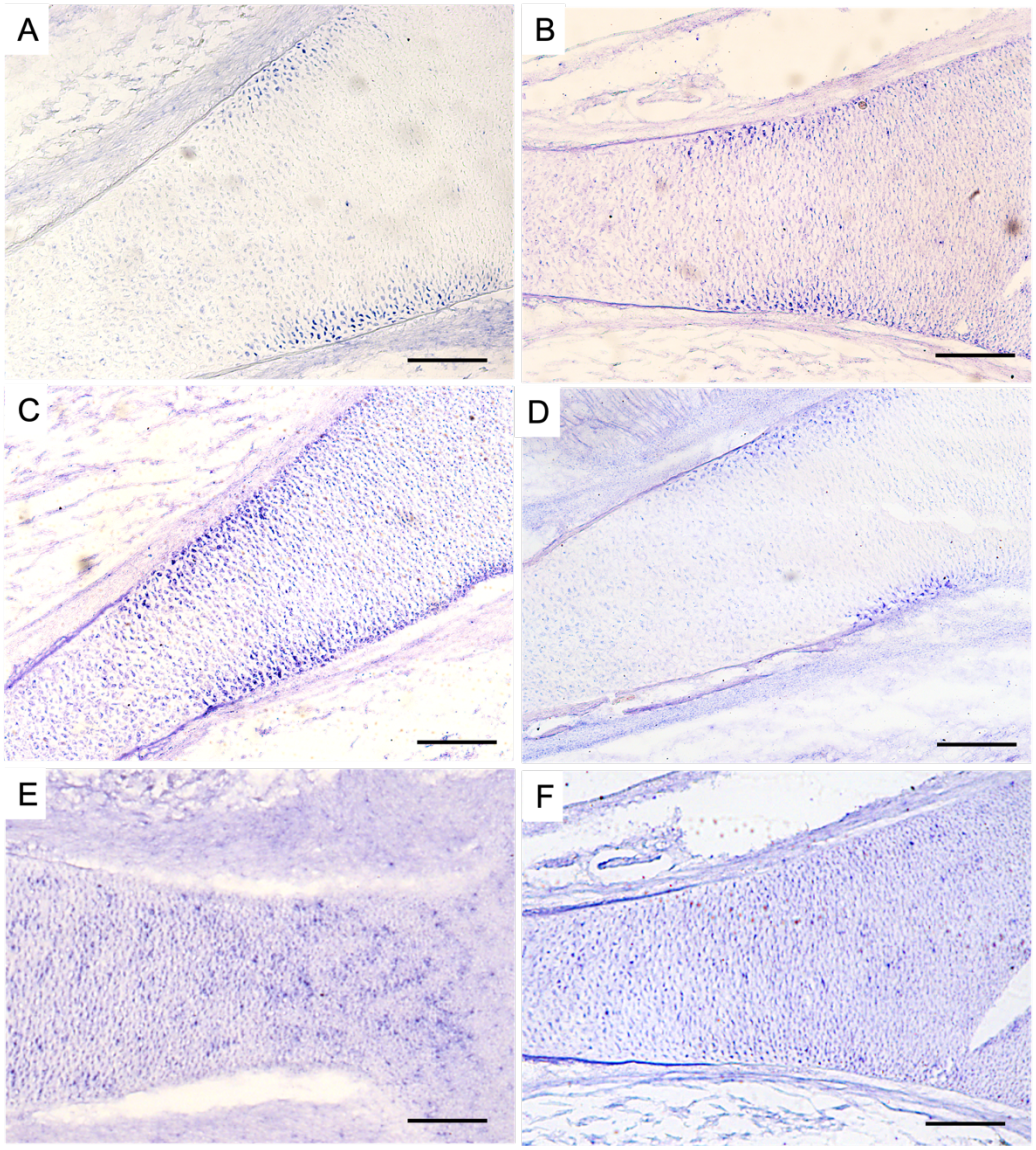


Figure 3.10 *ARSI* expression in chick humerus was consistent using different probes. Probes targeting 3 different regions of *ARSI* mRNA resulted in the same peripheral pattern of *ARSI* expression. (A) ISH with probe targeting *ARSI* CDS. (B) *ISH* with probe targeting *ARSI* CDS plus part of 3'UTR. (C) ISH with probe targeting *ARSI* 3'UTR. (D) *ISH* with hydrolyzed probe (targeting CDS). (E-F) *ISH* with *ARSI* sense probes (targeting CDS and 3'UTR, respectively) were used as controls and showed no specific pattern of expression. Scale: 200um

RNA expression domains do not always correlate with protein expression domains, so to investigate how proteins were expressed during cartilage maturation IHC for ARSB, GALNS, and ARSI was performed on sections of the HH36 chick humerus (N=10; Figure 3.11). The region of mature cartilage was defined by the presence of

hypertrophic chondrocytes under perichondral bone and by COL10 expression (Figure 3.11A-C). GALNS expression was not observed above background levels in any region of the humerus (Figure 3.11D). ARSB expression appeared at similar levels in both mature and immature cartilage regions (Figure 3.11E). By contrast, ARSI expression appeared to be relatively specific to the mature cartilage region, (Figure 3.11F). Staining with both antibodies on one section confirmed that both ARSI and COL10 were co-expressed in mature cartilage (Figure 3.11G). These results confirmed the specific expression of ARSI in mature cartilage, in a similar *in vivo* domain as the one identified with decreased sulfate esters using XRF imaging (Figure 3.4C). Further IHC for COL10 and ARSI was performed in mouse developing humerus (EE 14.5, N=3). Similarly to chick IHC, ARSI was observed specifically in mature cartilage, suggesting protein conservation across species (Figure 3.12)

ARSI IHC was also performed in serial sections of chick humerus and its specific expression in mature cartilage varied with tissue depth. Measurements of the humerus width indicated that slides closer to the humerus periphery had a larger domain of expression of ARSI in hypertrophic chondrocytes and slides closer to the center of the humerus had a more restricted domain of expression in those same cells (Figure 3.13). Interestingly, even though ISH results showed *ARSI* mostly in prehypertrophic chondrocytes and IHC showed ARSI mostly in hypertrophic chondrocytes, the expression of the RNA and protein happened in a similar ring-like shape pattern.

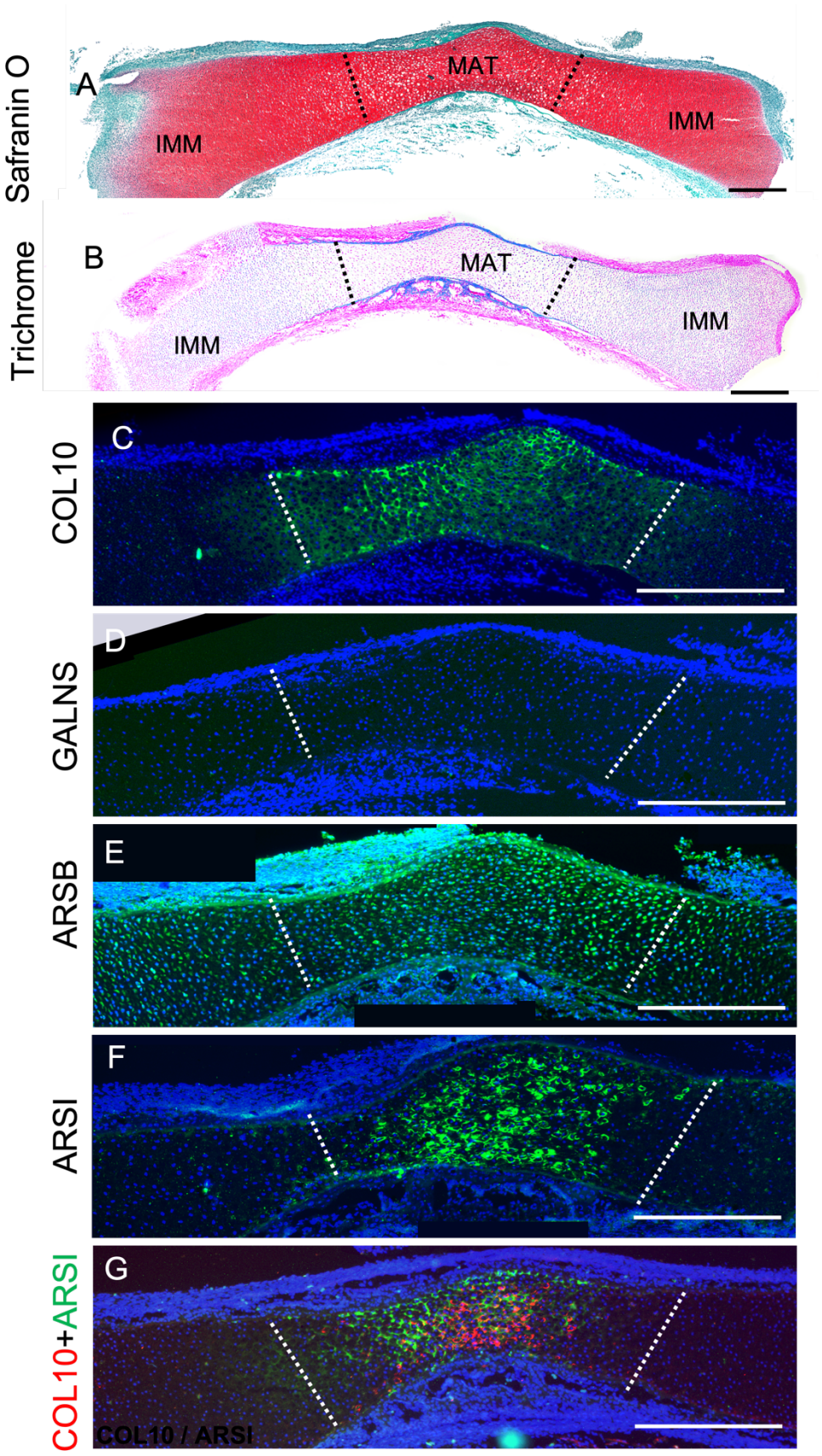

Safranin O
A
MAT
IMM
IMM
Trichrome
B
MAT
IMM
IMM
COL10
C
GALNS
D
ARSB
E
ARSI
F
COL10+ARSI
G
COL10 / ARSI

Figure 3.11 ARSI- but not GALNS and ARSB- is differentially expressed in the MAT of chick humerus (HH36). (A) Safranin O staining showing sulfated PGs in red. A high density of hypertrophic chondrocytes - observed in the central region of the chick humerus – was used to define the MAT and IMM cartilage regions defined by dotted lines. (B) Trichrome staining showing the perichondral bone in blue, which is typically found surrounding MAT. (C) MAT region was also confirmed by COL10 expression on cartilage ECM. (D) GALNS was not observed above background levels in any region of the humerus. (E) ARSB was observed in MAT and IMM regions at similar levels. (F) ARSI expression appeared to be relatively specific to the MAT region. (G) Both ARSI and COL10 were expressed in the same MAT domain, however, there was not much cellular co-localization. IMM: Immature cartilage, MAT: Mature cartilage. Scale bars 400um. N=10

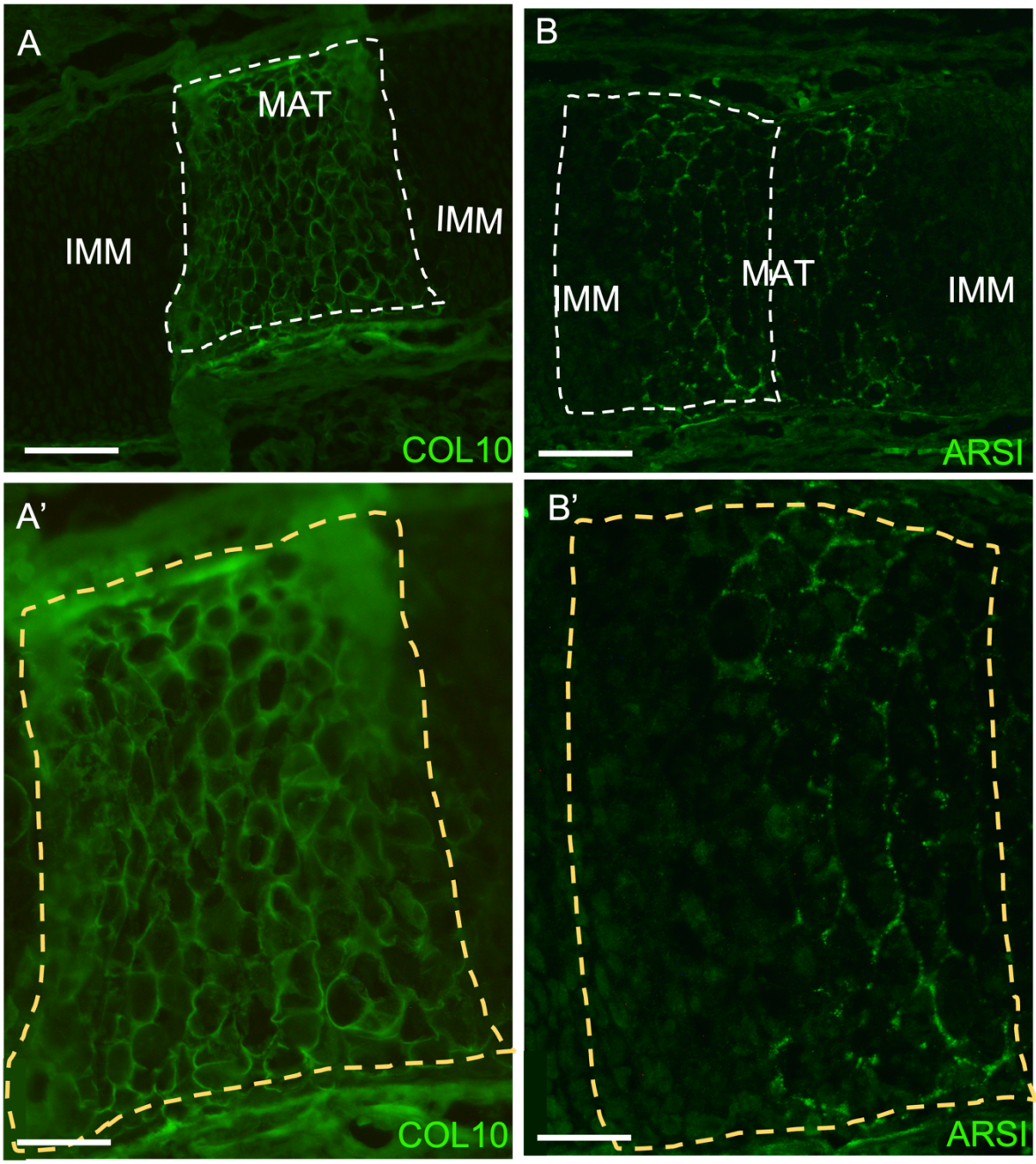


Figure 3.12 ARSI is expressed specifically in mouse mature cartilage. (A) IHC for COL10 in mouse sections was used to define (MAT) (white dotted line). (A') Higher magnification of the MAT region (surrounded by a yellow dotted line) shows COL10 expressed in the ECM. (B) IHC for ARSI showed that this protein is expressed in MAT but not in the IMM region. Region comprising half of the MAT and part of IMM cartilage is surrounded by a white dotted line (B') Higher magnification of the region surrounded by a white dotted line containing MAT and IMM. ARSI seems to be expressed in "dots" close to the cell membranes, which are not evenly distributed throughout MAT. Scale: A-B100um; A'-B' 50um

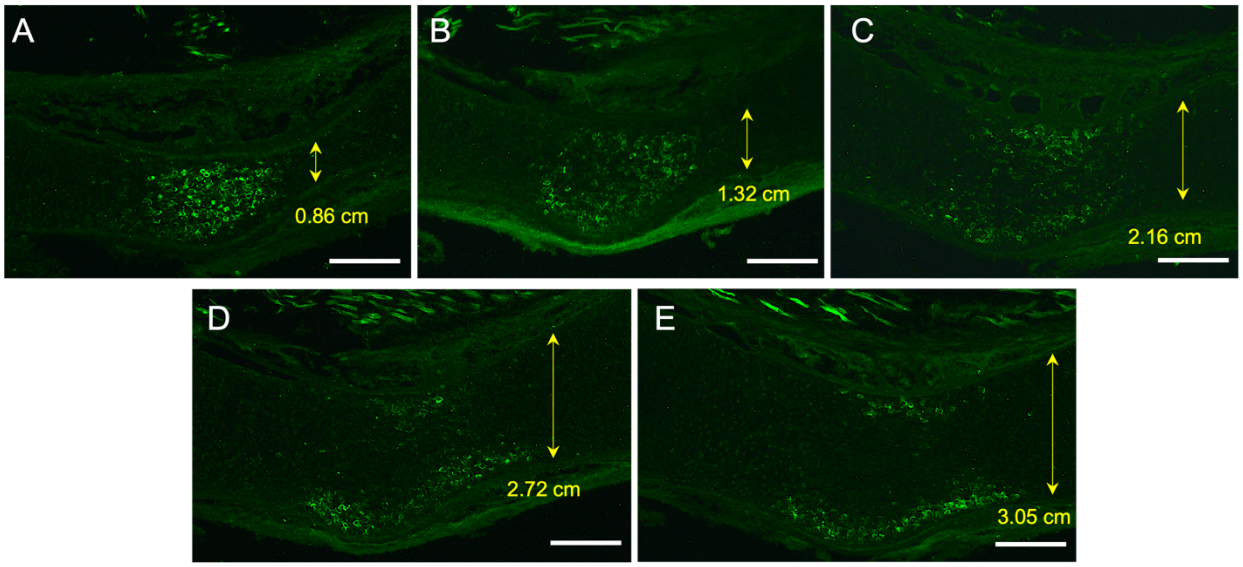


Figure 3.13 ARSI protein expression in chick humerus varies with tissue depth. (A-C) When using sections closer to the humerus periphery ARSI has a larger domain of expression in mature cartilage. (D-E) In deeper sections the ARSI domains of expression are more intense in some layers of hypertrophic chondrocytes closer to the perichondrium. Measurements in yellow indicate width of the humerus in approximate same regions, smaller widths are closer to tissue periphery and larger widths are closer to the center of the tissue. Scale: 200um

To further investigate the expression of ARSI protein during chondrocyte maturation, ATDC5 cells were assessed during *in vitro* chondrogenic differentiation (Figure 3.14). Cartilage maturation was confirmed in micromasses of ATDC5 cells by Alcian Blue staining and COL10 expression over 21 days of cell culture (N=3, Figure 3.14A-F). IHC (N=3) showed that ARSI levels increased over time, with a higher expression on days 14 and 21 as ATDC5 cells matured (Figure 3.14G-J). Western blot, performed with a pool of micromasses (7 for each time point), confirmed the IHC results (three technical replicates, Figure 3.14G-J). These *in vitro* data confirm that ARSI is upregulated during cartilage maturation.

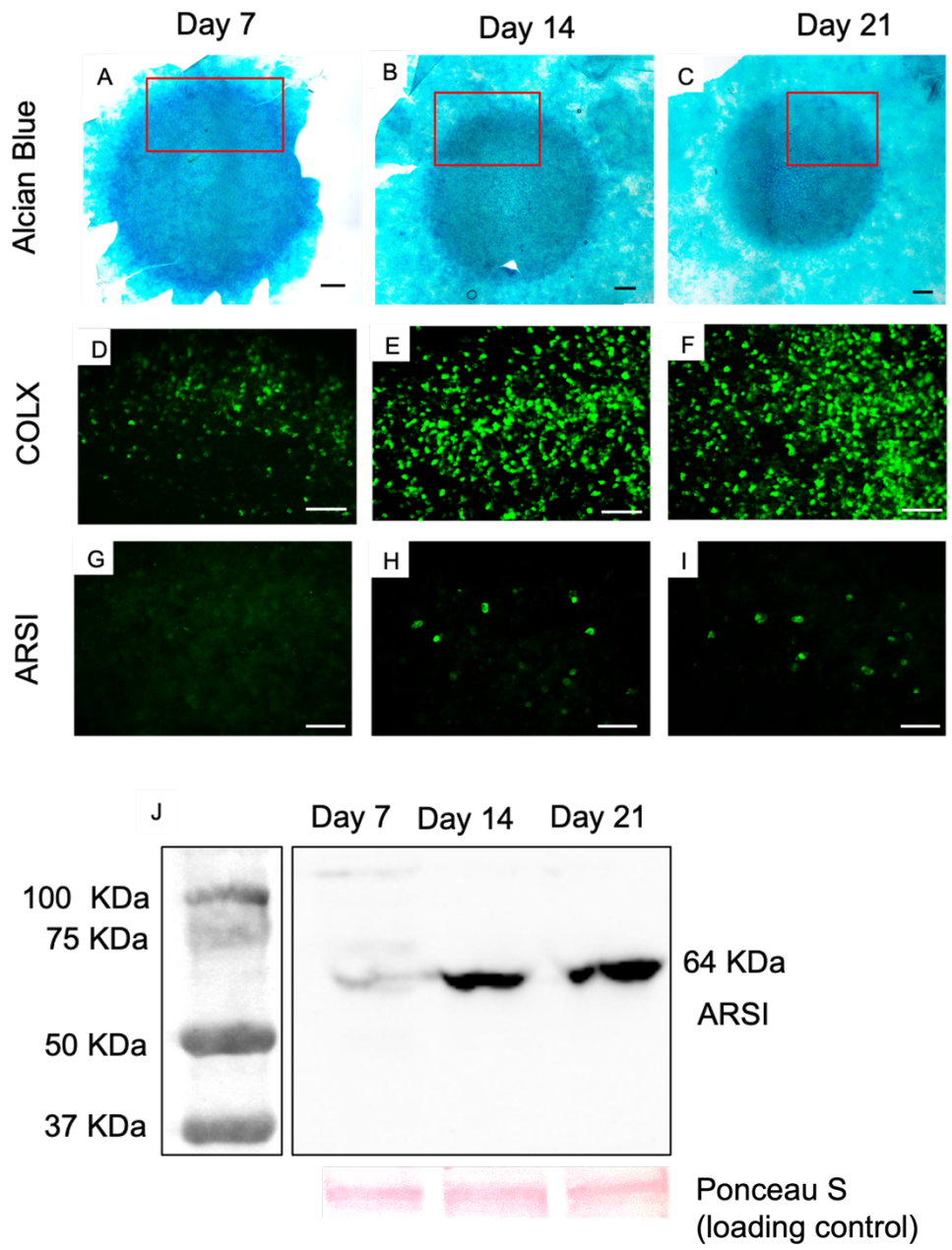


Figure 3.14 ARSI expression increases with ATDC5 chondrogenic differentiation.(A-C) Alcian blue staining of ATDC5 micromasses after 7, 14, and 21 days of culture in differentiation media shows that cells produce sulfated proteoglycans. (D-F) COL10 expression, observed by IHC, indicates ATDC5 are differentiating into chondrocytes. (G-I) IHC showed an increase in ARSI expression as ATDC5 cells matured (N=3). (J) Western blot with a pool of micromasses confirmed the increase in ARSI in ATDC5 over time. Ponceau S staining used as loading control (red bands) (N=1, 3 technical replicates). Red rectangles indicate approximate regions that were imaged through fluorescence. Scale bars: 500 μm (A-C) and 100 μm (D-I)

3.4.4 ARSI is a novel PG sulfatase

To determine the biochemical activity of the orphan sulfatase ARSI *in vitro*, human *ARSI* was overexpressed in HeLa cells. HeLa cell clones transfected with an empty vector expressed *ARSI* at very low levels, and *ARSI* was successfully overexpressed in *ARSI*-transfected HeLa cell clones, as shown by RT-PCR (Figure 3.15A). To investigate if ARSI was a PG sulfatase, CS and HS disaccharides harvested from empty vector-transfected and *ARSI*-transfected HeLa cell clones were analyzed by high-performance liquid chromatography. Treatment with bacterial eliminases converts the original uronic acid structure of GAGs into an artificial structure, the 4,5-unsaturated uronic acid, 4-deoxy-α-L-*threo*-hex-4-enepyranosyluronic acid (ΔHexA).

Harvested CS disaccharides were: ΔHexAα1-3GalNAc(4S) (C4S), ΔHexAα1-3GalNAc(6S) (C6S), ΔHexA(2S)α1-3GalNAc(6S) (C2S6S), ΔHexα1-3GalNAc(4S, 6S) (C4S6S), and ΔHexAα1-3GalNAc (C0S). Where, GalNAc 2S, 4S, and 6S represent *N*-acetylgalactosamine, 2-*O*-sulfate, 4-*O*-sulfate, and 6-*O*-sulfate respectively. The total CS amounts were reduced significantly in HeLa cells stably expressing *ARSI*, compared with empty vector controls (Figure 3.15B). Specifically, C4S and C6S significantly decreased in *ARSI*-expressing cells, compared to control-transfected cells (Figure 3.15C). *ARSI* expression caused the ratio of C4S to total CS to decrease, while the ratio of C0S to total CS increased (Figure 3.15D). Arguing against any complication from genetic compensation in this over-expression experiment, *ARSI*-transfected HeLa cell clones did not alter the expression of genes that promote CS sulfation. Expression of genes encoding the sulfotransferases CHST3, CHST11, CHST12, and B3GAT3 were unaffected, compared to empty vector-transfected HeLa cell clones (Figure 3.15E).

Harvested HS dissacharides were: ΔHexAα1-4GlcNAc (0S), ΔHexAα1-4GlcN(NS) (NS), ΔHexAα1-4GlcNAc(6S) (6S), ΔHexA(2S)α1-4GlcN(NS) (2SNS), ΔHexAα1-4GlcN(NS, 6S) (6SNS), and ΔHexA(2S)α1-4GlcN(NS, 6S) (2SNS6S). Where GlcNAc, GlcN, 2S, 6S, and NS represent *N*-acetylgalactosamine, *N*-acetylglucosamine, glucosamine, 2-*O*-sulfate, 6-*O*-sulfate, and 2-*N*-sulfate, respectively. The total HS amount was also reduced significantly in HeLa cells stably expressing *ARSI* compared

with empty vector controls (Figure 3.16A). More specifically, there were reductions in 0S and NS amounts in *ARSI*-transfected cells (Figure 3.16B). *ARSI* expression caused the ratio of 6S to total HS to increase, while the ratio of 2SNS to total HS decreased (Figure 3.16C). These *in vitro* biochemical data suggest that ARSI is a novel PG sulfatase that may be acting on CSPG and HSPG.

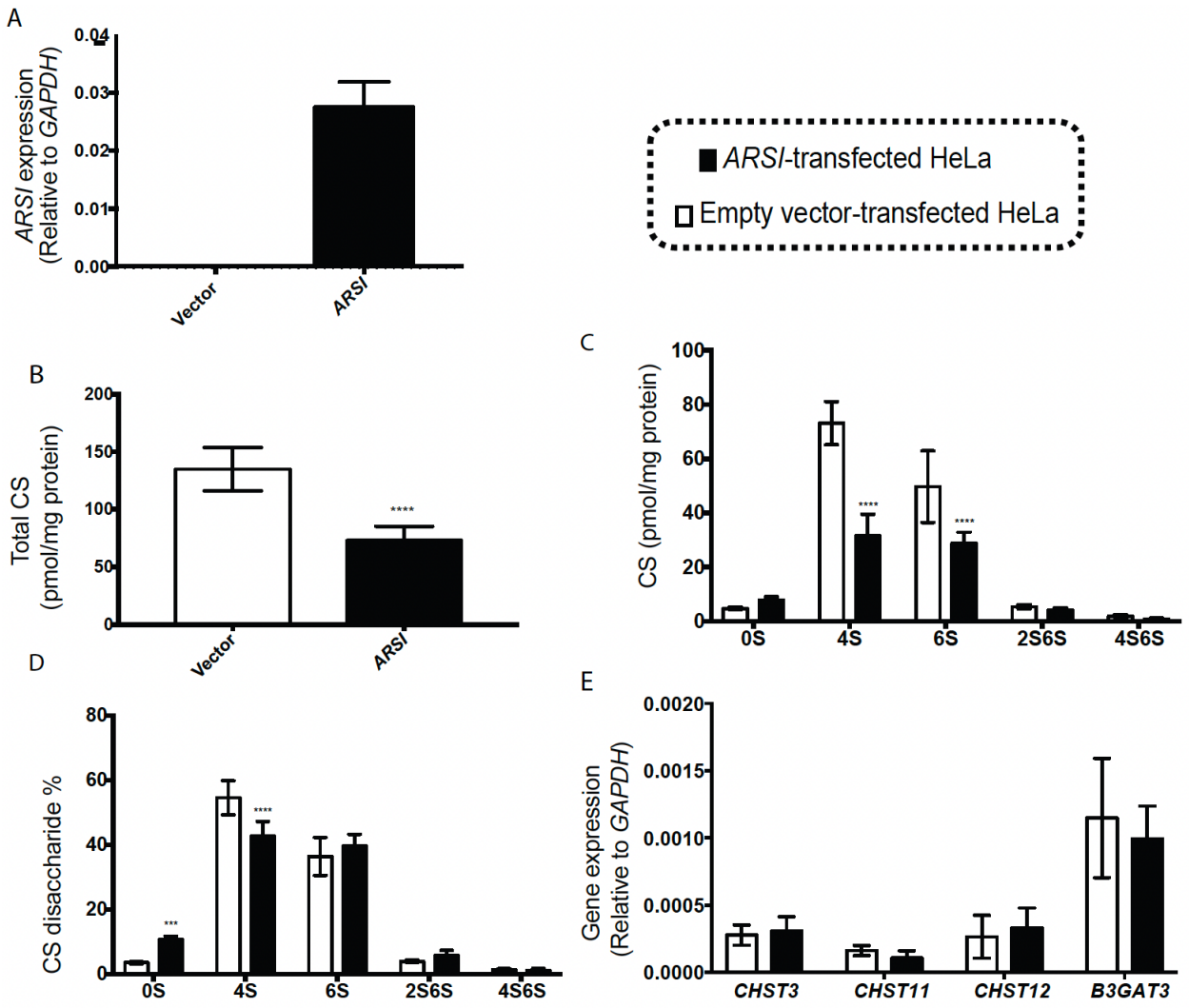


Figure 3.15. *ARSI* is a novel sulfatase of chondroitin sulfate proteoglycans.(A) RT-PCR shows that human *ARSI* was successfully overexpressed in transfected HeLa cells. (B) Liquid chromatography results show a decrease in total CS in *ARSI*-transfected HeLa cells, compared with empty vector controls. (C) C4S and C6S significantly decreased in *ARSI*-expressing cells, compared to control-transfected cells. (D) At the same time, the ratio of C4S decreased and the ratio of non-sulfated C0S increased compared to total CS. (E) The expression of enzymes that promote CS sulfation, like *CHST3, CHST11,* and *CHST12,* were not affected. Error bars represent the standard error of the mean. Statistics were calculated using the Unpaired T-test (B) or two-way ANOVA (C-E). *** p=0.0002, **** p<0.0001

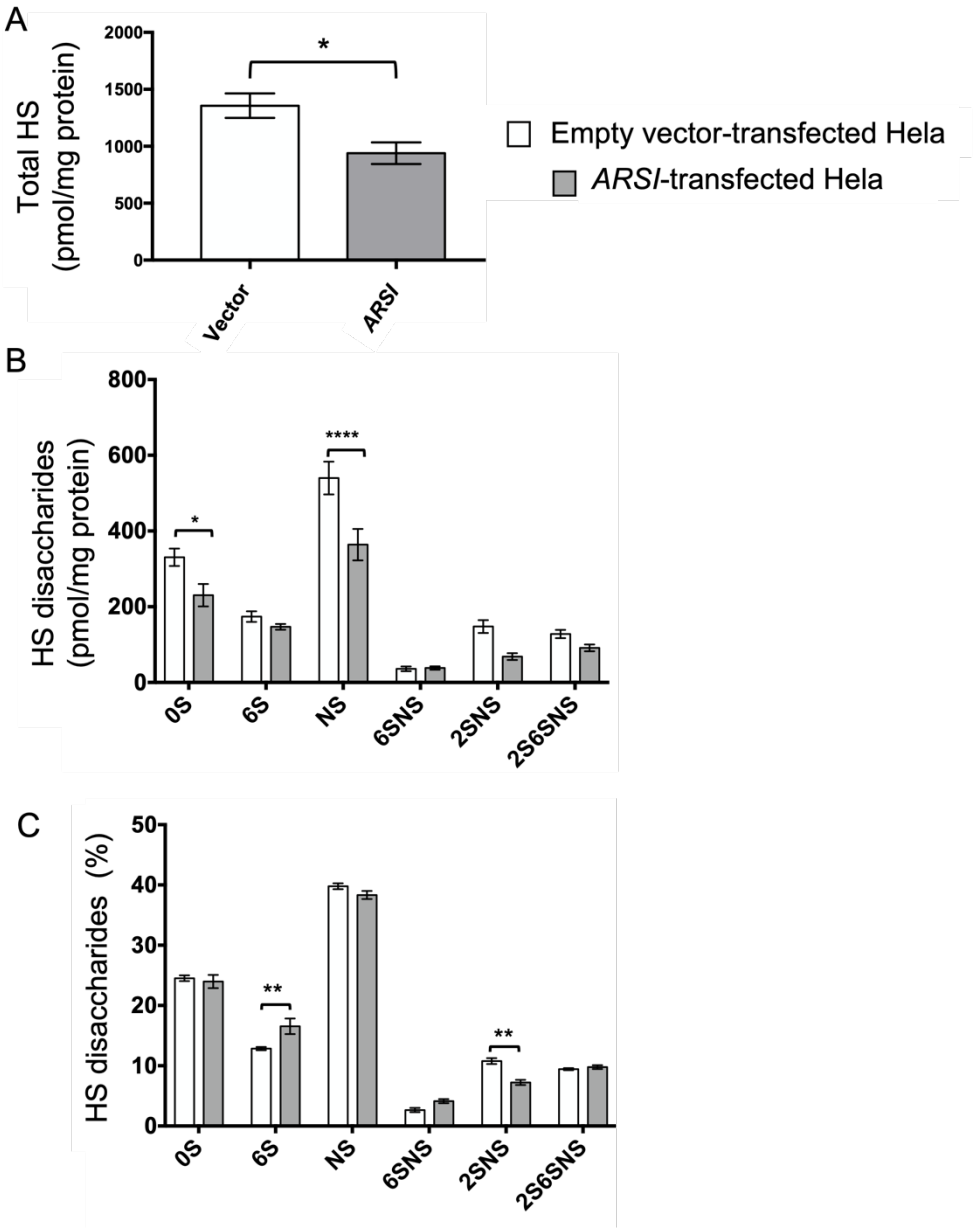


Figure 3.16 Heparan sulfate was reduced significantly in HeLa cells stably expressing ARSI.(A) Liquid chromatography results show a decrease in total HS in ARSI-transfected HeLa cells, compared with empty vector controls. (B) ΔHexAα1-4GlcNAc (0S) and ΔHexAα1-4GlcN(NS) significantly decreased in ARSI-expressing cells, compared to control-transfected cells. (C) At the same time, the ratio of ΔHexAα1-4GlcNAc(6S) to total HS increased, while the ratio of ΔHexA(2S)α1-4GlcN(NS) to total HS decreased. Error bars represent the standard error of the mean. Statistics were calculated using the Unpaired T-test (A) or two-way ANOVA (B-C). *p≤0.05, ** p≤ 0.01, p≤0.001.

3.5 Discussion

During endochondral ossification, chondrocyte maturation is accompanied by ECM biochemical modifications, remodelling, and reorganization (Behonick & Werb, 2003; Deutsch et al., 1995; Ortega et al., 2004). Highly sulfated PGs are one of the most abundant molecules in cartilage ECM, and their sulfation levels may affect PG turnover, GF signalling, and the expression of genes that regulate endochondral ossification (da Costa et al., 2017; Esko & Selleck, 2002; Kluppel et al., 2005; Settembre et al., 2008; Stelzer et al., 2007). Sulfatases act by removing sulfate esters from PGs, but their participation in cartilage maturation and endochondral ossification has not been fully understood. Here it was demonstrated that while total sulfur and sulfate esters decreased significantly in developing mature cartilage, ARSI was differentially expressed in this same region. This study showed the expression of *ARSI* mRNA expression in chick for the first time, as well as the expression of ARSI protein in biological tissues. Additionally, the first biochemical characterization of ARSI as a PG sulfatase was presented. Our results supplement the scarce literature about ARSI and potentiate future work uncovering its role in skeletal development and disease.

The reduction in total sulfur in chick humeri mature cartilage, measured by XRF imaging, was consistent with previous studies which showed reductions in sulfur in mature cartilage of growth plates of fetal calves and chick juveniles, and in the ceratohyal of zebrafish larvae using biochemical, histological and synchrotron analyses (Farquharson et al., 1994; Hackett et al., 2016; Wuthier, 1969). Conversely, studies measuring sulfur levels on the articular cartilage of horses and pigs using X-Ray emission techniques found higher sulfur levels in more mature chondrocytes in the middle and deep zones or articular cartilage respectively (Rizzo et al., 1995; T. Reinert, 2001). Further studies are needed to investigate if PG sulfation patterns differ in the growth plate and articular cartilage. Discrepancies in the measurements of sulfur levels in mature versus immature cartilage in different studies may be explained by the sensitivity of the methods used, but also by the developmental stage of the animals studied. In this work, embryonic chick humeri at stage HH36 were used because immature, mature cartilage, and perichondral bone were noticeably visible, and there

was still no vascular invasion of the cartilaginous template. At later developmental stages, blood vessel invasion, bone marrow formation, and progressive mineralization may contribute to changes in the levels of sulfur and other ECM components (Althoff et al., 1982; van Donkelaar et al., 2007)

The combination of XANES and chemically specific XRF imaging is a powerful tool to analyze which sulfur isoforms are present in biological tissues and their *in-situ* distribution (Castillo-Michel et al., 2016; Hackett et al., 2016). The results presented here show that sulfate-o-esters was the main sulfur isoform in the developing chick humerus and that its levels decreased in mature cartilage. Those findings support the idea that the majority of sulfur signal comes from PGs and that PG sulfation decreases during cartilage maturation. Similar results were observed by Hacket et al., in the developing cartilage of zebrafish larvae using the same methods described here (Hackett et al., 2016). Interestingly, using the same techniques, *fam20b* mutant zebrafish with accelerated chondrocyte maturation were shown to have fewer sulfate esters, but no less total sulfur levels in their cartilages (Eames et al., 2011; Hackett et al., 2016). The *fam20b* mutant zebrafish phenotype demonstrates that sulfate esters are indeed relevant to PG function and suggest a functional link between PG sulfation levels and chondrocyte maturation.

The degradation of sulfate esters in the lysosome results in the release of free sulfate as a by-product of PG breakdown. However, in tissue culture experiments, these free sulfate groups are generally not recycled back into PGs (Harper et al., 1993; Rome and Hill, 1986). Including a free sulfate standard in our analysis, we found no free sulfate contributing to the overlapping peaks at 2478 eV. These findings suggest that sulfate esters lost in mature cartilage do not persist as free sulfate for an extended period within the extracellular matrix (ECM). One possible explanation is that free sulfate groups might be removed during the sample preparation process.

The FTIR analysis presented here showed that there was no significant decrease in proteins or PG in mature cartilage, even though lower sulfation levels were observed in that region. This indicated that sulfate esters decrease in the mature cartilage ECM

independently of overall PG levels. Even though there is no specific band for detecting PGs, the two bands used here (985 – 1140 cm^{-1} and the 1374 cm^{-1} 2^{nd} derivative peak) were shown to be highly correlated with PG or GAG contents in cartilage cell culture, articular cartilage sections, and in standard mixtures of PGs and collagen (Boskey & Pleshko Camacho, 2007; Kim et al., 2005; Rieppo et al., 2012; Saarakkala & Julkunen, 2010). It is important to notice that the accuracy of PG measurements by FTIR can vary among the regions of the cartilage and species being analyzed and may be affected by the collagen absorption (Rieppo et al., 2012; Saarakkala & Julkunen, 2010). Here it was observed that the overall PG levels did not change in mature cartilage. However, since chondrocyte maturation involves mineralization and PGs tend to inhibit cartilage mineralization, a reduction in PG size or density seems to be needed for endochondral ossification to progress (Alini et al., 1992; Buckwalter et al., 1987; Poole et al., 1982). Since an early developmental stage was analyzed in this study, the observed sulfur reduction might be an initial step for the future degradation and reorganization of PGs.

A possible explanation for the reduction of sulfate esters in mature cartilage is the action of sulfatases, which hydrolyze sulfate esters from different substrates including GAGs (Hanson et al., 2004). Although ARSB and GALNS were the only known sulfatases to act on CSPGs, our RNA sequencing, ISH, and IHC results showed that these sulfatases were not differentially expressed in mouse (E14.5) or chick (HH36) mature cartilage. On the other hand, the orphan sulfatase ARSI was differentially expressed in mature cartilage, making it a promising candidate to regulate PG sulfation levels in this area. Data regarding sulfatase expression patterns are scarce, and here we showed for the first time that ARSI protein is present in the mature cartilage *in vivo* in developing chick and mouse humerus, and *in vitro* in ATDC5 cell culture. The only other study found to investigate sulfatase expression in developing bone, assessed the presence of 9 sulfatases in mouse embryos by ISH including *Arsb, Galns* and *Arsi.* Different from our results in the chick, *Arsb* expression was not observed in the skeletons of developing mice (E14.5). On the other hand, similar to our chick ISH and IHC data, mouse embryos had no expression of *Galns* in their skeleton and *Arsi* expression was not observed in bone or immature cartilage in chondral bones,

supporting our finding of *Arsi* differential expression in mature cartilage (Ratzka et al., 2010). Data from microarray analysis available at LifeMap discovery also suggests an *Arsi* increase with chondrocyte maturation. This data is not mentioned or discussed in any paper, but in the Lifemap database, it is possible to see that zeugopod chondrocytes (E13.5) had a 2.87 log2 fold change increase in *Ars*i levels compared to zeugopod mesenchymal condensate cells (E12.5) (Cameron et al., 2009; Edgar et al., 2013). This finding supports our *in vivo* and *in vitro* results showing an increase of ARSI with chondrocyte maturation. Later time points, where there is a clear distinction between mature and immature cartilage were not available for comparison.

Both *ARSI* mRNA and protein were expressed in chick mature cartilage, however, *ARSI* mRNA was mostly expressed in prehypertrophic chondrocytes and ARSI protein was mostly expressed in hypertrophic chondrocytes. This finding agrees with many papers showing that RNA and protein levels do not always correlate. In fact, the amount of mRNA can just predict the amount of proteins around 40% of the time (Liu et al., 2016; Maier et al., 2009; Payne, 2015). Various regulatory mechanisms may explain the differences in mRNA and protein expression, such as microRNAs that can destabilize mRNA or repress translation and different synthesis and decay rates for mRNA and proteins which can influence the final protein output despite having a certain amount of RNA present. (Payne, 2015). A discrepancy between the domain of mRNA and protein expression was also observed for other genes, such as COL10. Like other papers analyzing similar stages of chick skeleton, *COL10* mRNA was not observed in the center of the cartilaginous template where the late hypertrophic marker *SPP1* was expressed (Conen et al., 2009; Zhu et al., 2010), however, the COL10 protein was highly expressed in that same region.

No other publications showing *ARSI* expression in chick skeletons were found for comparison. However, the same pattern of *ARSI* expression was observed using three probes targeting different regions of this gene, and our sense probe did not show any specific expression pattern making our data consistent. Further studies are needed to understand why ARSI expression varies with the tissue depth and if sulfur levels accompany this variation. Different from previously published *Arsi* ISH in mice, chick

ARSI ISH did not show expression in hypertrophic chondrocytes, but in both animal models, *ARSI* was expressed in a domain that partially overlaps Ihh (an important regulator of cartilage maturation) (Ratzka et al., 2010). The mice and chick growth plates are organized similarly in zones containing chondrocytes in different maturation states, however, endochondral ossification happens more uniformly in mammals, which could be playing a role in the differences observed in the ISH results (Barreto & Wilsman, 1994; Farnum & Wilsman, 2001).

Although ARSB and GALNS were the only sulfatases to date known to act in CSPG, the results presented here suggest that ARSI is a novel CSPG sulfatase. Because of sequence similarities, ARSI is considered to be closely related to ARSB, supporting the argument that their functions are similar (Holmes, 2016; Obaya, 2006; Sardiello et al., 2005). The activity of ARSI as a sulfatase was previously described using 4-methylumbelliferyl sulfate, a general arylsulfatase substrate, but its biological substrate has not been described (Oshikawa et al., 2009). The ARSI function as a CSPG sulfatase described here, patterned with its differential expression, might underlie the decrease of sulfation observed in mature cartilage. Levels of C6S and C4S decreased in ARSI-transfected cells. Interestingly, C6S and C4S were previously reported to decrease in mature chondrocytes compared to immature chondrocytes (Farquharson 1994). It was also observed that while the proportion of C4S decreased in ARSI-transfected cells, C0S increased, and this same trend has been suggested as important for cartilage maturation and mineralization during endochondral ossification (Byers et al., 1997; Kluppel et al., 2005).

Even though sulfatases are usually specific for their substrates, the presented results indicate that besides CS, total HS levels also decreased in ARSI-transfected cells. Results showed a significant reduction of sulfated disaccharides (NS: ΔHexAα1-4GlcN) and unsulfated disaccharides (0S: ΔHexAα1-4GlcNAc) in transfected cells. It is unexpected that the removal of sulfur by a sulfatase would decrease the amount of unsulfated disaccharides, but that might suggest some compensatory effect is happening when HeLa cells overexpress *ARSI*. A simultaneous reduction of sulfated and unsulfated HS disaccharides was previously reported in $pic^{-/-}$ zebrafish (Clement et

al., 2008). The mechanisms of this simultaneous reduction are still unknown, but the authors suggested that loss of sulfation affects the process and stability of HS, affecting different disaccharides (Clement et al., 2008). Other sulfatases were reported to act in different PG kinds, however always in the same sugars, for example, ARSB hydrolyzes sulfate esters from the 4S position of GalNAc in DS and CS, while GALNS cleaves sulfate esters from 6-sulfated galactosides and 6-sulfated N-acetylgalactosaminides in KS and CS (Mathew et al., 2015; Rivera-Colon et al., 2012). Overall, our experiments showed that ARSI acted in CSPGs (positions 4 and 6 of GalNac sugar), and HSPGs (GlcN sugar). Even though recent work with a newly discovered sulfatase -ARSK- speculates that it acts in different sulfation positions (2-O-sulfate groups of glucuronic acids and 3-O-sulfate groups of glucosamine), and on different HS and CS (Wiegmann et all 2013, Dhamale et all 2017), no previous sulfatase was confirmed to have different sugars as natural substrates or to act in CS and HS simultaneously, so the mechanisms by which ARSI acts that way need to be further studied.

Appropriate sulfur levels in cartilage PGs are critical for the cartilage mechanical properties and the regulation of signalling pathways that can change gene expression and cell differentiation (Khatri and Schipani, 2008). Here it was demonstrated that cartilage maturation was related to sulfur levels during endochondral ossification. It is still not clear if ARSI is involved in signalling pathways responsible for cartilage maturation. However, based on the results showing the pattern of ARSI expression and its substrate, it is possible to speculate that while cartilage matures, ARSI levels increase, removing sulfur from PGs. GAG desulfation is a necessary step for PG degradation and turnover (Settembre et al 2008., Otsuki et al., 2009), and desulfation by ARSI could affect the stability of PGs and impact binding affinities and diffusion of growth factors that control cartilage maturation and bone formation, such as Ihh (Cortes et al., 2009). More experiments are needed to test for ARSI contribution in different steps of endochondral ossification.

In summary, we have shown that there are fewer sulfate esters in mature cartilage during endochondral ossification and that ARSI is a new PG sulfatase that may be controlling this process. As far as we know, we are the first group to show that ARSI

acts on CS and to show ARSI protein expression in biological tissues. The combination of chick synchrotron results, laser RNA sequencing data, ARSI expression patterns, and the biochemical data suggests that this enzyme removes sulfur from CSPGs during cartilage maturation. To further study ARSI's biological function in bone development, the timing of endochondral ossification and the associated signalling pathways involved with ARSI function should be investigated. The process of cartilage maturation into bone starts during embryonic development but is also witnessed postnatally on the growth plates of long bones, and pathological events like advanced cases of osteoarthritis. Finding new regulators of this process is particularly important to better understand skeletal development and disease mechanisms.

Acknowledgements

Part of the research described in this paper was performed at the Canadian Light Source, a national research facility of the University of Saskatchewan, which is supported by the Canada Foundation for Innovation (CFI), the Natural Sciences and Engineering Research Council (NSERC), the National Research Council (NRC), the Canadian Institutes of Health Research (CIHR), the Government of Saskatchewan, and the University of Saskatchewan.

CHAPTER 4: GENERATING TRANSGENIC ZEBRAFISH TO ANALYZE ARSI BIOLOGICAL FUNCTION[1]

Rafaela Grecco-Machado, Maya Berscheid, B. Frank Eames

During this project, I co-supervised two summer students who contributed to some results of this chapter as follows: Asawari Albal performed PCRs to amplify the *col2a1a/hARSI* junction in Tg zebrafish and sent samples for sequencing (Figure 4.5A and B); Maya Bersheid performed PCRs to amplify the *hARSI/EGFP* junction in Tg zebrafish and sent samples for sequencing (Figure 4.5C and D). Maya also injected zebrafish with SUMF1 RNA and scored their bone levels (Figure 4.17). I designed the summer students' experiments, the primers for the genotyping PCRs, and prepared the *SUMF1* RNA for injections. The *col2a1aEGFP* construct used in this chapter was previously generated by Tuanjie Chang. I injected zebrafish with this construct and generated a stable transgenic line. All other experiments described in this chapter were performed by me.

4.1 Abstract

Chondroitin sulfate proteoglycans (CSPGs) are a main component of the cartilage extracellular matrix (ECM). Alterations in CSPG sulfation can influence the timing of chondrocyte maturation, leading to changes in endochondral ossification. Arylsulfatase I (ARSI) is a novel chondroitin sulfate proteoglycan (CSPG) sulfatase expressed during cartilage maturation. However, its biological function remains unknown. The zebrafish is a valuable animal model for studying bone development due to its similarities to human skeletal biology. To investigate our hypothesis that ARSI

[1] Since ARSI overexpression was not proven in the putative Tg zebrafish, the results of this chapter are not going to be submitted to publication for now. Parts of this chapter may be incorporated to future publications.

decreases sulfate ester levels in mature cartilage thus promoting endochondral ossification, four Tg zebrafish lines designed to overexpress human ARSI (hARSI) were generated using the Gateway Tol2-system. All transgene sequences were confirmed by DNA sequencing before injection. Genomic insertion of transgenes was confirmed by the presence of green hearts and genomic DNA sequencing. RT-PCRs detected RNA transcription for some constructs, while protein expression was not confirmed by western blot or immunohistochemistry, potentially due to post-translational mechanisms. Compared to wild-type siblings, the putative Tg zebrafish had higher levels of potassium (K) and sulfur (S) in their cartilages. Furthermore, there was an increase in perichondral bone in the hyomandibula (hm) of the putative transgenic zebrafish. Those phenotypes may be related to ARSI function if this protein is in fact being overexpressed, but are more likely due to Tg fish degrading/counteracting overexpressed ARSI to go back to a homeostatic state.

4.2 Introduction

In the previous chapters, it was shown that ARSI is expressed during cartilage maturation *in vivo* in mouse and chick and *in vitro* in ATDC5 cells. Besides that, biochemical characterization revealed that ARSI is a CSPG sulfatase. CSPGs are one of the main components of cartilage ECM. Modifications in their sulfation could change the timing of chondrocyte maturation and consequent bone formation(Cortes et al., 2009; Kluppel et al., 2005; Orkin et al., 1977; Schwartz & Domowicz, 2002). However, how ARSI acts during these biological processes is still unknown. One significant way to test for gene function is to create transgenic and mutant animals for a given gene and correlate observed phenotypes to molecular pathways (Simmons, 2008). For this work, *Danio rerio* (zebrafish) was selected as the model organism to interrogate ARSI function. This is due to their rapid generation time, the possibility of using Gateway Tol2-system to generate lines expressing ARSI specifically in cartilage, and other advantages discussed at length in the review paper in chapter 2 (Machado & Eames, 2017).

A whole genome duplication event happened approximately 340 million years ago in an ancient common ancestor of all living teleost, which includes zebrafish (Postlethwait et al., 2000). However, it is estimated that only about 20% of zebrafish genes have two functional copies since the most common fate of duplicated genes (50–90%) is nonfunctionalization (Postlethwait et al., 2004). Other fates of duplicated genes include neofunctionalization and subfunctionalization, which might make it complex to investigate gene function in these animals (Glasauer & Neuhauss, 2014). Nonetheless, there is a high similarity between the zebrafish and the human genome (71% of human genes have orthologs in zebrafish) and a high homology of genes between those species. Therefore, zebrafish have been used to overexpress human genes and study their function (Kabashi et al., 2011; Sakai et al., 2018).

Zebrafish have two copies of the sulfatase gene *arsi* (*arsia* and *arsib*). Despite lacking a known biological function, their sequences are documented in GenBank (NCBI). Sulfatases are widely conserved across vertebrates, and comparisons between the full-length ARSI protein in zebrafish and humans revealed more than 65% of amino acid conservation (Figure 3.7B). This conservation is even more evident in the sulfatase domain, which contains the putative active sites, and is more than 75% similar between zebrafish and humans (Figure 3.7C). Considering zebrafish have been used to study human gene function, and the high degree of conservation of sulfatases across vertebrates, the overexpression of human ARSI (hARSI) in zebrafish could be a first step towards uncovering the biological function of this gene.

The zebrafish has emerged as a valuable model organism for studying bone development due to the presence of similar genes, cells, and developmental processes in the skeleton of this species and humans (Giovannone et al., 2019; Le Pabic et al., 2022; Machado & Eames, 2017). Although with some differences from mammals, endochondral ossification can be observed in certain zebrafish elements such as the ceratohyal (ch) and hyomandibula (hm) (Tonelli et al., 2020; Weigele & Franz-Odendaal, 2016). In these elements, in earlier developmental stages, the growth plate contains immature and mature cartilage divided into zones of resting, proliferative, pre-hypertrophic and hypertrophic chondrocytes, similar to mammals (Weigele & Franz-

Odendaal, 2016). Furthermore, in zebrafish chondral bones, as with other vertebrates, perichondrium cells differentiate into osteoblasts and deposit perichondral bone around the mature cartilage region (Eames et al., 2011). Thus, the presence of perichondral bone indirectly indicates cartilage maturation and progression of endochondral ossification.

This chapter aims to test our hypothesis that ARSI decreases sulfate ester levels in mature cartilage thus promoting endochondral ossification. Our main objectives are to generate Tg zebrafish lines overexpressing hARSI and to assess skeletal development and sulfur levels in the putative Tg zebrafish. ARSI Tg fish are expected to have lower amounts of sulfur and higher levels of perichondral bone compared to their WT siblings during endochondral ossification. Since no available Tg or mutant animal was previously described for ARSI, this chapter has a strong methodological component. Here, I describe the steps and outcomes obtained while generating putative hARSI Tg zebrafish using the Gateway Tol2-system.

4.3 Methods

4.3.1 Zebrafish

Wild-type AB adult zebrafish were maintained under controlled conditions at the Lab Animal Services Unit (LASU) at the University of Saskatchewan. The protocols used for this research were approved by the University Animal Care Committee Animal Research Ethics Board (UACC AREB) at the University of Saskatchewan (Animal Use Protocol; AUP# 20120068).

4.3.1 Gateway-Tol2 reactions

The GatewayTol2kit system employs site-specific recombination-based cloning to facilitate rapid and modular assembly of constructs, comprising [5'promoter]–[coding

sequence]–[3' tag], within a Tol2 transposon backbone. Five Gateway constructs were constructed and injected in zebrafish eggs: *col2a1a_EGFPpA*, *col2a1a_hARSI_EGFPpA*, *col2a1a_hARSI_pA*, β*actin_hARSI_EGFPpA*, and β*actin_hARSI_pA*. The *hARS*I gene was received from Dr. Hiroshi Kitagawa, and it was cloned as previously described in section 3.3.13.

PCRs were performed with Phusion® High-Fidelity DNA Polymerase (NEB-M0530S) and the following primers to add attB1 and attB2 sites to *hARSI*: *hARSI*-Fwd: GGGGACAAGTTTGTACAAAAAAGCAGGCTTCgccaccatgcacaccctcactg; *hARSI*-Rv: GGGGACCACTTTGTACAAGAAAGCTGGGTTtcagatccgttgggacattag. To generate the constructs with *EGFP* fused to *hARSI*, a different reverse primer was used to remove the stop codon from the *hARSI* sequence: *hARSI-EGFP*-Rev: GGGGACCACTTTGTACAAGAAAGCTGGGTTgatccgttgggacattagcc. Bands with the expected sizes were extracted from the agarose gel with a QIAquick Gel Extraction Kit (Qiagen). Then, the PCR products were sent for sequencing.

BP reactions

After sequence confirmation, PCR products containing attB sites were cloned into pDONR221 vectors (vectors used to capture a gene or gene fragment of interest containing attP sites) through BP reactions following the Gateway technology manufacturer's manual (Invitrogen). BP reactions are Gateway reactions that use BP clonase enzymes to recombine attB sites surrounding our gene on interest with attP sites present in the pDONOR221 vector. These enzymes allowed for recombination and generated attL sites surrounding our gene on interest in the two *hARSI* middle entry vectors (pME) (Figure 4.1 A). Each BP reaction contained 50 ng of a specific purified PCR product, 150ng of pDONR221, 1 µL of BP enzyme, and TE buffer to generate a 5µl reaction. BP reactions were left overnight at room temperature and treated with 1 µl of proteinase K (2 µg/µl) for 10 min at 37°C the next day. The product of the BP reactions was then transformed into One Shot™ TOP10 chemically competent *E. coli* (Invitrogen), plated on LB plates with kanamycin (50 µg/mL), and left upside down at 37°C overnight. On the next day, 5 colonies were collected in five separate tubes

containing liquid LB with kanamycin (50 µg/ml) and put in a shaking incubator at 37°C overnight (210 rpm). Plasmids were then purified with the QIAprep Spin Miniprep Kit (Qiagen), and enzymatic digestions and DNA sequencing were used to confirm sequences of the two entry clones.

LR reactions

The two entry clones generated by BP reactions were recombined with different 2 different enhancer (inserted in 5'-entry clones), and with a poly A tail or with an EGFP fused to a polyA tail (inserted into 3'-entry clones). these RECOMBINATIONS, between 5'-entry clones, middle entry clones, and 3' entry clones into a destination vector, were achieved through LR reactions. Destination vectors are commercially available vectors which contain the Gateway recombination sites, allowing rapid and efficient transfer of sequences using site-specific recombination. LR reactions are Gateway reactions where the Gateway™ LR Clonase™ II enzyme mix catalyzes the *in vitro* recombination between entry clones (containing a gene of interest flanked by *att*L sites) and a destination vector (containing *att*R sites) to generate an expression clone. The 5'-entry clone containing the β-actin enhancer (*p5E*-β*actin2*) and the 3'entry-clones containing EGFP plus poly A (*p3E-EGFPpA*) and just polyA (*p3E-polyA*) were obtained from Invitrogen and were described elsewhere (Kwan et al., 2007). In addition, the 5'-entry clone containing the *col2a1a* enhancer was obtained from Dr. Rodney Dale (Dale and Topczewski, 2011). Multi-site gateway LR reactions were performed using LR clonase and different combinations of the previously described entry clones, which were recombined into pDestTol2CG2 destination vectors according to the manufacturer's instructions (Invitrogen). Briefly, each LR reaction contained 1 µl of each entry clone (100-300 ng), 2 µl of the pDestTol2CG2vector (150 ng/µL), 2 µl of LR clonase, and 3 µl of TE buffer. LR reactions were left overnight at room temperature and treated with 1 µl of proteinase K (2 µg/µl) for 10 min at 37°C on the next day. The product of the LR reactions was then transformed into One Shot™ TOP10 chemically competent E. coli (Invitrogen), plated on LB plates with ampicillin (100 µg/mL), and left upside down at 37°C overnight. On the next day, 5 colonies were collected in five separate tubes containing liquid LB with ampicillin (50 µg/ml) and put in a shaking incubator at 37°C

overnight (210 rpm). Plasmids were then purified with the QIAprep Spin Miniprep Kit (Qiagen), and sequences of the 4 final constructs were confirmed by enzymatic digestions and DNA sequencing.

The *col2a1a_EGFPpA* construct was previously generated in the lab by recombination of the same 5' entry clone containing the *col2a1a* enhancer, a middle entry clone containing *EGPF*, and a p3E-polyA into a pDestTol2CG2 destination vector. The *in silico* BP and LR reactions were simulated using the Serial Cloner software, using the recombination tool. The *in silico* recombination produced the anticipated sequences present in the plasmids. Subsequently, the BP and LR products were sent for DNA sequencing, and the obtained DNA sequences were compared to the anticipated *in silico* results to assess accuracy. Also, in Serial Cloner, the virtual cut tool was used to select enzymes that would generate specific cutting sites in each construct. Enzymatic digestions were performed with different NEB enzymes following the manufacturer's instructions. The generated middle entry clones and final constructs were sent for DNA sequencing at Eurofins. Sequencing results were aligned with the *in silico* sequenced generated at Serial Cloner.

4.3.2 Injection of constructs, screening, and establishment of lineages

WT fish were incrossed, and fertilized one-cell-stage eggs were injected with the generated different DNA constructs (100 ng/µL) along with transposase mRNA (50 ng/µL) and phenol red (0.050%). Injected embryos and their WT siblings were kept in embryo media at 28.5°C. Zebrafish were screened for GFP expression in their hearts between 3 and 4 dpf, raised to adulthood, and outcrossed to WT fish to identify founders (fish able to transfer the transgene to next generations). The founders were outcrossed with WT fish, and positive F1 embryos were raised to establish non-mosaic stable Tg lines.

4.3.3 Genomic DNA extraction and genotyping

Zebrafish embryos were collected into PCR tubes in pools of 10 at 6 dpf days post fertilization. First, 50 µl of alkaline lysis buffer (25 mM NaOH, 0.2 mM EDTA, pH 12) was added to each PCR tube. Next, the tubes were placed into a thermocycler at 95℃ for 25 minutes, cooled down on ice for 15 minutes, and 50µl of neutralizing buffer (40mM Tris-HCl pH 5) was added to each tube. To check for genomic integration of the constructs, primers to amplify the junctions between *col2a1a* and *hARSI* and *hARSI* and *EGFP* sequences were designed, and PCRs were performed. PCR conditions were as follows: hot start at 95℃, one cycle at 95ºC for 3 min as initial denaturation, followed by 40 cycles of 30 seconds at 95℃ for denaturation, 30 seconds at 56℃-60℃ as annealing, and 40 seconds at 72℃ for extension, and a final extension at 72ºC for 10 min. The PCR product was visualized on a 1% agarose gel run with 1xTAE buffer. Bands of the expected size were extracted extracted and purified using the QIAquick® Gel Extraction kit (Qiagen). The purified product was cloned into the pCR™4-TOPO™ vector (Thermo Fisher), transformed into One Shot™ TOP10 cells (Invitrogen), and grown overnight at 37°C on LB agar plates containing 100 ug/ml ampicillin. The next day, colonies were selected and grown in 5 mL of liquid LB (100ug/ml of ampicillin) overnight in a shaker at 37°C. Plasmids were purified using the QIAprep® Spin Miniprep Kit (Qiagen) and sent to Eurofins for DNA sequencing. Sequencing results were analyzed by aligning the Tg sequencing product with the construct sequence predicted by Serial Cloner.

4.3.4 RT-PCRs

Reverse transcription PCR (RT-PCR) was used to detect transcripts derived from the constructs in the Tg fish. For that, zebrafish were screened by the presence of green hearts, and RNA was extracted from putative Tg and WT siblings using the RNeasy Mini Kit (Qiagen). Reverse transcription was then performed at 42ºC using RevertAid H Minus First Strand cDNA Synthesis Kit (Thermo Fisher). Specific primers for construct regions were designed based on the sequencing results obtained after LR

reactions. RT-PCR conditions were as follows; hot start at 95ºC, one cycle at 95ºC for 3 min as initial denaturation, followed by 40 cycles of 30 seconds at 95℃ for denaturation, 30 seconds at 56℃-60℃ as annealing, and 40 seconds at 72℃ for extension, and a final extension at 72ºC for 10 min. The PCR product was visualized on a 1% agarose gel run with 1xTAE buffer. DNA fragments were purified and cloned into a pCR™4-TOPO® vector (Thermofisher). Cloned DNA fragments were transformed into E.coli cells (OneShot, Invitrogen), selected for with ampicillin and cultured overnight in LB liquid medium. The plasmid DNA was extracted and purified to produce minipreps, which were sequenced and aligned with expected sequences.

4.3.5 Zebrafish skeletal staining

Alcian blue and Alizarin red staining of zebrafish skeletons was performed as previously described (Eames et al., 2011). Alcian blue stains the sulfated PGs, and Alizarin red stains mineralized tissues. Following the staining, zebrafish ceratohyal (ch) and hyomandibula (hm) were scored on a scale of 0-3 based on the amount of perichondral bone (Figure 4.16B) (DeLaurier et al., 2019; Zare Mirakabad et al., 2019; Zare Mirakabad & Khorramizadeh, 2022). The bone scores were compared between Tg and WT siblings at 8 dpf.

4.3.6 Alcian Blue quantitation

To assess if levels of sulfated glycosaminoglycans changed in Tg zebrafish, Alcian blue levels were quantified in Tg and WT siblings as described previously with small modifications (Eames et al., 2011). Briefly, 8 dpf zebrafish whole skeletons were stained with Alcian blue, lysed (DNeasy kit; QIAGEN Inc.), spun briefly, and measured spectrometrically at 620 nm (BioTek Synergy HT Plate Reader). Each well contained lysate from 10 fish. The experiment was performed in triplicate, with fish from 3 different clutches analyzed for the total number of replicates.

4.3.7 Synchrotron imaging and data processing

XRF data collection at the VESPERS beamline was performed as described in section 3.3.4. Each ch slide (8 thick) was raster-scanned with a 5 µm step size with a 10-second dwell time per pixel, generating XRF maps. Ch maps had approximately 200x100 µm in size, resulting in scans of approximately 3 hours per sample. Besides Sulfur, other channels for different elements, like Ca, P, K, Cl, Fe, Cu, and Zn, were collected for each pixel of the sample. In total, 15 ch samples were measured at this beamline. Chemically specific XRF data was collected at SXRMB as described in 3.3.4. Each ch was raster-scanned with a 7 µm step size with a 1-second dwell time per pixel. Ch maps had approximately 200 x100 µm in size, taking around 2 hours each, totalling 8 hours/sample. In total, 14 ch were imaged at this beamline, all of which were also imaged at the VESPERS beamline. Data processing was performed as previously described in 3.3.4. But images were quantified initially at 16-bit and then at 32-bit, and measurements were compared.

4.3.8 Immunohistochemistry (IHC)

IHC with anti-ARSI (Sigma - HPA038398) and anti-ColII (II-II6B3 - DSHB) antibodies was performed in whole-mount and sectioned WT and putative Tg zebrafish embryos at different time points. Whole-mount IHC was performed in 2, 3, 4, 5, 6, and 8 dpf zebrafish similarly to previously described protocols (Eames et al., 2010). Zebrafish embryos were fixed overnight at 4°C in 4% PFA/1X PBS and washed three times in PBST for 5 min. Fish were then digested in 0.1% Trypsin (MP Biomedicals; Santa Ana, CA)/1mM EDTA/1X PBS for 10 (2 and 3 dpf), 20 min (4 and 5 dpf), and 30 min (6 and 8 dpf) at 37°C, washed three times in PBST for 5 min, digested with 0.5% hyaluronidase (Worthington Biochemical Corporation; Lakewood, NJ)/PBST for 10 (2 and 3 dpf), 20 min (4 and 5 dpf), and 30 min (6 and 8 dpf) at 37°C, and then put in block solution (4% Goat Serum/2% Sheep Serum/0.5% TritonX-100/1X PBS) for 2 hours. Samples were incubated overnight at 4°C with anti-ARSI (1:300) or anti-ColII (1:100) antibodies in

block solution, washed 6 times for 15 min in block solution and incubated with the secondary antibodies Alexa Fluor 488 goat anti-rabbit IgG (for ARSI) or Alexa Fluor 488 goat anti-mouse IgG (for ColII) (1:1000) (Thermofisher; Waltham, MA) overnight at 4°C. Lastly, samples were washed 6 times for 20 min in PBST and imaged under fluorescence light. This same whole-mount protocol was also performed in dissected jaws of 8 dpf zebrafish.

For section IHC, 8 dpf zebrafish were embedded in OCT (optimum cutting temperature) compound (Tissue-Tek; Sakura Finetek USA, Torrance, CA), using standard cryotemplates (Tissue-Tek; Sakura Finetek USA, Torrance, CA). Samples were flash-frozen in liquid nitrogen-cooled isopentane and sectioned (5 µm). Frozen sections were allowed to air dry for 10 min, post-fixed in 4% PFA/1xPBS at room temp for 20 mins and rinsed twice in PBST for 5 mins. Slides were then digested with 0.1% trypsin/1 mM EDTA/1X PBS for 30 mins at 37 °C in a humidity chamber, rinsed twice in PBST for 5 mins each, digested with 0.5% hyaluronidase/PBST for 30 mins, rinsed twice in PBST for 5 mins each, blocked with blocking solution for 1 hour, and then incubated overnight at 4°C with anti-ColII (1:100) or hARSI (1:300) antibodies diluted with blocking solution. Slides were then rinsed three times with PBST for 2 mins each, incubated in the dark with fluorescent-labelled goat anti-rabbit or goat anti-mouse secondary antibodies in blocking solution (1:1000) overnight at 4°C, rinsed three times with PBST for 2 mins each, stained with 300 nM DAPI/PBST for 10 mins, rinsed twice with PBST for 5 mins each, and then mounted for fluorescent microscopy imaging.

4.3.9 Western Blot

Fish injected with all 4 hARSI constructs and their WT siblings were collected for Western blot. 15 fish were pooled into each tube and were dissociated on Laemmli buffer after the addition of Halt™ Protease and Phosphatase Inhibitor Cocktail (Thermo Fisher) at 95°C with the help of a syringe. Samples were treated as described in item 3.3.12. ATDC5 cells were used as a positive control to compare to zebrafish samples.

4.3.10 Bafilomycin treatments

Putative *col2a1a_hARSI_EGFPpA* Tg fish and their WT siblings were treated with two concentrations of Bafilomycin (Sigma B1793). WT and Tg fish were screened for green hearts at 2 dpf and divided into three 75mm plates containing around 20 Tg and 20 WT fish each. At 55hpf, the embryo media in the 3 plates was replaced with solutions containing 50nM bafilomycin, 20nM Bafilomycin, or embryo media with 0.1% DMSO (control group). The solutions were refreshed every day for three days, and fish were observed under fluorescence to check for changes in GFP expression.

4.3.11 SUMF1 injections

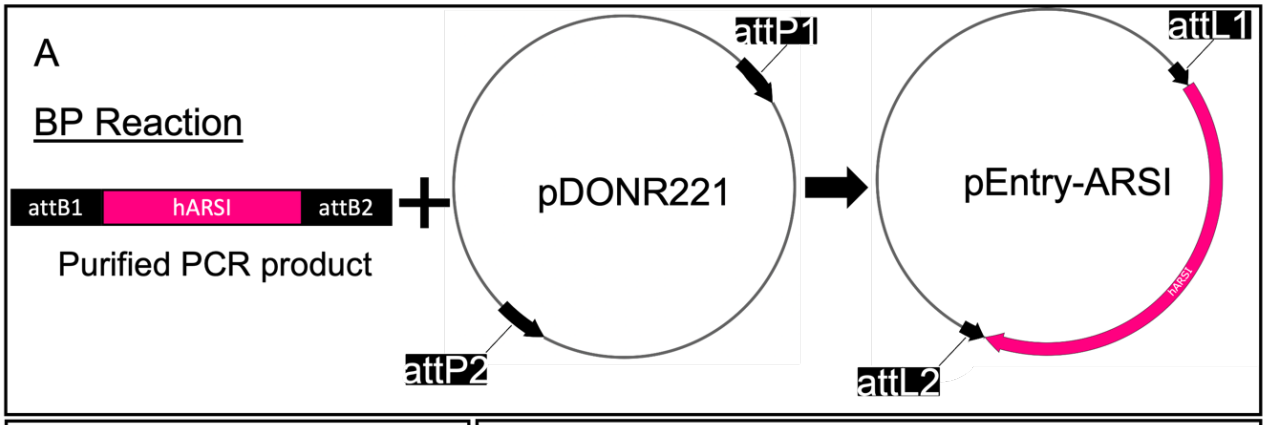
The *SUMF1* CDS sequence (transcript variant 1 - NM_182760.4) was obtained from, Biobasics in a pBluescript II-KS(+) plasmid. *SUMF1* capped RNA was generated by *in vitro* transcription using the mMESSAGE mMACHINE™ T3 Transcription Kit (AM 1348 – Thermofisher). RNA was purified with the RNeasy Mini kit (Qiagen), RNA concentration was measured in the nanodrop, and samples were stored at -80ºC. On the day of the injection, the RNA sample (100ng/µl of *SUMF1)* was vortexed, spun down briefly and mixed with sterile water and phenol red (0.050%). Putative *col2a1a_hARSI_EGFPpA* Tg were outcrossed with WT fish, and 1-cell stage fertilized eggs were injected with the mixture. Tg and WT uninjected eggs from the same crossings were used as uninjected controls.

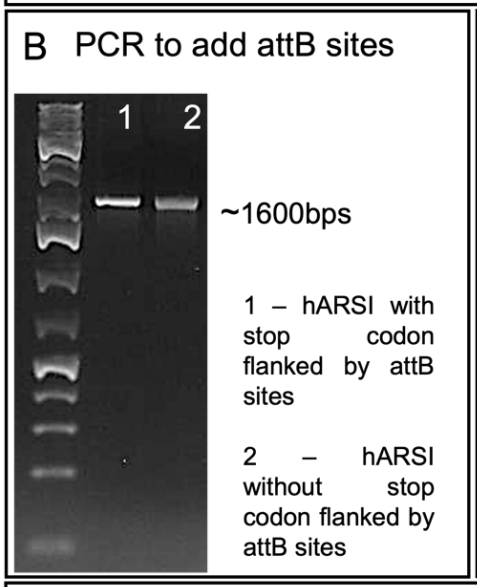
4.4 RESULTS

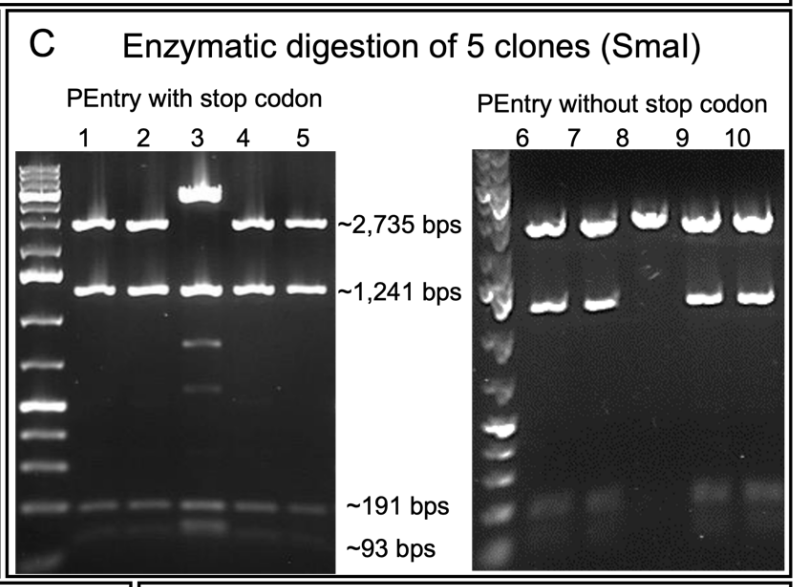
4.4.1 Generation of hARSI expression constructs using the Gateway Tol2-system.

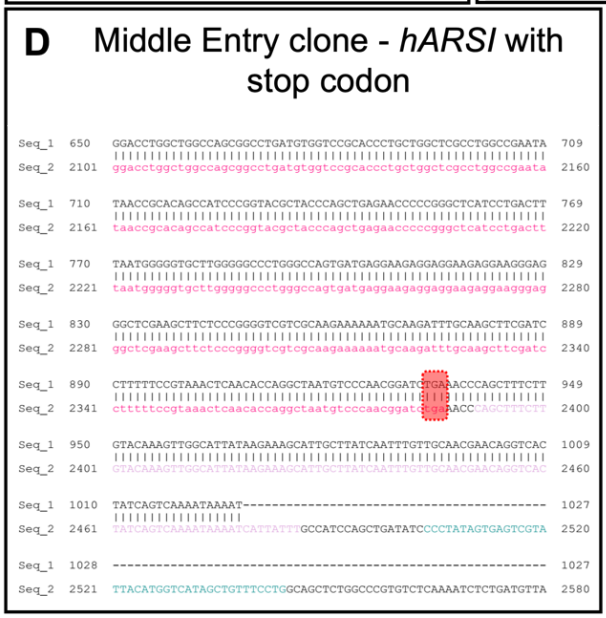
To study ARSI biological function, the present experiments focused on overexpressing hARSI in zebrafish using the Gateway-Tol2 system. Overexpression was attempted in WT zebrafish with endogenous *arsia* and *arsib* described in their genomes. Protein alignments showed that the ARSI catalytic core and 100% putative catalytic sites were conserved in both zebrafish and humans. Moreover, the sulfatase domain of ARSI showed over 74% conservation between the two copies of zebrafish arsi and human ARSI (Figure 3.7), indicating that protein function is likely conserved between species.

To create *hARSI* entry clones for zebrafish overexpression studies, two sets of primers were designed to add recombination sites (attB sites) surrounding *hARSI* cDNA. The first set of primers maintained the original *hARSI* sequence, while the second set removed a stop codon at the end of the *hARSI*. Amplification of the *hARSI* using both sets of primers resulted in a single specific band on a 1% agarose gel. The *attB-hARSI* purified PCR products were then recombined with a donor vector (pDONR221) in the BP recombination reaction step of the Gateway protocol, generating two *hARSI* entry clones (Figure 4.1A and B). The presence of *hARSI* in the entry clones was confirmed by enzymatic digestions, where the expected band pattern (2,735 bps,1,241 bps, 191 bps, and 93 bps) was observed in 4 out of 5 clones digested with SmaI enzyme (Figure 4.1C). Entry clones confirmed by enzymatic digestion were sent for DNA sequencing. Results showed that two middle entry clones containing *hARSI* with and without a stop codon were successfully generated through Gateway BP reactions (Figure 4.1D and E).


A
BP Reaction
attB1
hARSI
attB2
Purified PCR product
pDONR221
attP1
attP2
attL1
attL2
pEntry-ARSI
hARSI


B PCR to add attB sites
1
2
~1600bps
1 – hARSI with stop codon flanked by attB sites
2 – hARSI without stop codon flanked by attB sites


C Enzymatic digestion of 5 clones (SmaI)
PEntry with stop codon
1 2 3 4 5
PEntry without stop codon
6 7 8 9 10
~2,735 bps
~1,241 bps
~191 bps
~93 bps


D Middle Entry clone - hARSI with stop codon

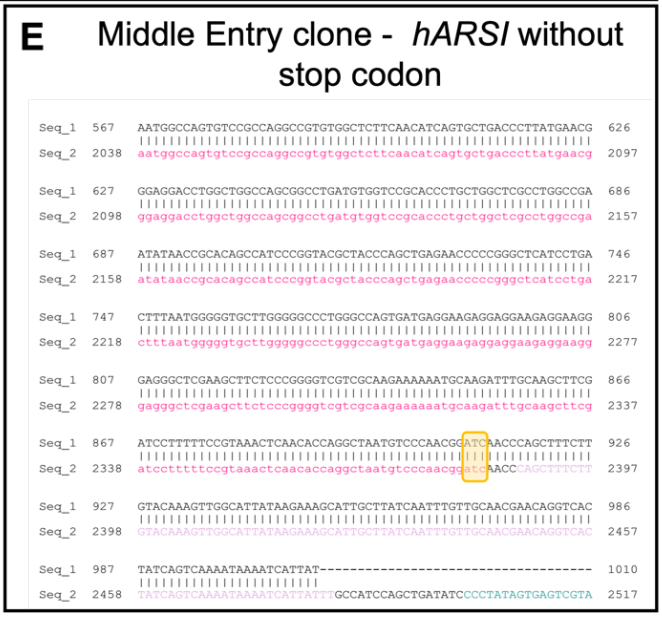

E Middle Entry clone - hARSI without stop codon

Figure 4.1 Two entry clones containing *hARSI* were generated through Gateway BP reactions. (A) Purified PCR products containing recombination sites surrounding the coding region of the *hARSI* gene were recombined with a pDONR vector through BP reactions. (B) High-fidelity PCR was performed to add attB recombination sites to hARSI. Bands in the agarose gel show amplification of DNA in the expected sizes. (C) Enzymatic digestion of minipreps from bacteria transformed with the 2 middle entry clones. 4 out of 5 clones with the stop codon (1, 2, 4, and 5) and 4 out of 5 clones without the stop codon (6,7,9 and 10) had the expected fragment sizes. (D-E) Sequencing results of two different middle entry clones. Seq_1 is the sequence obtained from the sample, and Seq_2 is the presumptive sequence. ARSI sequence (in pink) is followed by part of the PDONr sequence (light purple). (D) The first middle entry clone had a stop codon at the end of the *hARSI* (TGA sequence in the red box). (E) The second middle entry clone was generated without the stop codon (yellow box shows the codon preceding the original stop codon), so it could be fused with an *EGFP* sequence.

Gateway technology allows for a modular assembly of different expression constructs. Using LR reactions, the two hARSI entry clones generated in the previous step were recombined with 5' entry clones containing either a col2a1a or a ß-actin enhancer, and with 3' entry clones containing an EGFP sequence fused to a polyA tail or a polyA tail alone. These sequences were recombined into a destination vector (pDestTol2CG2) (Figure 4.2A). Four expression constructs were generated, two containing a col2a1a enhancer which drives gene expression in cartilage, and two with a ß-actin enhancer which drives gene expression in all zebrafish cells. Constructs containing hARSI without a stop codon were recombined with a 3' entry clone containing an EGFP sequence fused to a polyA tail for the visualization of EGFP in the same regions as hARSI expression (Figure 4.2B). Enzymatic digestions were performed to confirm the correct generation of expression constructs. Two enzymes were used to digest each construct, and the expected banding pattern was observed for all constructs in multiple clones (Figure 4.2C-F). Expression constructs confirmed by enzymatic digestion were sent for DNA sequencing to check for the recombination sites (Figure 4.2D and E). Results confirmed that the four expression constructs had the expected DNA sequences after LR reactions.

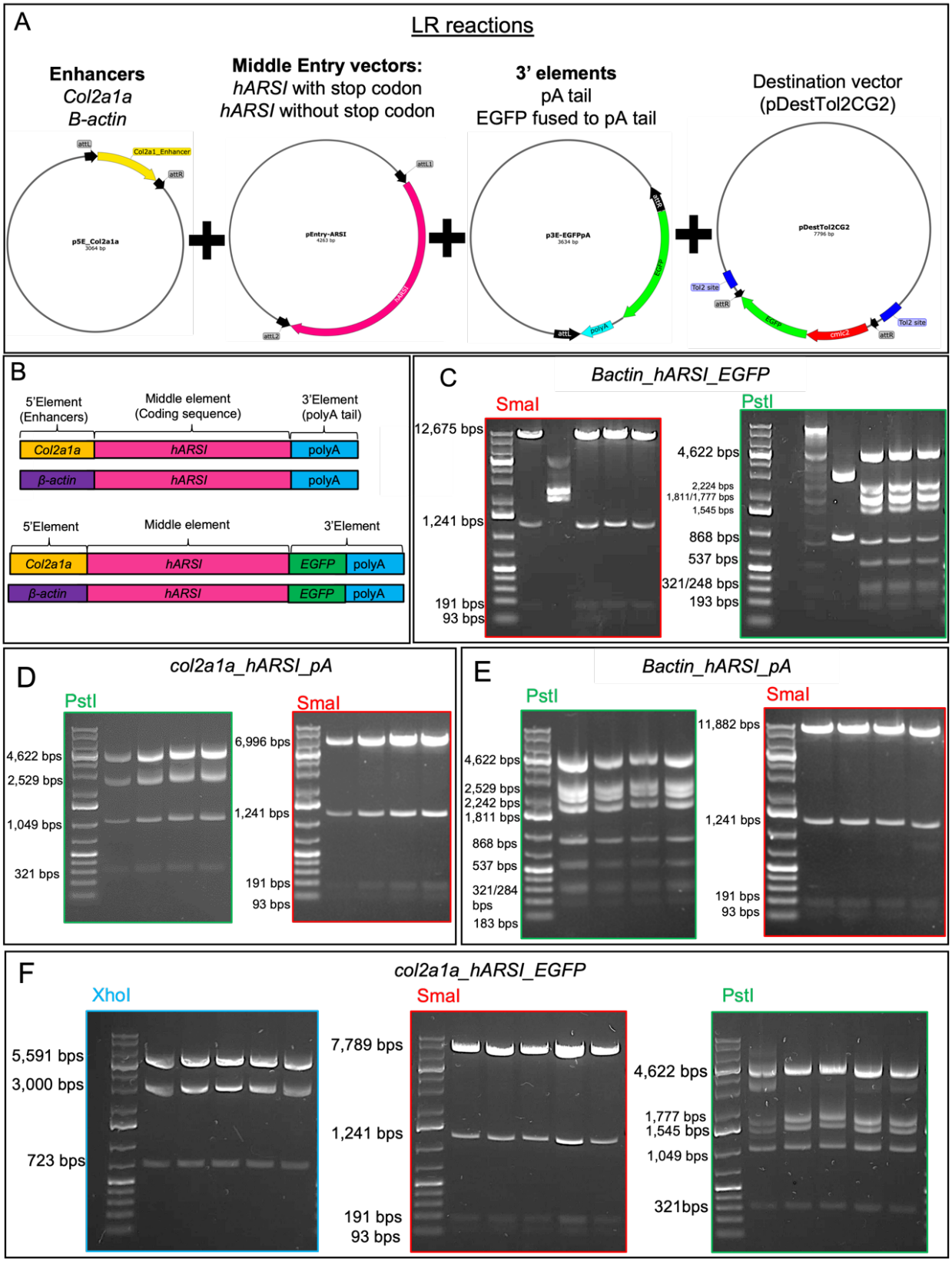

A
LR reactions
Enhancers
Col2a1a
B-actin
Middle Entry vectors:
hARSI with stop codon
hARSI without stop codon
3' elements
pA tail
EGFP fused to pA tail
Destination vector
(pDestTol2CG2)
B
5'Element
(Enhancers)
Middle element
(Coding sequence)
3'Element
(polyA tail)
Col2a1a
hARSI
polyA
β-actin
hARSI
polyA
5'Element
Middle element
3'Element
Col2a1a
hARSI
EGFP
polyA
β-actin
hARSI
EGFP
polyA
C
Bactin_hARSI_EGFP
SmaI
12,675 bps
1,241 bps
191 bps
93 bps
PstI
4,622 bps
2,224 bps
1,811/1,777 bps
1,545 bps
868 bps
537 bps
321/248 bps
193 bps
D
col2a1a_hARSI_pA
PstI
4,622 bps
2,529 bps
1,049 bps
321 bps
SmaI
6,996 bps
1,241 bps
191 bps
93 bps
E
Bactin_hARSI_pA
PstI
4,622 bps
2,529 bps
2,242 bps
1,811 bps
868 bps
537 bps
321/284 bps
183 bps
SmaI
11,882 bps
1,241 bps
191 bps
93 bps
F
col2a1a_hARSI_EGFP
XhoI
5,591 bps
3,000 bps
723 bps
SmaI
7,789 bps
1,241 bps
191 bps
93 bps
PstI
4,622 bps
1,777 bps
1,545 bps
1,049 bps
321bps

Figure 4.2 Four expression constructs were generated to overexpress *hARSI* in zebrafish.(A) Gateway LR reactions were used to recombine the *hARSI* entry clones with two different enhancers (*col2a1a* and *ß-actin*) and two different 3' elements (poly A tail and EGFP fused to poly A tail) into a destination vector. (B) Summary of the four generated constructs after LR reactions. (C) Enzymatic digestion of the final *ß-actin_hARSI_EGPF* construct with two restriction enzymes (SmaI and PstI). Four out 5 minipreps had the expected band pattern after SmaI digestion, and 3 out of 5 had the expected band pattern after PstI digestion. (D) Enzymatic digestion of the final *col2a1a_hARSI_pA* construct with SmaI and PstI. All 5 minipreps had the expected band pattern after enzymatic digestion. (E) Enzymatic digestion of the final *ß-actin_hARSI_EGPF.* construct with SmaI and PstI. All 5 minipreps had the expected band pattern after enzymatic digestion. (F) Enzymatic digestion of the final *col2a1a_hARSI_pA* construct with SmaI, PstI, and XhoI. All 5 minipreps showed the expected band pattern after enzymatic digestions, except for miniprep 1 with PstI (first gel lane - probably due to incomplete degistion).

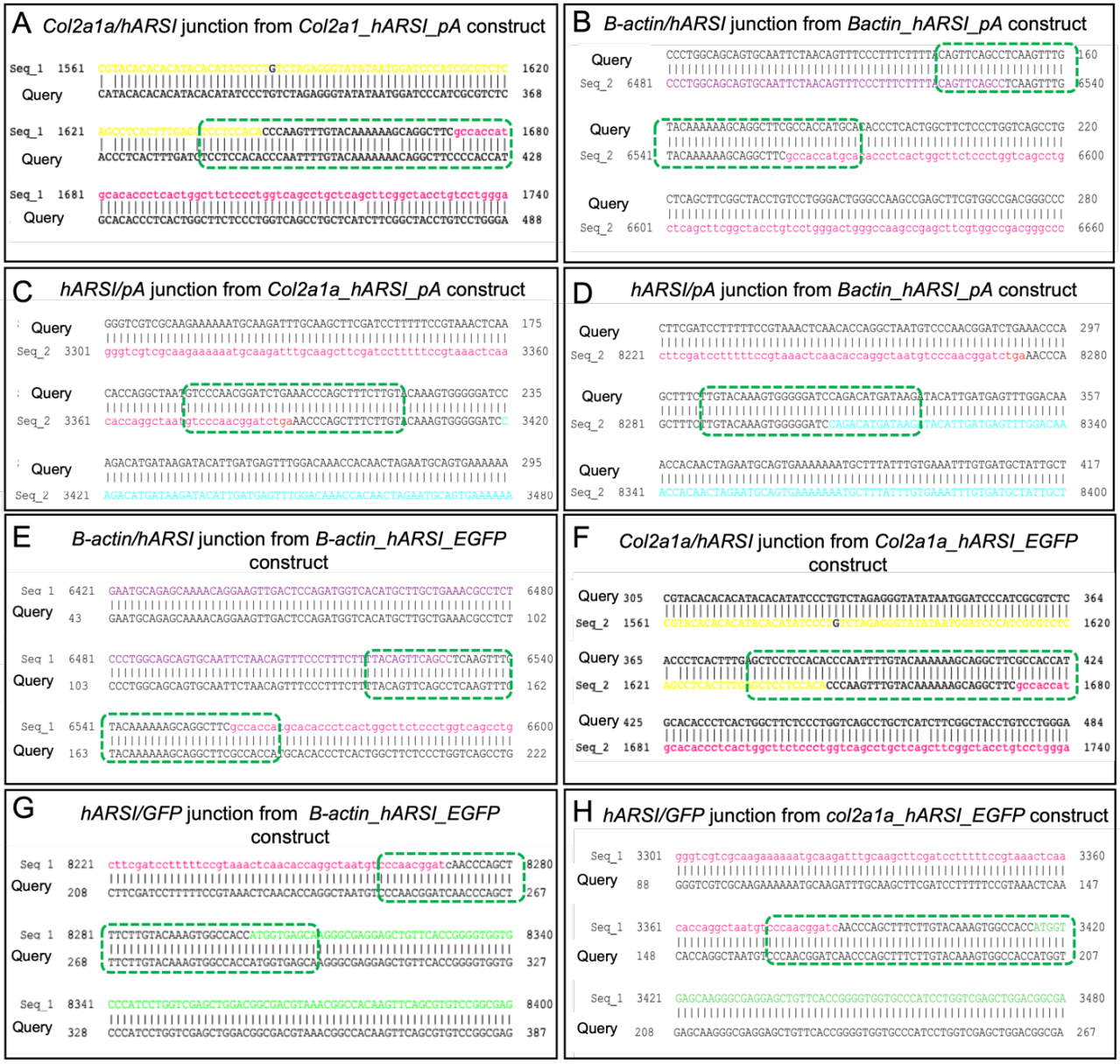


Figure 4.3 Sequencing results of recombination sites of constructs before injection. (A-H) Primers were designed to sequence recombination sites between the enhancers and the *hARSI* gene, the *hARSI* and the pA tail, and between *hARSI* and *EGFP*. Final constructs that showed the expected band pattern after enzymatic digestion were sent for sequencing. All the junctions between enhancers, hARSI, and eGFP 3' elements were observed in the sequencing results (green rectangles). Colours in the sequencing results: *hARSI* in pink; *col2a1a* enhancer in yellow; *ß-actin* enhancer in purple; and *EGFP* in green.

4.4.2 hARSI expression constructs were integrated into the zebrafish genome

All four gateway expression constructs were individually microinjected in zebrafish fertilized eggs at the one-cell stage, together with a transposase enzyme, which allows for the integration of the constructs into the zebrafish genome. It is important to notice that the construct sequences were followed by a secondary EGFP sequence controlled by a cmlc2 enhancer. This enhancer is activated only in cardiomyocytes and is commonly used to screen for genomic integration of the transgenes. The *cmlc2_EGFP* sequence was present in the backbone of the destination vector, and it was located within the transposase cutting sites (Figure 4.4A). After the integration of the constructs with the *cmlc2_EGFP* sequence into the zebrafish genome, mosaic EGFP fluorescence was observed in the zebrafish hearts (Figure 4.4 B and C). Around 15-20% of the injected embryos had mosaic EGFP expression in their hearts and were raised into adults. Those zebrafish that had the insertion of the transgene into their germ cells and could transfer it to the next generation were considered founders. At least 3 founders of each construct were outcrossed with WT zebrafish, and larvae with a non-mosaic expression of the transgene (full green hearts) were raised to create stable F1 lineages (Figure 4.4C and D). The *col2a1a_hARSI_EGFP* and *ßactin_hARSI_EGFP* lines were designed to overexpress *hARSI* fused with *EGFP*, so we anticipated seeing EGPF fluorescence in the zebrafish cartilages and their whole body, respectively. Unexpectedly, EGFP expression was just observed in the zebrafish hearts in all four lines.

To confirm the integration of the *col2a1a_hARSI_EGFP* construct into the zebrafish genome, PCR and DNA sequencing were conducted with the second generation of putative Tg zebrafish (F2). PCR with primers designed to amplify the construct region between *col2a1a* and *hARSI* showed a band in the expected size (510 bps) in Tg fish from three different founders but not in their WT siblings (Figure 4.5A). DNA sequencing of the cloned PCR products confirmed the presence of the *col2a1a_hARSI* junction in the genomic DNA of the Tg zebrafish (Figure 4.5B). Other primers were designed to amplify the *hARSI_EGFP* junction, and again a band in the

expected band size (around 300 bps) was observed in the Tg fish but not in their WT siblings (Figure 4.5C). DNA sequencing results confirmed that the *hARSI_EGFP* junction was present in the Tg zebrafish genomic DNA (Figure 4.5D).

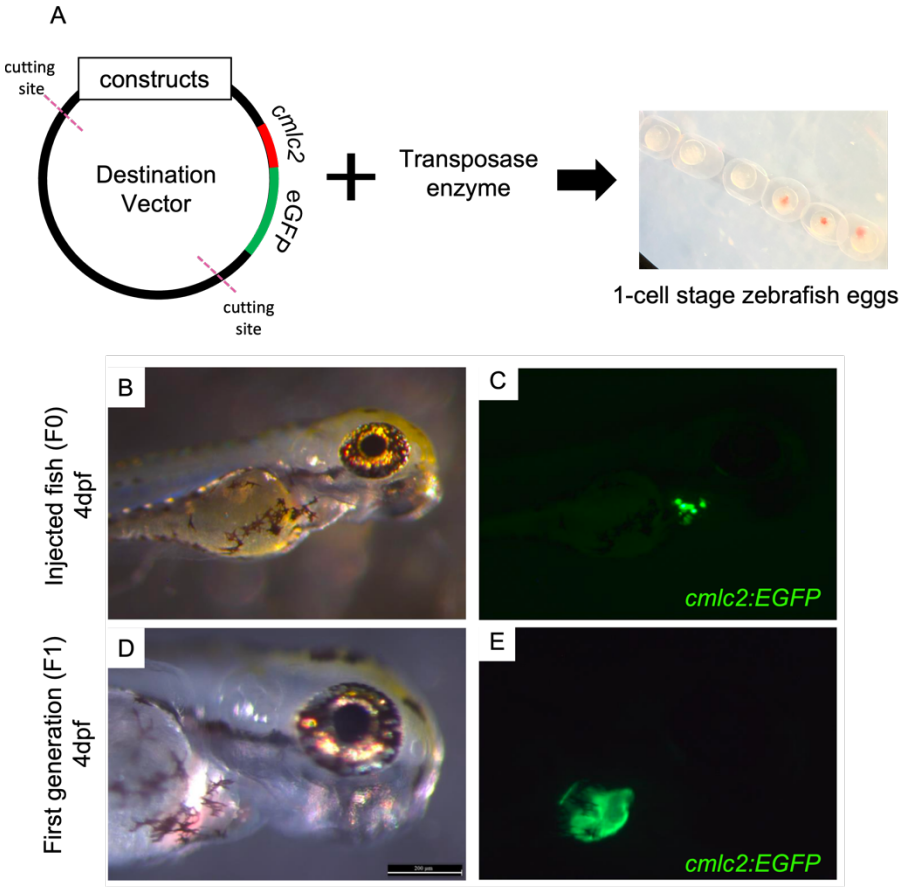


Figure 4.4 Four different putative zebrafish transgenic lines were created. A) All four constructs were injected individually with a transposase enzyme in one cell stage WT zebrafish eggs. Besides *col2a1a* or *ß-actin* enhancers, all constructs had a *cmlc2* enhancer in their destination vector which drives *EGFP* expression in the zebrafish hearts. The transposase enzyme catalyzes excision of the construct plus the *cmlc2_EGFP* sequence from the plasmid. (B-C) Injected fish with a mosaic *EGFP* expression in their hearts by 3 dpf were raised and crossed to WT to create the F1 of the four transgenic lines. (D-E) First generation of zebrafish show a complete green heart, indicating a non-mosaic expression. All four zebrafish lines had green hearts, but expression of EGFP in other regions was not observed.

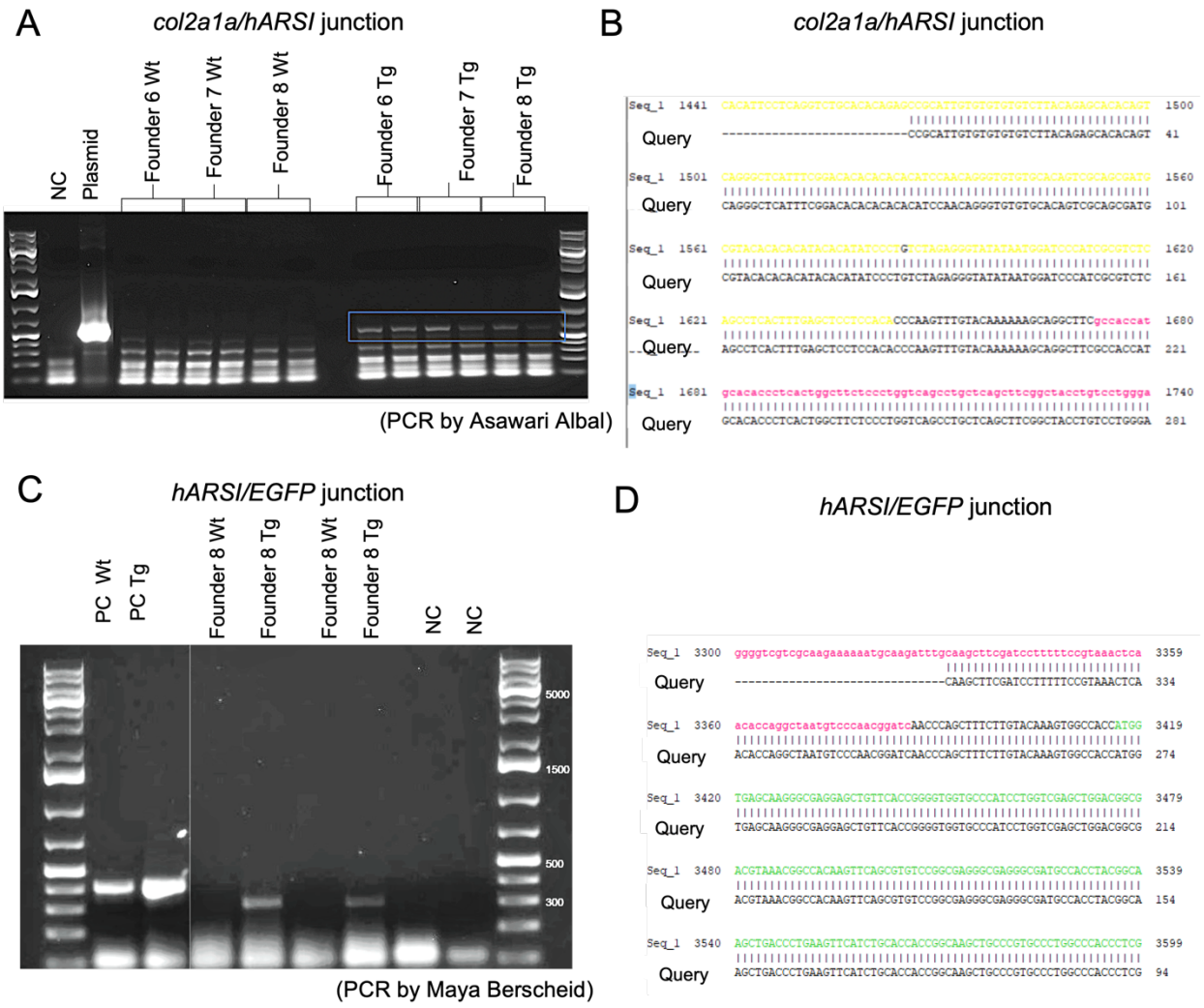


Figure 4.5 Genomic insertion of the *col2a1a_hARSI_GFP* construct was confirmed by PCR. (A) PCR results using genomic DNA and primers designed for the junction between the col2a1a enhancer and the hARSI gene. The amplified fragment (510 bps) was observed in Tg zebrafish and not in their WT siblings, coming from 3 different founders. (B) Sequencing results confirmed that the fragment had the expected sequence. The col2a1a enhancer is shown in yellow and hARSI in pink. (C) PCR results using genomic DNA and primers designed for the junction between the hARSI gene and EGFP. The amplified fragment (300 bps) was observed in Tg zebrafish and not in their WT siblings in Founder 8. (D) sequencing results confirmed that the fragment had the expected sequence. The hARSI sequence is shown in pink and the EGFP in green. PCRs were done by summer students I co-supervised, and I designed primers. NC: negative control; PC: positive control (ß-actin).

To further investigate the absence of EGFP expression in the *col2a1a_hARSI_EGFP* cartilages, another transgenic fish line was generated to examine the function of the *col2a1a* enhancer. Previously made *col2a1aEGFP* constructs, containing the same enhancer and destination vector as the *hARSI* transgenic fish lines, were injected by me, and fish with green hearts were raised to adulthood. Injected fish were then outcrossed with WT zebrafish, and larvae with a non-mosaic expression of the transgene (full green hearts) were raised to create a stable F1 lineage. EGPF expression was observed in head cartilages from 50hpf, which increased over time along with the development of other cartilage elements. These findings indicated that the *col2a1a* enhancer sequence was working as expected (Figure 4.6). This is the same expression pattern expected for the *hARSI* gene fused with this same enhancer.

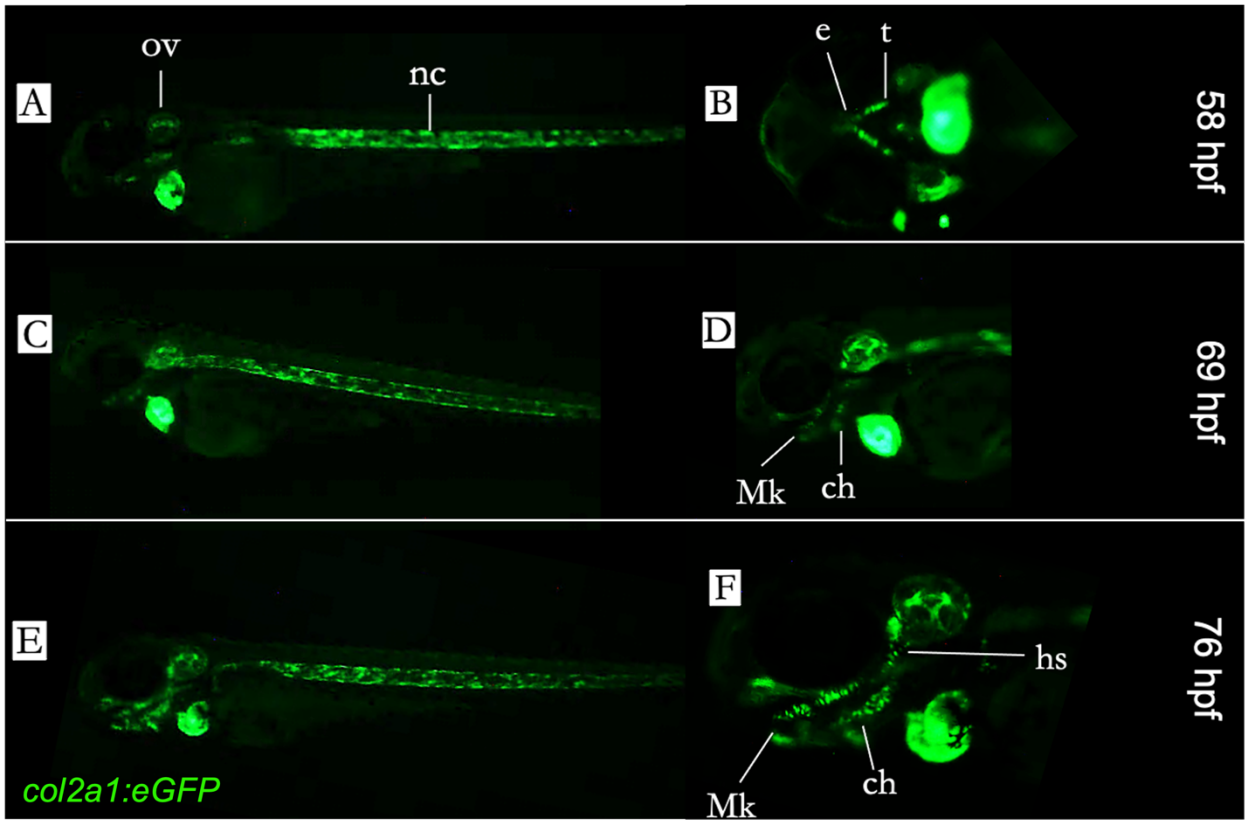


Figure 4.6 *col2a1aEGFP* transgenic fish expressed EGPF in their cartilages. A zebrafish line was created by injecting a gateway construct containing a *col2a1a* enhancer followed by an *EGFP* gene. The same enhancer and destination vector were used to create *hARSI* transgenic fish lines. EGFP expression was observed in transgenic zebrafish head cartilages since 50hpf. (A-B) Pictures were taken 58hpf show EGFP at the notochord (nc), otic vesicle (ov), ethmoid (e), and trabecula (t). (C-D) The EGFP was also evident at the ceratohyal (ch) and Merkel's cartilage (Mk). (E-F) EGFP expression increased over the hours and was visible in more cartilaginous elements, including the hyosymplectic (hs).

4.4.3 Checking for the presence of hARSI in the Tg lines

To assess the transcription of *hARSI* in putative Tg zebrafish, RT-PCRs were performed. Primers were designed to amplify 6 junctions present in the final constructs but absent in the WT zebrafish genome (Figure 4.7A). The RT-PCR analysis showed that products with expected band sizes were amplified from junctions 1, 3, 4, and 5 in the transgenic fish but not in their WT siblings (Figure 4.7B). In addition, all products were subjected to DNA sequencing, and two samples yielded clear sequencing data for the junctions. These results show that RNA was transcribed at least in some of the fish lines (Figure 4.7C and D).

To understand the typical expression pattern of ARSI during zebrafish development. The presence of the ARSI protein was assessed in WT by IHC. IHC was performed in 2, 3, 4, 5, 6, and 8 dpf zebrafish larvae. At 2 dpf, ARSI expression was restricted to the somites and eye lenses (Figure 4.8A and B); it extended to the jaw muscles at 3 dpf (Figure 4.8C and D) and to all head cartilages at 5 dpf (Figure 4.8E and F). From 5 dpf to 8 dpf, ARSI expression appeared ubiquitous and did not change significantly (Figure 4.8G-L).

Subsequently, whole mount IHC was performed in all putative Tg zebrafish lines and their WT siblings to check for increased ARSI expression in Tg lines. After visual inspection, the levels and regions of ARSI fluorescence did not seem to change between any Tg and WT fish (Figure 4.9). Because ARSI expression was widespread, it was challenging to make a specific comparison of ARSI levels in cartilage. Unis

antibodies against GFP could be a future direction to check for hARSI fused to EGFP in these putative Tg fish.

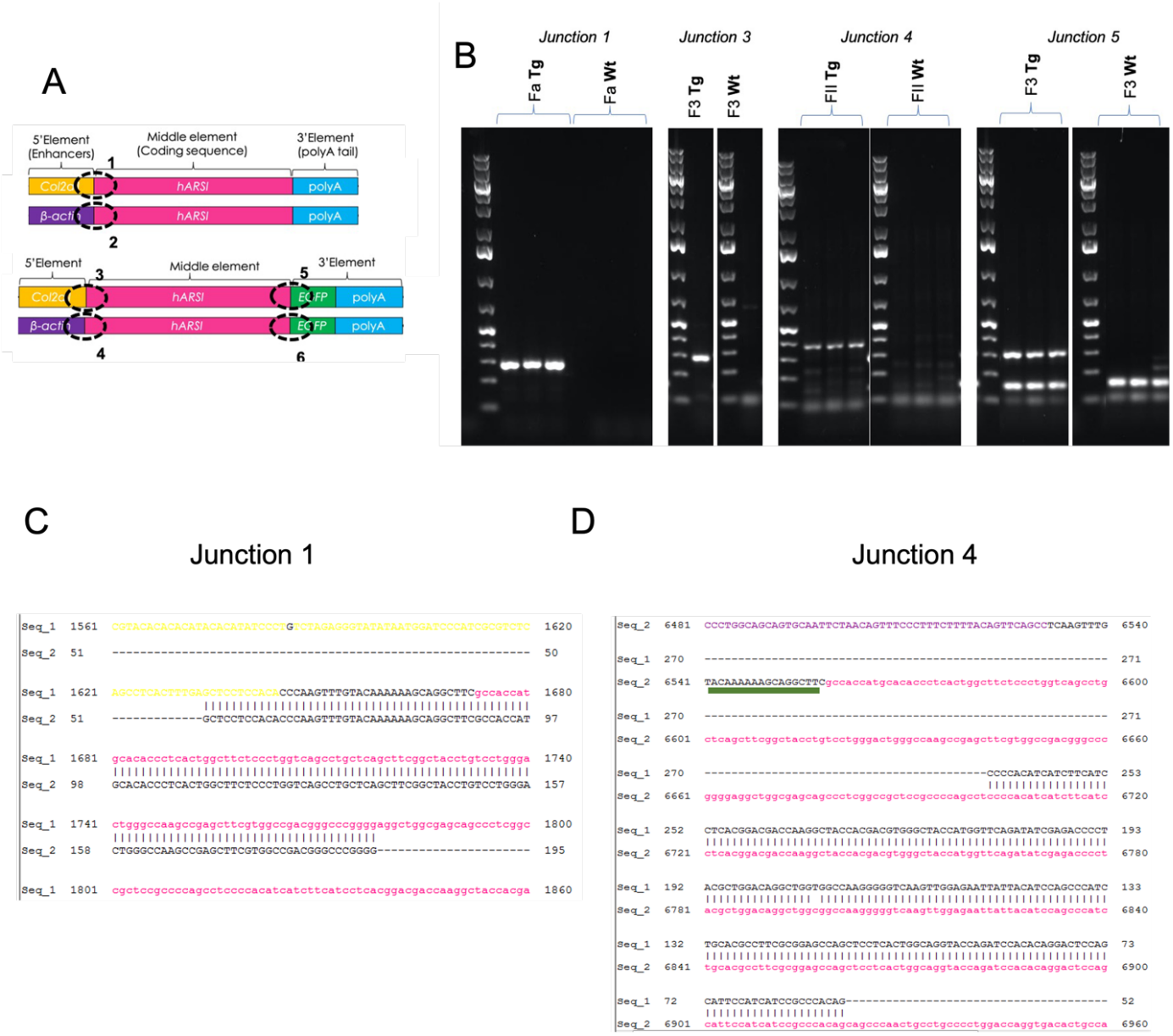


Figure 4.7 RT-PCR results to assess transcription of transgenes.

(A) Different primers were designed to amplify the cDNA junctions, which would be present in Tg fish but not in WT siblings. (B) Expected band sizes were observed in 4 out of the 6 junctions. (C-D) Two samples produced readable sequencing results from junction 1 and junction 4. The *col2a1a* enhancer is shown in yellow, the ß-actin enhancer in purple and *hARSI* in pink. In C, Seq_1 represents the predicted sequence, and Seq_2 is the obtained sequencing result. In D, Seq_2 represents the predicted sequence, and Seq_1 is the obtained sequencing result.

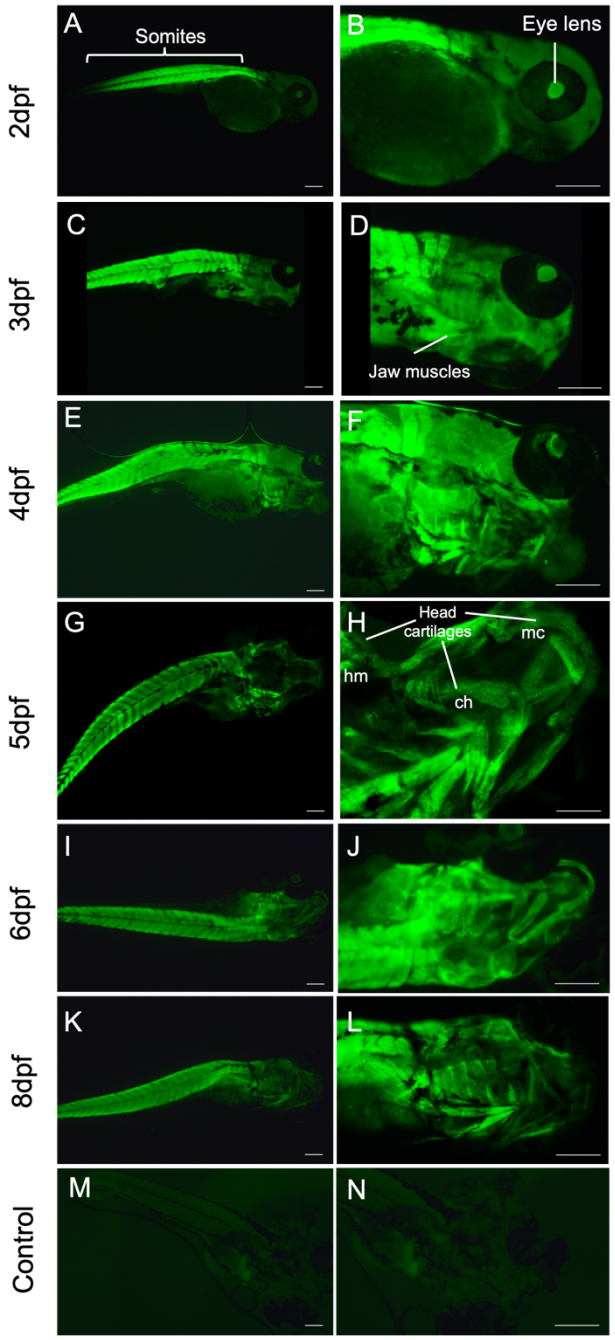


Figure 4.8 ARSI IHC over time in WT zebrafish. (A-B) at 2 dpf there was a strong ARSI signal at the zebrafish somites and eye lenses.(C-D) at 3 dpf ARSI expression extended to the jaw muscles. (E-F) at 4 dpf ARSI expression increased all over the head region. (G-H) at 5 dpf all head cartilages seemed to express arsi. (I-L) From 5 dpf to 8 dpf, ARSI expression seemed ubiquitous and did not change over time. (M-N) 8 dpf zebrafish stained just with secondary antibody were used as an experimental control and did not show fluorescence. Abbreviations: ch: ceratohyal; hm: hyomandibula; mc: Meckel's cartilage. Scale bars: 200um

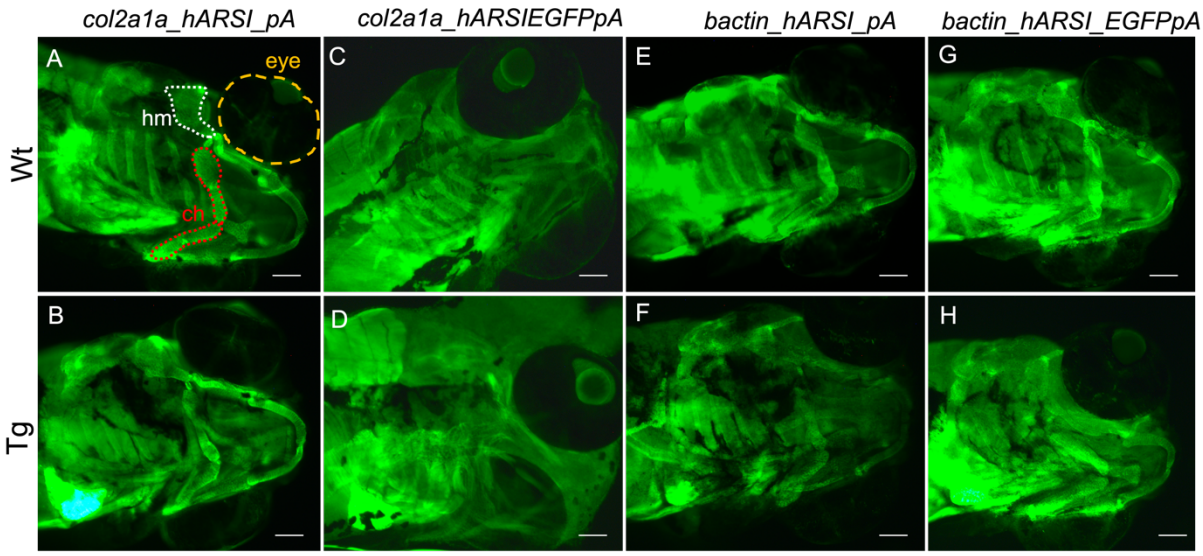


Figure 4.9 ARSI whole mount IHC for WT and Tg zebrafish. There was no difference observed in ARSI fluorescence levels between WT fish and Tg siblings for *col2a1a_hARSI_pA* (A-B), *col2a1a_hARSI_EGFPpA* (C-D), *ßactin_hARSI_pA* (E-F), and *ßactin_hARSI_EGPFpA*. Abbreviations: ch: ceratohyal; hm: hyomandibula. Scale bars: 100 µm

To confirm that the expression of ARSI in non-target tissues did not interfere with the visualization of ARSI in the ch, our main target, jaws from Tg and WT zebrafish were dissected before the whole mount protocol (Figure 4.10). Dissected jaws allowed a clearer visualization of the cartilages, including the ch, which presented a strong ARSI fluorescence signal (Figure 4.10A-D). ARSI was expressed in a very similar pattern to col2, a known marker of cartilage (Figure 4.10E-F). No apparent differences were observed in fluorescence levels between WT and Tg groups. Additional zebrafish samples were sectioned and posteriorly stained with anti-ARSI antibody. Consistent with the whole mount results, there was no difference in ARSI levels between Tg and WT zebrafish, and ARSI was observed in the same domains as col2 (Figure 4.10G-I).

The levels of ARSI protein were also assessed by Western blot in 8 dpf zebrafish larvae from all 4 putative Tg lines (Figure 4.11). F1 embryos from 6 different founders

were tested by Western blot (3 founders from the *col2a1a_hARSI_pA*, 1 founder from the *col2a1a_hARSI_EGFPpA*, 1 founder from the *ßactin_hARSI_pA*, and one 1 founder from the *ßactin_hARSI_EGFPpA* line). While the expected band for human ARSI (~64Kda) was not observed in any zebrafish samples, many extra bands were observed in all samples. These bands also did not correspond to the endogenous zebrafish arsia and arsib proteins, which are also approximately 64Kda in size (Figure 4.11A). ATDC5 cells at 21 days of culture were used as a positive control for ARSI expression and showed one band at the expected size (Figure 4.11B). Together with the ubiquitous ARSI expression observed by IHC, Western blot results indicate that the ARSI antibody binds non-specifically to zebrafish samples, making it challenging to assess the overexpression of ARSI in Tg fish.

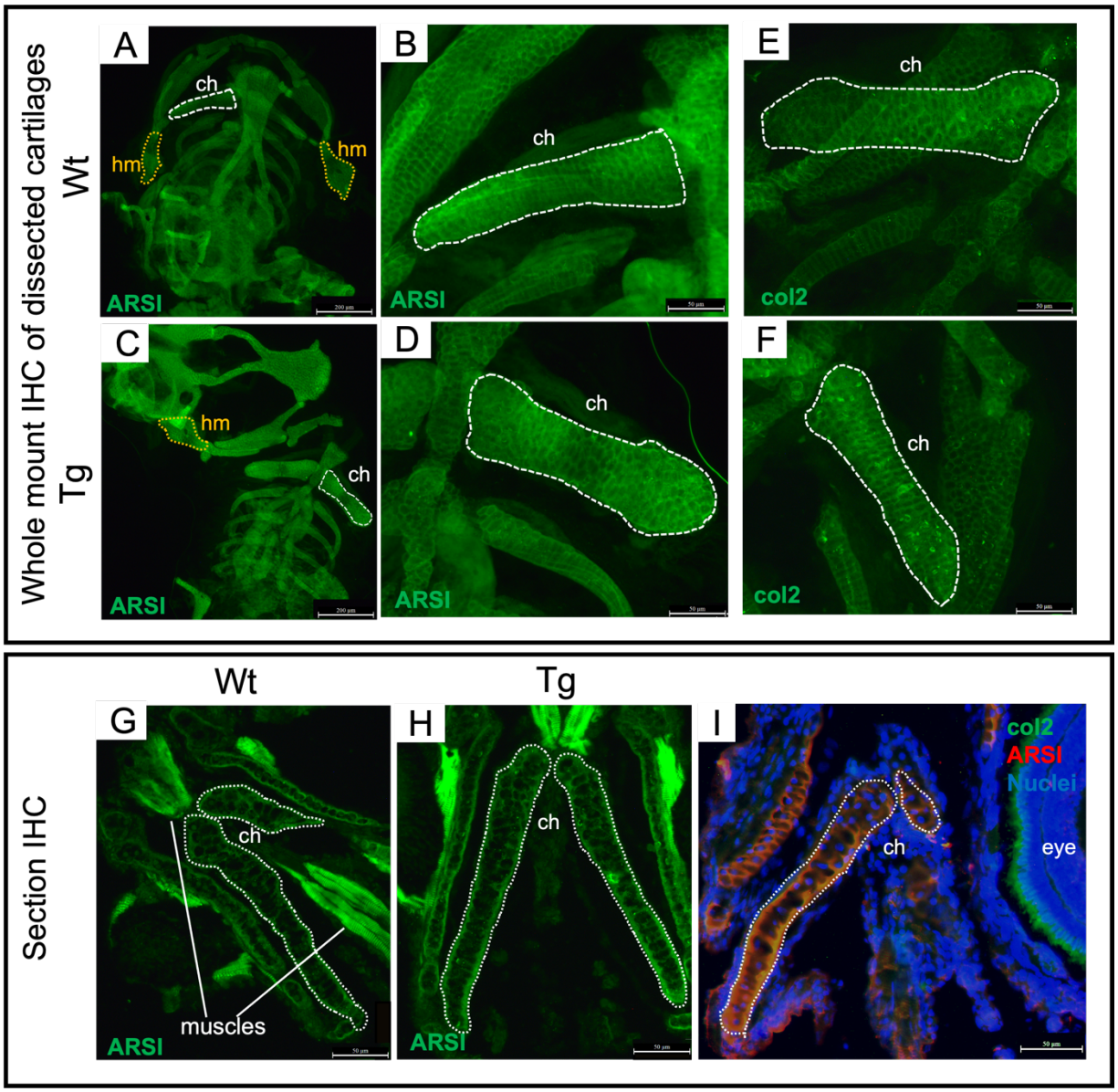
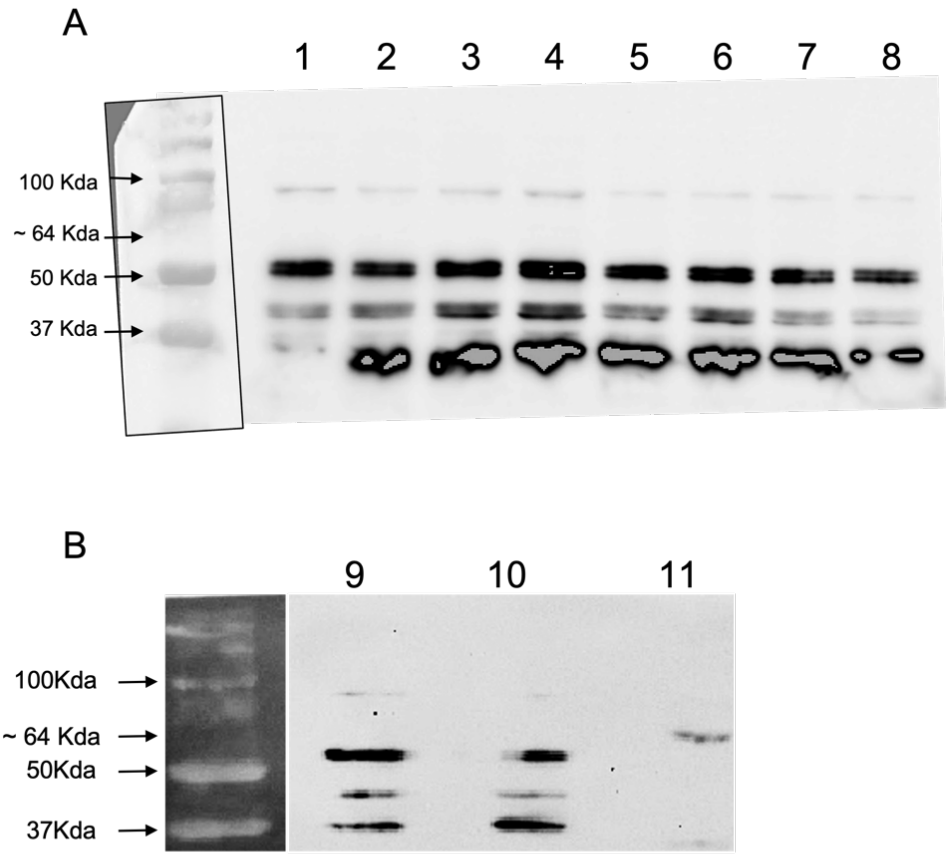


Figure 4.10 - ARSI fluorescence did not change between Tg and WT fish in IHC.

(A-D) Dissected jaws of WT and Tg 8 dpf zebrafish injected with *col2a1a_hARSI_GFP* construct were used to check for ARSI protein levels by IHC. (A, C) Fluorescence was observed in all WT and Tg dissected cartilages at similar levels. (B, D) Higher magnification of the ch element (delimited by a white dotted line) shows less fluorescence in the center of the cartilage. A similar pattern was observed in WT and Tg. (E-F) col2, a known marker of immature cartilage, was expressed in a very similar pattern to ARSI in all WT and Tg zebrafish cartilages at similar levels. (G-H) ARSI IHC was also performed in 7 µm ch sections of WT and Tg fish. There were no apparent changes in fluorescence levels between groups. Cartilage neighbouring tissues also reacted with ARSI antibody. (I) Overlay of col2 and ARSI antibodies showed that both are expressed in the same regions in cartilage. Abbreviations: ch: ceratohyal; hm:hyomandibula.



1- *Col2a1a_hARSI_pA* (GFP+)
2- *Col2a1a_hARSI_pA* (GFP-)
3- *Col2a1a_hARSI_EGFPpA* (GFP+)
4- *Col2a1a_hARSI_EGFPpA* (GFP-)
5- *Bactin_hARSI_pA* (GFP+)
6- *Bactin_hARSI_pA* (GFP-)
7- *Bactin_hARSI_EGFPpA* (GFP+)
8- *Bactin_hARSI_EGFPpA* (GFP-)
9- *Col2a1a_hARSI_pA* (GFP-)
10- *Col2a1a_hARSI_pA* (GFP+)
11-ATDC5 non-transfected 21 days (+control)

Figure 4.11 ARSI antibody binds non-specifically to zebrafish samples. (A) Western blot for ARSI in putative Tg fish and in their WT siblings showed multiple unspecific bands. There was no difference between WT and Tg fish and no band in the expected ARSI size (~64kda) was observed. (B) Tg and WT zebrafish samples were run together with ATDC5 cells that express ARSI after 21 days in culture. One band around 64Kda was observed in line 11, which was not present in the Tg or WT zebrafish.

4.4.4 On the ARSI antibody specificity

To check for antibody specificity, the ARSI antibody (Anti-ARSI HPA038398) immunogen sequence was compared with ARSI sequences from different species used in this study. The immunogen sequence was 100% comprised in the human ARSI sequence and showed decreasing levels of similarity when compared to mouse (95.23%), chick (79.36%), and zebrafish arsia (66.7%) and arsib (63.5%) (Figure 4.12A). Because of high sequence similarities among sulfatases, the same immunogen sequence was compared against all proteins in the different species for any hits different than ARSI. In all species, the immunogen was more similar to ARSI than any other sulfatase. Other sulfatases, mainly ARSJ, presented some similarities but with a smaller query cover and number of identities (Figure 4.12B-E). Overall, these findings indicate that antibody used should be specific for ARSI.

Further experiments were performed to confirm the specificity of the ARSI antibody. IHC of ATDC5 cells transfected with a construct containing hARSI fused with mCherry showed a clear increase in fluorescence signal compared to the non-transfected ATCD5 cells (Figure 4.12F-I). Western blot with additional ATDC5 cells samples showed an increase in the ARSI protein in transfected cells compared to non-transfected cells, where the ARSI fused to mCherry band was observed in the estimated size (85 kDa) (Figure 4.12J). Lastly, further experiments by our collaborators utilizing the same hARSI antibody showed that ARSI was highly expressed in Swarm rat chondrosarcoma (RCS) cells in the expected size (64 Kda) but disappeared when ARSI was knocked out in that same cell line (Figure 4.12K). In summary, the hARSI antibody specificity was observed by Western blot for human (overexpressed in ATDC5 cells), rat (RCS cell line), and mouse (non-transfected ATDC5 cells) samples. Also, ARSI expression by IHC seemed specific in chick and mouse humerus. Possibly the differences in the epitopes of zebrafish proteins are too large for the antibody to work properly.

A

Organism		
H.sapiens	456	A-SVRQAVW----------LFNISADPYEREDL--AGQRPDV VRTLLARLAEYNRTAIPVRYPAENPRAHPDFNGGAWGPWASDE-EEEEEE
M.musculus	456	A-SIRQAVW----------LFNISADPYEREDL--AGQRPDV VRTLLARLADYNRTAIPVRYPAANPRAHPDFNGGAWGPWASEEEEEEEEEE
G.gallus	459	TDGLRKSVW----------LFNITADPYERYDL--SEQRPDV VRALLTRLVHYNRTAIPVRYPAENPRAHPDFNGGAWGPWASE----DEVE
D.rerio	448	T-EAKKSLW----------LFNVADDPYERFDL--AERRPDV VKQLLARLSFYNRTAVPVRYPSEDPRADPSRNGGAWGPWEGDG------E
D.rerio	453	IGEKRKSVW----------LFNVTGDPCERHDL--AVHRPDV VRELLARLAFHNRTAIPVRYPPDDTRANPSANGGAWRPWVGED---DEEE

B

Human

	Max Score	Total Score	Query Cover	E value	Per. Ident
arylsulfatase family, member I [Homo sapiens]	131	131	100%	9e-38	100.00%
unnamed protein product [Homo sapiens]	131	131	100%	1e-37	100.00%
arylsulfatase I precursor [Homo sapiens]	132	132	100%	4e-37	100.00%
unnamed protein product [Homo sapiens]	83.2	83.2	100%	4e-21	52.38%
hCG20717, isoform CRA_b [Homo sapiens]	83.2	83.2	100%	5e-20	52.38%
arylsulfatase J isoform X2 [Homo sapiens]	83.6	83.6	100%	7e-20	52.38%

C

Mouse

	Max Score	Total Score	Query Cover	E value	Per. Ident
Arsi protein [Mus musculus]	127	127	100%	3e-38	96.83%
PREDICTED: arylsulfatase I isoform X1 [Mus musculus]	119	119	92%	5e-33	96.55%
arylsulfatase I precursor [Mus musculus]	119	119	92%	1e-32	96.55%
arylsulfatase J [Mus musculus]	80.5	80.5	88%	5e-19	57.14%

D

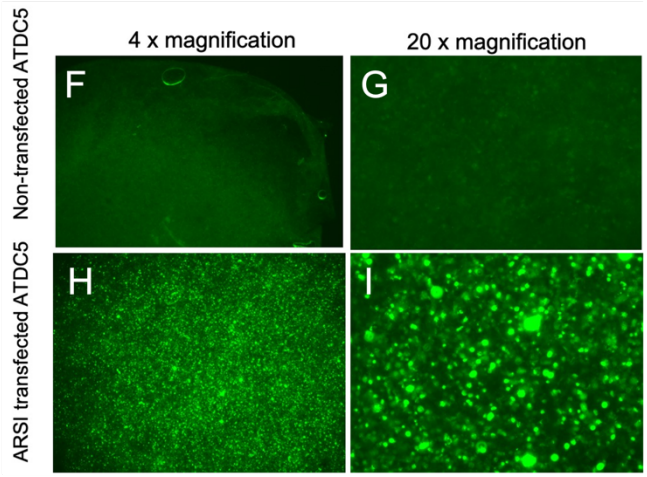
Chick

	Max Score	Total Score	Query Cover	E value	Per. Ident
arylsulfatase I [Gallus gallus]	108	108	92%	2e-29	86.21%
arylsulfatase J [Gallus gallus]	84.0	84.0	88%	1e-20	60.71%

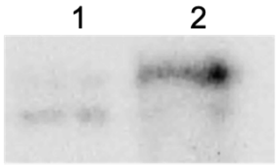
E

Zebrafish

	Max Score	Total Score	Query Cover	E value	Per. Ident
arylsulfatase I-like [Danio rerio]	91.3	91.3	95%	5e-23	71.67%
arylsulfatase I [Danio rerio]	85.9	85.9	100%	4e-21	65.08%
arylsulfatase I-like [Danio rerio]	75.5	75.5	98%	1e-17	56.45%
arylsulfatase J [Danio rerio]	73.2	73.2	88%	1e-16	51.79%



J



1 - ATDC5 GFP-transfected
2 - ATDC5 ARSI fused mCherry transfected

K

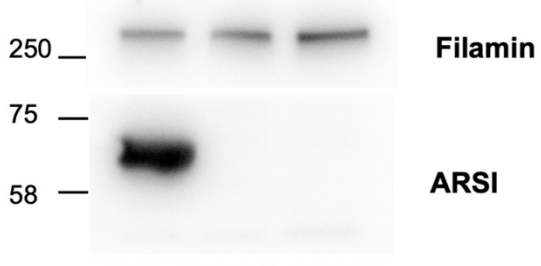


Figure 4.12 ARSI antibody specificity. (A) Alignment of ARSI antibody immunogen sequence with ARSI sequences in different species. The immunogen sequence is SADPYEREDLAGQRPDVVRTLLARLAEYNRTAIPVRYPAENPRAHPDFNGGAWGPWASDEEEE. Purple highlights show the positive matches, which are 100% for human, 95.23% for mice, 79.36% for chick, and 63.5-66.7% for zebrafish sequences. (B-E) The immunogen sequence was compared against all proteins in the different species. ARSI was always the top hit in all species, followed by ARSJ. (F-I) ATDC5 cells transfected with an hARSI construct and stained by IHC with the ARSI antibody showed an increased fluorescence compared to non-transfected cells. (J) Western blot with ATDC5 cells showed an increase in the ARSI protein expression in the transfected cells compared to non-transfected ATDC5 cells. An ARSI plus mCherry band was observed in the estimated size of 85Kda. (K) ARSI was highly expressed in Swarm rat chondrosarcoma (RCS) cells in the expected size but disappeared when ARSI was knocked out in that cell line.

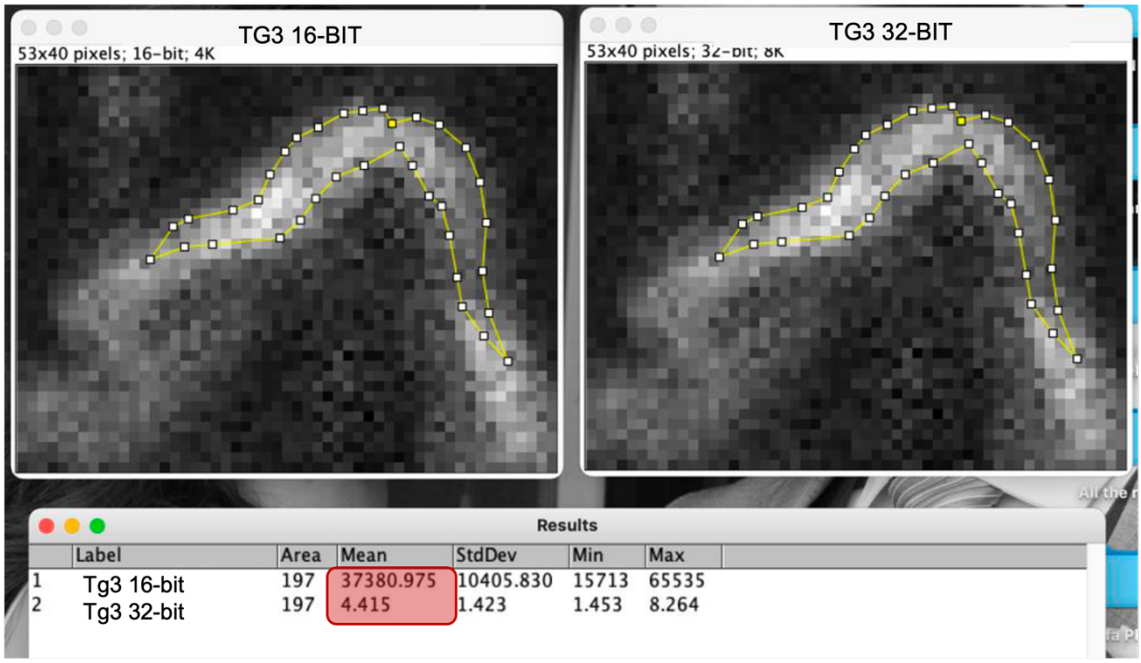
4.4.5 Quantification of X-ray fluorescence sulfation maps changed in 16 vs. 32-bit images.

To evaluate the biochemical function of ARSI, X-ray fluorescence (XRF) imaging was used to measure the sulfur content in the ch of putative Tg fish at 6 dpf. Due to limited beamtime and time constraints, the experiments were focused on one founder (Founder 8) of the *col2a1a_hARSI_EGFP* line. This line was chosen because *hARSI* was supposed to be expressed just in cartilage, causing fewer secondary effects during development. Additionally, the potential to see EGFP expression in the same regions as hARSI further influenced the decision to use this line.

XRF map quantifications were initially done after converting the images to 16-bit using ImageJ as previously described (Hackett et al., 2016). However, during the quantification process, the pixel values in the converted images significantly differed from those in the original 32-bit images (Figure 4.13A). While a difference in values was expected between 16 and 32-bit, the observed differences were not relative (e.g., in 16-bit images, WT2 and Tg2 fish had the highest sulfate ester levels, while in 32-bit images, WT1 and Tg4 had the highest sulfate ester levels) (Figure 4.13 B and C). These findings indicate that the differences in measurements were not just related to a scaling difference. It is important to make this point because our initial quantification results at 16-bit – which will be further discussed in this chapter- indicated a significant

decrease in sulfate esters in the Tg fish. A sulfate ester decrease in Tg fish would support our hypothesis that ARSI works by decreasing sulfur levels. However, this difference was non-existent in quantifications of the same images using 32-bit.

A



B

Examples of sulfate ester measurements (16-bit)

	Wt Averages	Tg averages	
Wt1	46826.4623	38888.2287	Tg1
Wt2	47713.8203	39411.7323	Tg2
Wt3	43124.8058	37757.7503	Tg3
Wt4	46611.1393	37615.8567	Tg4

C

Examples of sulfate ester measurements (32-bit)

	Wt	Tg	
Wt1	5.291	4.249333333	Tg1
Wt2	4.541666667	2.694666667	Tg2
Wt3	3.063666667	3.938	Tg3
Wt4	4.113666667	4.481333333	Tg4

Figure 4.13 16-bit and 32-bit images from the same sample presented largely different pixel intensities. (A) XRF maps were transformed into images, opened in ImageJ and converted to 16-bit or not converted at all (32-bit). The same region of interest (circled in yellow) was measured in both images resulting in very different mean pixel intensities (red box). (B-C) Examples of sulfate ester measurements in 4 WT and 4 Tg samples in 16-bit and 32-bit images showed no direct correlation between sulfate ester levels in 16 and 32-bit images.

4.4.6 Potassium and total sulfur increased in Tg zebrafish cartilage.

To visualize *in situ* the distribution of sulfur in the Tg and WT fish, XRF maps of their ch were generated. Sulfur (S), potassium (K), calcium (Ca), phosphorus (P), and chlorine (Cl) maps were collected simultaneously from sections of zebrafish ch at the VESPERS beamline at the Canadian Light Source (CLS) (N=9 WT and 8 Tg) (Figure 4.14A-L). Total sulfur levels did not change between putative Tg and WT fish, as determined by 16 or 32-bit quantifications (Figure 4.14M and N). However, K levels significantly increased in putative Tg zebrafish compared to WT as determined by 32-bit quantifications (Figure 4.14O and P). There were no obvious differences in the Ca, P, and Cl levels between WT and Tg zebrafish, although these were not quantified.

To assess the distribution of different chemical forms of sulfur in developing cartilage, chemically specific XRF maps were generated using the SXRMB at the CLS. These maps allowed for the visualization of background plus lower oxidation levels of sulfur, total sulfur, sulfate esters, and sulfonic acids in the zebrafish ch (N=7 for total sulfur, and 6 for sulfate esters and sulfonic acids; Figure 4.15A-E). Like what was seen for VESPERS results, the statistical analysis of the SXRMB data was affected by the quantification of images at different bit depths. Specifically, analysis of 16-bit chemically specific XRF maps revealed a significant decrease in sulfate esters in Tg fish compared to their WT siblings. At the same time, there was no difference in the total sulfur or sulfonic acid maps (Figure 4.15F-H). On the other hand, analysis of the 32-bit images of the same maps showed an increase in total sulfur and a slight non-significative decrease of sulfate esters in Tg fish compared to their WT siblings (Figure 4.15I-K).

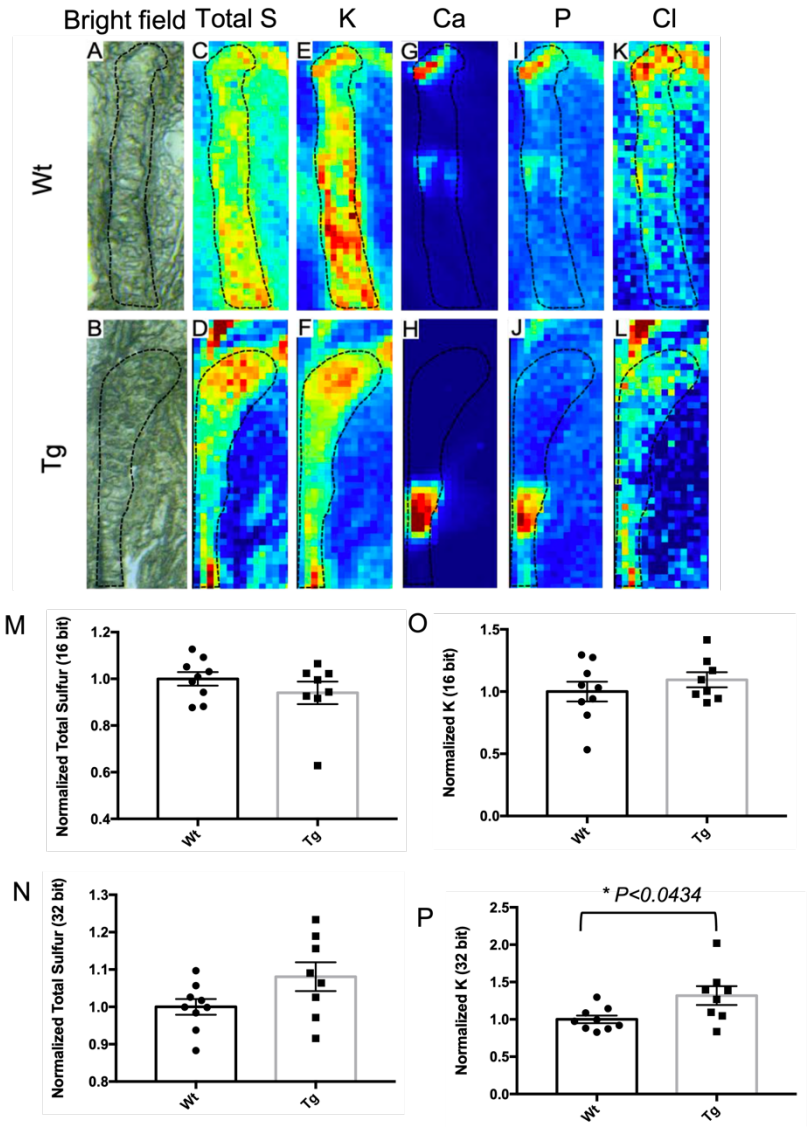


Figure 4.14 : XRF imaging showed no change in sulfur and an increase in K in transgenic cartilage.

(A-B) Sections of 7 dpf zebrafish ch used for XRF imaging. (C-D) XRF maps of total sulfur in WT and Tg fish. (E-F) XRF maps of potassium in WT and Tg fish. (G-H) XRF maps of calcium in WT and Tg fish. (I-J) XRF maps of phosphorus in WT and Tg fish. (K-L) XRF maps of chlorine in WT and Tg fish. (M) Total sulfur levels trended towards a slight, non-significant decrease in ARSI-expressing transgenic cartilage, compared to wild-type controls. (N) Levels of potassium trended towards a slight, non-significant increase in ARSI-expressing transgenic cartilage. N=9 for WT and N=8 for Tg. Student T-test was used to compare between WT and Tg groups, error bars represent standard deviation. $P<0.05$ Ch: ceratohyal; WT: wild type; Tg: Transgenic

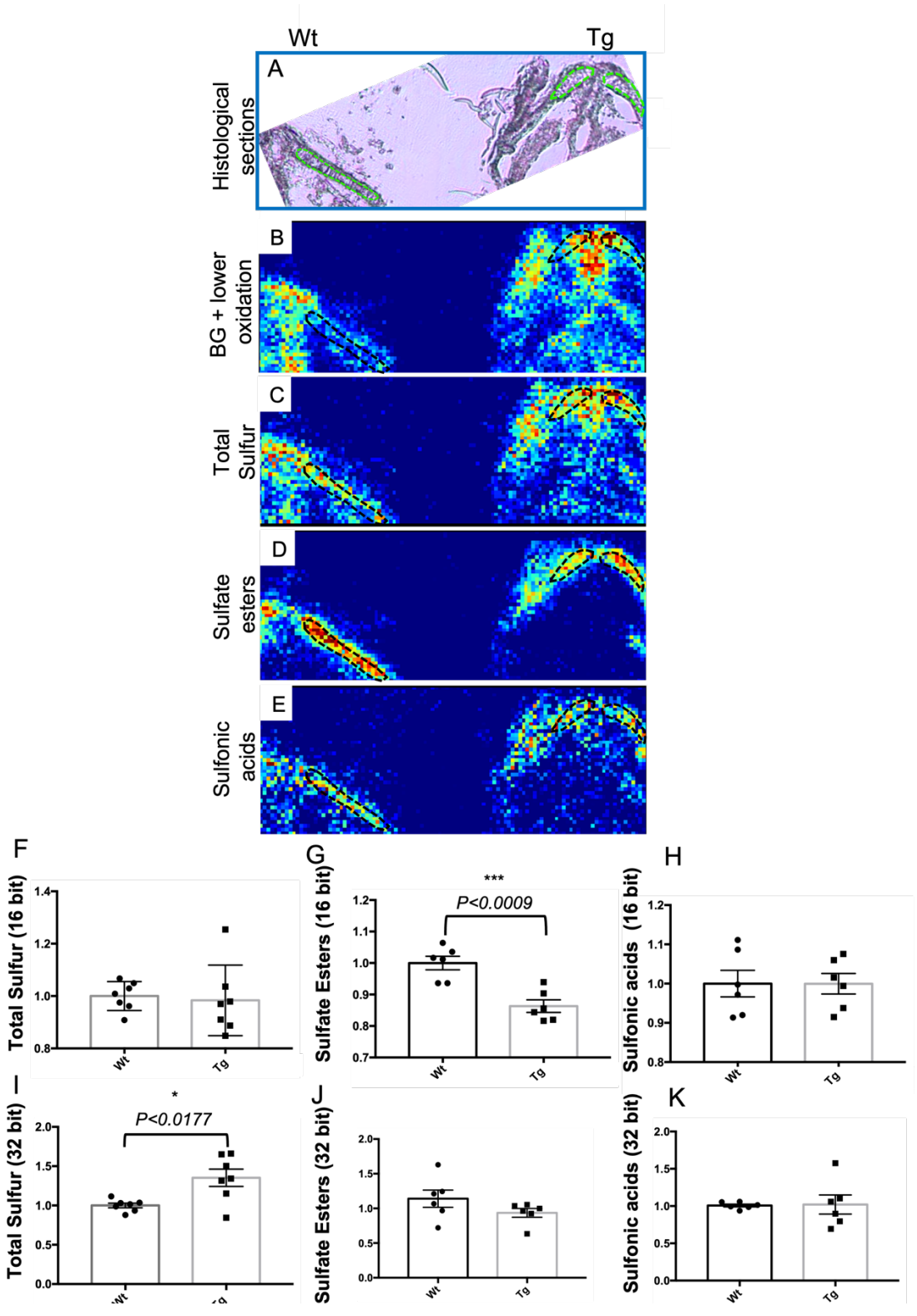

Wt
Tg
A
Histological sections
B
BG + lower oxidation
C
Total Sulfur
D
Sulfate esters
E
Sulfonic acids
F
Total Sulfur (16 bit)
G
Sulfate Esters (16 bit)
***
P<0.0009
H
Sulfonic acids (16 bit)
I
Total Sulfur (32 bit)
*
P<0.0177
J
Sulfate Esters (32 bit)
K
Sulfonic acids (32 bit)

Figure 4.15 Chemically specific XRF of Tg and WT zebrafish ch.(A) Histological sections of WT and Tg (*col2a1a_hARSI_EGFP*) zebrafish ch used for chemically specific XRF imaging. The dotted lines show regions where sulfur levels were quantified. (B-E) Chemically specific XRF maps for WT and Tg zebrafish ch for lower oxidation sulfur and background fluorescence, total sulfur, sulfate esters, and sulfonic acids. (F-H) Quantification of 16-bit images showed that total sulfur and sulfonic acid levels did not change in Tg fish, while sulfate esters significantly decreased. (I-K) When the same maps were quantified in the 32-bit format, there was an increase in total sulfur in Tg zebrafish, sulfate ester levels trended towards a slight, non-significant decrease in Tg zebrafish, and sulfonic acid levels didn't change between the two groups. N=6-7; student T-test was used to compare between WT and Tg groups. BG: background fluorescence; WT: wild type; Tg: Transgenic.

4.4.7 Putative transgenic zebrafish had more perichondral bone compared to their WT siblings.

No gross phenotypical differences were observed in the injected fish (F0) and F1 transgenic fish lines compared to their WT siblings. To assess if hARSI overexpression accelerated bone development, putative F1 Tg zebrafish skeletons were analyzed by Alcian blue and alizarin red staining. Perichondral bone levels were assessed in the ch and hm of *col2a1a_hARSI_EGFP* putative Tg fish and their WT siblings at 8 dpf following a perichondral bone scoring system (Figure 4.16A and B). The results showed that the putative Tg zebrafish had more perichondral bone in the hm than their wild-type siblings, but there was no difference in the amount of perichondral bone in their ch (Figure 4.16 C and D). Alcian blue levels, which reflect sulfated PG levels in cartilage ECM, were also compared between putative Tg and WT fish. Spectrometric quantification of lysates from three Alcian blue-stained clutches showed a slight but non-significant decrease in putative Tg compared to WT zebrafish (Figure 4.16 E).

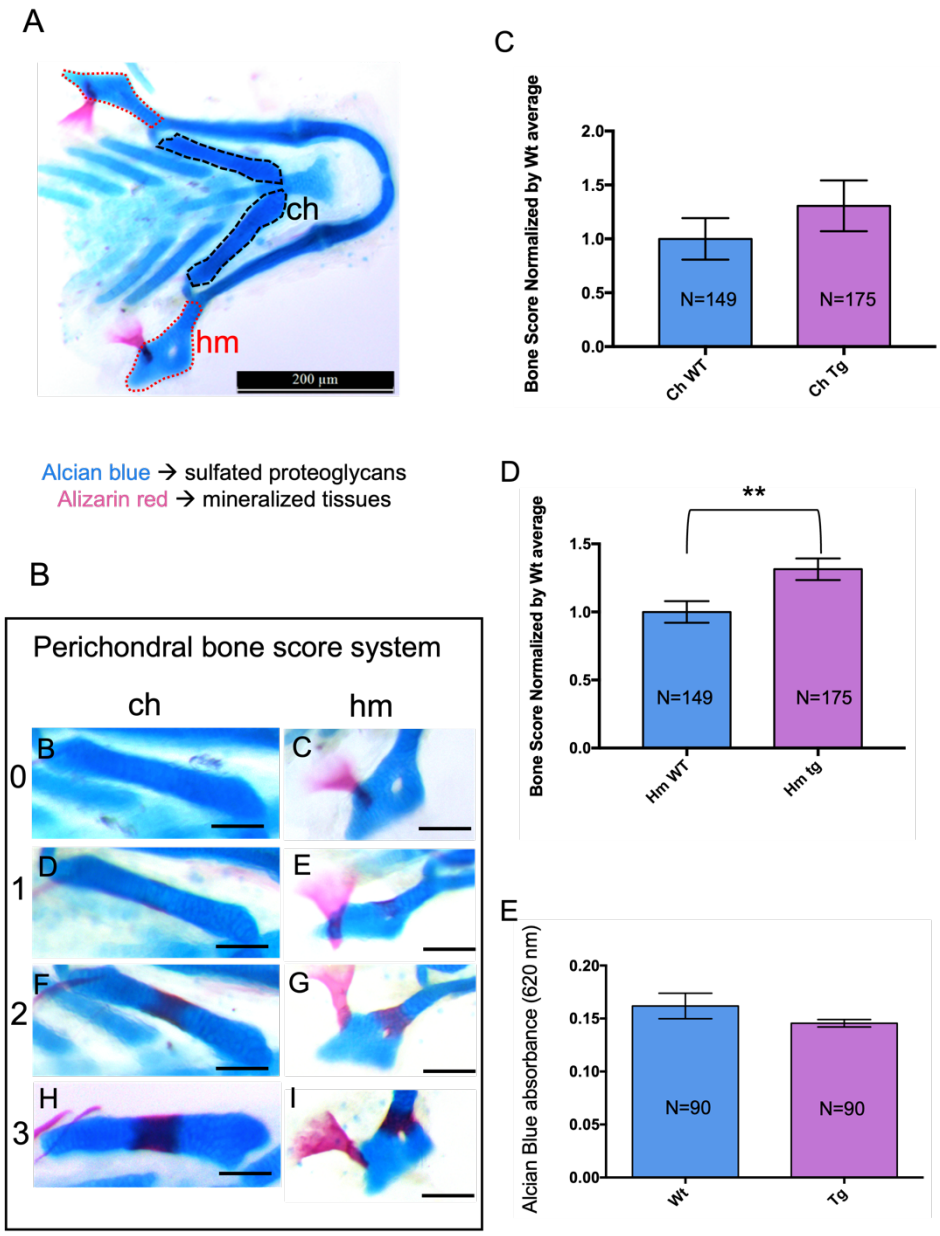


Figure 4.16 Assessment of Tg (*col2a1a_hARSI_EGFP*) zebrafish and WT skeleton by Alcian blue and Alizarin red staining at 7 dpf.(A) Zebrafish larvae head skeleton showing cartilage stained with Alcian blue and bone stained with Alizarin red. The analyzed elements were the hm and ch, highlighted in red and black, respectively. (B) The amount of perichondral bone was scored based on a scoring system where 0 is no bone and 3 is a highly mineralized perichondral bone surrounding all the ch or hm element. (C) There was a slight non-significant increase in the amount of perichondral bone in the ch between WT and Tg fish. (D) There was a significant increase in perichondral bone in the hm in the Tg zebrafish compared to the WT siblings (**p<0.0058). (E) There was a slight non-significant decrease in Alcian blue in Tg compared to WT fish.

4.4.8 Bafilomycin treatment did not increase EGFP visualization in Tg fish.

One possibility for the lack of EGFP visualization in the transgenic zebrafish cartilages could be due to the rapid degradation of the overexpressed hARSI linked to the EGFP. Autophagy is one of the major degradation pathways in eukaryotes, which functions by capturing unwanted material in autophagosomes that subsequently fuse with lysosomes and degrade the contents (Lorincz & Juhasz, 2020). To investigate the possibility of protein degradation through this pathway, Bafilomycin was used. Bafilomycin works by inhibiting vacuolar ATPase, preventing acidification and degradation in lysosomes and autophagosome-lysosome fusion (Yamamoto et al., 1998; Yoshimori et al., 1991). *col2a1a_hARSI_EGFP* fish were treated with 2 concentrations of Bafilomycin (20nM and 50nM) for 3 days, but no difference in EGFP expression between treated and untreated groups was observed (data not shown).

4.4.9 SUMF1 injections did not intensify Tg zebrafish skeletal phenotype

To evaluate if the endogenous SUMF1 levels were not enough to activate the overexpressed hARSI, *SUMF1* mRNA was injected in the eggs of Tg zebrafish and their WT siblings. The amount of perichondral bone in 8 dpf zebrafish was assessed in their ch and hm according to the table in Figure 4.16B. The *SUMF1* injections led to large death rates in injected embryos and did not seem to potentiate the phenotype of putative Tg fish compared to their WT siblings (Figure 4.17).

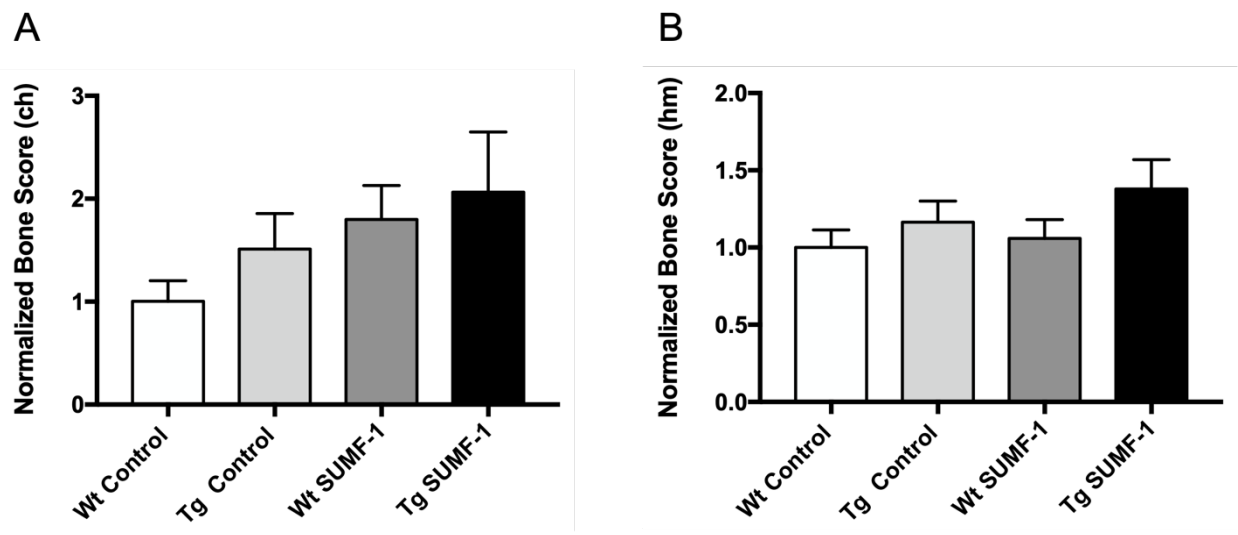


Figure 4.17 - *SUMF1* injections did not potentiate the Tg phenotype. (A-B) There was no significant difference in the amount of perichondral bone in Tg zebrafish ch and hm compared to WT or Tg uninjected fish. Fish from 2 clutches (N=44 WT control; N=45 Tg control; N=40 WT *SUMF1*; N=25 Tg *SUMF1*). One-way ANOVA was used to compare all groups WT: wild type; Tg: Transgenic.

4.5 Discussion

There are various animal models with impaired sulfated PGs and skeletal phenotype defects. These models typically involve mutations in genes associated with PG synthesis and sulfation pathways, such as *Papss2*, *uxs1*, *fam20b*, and *Sumf1* (Cortes et al., 2009; Eames et al., 2010; Eames et al., 2011; Settembre et al., 2008). Even though many sulfatases are known to control PG sulfation levels, to the best of our knowledge, just one transgenic animal model for sulfatases has been reported in the literature: the Arsa transgenic mice, which was used for the study of metachromatic leukodystrophy (Capotondo et al., 2007). Considering the existence of 17 sulfatases, and their importance for health and disease, it is not clear why transgenic animals for sulfatases are not abundant. There may be some intrinsic difficulties in overexpressing sulfatases in animals, but the lack of literature reporting these challenges hinders the process of generating new animal models.

In this study, genomic integration of all Gateway constructs and RNA transcription for some of them was observed. However, the overexpression of ARSI protein was not confirmed in the putative Tg fish. The Gateway-tol2 system has some limitations. For example, the random insertions of the constructs driven by the Tol2 transposon sites could lead to insertions in genomic regions that are not transcribed (Rafferty & Quinn, 2018). There was no EGFP expression or differences in IHC and western blot results for different constructs and founders. Most likely, ARSI protein was not observed due to post-transcriptional and post-translational processes. Post-transcriptional regulation, which may control RNA decay, transcript elongation, and translation initiation, is usually controlled by small RNAs and RNA-binding proteins and might explain why it was not possible to show amplification for all junctions in the RT-PCRs and may be a reason for the lack of protein expression (Van Assche et al., 2015). Another possibility is that *hARSI* mRNA is being degraded due to being recognized as a foreign sequence (human), in zebrafish.

The absence of hARSI_EGPF could also happen due to post-translational modifications (PTMs). There are hundreds of described PTMs (e.g. phosphorylation, acetylation, glycosylation and ubiquitination), which normally increase protein diversity, and determine protein structure, destination, and function (Karve & Cheema, 2011). Due to their importance, PTMs are usually tightly regulated processes. However, non-specific PTMs might occur when cellular stress happens, such as when proteins are overexpressed (Wani et al., 2015). In this sense, hARSI folding or stability in transgenic animals might be affected by PTMs targeting the extra proteins for degradation to maintain the physiological homeostasis (Li et al., 2021).

No other studies overexpressing ARSI in animals were found for comparison. However, previous work overexpressing ARSI in retinal pigment epithelium cell line ARPE-19 concluded that overexpressed ARSI was unstable and rapidly degraded (Oshikawa et al., 2009). This phenomenon could be occurring in our putative Tg zebrafish, leading to the absence of EGFP and mild phenotypes. There are two main pathways for protein degradation in eukaryotic cells: the ubiquitin-proteasome and autophagy-lysosome pathways (Rubinsztein, 2006). According to Oshikawa's cell

culture experiments, ARSI was degraded in the ER, and a portion of the overproduced ARSI remained in its denatured form in the ER (Oshikawa et al., 2009). As a result, ARSI is likely degraded by proteasome complexes rather than autophagosomes, which may explain why our bafilomycin treatments did not allow us to visualize EGFP.

Due to the novelty of working with ARSI, limited commercial antibodies are available for investigating its protein expression. The anti-ARSI antibody used in this study (Anti-ARSI HPA038398) was purchased from Sigma and produced against a human ARSI immunogen sequence. The antibody was used to assess ARSI expression in chick (Figure 3.11F-G) and mouse humerus sections (Figure 3.12); transfected (Figure 4.12F-J) and non-transfected mouse ATDC5 cells (Figure 3.14); and zebrafish. Except for zebrafish experiments, the antibody seemed to recognize ARSI in mature chondrocytes (more about the anti-ARSI antibody specificity is described in section 4.4.4). Since I overexpressed human *ARSI* in zebrafish, I expected the antibody to be specific for this protein. However, a ubiquitous expression was observed even in WT fish by IHC, and many extra bands were observed in WT and Tg Western blot, suggesting that the antibody is binding to unspecific epitopes in zebrafish. Notably, no commercial antibody was available to investigate endogenous arsi expression in zebrafish. As highlighted by prior studies, one of the significant limitations of working with this animal model has been the lack of appropriate antibodies. This is because zebrafish proteins are more diverged than other vertebrates and may not be specifically recognized by antibodies designed for other species (Dhar et al., 2020; Staudt et al., 2015; Tonelli et al., 2020). One option to check if any of the bands observed in the zebrafish Western blots corresponded to ARSI would be to extract specific bands and perform mass spectroscopy in the samples.

The conversion of synchrotron images and other large images from 32-bit to 16-bit is a relatively common practice to reduce file size in the field of imaging analyses (Brown, 2017; Hackett et al., 2016; Power et al., 2022; Singh et al., 2018). However, our quantifications showed that such a conversion could significantly affect statistical results. Specifically, while 16-bit quantifications indicated that sulfate esters decreased in putative Tg zebrafish, quantifications in 32-bit showed no difference in sulfate esters

and increased total sulfur levels in this same group. One possible explanation for these differences is that when images are converted from 32-bit to 16-bit, not all values from a higher bit-depth image fit into an image with a lower bit-depth. To convert a 32-bit pixel value to a 16-bit pixel value, ImageJ adds or subtracts a constant from the image pixels, which are further divided by another constant. After these calculations, the converted numbers are assigned to a valid value within the 16-bit scale, and clipping or rounding of numbers may occur. The constants used to perform this rescaling are determined by the original minimum and maximum values in each 32-bit image (Bankhead, 2022). However, the 32-bit images generated at the synchrotron did not have the same minimum and maximum values when compared to each other. For example, each sample had different minimum and maximum sulfur levels, resulting in different constants generated by the program for rescaling each image. Using different constants could place numbers that were the same in 32-bit images into different assigned 16-bit values, making images not comparable. In summary, using 32-bit images for quantification leads to more accurate results since the data is not clipped or converted. Using non-converted images would avoid skewing the results observed by the very sensitive technique used in this research.

The quantification of 32-bit XRF images revealed a significant increase in potassium and total sulfur in putative Tg fish compared to their WT siblings. Levels of sulfur and potassium were previously shown to be correlated in pig and chick cartilages and even in plants (Hargest et al., 1985; Reich et al., 2016; van Donkelaar et al., 2007). Thus, the observed correlation between S and potassium levels increases our data reliability. Although the specific function of potassium during cartilage maturation remains unclear, it is known to play a role in balancing the negative charges of sulfate anions (Hargest et al., 1985; Reich et al., 2016; van Donkelaar et al., 2007). The increase in sulfur in Tg fish is puzzling given that sulfatase typically hydrolyzes sulfate esters, and overexpression of these proteins would intuitively lead to a sulfur reduction. Interestingly, prior research has shown that the overexpression of ARSI in cells decreased the endogenous arylsulfatase activity, possibly because of the degradation of endogenous SUMF1 due to the overproduction of ARSI (Oshikawa et al., 2009). The overexpressed ARSI in putative Tg zebrafish may be causing a decrease in SUMF1 and

overall arylsulfatase activity, which could account for the increased sulfur levels observed in Tg cartilage.

In agreement with what was reported by Oshikawa et al. (2019), further investigation is required to obtain large amounts of ARSI *in vitro* and *in vivo* to study its biological function. To achieve this, it may be necessary to co-overexpress ARSI with other components, such as SUMF1. While our Tg zebrafish injected with *SUMF1* mRNA did not exhibit a significant increase in bone density, the simultaneous expression of ARSI and SUMF1 has been demonstrated to enhance ARSI function (Cosma et al., 2003; Oshikawa et al., 2009). The injection of capped mRNA in zebrafish eggs is a standard method to study gene function; however, it does not generate cell-type-specific phenotypes. Furthermore, injected mRNA tends to be stable just during early developmental stages, and the short life of the generated proteins might not be enough to cause a phenotypical change (Köster & Sassen, 2015). An option to investigate changes in skeletal phenotypes could be to outcross a stable zebrafish line overexpressing *SUMF1* under the control of a *col2a1a* enhancer to our putative Tg fish.

Using skeletal staining, putative transgenic zebrafish presented more perichondral bone in their hyomandibula (hm) compared to their WT siblings. This increase in bone supports our hypothesis that ARSI would promote endochondral ossification. Notably, this increase was not observed in the ceratohyal (ch). We speculate that these differences may be attributed to the timing of ossification, as the hm element undergoes ossification earlier than the ch element. Consequently, the differences seen in hm could be a consequence of more mature bone being present at this time point. Further investigations using skeletal staining at later time points, when perichondral bone in the ch element becomes more evident, might also reveal an increase in that element. Another possibility for those differences is that even though the ch and hm ossify through endochondral ossification, their development has some differences. For example, while in the ch, the growth plate is followed by a zone of calcification and a degradation zone with chondroclasts and osteoblasts, in the hm, the hyperthophic zone of the growth plate directly adjacent to adipose tissue, and because of that is called tubular bone (Weigele & Franz-Odendaal, 2016). Although the direct

influence of these differences on perichondral bone formation and timing is unclear, it raises the possibility that an increase in ARSI may produce different results depending on the origin and development of these elements.

In summary, putative Tg zebrafish were generated for all 4 constructs. It was confirmed that the constructs had the correct sequence and were integrated into the zebrafish genome by the presence of green hearts and by DNA sequencing. It was demonstrated that RNA transcription was happening in the Tg zebrafish, but protein overexpression was not ultimately observed. Nonetheless, XRF imaging and skeletal preps revealed phenotypic differences between the putative Tg and WT fish. Those phenotypes may be related to ARSI function, but most likely, they are due to the Tg fish degrading overexpressed ARSI or counteracting ARSI overexpression through unknown mechanisms. The lack of animal models overexpressing sulfatases makes it hard to demonstrate how they act *in vivo*. It would be beneficial to have literature reporting possible pitfalls in generating Tg zebrafish by Gateway and problems in the overexpression of sulfatases. As the overexpression of ARSI protein in the putative Tg fish could not be confirmed, we partially addressed the hypothesis and objectives for Chapter 4, and more studies are needed to assess the biological function of ARSI.

Acknowledgements

Part of the research described in this paper was performed at the Canadian Light Source, a national research facility of the University of Saskatchewan, which is supported by the Canada Foundation for Innovation (CFI), the Natural Sciences and Engineering Research Council (NSERC), the National Research Council (NRC), the Canadian Institutes of Health Research (CIHR), the Government of Saskatchewan, and the University of Saskatchewan.

CHAPTER 5: SUMMARY, LIMITATIONS, AND FUTURE DIRECTIONS

5.1 Summary

Many examples discussed throughout this thesis show that inadequate sulfation of PGs can cause various skeletal defects (Cortes et al., 2009; Gualeni et al., 2013; Settembre et al., 2007; Settembre et al., 2008). However, how PG sulfation levels are modulated during endochondral ossification and how they influence skeletal phenotypes needs to be clarified. Sulfatases are promising target genes to control those processes since many of them remove sulfate esters from PGs, and sulfatase mutations lead to a variety of skeletal phenotypes (Otsuki et al., 2008; Sardiello et al., 2005; Settembre et al., 2007). This thesis aimed to investigate the role of a novel sulfatase (ARSI) in cartilage maturation during endochondral ossification. I used various techniques to achieve this goal, including synchrotron XRF, immunohistochemistry, *in situ* hybridization, cell culture, Western blot, and Gateway-tol2 transgenesis.

The results presented here confirmed that sulfate esters decreased in mature cartilage compared to immature cartilage during endochondral ossification. Sulfate esters reflect sulfate groups linked to PGs, which in cartilage are mainly CSPGs (Hanson et al., 2004). ARSB and GALNS, the only two sulfatases previously reported to act on CSPGs, were not differentially expressed in mature cartilage, making them less likely to decrease sulfate esters in that region (Bhattacharyya et al., 2015; Tomatsu et al., 2005). Here I showed increased levels of ARSI in mature cartilage compared to immature cartilage in chick and mouse embryos and in cell culture experiments. Further biochemical experiments by our collaborators revealed that ARSI was a novel CSPG sulfatase. To our knowledge, this is the first report of ARSI as a CSPG sulfatase. Also, no previous papers showed *ARSI* mRNA and protein expression in chick or ARSI protein expression in mice and ATDC5 cells. Together, the expression and biochemical results strongly suggest that ARSI decreases sulfate ester levels in mature cartilage.

Just two papers reported experimental work with ARSI, and one additional paper mentioned this gene expression in skeletal development (Obaya, 2006; Oshikawa et al., 2009; Ratzka et al., 2010). The findings presented here regarding ARSI biochemical activity and expression patterns in developing bones and cells significantly contribute to current knowledge about this sulfatase. The specific expression of ARSI in mature cartilage in different animal and cell models also raises questions regarding the role of ARSI in cartilage maturation. I attempted to test ARSI function by generating the first ARSI Tg animal model in the literature. *ARSI* DNA and mRNA were observed in the putative Tg zebrafish, but ultimately, protein overexpression was not proven due to possible post-translational regulatory mechanisms. More experiments should be done to investigate ARSI's biological function, and suggestions will be discussed further in this chapter.

5.2 Limitations and future directions

5.2.1 ATDC5 cells to study ARSI biological function.

One limitation of this study was that ARSI overexpression was not confirmed in our putative Tg zebrafish, probably due to post-translational mechanisms. Cell culture experiments are usually a simpler alternative to genetically modified animals to test gene function. Therefore, they could be used in preliminary studies to generate data on ARSI localization and function. One option would be the use of ATDC5 cells, which are a well-established cell line, to study chondrocyte differentiation (Yao & Wang, 2013). To test if ATDC5 cells are a good model to study ARSI function, I performed initial experiments transfecting those cells with constructs previously generated in Eames Lab. ATDC5 cells were transfected with three different constructs: (1) *GFP* fused to *mCherry* by a 2A peptide region (P2A), (2) hARS*I* fused to *mCherry* by P2A, and (3) hARSI fused with mCherry (Figure 5.1A). The P2A sequence consists of a coding sequence of a self-cleaving peptide that links our "protein of interest" (ARSI) to a fluorescent protein (mCherry). During translation, the P2A sequence leads to ribosome

“skipping, " resulting in a missing peptide bond and two separated proteins (Liu et al., 2017). This avoids potential folding issues when a fluorescent protein is fused to the gene of interest. Construct (1) was used as a transfection control and allowed for the visualization of GFP and mCherry in the same cells and location. Both constructs (2) and (3) lead to ARSI overexpression, which was observed by mCherry presence in the transfected cells (Figure 5.1B). Using these similar constructs in zebrafish would be interesting, where *hARSI* could be fused to *mCherry* or *GFP* through a P2A protein, avoiding potential folding issues.

The transfection worked with all three constructs, and even without selection, expression was still observed after ten days (Figure 5.1B). ARSI-transfected and non-transfected cells were also stained with anti-ARSI antibody, and an increase in the fluorescence was observed in transfected micromasses (Figure 4.12H and I). Western blots performed ten days after transfection also showed increased ARSI band intensity when cells were transfected with *ARSI* fused with *mCherry* (Figure 5.1C). To confirm that cells could still differentiate into cartilage, Alcian blue staining was performed after transfection showing sulfated PGs in all groups (Figure 5.1D). Some options to obtain stronger or more specific bands could include to use anti-mCherry or anti-GFP antibodies instead of anti-ARSI for the Western blot experiments.

With this system established, further experiments include selecting cells that express ARSI, quantifying and comparing Alcian blue levels and markers of chondrocyte maturation among the different groups. If ARSI promotes cartilage maturation, a decrease in Alcian blue and an increase in mature chondrocyte markers are expected in *ARSI*-transfected cells.

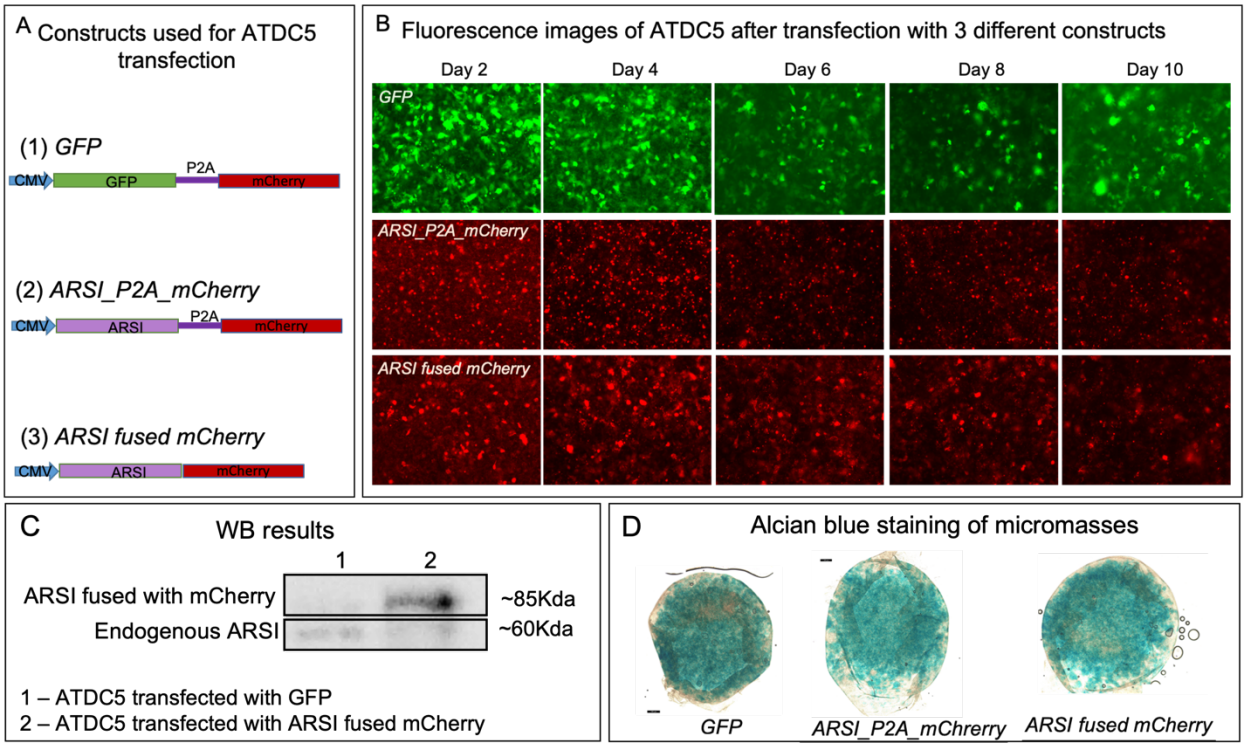


Figure 5.1 Preliminary ATDC5 transfection results. (A) ATDC5 cells were transfected with three different constructs. (B) Transfected cells were imaged every other day for ten days to check for transfection efficiency. (C) At ten days of culture, western blot results show increased ARSI expression in cells transfected with a construct containing *ARSI* fused with *mCherry*. (D) Alcian blue staining showed the presence of sulfated PGs in all transfected groups.

5.2.2 Mutant zebrafish to study ARSI biological function

Another possibility to test for ARSI function is the generation of mutant zebrafish using the CRISPR/Cas9 system. With that in mind, I designed and produced guide RNAs (gRNAs) targeting both copies of ARSI in zebrafish (*arsia* and *arsib*) and started implementing a CRISPR protocol in the lab. Maya Berscheid – a summer student I co-supervised- injected those gRNAs, with *Cas9* mRNA in zebrafish eggs, genotyped putative crispants, and performed preliminary skeletal preps to check for perichondral

bone. Injections with *arsia* gRNA were promising generating crispants. The putative mutated site was amplified with custom-made primers and sequenced. Two potential founders were discovered with mutations leading to early stop codons. (Figure 5.2A). These mutations were confirmed in intected fish (F0), asn dare currently under investigation in F1 (confirmed F0 outcrossed with WT).

To remove effects related to potential mosaicism in injected fish (F0), *arsia* putative mutant fish were raised into adulthood and in-crossed to obtain the first generation (F1). F1 fish were fixed at 8 dpf, stained with Alcian blue and alizarin red, and the amount of perichondral bone was scored and compared to WT fish. There was significantly less perichondral bone in the hm of the F1 mutants compared to WT (Figure 5.2C). This result suggests that *arsia* promotes endochondral ossification. However, more experiments need to be performed. Since putative, but not confirmed, mutants were incrossed to generate F1, we are not sure which mutations exist in those fish. Also, even if there are mutations, when the fish are incrossed, it is possible that part of the progeny will not have them. Even with this caveat, the skeletal prep results are promising. This is because there was a significant decrease in perichondral bone in the pool of fish regardless of whether that included some WT siblings. The observed mutations were just 1 bp, so those differences were not clearly visualized on an agarose gel. Future directions include finding a better protocol to genotype embryos and selecting mutants before scoring their bone levels.

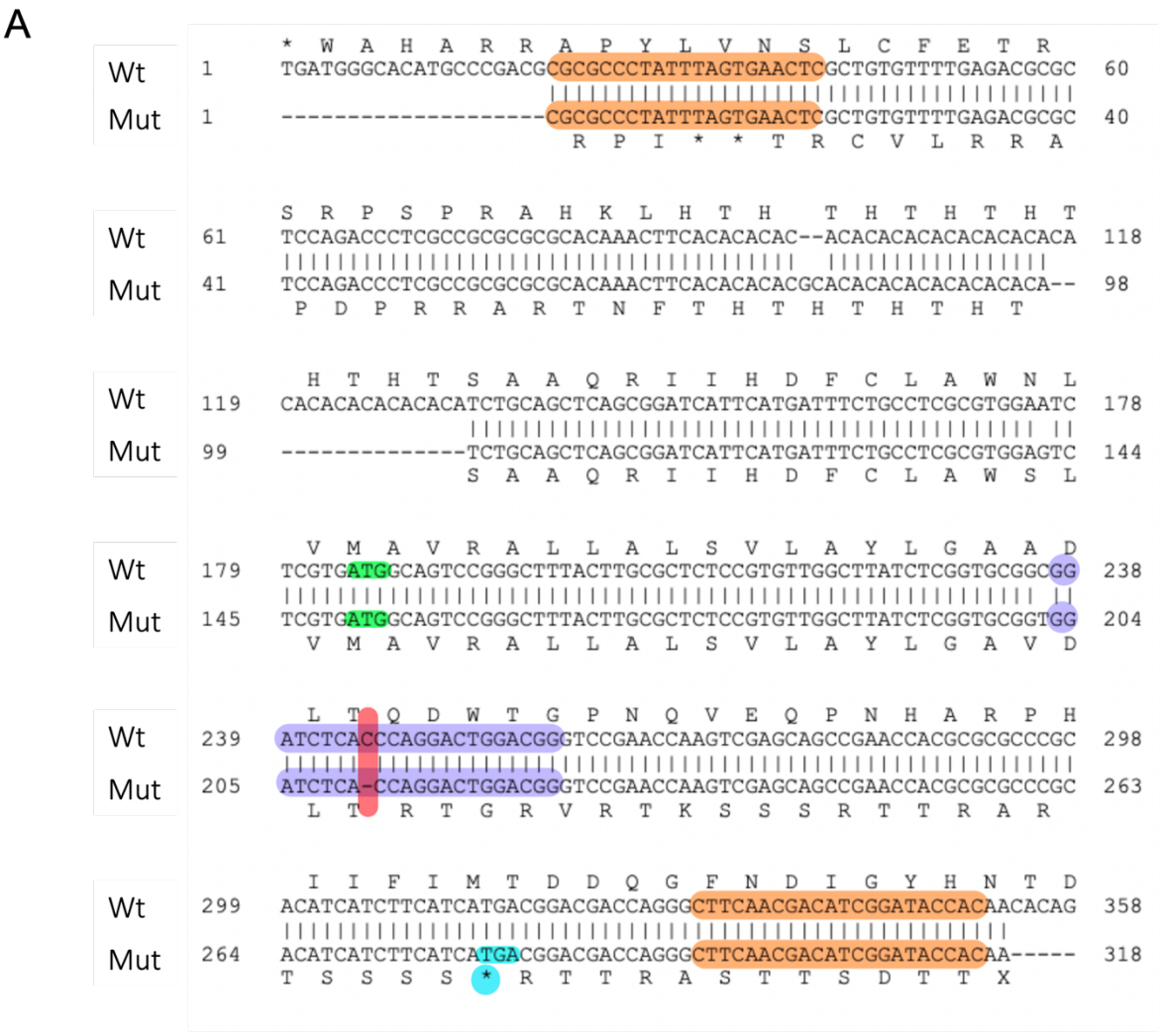

A
Wt
Mut
* W A H A R R A P Y L V N S L C F E T R
TGATGGGCACATGCCCGACGCGCGCCCTATTTAGTGAACTCGCTGTGTTTTGAGACGCGC
CGCGCCCTATTTAGTGAACTCGCTGTGTTTTGAGACGCGC
R P I * * T R C V L R R A
S R P S P R A H K L H T H T H T H T H T
TCCAGACCCTCGCCGCGCGCGCACAAACTTCACACACAC--ACACACACACACACACACA
TCCAGACCCTCGCCGCGCGCGCACAAACTTCACACACACGCACACACACACACACACA--
P D P R R A R T N F T H T H T H T H T
H T H T S A A Q R I I H D F C L A W N L
CACACACACACACATCTGCAGCTCAGCGGATCATTCATGATTTCTGCCTCGCGTGGAATC
TCTGCAGCTCAGCGGATCATTCATGATTTCTGCCTCGCGTGGAGTC
S A A Q R I I H D F C L A W S L
V M A V R A L L A L S V L A Y L G A A D
TCGTGATGGCAGTCCGGGCTTTACTTGCGCTCTCCGTGTTGGCTTATCTCGGTGCGGCGG
TCGTGATGGCAGTCCGGGCTTTACTTGCGCTCTCCGTGTTGGCTTATCTCGGTGCGGTGG
V M A V R A L L A L S V L A Y L G A V D
L T Q D W T G P N Q V E Q P N H A R P H
ATCTCACCCAGGACTGGACGGGTCCGAACCAAGTCGAGCAGCCGAACCACGCGCGCCCGC
ATCTCA-CCAGGACTGGACGGGTCCGAACCAAGTCGAGCAGCCGAACCACGCGCGCCCGC
L T R T G R V R T K S S S R T T R A R
I I F I M T D D Q G F N D I G Y H N T D
ACATCATCTTCATCATGACGGACGACCAGGGCTTCAACGACATCGGATACCACAACACAG
ACATCATCTTCATCATGACGGACGACCAGGGCTTCAACGACATCGGATACCACAA-----
T S S S S * R T T R A S T T S D T T X
B
8dpf ch
Normalized Bone Score
2.0
1.5
1.0
0.5
0.0
WT
MUT
C
8dpf hm
****
Normalized Bone Score
1.5
1.0
0.5
0.0
WT
MUT
B'
B''
C'
C''

Figure 5.2 *arsi* putative crispants. (A) Alignment of *arsia* DNA and amino acids sequences of WT and Mut zebrafish. The putative mutated site was amplified with custom-made primers (sequences in orange). Start codon in green and gRNA binding site in purple. DNA sequencing of the PCR product showed a 1 bp deletion in the gRNA binding region (highlighted in red). The sequence from the Mut was translated using Serial Cloner 2.6, and an early stop codon (highlighted in blue) was observed as the result of a 1bp deletion on the gRNA binding domain. (B) there was no statistical difference between the amount of perichondral bone in ch of mutants and WT. (B' - B'') Representative images of the scored ch of WT and mutant fish at 8 dpf. (C) perichondral bone was significantly reduced in mutant hyomandibula compared to WT. (C'-C'') Representative images of the scored hm of WT and mutant fish at 8 dpf. One clutch was analyzed (N = 19 WT, 28 MUT). Error bars represent the standard error of the mean. Statistics were calculated using the Unpaired T-test with Welch's correction. Abbreviations: ch: ceratohyal; hm: hyomandibula; MUT: mutant; WT: wild type.

While looking for human ARSI function using a zebrafish model, it is important to point out that ARSI biological function is yet not described for any animal. To better understand results in mutant zebrafish, it would be interesting to check for the expression and function of ARSI in these animals. Zebrafish have two copies of the ARSI (*arsia* and *arsib*), there are no commercial antibodies against zebrafish arsi proteins, and the antibody used in this thesis was unspecific in zebrafish samples Figure 4.11. An alternative and more specific option to assess *arsia* and *arsib* expression is to perform *ISH* for both genes. To do that, I designed probes for specific regions of *arsia* and *arsib* and made preliminary experiments to check for their expression in zebrafish ch. Preliminary results showed *arsia* expression in different zebrafish tissues, including chondrocytes in the ceratohyal (Figure 5.3 B-B'). On the other hand, *arsib* did not seem to be expressed above background levels in any of the tissues observed (Figure 5.3 C-C'). Optimization of the ISH protocol is needed to reduce background staining and confirm the location of *arsi* mRNA in zebrafish cartilages.

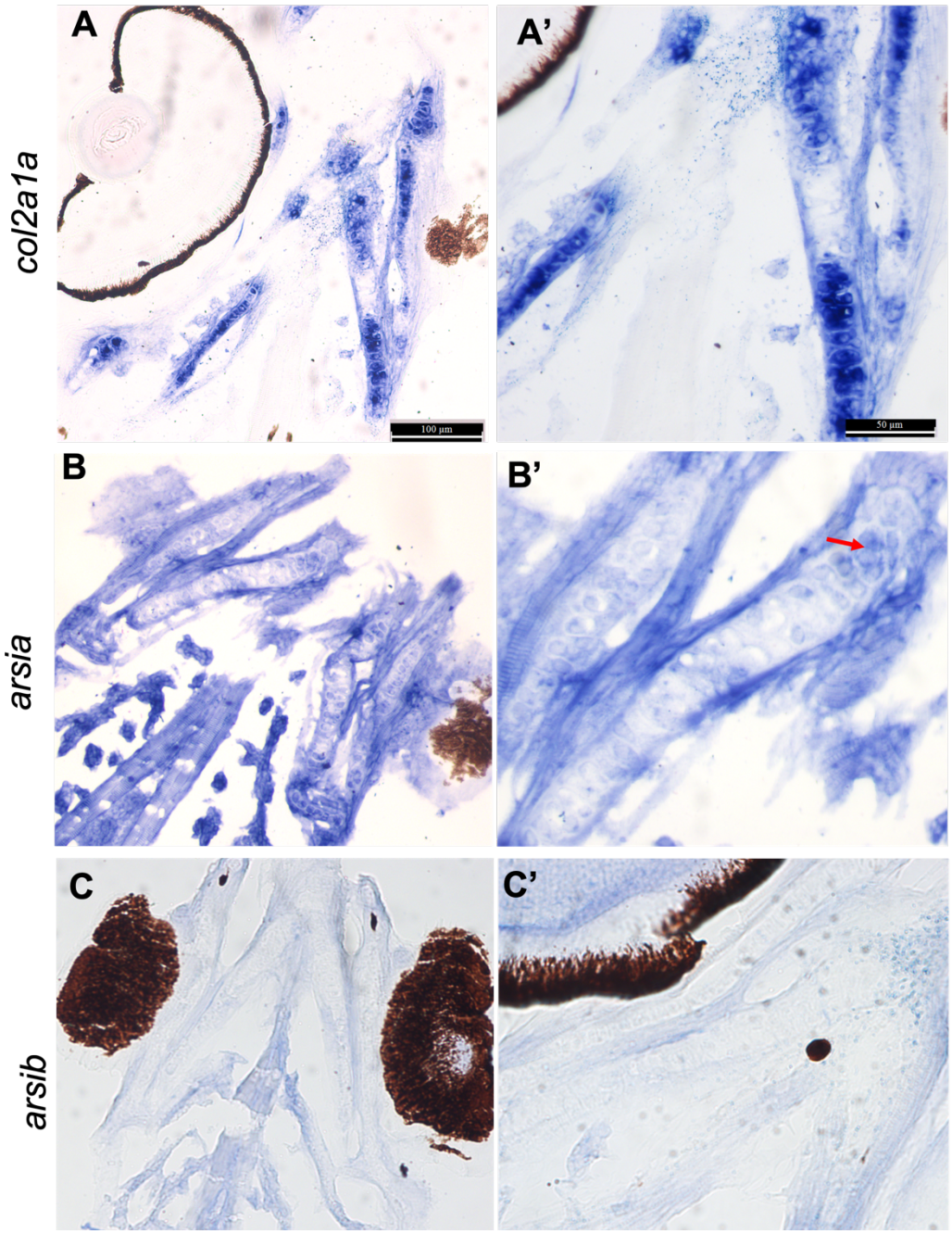


Figure 5.3 Preliminary *ISH* for zebrafish *arsia* and *arsib*. (A) *ISH* for *col2a1a* was used as a control to stain zebrafish cartilage at 7 dpf. (A') higher magnification shows *col2a1a* is expressed in immature cartilage of zebrafish ch. (B) *arsia ISH* shows signal in different zebrafish tissues. (B') Some chondrocytes inside the ch seem to have stronger *arsia* expression (red arrow). (C-C') *arsib ISH* didn't show signals above background levels in fish tissues.

5.2.3 Speculations about ARSI function in cartilage maturation

The finding that ARSI is mainly expressed in mature cartilage in different models suggests that this enzyme is involved in cartilage maturation. Our preliminary results with crispants indicate that ARSI reduction decreases perichondral bone, an indirect measurement of cartilage maturation. One possible mechanism by which ARSI might control the timing of cartilage maturation is by changing CSPG sulfation levels. A reduction of PG sulfation is one of the first steps for PG degradation, and at the same time, lower levels of PGs accelerate perichondral bone formation (Eames et al., 2011; Settembre et al., 2008). PG sulfation also affects binding affinities and the signalling activity of growth factors in the cartilage matrix, which might influence the expression of cartilage maturation genes (Brown & Eames, 2016).

The growth factor IHH is a central regulator of chondrocyte differentiation, expressed mainly by prehypertrophic chondrocytes (mature cartilage) (Long & Ornitz, 2013). One of the functions of Ihh is to stimulate the production of parathyroid hormone-related peptide (PTHrP) in proliferating chondrocytes (immature cartilage). PTHrP stimulates chondrocyte proliferation, inhibits chondrocyte hypertrophy, and delays Ihh production, keeping immature cartilage cells in an undifferentiated stage (Yang et al., 2015). CSPGs are needed to form an IHH gradient in the growth plate. These PGs affect Ihh diffusion by aiding its trafficking to its target cells or by protecting these molecules from degradation (Cortes et al., 2009). Since *ARSI* is a CSPG sulfatase expressed in partially overlapping *Ihh* domains, it is tempting to speculate that ARSI acts on chondrocyte maturation by changing IHH diffusion. One hypothesis is that, as chondrocytes mature, ARSI is expressed and removes sulfur from CSPG, stimulating its degradation or altering its binding affinity for IHH. A decrease in IHH diffusion reduces PTHrP and allows the chondrocytes to enter hypertrophy and endochondral ossification to progress. Interestingly, Ihh has a higher affinity to CS4, which seems to be the main ARSI target in our biochemical analysis (Cortes et al., 2009). More experiments need to be done to test if IHH diffusion and cartilage maturation are affected by ARSI.

Defects in the timing of cartilage maturation and bone formation may contribute to dwarfism, ectopic bone formation, and other severe conditions. If ARSI removes sulfate esters from mature cartilage and promotes endochondral ossification, this gene could be used as a target to regulate skeletal development and disease. For example, manipulating ARSI might be of great value to understanding osteoarthritis (OA) etiology and proposing new treatments. Changes in OA cartilages, such as decreased levels of sulfated PGs, increased collagen 10, and chondrocyte hypertrophy, resemble cartilage maturation and bone formation steps during endochondral ossification (Brown & Eames, 2016). Finding a way to delay cartilage maturation could help to slow disease progression in those cases. Controlling the timing of endochondral ossification could also be a key to preventing it from happening prematurely, which might be the case for short-stature phenotypes observed in many skeletal diseases, such as MPSs.

5.2.4 Conservation of molecular pathways vs. intrinsic differences between species

The conservation of sulfatases and molecular pathways across different organisms allows us to make inferences about gene function in different models (Sardiello et al., 2005). In this study, samples from various sources were utilized: human (HeLa cells), mouse (ATDC5 cells), zebrafish, and chick, to build up the knowledge about ARSI. This was possible because sulfatases are highly conserved across vertebrates, especially at their catalytic core, which suggests their function is also conserved. Also, chondrocytes from teleosts, birds, and mammals express a highly conserved set of genes that play key roles in endochondral ossification, such as *Col2a1*, *Col10a1*, *Sox9,* and *Runx2* (Eames et al., 2012), indicating it is possible to use one animal model to make inferences about this process in other vertebrates.

To better compare results among species, similar developmental stages of chick, mouse, and zebrafish embryos were used to look for ARSI expression and cartilage maturation. Although the overall chondrogenesis and osteogenesis pattern is conserved, one must be careful when comparing findings in the different animal models

because of intrinsic differences between species. One important point is that because of skeletal heterochrony, the timing of ossification events tends to vary among species (Hall, 1984). For example, previous results from our lab show that even in similar developmental stages, terminal differentiation of chondrocytes (characterized by a reduction in *COL10A1* expression) occurs faster in chick then in mouse (Gomez-Picos, 2020). These differences are significant because they could reflect mRNA and protein expression domain changes in those animals. It is, therefore, very challenging to draw an exact comparison among species, even if the stages and cells look similar.

Growth plates also vary significantly between species based on their gross morphology and cellular characteristics. Hypertrophic chondrocytes, for example, are much larger and more distinct in mice humerus than in chick humerus and zebrafish ceratohyal (Mangiavini et al., 2016). Variations in chondrocyte hypertrophy might be related to different mechanisms of bone elongation across animals. While chick and zebrafish growth plates mainly grow based on the rate of cell proliferation, mouse growth plates elongate mostly because of cell hypertrophy and matrix deposition (Barreto & Wilsman, 1994; Heubel et al., 2021). Maybe because of these cellular differences, mature and immature cartilage regions are well-defined in mice but not in chicks. For example, in mouse humeri, cells were organized in columns, and prehypertrophic and hypertrophic cells were clearly in the center of the element. On the other hand, the chick growth plate had cells in different stages of differentiation located next to each other in different zones (Leach & Gay, 1987; Pines & Hurwitz, 1991). These characteristics made it harder to define a limit between immature and mature cartilage in the chick humerus. Thus, careful consideration must be taken while choosing regions to compare between species.

Mouse experiments were less explored than chick experiments in this thesis. This happened because I had easier access to chick embryos, and the larger organ size of chick humerus was appealing for our experiments. It would be interesting to perform IHC and ISH in different depths of mouse cartilage to check for conservation of the ARSI expression pattern. XRF imaging could also be performed in mice samples to

confirm a sulfur reduction in mature cartilage vs. immature cartilage, as observed in chicks.

5.2.5 Gene expression levels and location may vary in different developmental stages.

A big challenge when working with developmental biology is to observe what happens in one or a few stages of development and extrapolate that for the rest of the developmental process. This is especially true when investigating gene expression during development. Regions of expression and abundance of proteins and mRNAs commonly change in different developmental stages, which may lead to different functions for the same protein at different time points. For example, ARSB is not detected in E12.5 and 14.5 in mice; however, there is a strong signal at E16.5 in the mice developing bones (Ratzka et al., 2010). ARSJ, on the other hand, is expressed in some mice cartilages at E14.5 but not at E16.5 (Ratzka et al., 2010).

To answer the question: “Is ARSI responsible for the decrease of sulfur observed in mature cartilage?” we used a defined developmental stage where mature cartilage, immature cartilage, and perichondral bone were present. However, a broader study involving other developmental time points or live imaging would be beneficial to answer the bigger question of “What is ARSI’s biological function during endochondral ossification?”. Experiments conducted at other time points could determine when gene and protein expression are turned on/off and how that expression relates to key points of endochondral ossification (e.g., Is ARSI expressed in other cells before they become mature chondrocytes? Does it have a role in establishing the cartilage template? Can we see co-localization of mRNA and protein at any time point?). With that in mind, I performed some preliminary ISH in HH32 chick humerus. There were almost no morphological differences between immature and mature chondrocytes (Figure 5.4A) and a thin layer of forming perichondral bone (Figure 5.4B). At this stage, *COL2A1* was expressed in the whole humerus, and *IHH* was highly expressed in the center of the humerus (Figure 5.4C and D). *ARSI* was expressed in a couple of layers of

chondrocytes all around the cartilaginous template, and *GALNS* seemed to be a little more intense in the central region of the template (Figure 5.4E and F). Interestingly, at HH32, *ARSI* expression was also in superficial chondrocytes, similar to the expression shown at HH36 (Figure 3.8G-G'). However, at HH32, the cartilage is mostly immature. This might suggest that ARSI has other functions during endochondral ossification (not strictly related to cartilage maturation) which will be further discussed in the next section. ISH and IHC from different time points and tissue depths would be helpful to determine how ARSI expression changes over time and which process it might be involved during endochondral ossification.

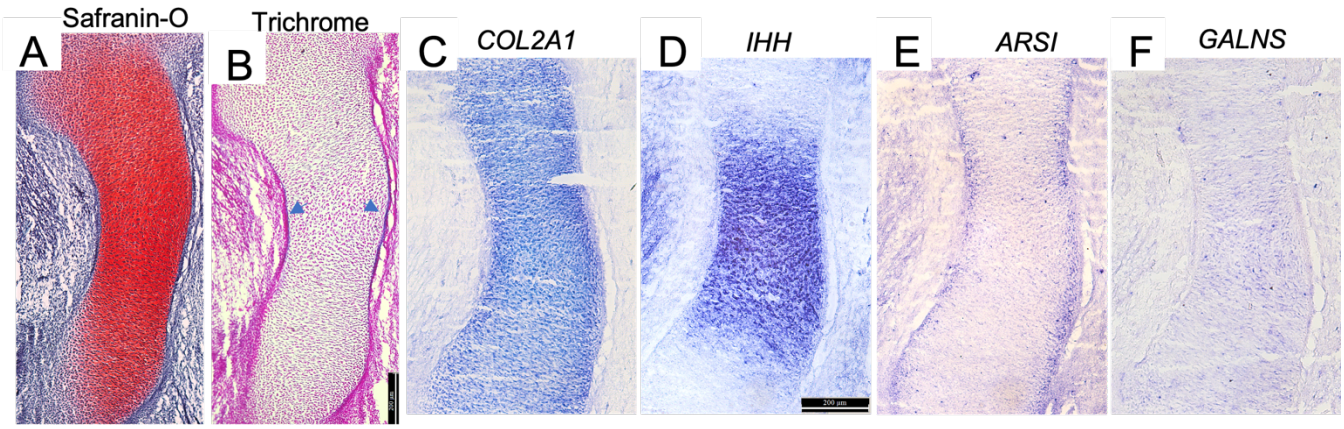


Figure 5.4 Characterization of developing chick humerus (HH32). (A) Safranin-O staining showing sulfated proteoglycans in the developing cartilage. (B) Trichrome staining showing a very thin layer of forming perichondral bone (arrowheads). (C) *COL2A1* is expressed in the whole humerus. (D) *IHH* is expressed towards the centre of the humerus. (E) *ARSI* was expressed in a couple of layers of cells surrounding the humerus. (F) *GALNS* expression is diffuse but a little more concentrated in the center of the humerus.

5.2.6 Why does ARSI expression vary with tissue depth?

A distinctive characteristic of *ARSI* compared to the other genes analyzed in skeletal development was its ring-like pattern of expression, which was observed with three different probes by *ISH* and with the anti-ARSI antibody by IHC. It is well-known that specific spatial gene expression is required for accurate tissue patterning (e.g., anterior-posterior axis and digits patterning) (Rosenberg et al., 2009; Sheth et al., 2012). However, many developmental biology analyses focus on longitudinal sections of the core region of tissues. This 2D information may hide significant portions of the developmental process or lead to wrong assumptions about gene function (Waylen et al., 2020).

3D imaging would be a great asset to visualizing mRNA and protein expression in the whole humerus. Whole mount *ISH* and IHC can be useful in studying tridimensional gene expression in young embryos. However, many challenges arise when applying those techniques to older embryos. For example, probes and antibodies cannot penetrate their cellular targets when embryos grow beyond a critical size. They may also be trapped in thick tissues, leading to false-positive results (Vauti et al., 2020). One option to get 3D images in those cases would be to perform a 3D reconstruction of gene and protein expression patterns obtained from slides from different depths. More experiments should be performed to assess if sulfation levels visualized in the sections in this thesis also vary in different tissue depths and if that is correlated to ARSI expression.

The reason why ARSI was expressed mainly in the superficial chondrocytes of the humerus still needs to be clarified. One possibility is that chondrocytes may mature at different times in the center and periphery of the cartilaginous template, which is reflected by ARSI expression. While performing the experiments described here, it was clear that COL10, a marker of mature chondrocytes, also varied between sections. In slides with adjacent sections, some of those had a larger domain of COL10 expression and more intense fluorescence than others. However, a parallel between the depth of

the tissue, ARSI expression, COL10 expression, and the state of individual chondrocyte maturation was not traced.

Another possibility to explain the ARSI distinctive expression pattern could be related to cartilage appositional growth. One study using human fetal endochondral bones at a similar stage to the bones used in the present study (20^{th}-22^{nd} weeks) showed differences in the chondrocytes in the periphery and center of the cartilaginous template (Pazzaglia et al., 2019). Different lacunae morphology and proliferation patterns indicated that peripheral chondrocytes participate mainly in cartilage appositional growth, whereas central chondrocytes participate primarily in interstitial growth (Pazzaglia et al., 2019). They focused this study on a transition zone between the mineralizing mature cartilage and the region where the epiphyses start to enlarge. This is the same region where I observed expression of *ARSI* mRNA in chick humerus and could also see differences in morphology of chondrocytes in the periphery and the centre of the humerus.

If ARSI has a role in cartilage appositional growth, we might look at the ISH result differently. The results showing *ARSI* surrounding the whole cartilage template at HH32 could indicate that this template is under intense appositional growth at this stage but not in HH36. In HH36, ARSI was mostly expressed right where the epiphyses start to enlarge and could be responsible for greater appositional growth in that region compared to the rest of the developing humerus. In fact, the existence of two patterns of chondrocyte participating in interstitial and appositional growth could be necessary for the clear variation between the epiphyseal and diaphyseal diameters present in most long bones (Pazzaglia et al., 2019). It is interesting to notice that ARSI transgenic and mutant fish had phenotypic changes just in their hyomandibula and not in their ceratohyal. While the hm has a triangular shape - similar to a humerus head - the ceratohyal does not have much diameter variation. Suppose ARSI is operating on those chondrocytes that contribute to the differences in diameter between the epiphysis and the diaphysis. In that case, it could be involved in a similar process to shape the hyomandibula.

Lastly, ARSI ring-like expression might be related to attachment points for ligaments and tendons. These regions, called entheses, comprise fibrous (perforating mineralized collagen fibres) or fibrocartilage tissues (Lu & Thomopoulos, 2013). Little is known about the molecular control of cell maturation at entheses (Benjamin et al., 2006). The specific expression of ARSI in this region may indicate its participation in enthesis development. Its expression could also be a response of those cells to mechanical load by the developing entheses. Tendons and ligaments were not the focus of our study, so they were not closely analyzed.

5.2.7 ARSI localization

The ARSI localization is still up for discussion, as there are different reports of this sulfatase being present in the ECM, lysosomes, or the Golgi (Obaya, 2006; Oshikawa et al., 2009; Ponten et al., 2008). The IHC results presented here indicate that ARSI was expressed close to the plasma membrane in mature chondrocytes of chick and mouse humeri and ATDC5 cells (Figure 5.5). Higher magnification of mature chick cartilage showed some dot-like expression of ARSI - which could correspond to lysosomal expression – but most of the fluorescence signal was in the cell periphery (Figure 5.5A). Co-staining of ARSI and COL10 antibodies in chick showed that while both are expressed in mature cartilage, for the most part, they do not have the same subcellular localization. As expected, COL10 was expressed in the ECM, but ARSI was mostly intracellular (Figure 5.5B). ARSI expression was also observed mostly close to the cell membranes in mature mouse cartilage and ATDC5 cells (Figure 5.5C and D).

Our biochemical results showed that ARSI acts on cartilage PGs, so it likely plays a role in remodelling the ECM during the transition from a cartilaginous into a bony skeleton. However, the lack of ARSI expression in the ECM makes it challenging to explain how this sulfatase could remodel the ECM *in vivo*. One possibility is that the used techniques - IHC plus fluorescence microscopy - might not be enough to detect ARSI in the ECM. Another possibility is that the highly sulfated PGs in immature cartilage are degraded and replaced with less sulfated PGs in mature cartilage. ARSI

could be acting removing sulfur fromÏÏ the newly produced PGs before secretion. To better understand ARSI subcellular localization, further experiments include co-staining of ARSI and fluorescence markers specific for lysosomes, Golgi, and pericellular matrix, visualization of cells by confocal microscopy to obtain a better resolution, and experiments involving co-immunoprecipitation.

Our XRF maps and quantifications showed a reduction of sulfur and more specific sulfate esters in mature cartilage compared to immature cartilage. To improve the specificity of our results, it would be interesting to see how the decrease of sulfur in mature cartilage happens at a subcellular level using beamlines that allow for this type of imaging. Those results, coupled with a better understanding of where ARSI is expressed, would lead to better inferences about ARSI's biological function.

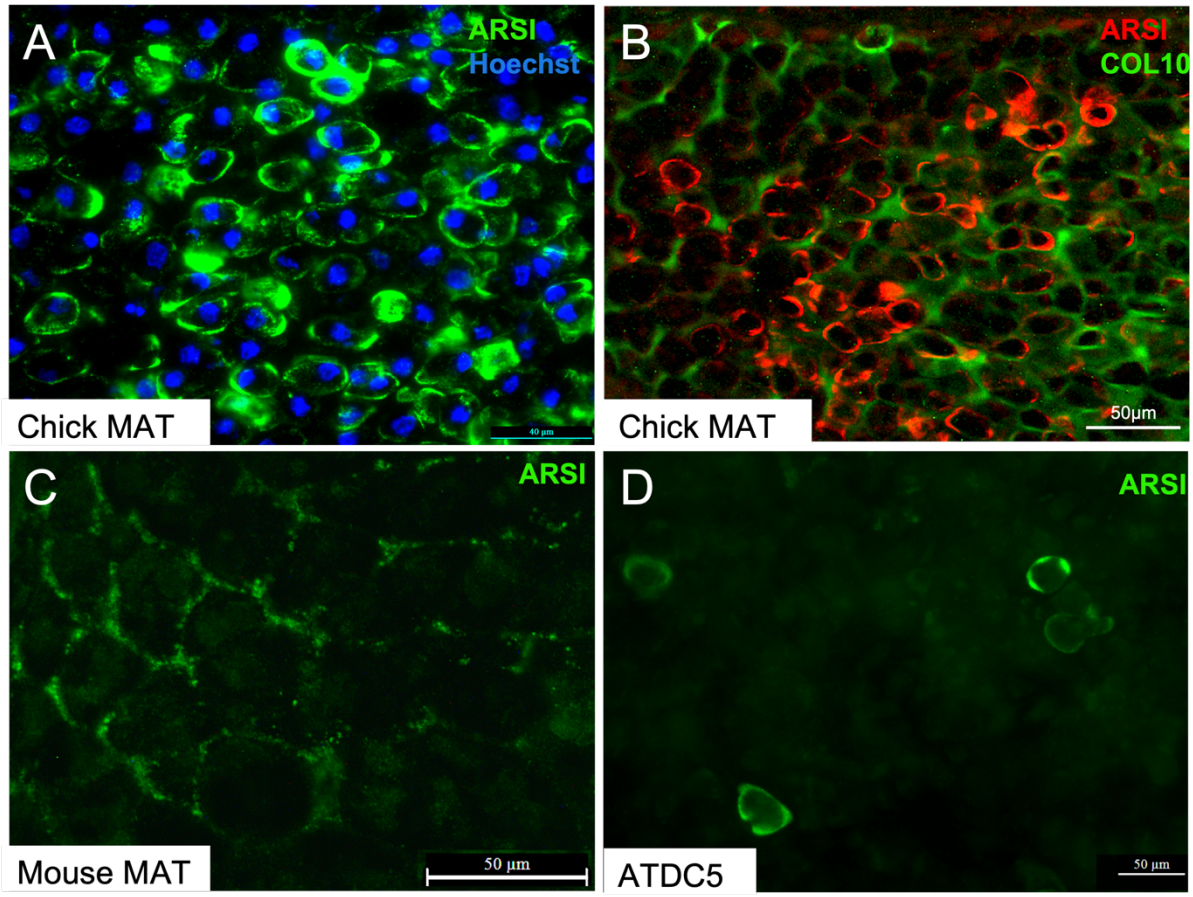


Figure 5.5 ARSI localization. (A) IHC in chick mature chondrocytes show dot-like expression in the cell cytoplasm, but most of the expression comes from the cell periphery. (B) Co-staining with ARSI and COL10 showed that most ARSI is not in the ECM. (C) IHC in mouse mature chondrocytes shows that ARSI is mainly expressed in cell periphery like chick results. (D) ARSI expression *in vitro,* in differentiating ATDC5 cells, showed a similar pattern to *in vivo* expression. Abbreviations: MAT: mature cartilage

In closing, this work showed that a reduction of sulfur happens in mature cartilage during endochondral ossification, and that a novel sulfatase -ARSI- is differentially expressed in this region, potentially controlling this reduction. Working with this gene was very challenging, considering the lack of data in the literature. Nonetheless, this research was extremely important for my training as a scientist. During this work, I could generate new data and compile and interpret data from collaborators considering my own hypothesis and experiments. I learned various techniques (XRF synchrotron imaging, Gateway-Tol2 transgenesis, cloning, PCR, immunocytochemistry, *in situ* hybridization, ATDC5 cell culture, among others), worked

with different animal models, and designed new lab protocols (such as Western Blot and CRISPR-Cas9). Besides that, I was deeply involved in troubleshooting the experiments with the transgenic zebrafish, which did not work as predicted. While this thesis propelled the study of a novel sulfatase, there is much to be investigated about the potential ARSI roles in skeletal development. The points discussed in this chapter are just some of the possibilities for future research that can give us insight into cartilage maturation, endochondral ossification, and potential applications in health.

REFERENCES

Aghajanian, P., & Mohan, S. (2018). The art of building bone: emerging role of chondrocyte-to-osteoblast transdifferentiation in endochondral ossification. *Bone Res*, *6*, 19. https://doi.org/10.1038/s41413-018-0021-z

Alini, M., Matsui, Y., Dodge, G. R., & Poole, A. R. (1992). The extracellular matrix of cartilage in the growth plate before and during calcification: changes in composition and degradation of type II collagen. *Calcif Tissue Int*, *50*(4), 327-335. http://www.ncbi.nlm.nih.gov/pubmed/1571844

Althoff, J., Quint, P., Krefting, E. R., & Hohling, H. J. (1982). Morphological studies on the epiphyseal growth plate combined with biochemical and X-ray microprobe analyses. *Histochemistry*, *74*(4), 541-552. http://www.ncbi.nlm.nih.gov/pubmed/7107329

Anthwal, N., Joshi, L., & Tucker, A. S. (2013). Evolution of the mammalian middle ear and jaw: adaptations and novel structures. *J Anat*, *222*(1), 147-160. https://doi.org/10.1111/j.1469-7580.2012.01526.x

Ashique, A. M., Atake, O. J., Ovens, K., Guo, R., Pratt, I. V., Detrich, H. W., 3rd, Cooper, D. M. L., Desvignes, T., Postlethwait, J. H., & Eames, B. F. (2022). Bone microstructure and bone mineral density are not systemically different in Antarctic icefishes and related Antarctic notothenioids. *J Anat*, *240*(1), 34-49. https://doi.org/10.1111/joa.13537

Attanasio, C., Nord, A. S., Zhu, Y., Blow, M. J., Li, Z., Liberton, D. K., Morrison, H., Plajzer-Frick, I., Holt, A., Hosseini, R., Phouanenavong, S., Akiyama, J. A., Shoukry, M., Afzal, V., Rubin, E. M., FitzPatrick, D. R., Ren, B., Hallgrimsson, B., Pennacchio, L. A., & Visel, A. (2013). Fine tuning of craniofacial morphology by distant-acting enhancers. *Science*, *342*(6157), 1241006. https://doi.org/10.1126/science.1241006

Baker, M. J., Trevisan, J., Bassan, P., Bhargava, R., Butler, H. J., Dorling, K. M., Fielden, P. R., Fogarty, S. W., Fullwood, N. J., Heys, K. A., Hughes, C., Lasch, P., Martin-Hirsch, P. L., Obinaju, B., Sockalingum, G. D., Sule-Suso, J., Strong, R. J., Walsh, M. J., Wood, B. R., Gardner, P., & Martin, F. L. (2014). Using Fourier transform IR spectroscopy to analyze biological materials. *Nat Protoc*, *9*(8), 1771-1791. https://doi.org/10.1038/nprot.2014.110

Bankhead, P. (2022). *Introduction to Bioimage Analysis* https://bioimagebook.github.io/chapters/1-concepts/3-bit_depths/imagej.html

Barreto, C., & Wilsman, N. J. (1994). Hypertrophic chondrocyte volume and growth rates in avian growth plates. *Res Vet Sci*, *56*(1), 53-61. https://doi.org/10.1016/0034-5288(94)90196-1

Bedell, V. M., & Ekker, S. C. (2015). Using engineered endonucleases to create knockout and knockin zebrafish models. *Methods Mol Biol*, *1239*, 291-305. https://doi.org/10.1007/978-1-4939-1862-1_17

Bedell, V. M., Wang, Y., Campbell, J. M., Poshusta, T. L., Starker, C. G., Krug, R. G., 2nd, Tan, W., Penheiter, S. G., Ma, A. C., Leung, A. Y., Fahrenkrug, S. C., Carlson, D. F., Voytas, D. F., Clark, K. J., Essner, J. J., & Ekker, S. C. (2012). In vivo genome editing using a high-efficiency TALEN system. *Nature*, *491*(7422), 114-118. https://doi.org/10.1038/nature11537

Behonick, D. J., & Werb, Z. (2003). A bit of give and take: the relationship between the extracellular matrix and the developing chondrocyte. *Mechanisms of Development*, *120*(11), 1327-1336. https://doi.org/10.1016/j.mod.2003.05.002

Benjamin, M., Toumi, H., Ralphs, J. R., Bydder, G., Best, T. M., & Milz, S. (2006). Where tendons and ligaments meet bone: attachment sites ('entheses') in relation to exercise and/or mechanical load. *J Anat*, *208*(4), 471-490. https://doi.org/10.1111/j.1469-7580.2006.00540.x

Benusiene, E., & Kucinskas, V. (2003). COL1A1 mutation analysis in Lithuanian patients with osteogenesis imperfecta. *J Appl Genet*, *44*(1), 95-102. http://www.ncbi.nlm.nih.gov/pubmed/12590186

Bhatia, S., Gordon, C. T., Foster, R. G., Melin, L., Abadie, V., Baujat, G., Vazquez, M. P., Amiel, J., Lyonnet, S., van Heyningen, V., & Kleinjan, D. A. (2015). Functional assessment of disease-associated regulatory variants in vivo using a versatile dual colour transgenesis strategy in zebrafish. *PLoS Genet*, *11*(6), e1005193. https://doi.org/10.1371/journal.pgen.1005193

Bhatia, S., & Kleinjan, D. A. (2014). Disruption of long-range gene regulation in human genetic disease: a kaleidoscope of general principles, diverse mechanisms and unique phenotypic consequences. *Hum Genet*, *133*(7), 815-845. https://doi.org/10.1007/s00439-014-1424-6

Bhattacharyya, S., Feferman, L., & Tobacman, J. K. (2014). Arylsulfatase B regulates versican expression by galectin-3 and AP-1 mediated transcriptional effects. *Oncogene*, *33*(47), 5467-5476. https://doi.org/10.1038/onc.2013.483

Bhattacharyya, S., Feferman, L., & Tobacman, J. K. (2015). Regulation of chondroitin-4-sulfotransferase (CHST11) expression by opposing effects of arylsulfatase B on BMP4 and Wnt9A. *Biochim Biophys Acta*, *1849*(3), 342-352. https://doi.org/10.1016/j.bbagrm.2014.12.009

Bielczyk-Maczynska, E., Serbanovic-Canic, J., Ferreira, L., Soranzo, N., Stemple, D. L., Ouwehand, W. H., & Cvejic, A. (2014). A loss of function screen of identified genome-wide association study Loci reveals new genes controlling hematopoiesis. *PLoS Genet*, *10*(7), e1004450. https://doi.org/10.1371/journal.pgen.1004450

Boskey, A., & Pleshko Camacho, N. (2007). FT-IR imaging of native and tissue-engineered bone and cartilage. *Biomaterials*, *28*(15), 2465-2478. https://doi.org/10.1016/j.biomaterials.2006.11.043

Briggs, D. C., & Hohenester, E. (2018). Structural Basis for the Initiation of Glycosaminoglycan Biosynthesis by Human Xylosyltransferase 1. *Structure*, *26*(6), 801-809 e803. https://doi.org/10.1016/j.str.2018.03.014

Brown, D. (2017). *UNDERSTANDING THE ROLE OF SULFUR IN CARTILAGE DEVELOPMENT: HOW STRUCTURAL PROTEOGLYCANS ARE IMPLICATED IN CARTILAGE MATURATION.* [University of Saskatchewan].

Brown, D. S., & Eames, B. F. (2016). Emerging tools to study proteoglycan function during skeletal development. *Methods Cell Biol*, *134*, 485-530. https://doi.org/10.1016/bs.mcb.2016.03.001

Buckwalter, J. A., Rosenberg, L. C., & Ungar, R. (1987). Changes in proteoglycan aggregates during cartilage mineralization. *Calcif Tissue Int*, *41*(4), 228-236. http://www.ncbi.nlm.nih.gov/pubmed/3119178

Bulow, H. E., & Hobert, O. (2006). The molecular diversity of glycosaminoglycans shapes animal development. *Annu Rev Cell Dev Biol*, *22*, 375-407. https://doi.org/10.1146/annurev.cellbio.22.010605.093433

Byers, S., van Rooden, J. C., & Foster, B. K. (1997). Structural changes in the large proteoglycan, aggrecan, in different zones of the ovine growth plate. *Calcif Tissue Int*, *60*(1), 71-78. https://doi.org/10.1007/s002239900188

Cade, L., Reyon, D., Hwang, W. Y., Tsai, S. Q., Patel, S., Khayter, C., Joung, J. K., Sander, J. D., Peterson, R. T., & Yeh, J. R. (2012). Highly efficient generation of heritable zebrafish gene mutations using homo- and heterodimeric TALENs. *Nucleic Acids Res*, *40*(16), 8001-8010. https://doi.org/10.1093/nar/gks518

Cameron, T. L., Belluoccio, D., Farlie, P. G., Brachvogel, B., & Bateman, J. F. (2009). Global comparative transcriptome analysis of cartilage formation in vivo. *BMC Dev Biol*, *9*, 20. https://doi.org/10.1186/1471-213X-9-20

Capotondo, A., Cesani, M., Pepe, S., Fasano, S., Gregori, S., Tononi, L., Venneri, M. A., Brambilla, R., Quattrini, A., Ballabio, A., Cosma, M. P., Naldini, L., & Biffi, A. (2007). Safety of arylsulfatase A

overexpression for gene therapy of metachromatic leukodystrophy. *Hum Gene Ther*, *18*(9), 821-836. https://doi.org/10.1089/hum.2007.048

Castillo-Michel, H. A., Diaz-Sanchez, A. G., Martinez-Martinez, A., & Hesse, B. (2016). Investigations of Sulfur Chemical Status with Synchrotron Micro Focused X-ray fluorescence and X-ray Absorption Spectroscopy. *Protein Pept Lett*, *23*(3), 291-299. https://doi.org/10.2174/0929866523666160108120117

Chahine, N. O., Chen, F. H., Hung, C. T., & Ateshian, G. A. (2005). Direct measurement of osmotic pressure of glycosaminoglycan solutions by membrane osmometry at room temperature. *Biophys J*, *89*(3), 1543-1550. https://doi.org/10.1529/biophysj.104.057315

Clement, A., Wiweger, M., von der Hardt, S., Rusch, M. A., Selleck, S. B., Chien, C. B., & Roehl, H. H. (2008). Regulation of zebrafish skeletogenesis by ext2/dackel and papst1/pinscher. *Plos Genetics*, *4*(7), e1000136. https://doi.org/10.1371/journal.pgen.1000136

Conen, K. L., Nishimori, S., Provot, S., & Kronenberg, H. M. (2009). The transcriptional cofactor Lbh regulates angiogenesis and endochondral bone formation during fetal bone development. *Developmental Biology*, *333*(2), 348-358. https://doi.org/10.1016/j.ydbio.2009.07.003

Cortes, M., Baria, A. T., & Schwartz, N. B. (2009). Sulfation of chondroitin sulfate proteoglycans is necessary for proper Indian hedgehog signaling in the developing growth plate. *Development*, *136*(10), 1697-1706. https://doi.org/10.1242/dev.030742

Cosma, M. P., Pepe, S., Annunziata, I., Newbold, R. F., Grompe, M., Parenti, G., & Ballabio, A. (2003). The multiple sulfatase deficiency gene encodes an essential and limiting factor for the activity of sulfatases. *Cell*, *113*(4), 445-456. https://doi.org/10.1016/s0092-8674(03)00348-9

Cubbage, C. C., & Mabee, P. M. (1996). Development of the cranium and paird fins in the zebrafish Danio rerio (Ostariophysi, Cyprinidae). *Journal of Morphology*, *229*, 121-160.

da Costa, D. S., Reis, R. L., & Pashkuleva, I. (2017). Sulfation of Glycosaminoglycans and Its Implications in Human Health and Disorders. *Annual Review of Biomedical Engineering, Vol 19*, *19*, 1-26. https://doi.org/10.1146/annurev-bioeng-071516044610

Dale, R. M., & Topczewski, J. (2011). Identification of an evolutionarily conserved regulatory element of the zebrafish col2a1a gene. *Dev Biol*, *357*(2), 518-531. https://doi.org/10.1016/j.ydbio.2011.06.020

DeLaurier, A., Alvarez, C. L., & Wiggins, K. J. (2019). hdac4 mediates perichondral ossification and pharyngeal skeleton development in the zebrafish. *PeerJ*, *7*, e6167. https://doi.org/10.7717/peerj.6167

DeLaurier, A., Eames, B. F., Blanco-Sanchez, B., Peng, G., He, X., Swartz, M. E., Ullmann, B., Westerfield, M., & Kimmel, C. B. (2010). Zebrafish sp7:EGFP: a transgenic for studying otic vesicle formation, skeletogenesis, and bone regeneration. *Genesis*, *48*(8), 505-511. https://doi.org/10.1002/dvg.20639

Deml, B., Reis, L. M., Muheisen, S., Bick, D., & Semina, E. V. (2015). EFTUD2 deficiency in vertebrates: Identification of a novel human mutation and generation of a zebrafish model. *Birth Defects Res A Clin Mol Teratol*, *103*(7), 630-640. https://doi.org/10.1002/bdra.23397

Demydchuk, M., Hill, C. H., Zhou, A., Bunkoczi, G., Stein, P. E., Marchesan, D., Deane, J. E., & Read, R. J. (2017). Insights into Hunter syndrome from the structure of iduronate-2-sulfatase. *Nat Commun*, *8*, 15786. https://doi.org/10.1038/ncomms15786

Dermitzakis, E. T., & Clark, A. G. (2009). Life after GWA Studies. *Science*, *362*(5950), 239-240. https://doi.org/ 10.1126/science.1182009

Deutsch, A. J., Midura, R. J., & Plaas, A. H. (1995). Structure of chondroitin sulfate on aggrecan isolated from bovine tibial and costochondral growth plates. *J Orthop Res*, *13*(2), 230-239. https://doi.org/10.1002/jor.1100130211

Dhar, P., Samarasinghe, R. M., & Shigdar, S. (2020). Antibodies, Nanobodies, or Aptamers-Which Is Best for Deciphering the Proteomes of Non-Model Species? *Int J Mol Sci*, *21*(7). https://doi.org/10.3390/ijms21072485

Dick, G., Akslen-Hoel, L. K., Grondahl, F., Kjos, I., & Prydz, K. (2012). Proteoglycan Synthesis and Golgi Organization in Polarized Epithelial Cells. *Journal of Histochemistry & Cytochemistry*, *60*(12), 926-935. https://doi.org/10.1369/0022155412461256

Diez-Roux, G., & Ballabio, A. (2005). Sulfatases and human disease. *Annu Rev Genomics Hum Genet*, *6*, 355-379. https://doi.org/10.1146/annurev.genom.6.080604.162334

Domowicz, M. S., Cortes, M., Henry, J. G., & Schwartz, N. B. (2009). Aggrecan modulation of growth plate morphogenesis. *Developmental Biology*, *329*(2), 242-257. https://doi.org/10.1016/j.ydbio.2009.02.024

Doyon, Y., McCammon, J. M., Miller, J. C., Faraji, F., Ngo, C., Katibah, G. E., Amora, R., Hocking, T. D., Zhang, L., Rebar, E. J., Gregory, P. D., Urnov, F. D., & Amacher, S. L. (2008). Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. *Nat Biotechnol*, *26*(6), 702-708. https://doi.org/10.1038/nbt1409

Driever, W., Solnica-Krezel, L., Schier, A. F., Neuhauss, S. C., Malicki, J., Stemple, D. L., Stainier, D. Y., Zwartkruis, F., Abdelilah, S., Rangini, Z., Belak, J., & Boggs, C. (1996). A genetic screen for mutations affecting embryogenesis in zebrafish. *Development*, *123*, 37-46. http://www.ncbi.nlm.nih.gov/pubmed/9007227

Dzobo, K., Turnley, T., Wishart, A., Rowe, A., Kallmeyer, K., van Vollenstee, F. A., Thomford, N. E., Dandara, C., Chopera, D., Pepper, M. S., & Parker, M. I. (2016). Fibroblast-Derived Extracellular Matrix Induces Chondrogenic Differentiation in Human Adipose-Derived Mesenchymal Stromal/Stem Cells in Vitro. *Int J Mol Sci*, *17*(8). https://doi.org/10.3390/ijms17081259

Eames, B. F., Amores, A., Yan, Y. L., & Postlethwait, J. H. (2012). Evolution of the osteoblast: skeletogenesis in gar and zebrafish. *BMC Evol Biol*, *12*, 27. https://doi.org/10.1186/1471-2148-12-27

Eames, B. F., de la Fuente, L., & Helms, J. A. (2003). Molecular ontogeny of the skeleton. *Birth Defects Res C Embryo Today*, *69*(2), 93-101. https://doi.org/10.1002/bdrc.10016

Eames, B. F., DeLaurier, A., Ullmann, B., Huycke, T. R., Nichols, J. T., Dowd, J., McFadden, M., Sasaki, M. M., & Kimmel, C. B. (2013). FishFace: interactive atlas of zebrafish craniofacial development at cellular resolution. *BMC Dev Biol*, *13*, 23. https://doi.org/10.1186/1471-213X-13-23

Eames, B. F., Singer, A., Smith, G. A., Wood, Z. A., Yan, Y. L., He, X., Polizzi, S. J., Catchen, J. M., Rodriguez-Mari, A., Linbo, T., Raible, D. W., & Postlethwait, J. H. (2010). UDP xylose synthase 1 is required for morphogenesis and histogenesis of the craniofacial skeleton. *Developmental Biology*, *341*(2), 400-415. https://doi.org/10.1016/j.ydbio.2010.02.035

Eames, B. F., Yan, Y. L., Swartz, M. E., Levic, D. S., Knapik, E. W., Postlethwait, J. H., & Kimmel, C. B. (2011). Mutations in fam20b and xylt1 reveal that cartilage matrix controls timing of endochondral ossification by inhibiting chondrocyte maturation [Research Support, N.I.H., Extramural]. *PLoS Genet*, *7*(8), e1002246. https://doi.org/10.1371/journal.pgen.1002246

Eberhart, J. K., He, X., Swartz, M. E., Yan, Y. L., Song, H., Boling, T. C., Kunerth, A. K., Walker, M. B., Kimmel, C. B., & Postlethwait, J. H. (2008). MicroRNA Mirn140 modulates Pdgf signaling during palatogenesis. *Nat Genet*, *40*(3), 290-298. https://doi.org/10.1038/ng.82

Edgar, R., Mazor, Y., Rinon, A., Blumenthal, J., Golan, Y., Buzhor, E., Livnat, I., Ben-Ari, S., Lieder, I., Shitrit, A., Gilboa, Y., Ben-Yehudah, A., Edri, O., Shraga, N., Bogoch, Y., Leshansky, L., Aharoni, S., West, M. D., Warshawsky, D., & Shtrichman, R. (2013). LifeMap Discovery: the embryonic development, stem cells, and regenerative medicine research portal. *PLoS One*, *8*(7), e66629. https://doi.org/10.1371/journal.pone.0066629

Edwards, S. L., Beesley, J., French, J. D., & Dunning, A. M. (2013). Beyond GWASs: illuminating the dark road from association to function. *Am J Hum Genet*, *93*(5), 779-797. https://doi.org/10.1016/j.ajhg.2013.10.012

Egawa, S., Miura, S., Yokoyama, H., Endo, T., & Tamura, K. (2014). Growth and differentiation of a long bone in limb development, repair and regeneration. *Dev Growth Differ*, *56*(5), 410-424. https://doi.org/10.1111/dgd.12136

Eisen, J. S., & Smith, J. C. (2008). Controlling morpholino experiments: don't stop making antisense. *Development*, *135*(10), 1735-1743. https://doi.org/10.1242/dev.001115

Elefteriou, F., & Yang, X. (2011). Genetic mouse models for bone studies--strengths and limitations. *Bone*, *49*(6), 1242-1254. https://doi.org/10.1016/j.bone.2011.08.021

Esko, J. D., & Selleck, S. B. (2002). Order out of chaos: assembly of ligand binding sites in heparan sulfate. *Annu Rev Biochem*, *71*, 435-471. https://doi.org/10.1146/annurev.biochem.71.110601.135458

Farnum, C. E., & Wilsman, N. J. (2001). Converting a differentiation cascade into longitudinal growth: Stereology and analysis of transgenic animals as tools for understanding growth plate function. *Current Opinion in Orthopaedics 12*(5), 428-433. https://doi.org/10.1097/00001433-200110000-00011

Farquharson, C., Whitehead, C. C., & Loveridge, N. (1994). Alterations in glycosaminoglycan concentration and sulfation during chondrocyte maturation. *Calcif Tissue Int*, *54*(4), 296-303. https://doi.org/10.1007/BF00295954

Ferrandiz, C., & Sessions, A. (2008). Preparation and hydrolysis of digoxygenin-labeled probes for in situ hybridization of plant tissues. *CSH Protoc*, *2008*, pdb prot4942. https://doi.org/10.1101/pdb.prot4942

Fisher, S., Jagadeeswaran, P., & Halpern, M. E. (2003). Radiographic analysis of zebrafish skeletal defects. *Dev Biol*, *264*(1), 64-76. https://doi.org/10.1016/s0012-1606(03)00399-3

Gaj, T., Gersbach, C. A., & Barbas, C. F., 3rd. (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol*, *31*(7), 397-405. https://doi.org/10.1016/j.tibtech.2013.04.004

Gama, C. I., Tully, S. E., Sotogaku, N., Clark, P. M., Rawat, M., Vaidehi, N., Goddard, W. A., 3rd, Nishi, A., & Hsieh-Wilson, L. C. (2006). Sulfation patterns of glycosaminoglycans encode molecular recognition and activity. *Nat Chem Biol*, *2*(9), 467-473. https://doi.org/10.1038/nchembio810

Gao, Y., Liu, S., Huang, J., Guo, W., Chen, J., Zhang, L., Zhao, B., Peng, J., Wang, A., Wang, Y., Xu, W., Lu, S., Yuan, M., & Guo, Q. (2014). The ECM-cell interaction of cartilage extracellular matrix on chondrocytes. *Biomed Res Int*, *2014*, 648459. https://doi.org/10.1155/2014/648459

Ghassibe-Sabbagh, M., Desmyter, L., Langenberg, T., Claes, F., Boute, O., Bayet, B., Pellerin, P., Hermans, K., Backx, L., Mansilla, M. A., Imoehl, S., Nowak, S., Ludwig, K. U., Baluardo, C., Ferrian, M., Mossey, P. A., Noethen, M., Dewerchin, M., Francois, G., Revencu, N., Vanwijck, R., Hecht, J., Mangold, E., Murray, J., Rubini, M., Vermeesch, J. R., Poirel, H. A., Carmeliet, P., & Vikkula, M. (2011). FAF1, a gene that is disrupted in cleft palate and has conserved function in zebrafish. *Am J Hum Genet*, *88*(2), 150-161. https://doi.org/10.1016/j.ajhg.2011.01.003

Gilbert, S. (2000). Osteogenesis: The development of bones. In *Developmental Biology* (6th ed.). Sinauer Associates. https://www.ncbi.nlm.nih.gov/books/NBK10056/

Giovannone, D., Paul, S., Schindler, S., Arata, C., Farmer, D. T., Patel, P., Smeeton, J., & Crump, J. G. (2019). Programmed conversion of hypertrophic chondrocytes into osteoblasts and marrow adipocytes within zebrafish bones. *Elife*, *8*. https://doi.org/10.7554/eLife.42736

Glasauer, S. M., & Neuhauss, S. C. (2014). Whole-genome duplication in teleost fishes and its evolutionary consequences. *Mol Genet Genomics*, *289*(6), 1045-1060. https://doi.org/10.1007/s00438-014-0889-2

Gomez-Picos, P. (2020). *COMPARATIVE TRANSCRIPTOMICS REVEALS THE GENE REGULATORY NETWORK UNDERLYING THE DIFFERENTIATION AND EVOLUTION OF SKELETAL CELLS* [University of Saskatchewan].

Gomez-Picos, P., & Eames, B. F. (2015). On the evolutionary relationship between chondrocytes and osteoblasts. *Frontiers in Genetics*, *6*. https://doi.org/ARTN29710.3389/fgene.2015.00297

Gomez-Picos, P., Ovens, K., & Eames, B. F. (2022). Limb Mesoderm and Head Ectomesenchyme Both Express a Core Transcriptional Program During Chondrocyte Differentiation. *Frontiers in Cell and Developmental Biology*, *10*, 876825. https://doi.org/10.3389/fcell.2022.876825

Gualeni, B., de Vernejoul, M. C., Marty-Morieux, C., De Leonardis, F., Franchi, M., Monti, L., Forlino, A., Houillier, P., Rossi, A., & Geoffroy, V. (2013). Alteration of proteoglycan sulfation affects bone growth and remodeling. *Bone*, *54*(1), 83-91. https://doi.org/10.1016/j.bone.2013.01.036

Habuchi, H., Habuchi, O., & Kimata, K. (2004). Sulfation pattern in glycosaminoglycan: does it have a code? *Glycoconj J*, *21*(1-2), 47-52. https://doi.org/10.1023/B:GLYC.0000043747.87325.5e

Hackett, M. J., George, G. N., Pickering, I. J., & Eames, B. F. (2016). Chemical Biology in the Embryo: In Situ Imaging of Sulfur Biochemistry in Normal and Proteoglycan-Deficient Cartilage Matrix. *Biochemistry*, *55*(17), 2441-2451. https://doi.org/10.1021/acs.biochem.5b01136

Haffter, P., Granato, M., Brand, M., Mullins, M. C., Hammerschmidt, M., Kane, D. A., Odenthal, J., van Eeden, F. J., Jiang, Y. J., Heisenberg, C. P., Kelsh, R. N., Furutani-Seiki, M., Vogelsang, E., Beuchle, D., Schach, U., Fabian, C., & Nusslein-Volhard, C. (1996). The identification of genes with unique and essential functions in the development of the zebrafish, Danio rerio. *Development*, *123*, 1-36. http://www.ncbi.nlm.nih.gov/pubmed/9007226

Hall, B. K. (1984). Developmental processes underlying heterochrony as an evolutionary mechanism. *Canadian Journal of Zoology*, *62*(1), 1-7. https://doi.org/10.1139/z84-001

Hall, B. K., & Miyake, T. (2000). All for one and one for all: condensations and the initiation of skeletal development. *Bioessays*, *22*(2), 138-147. https://doi.org/10.1002/(SICI)1521-1878(200002)22:2<138::AID-BIES5>3.0.CO;2-4

Hamburger, V., & Hamilton, H. L. (1951). A series of normal stages in the development of the chick embryo. *Developmental Dynamics*, *195*(4), 231-272. http://onlinelibrary.wiley.com/doi/10.1002/aja.1001950404/abstract;jsessionid=D0AA7DE88D13714C58BCA1E3371B448A.f04t04

Hammond, C. L., & Moro, E. (2012). Using transgenic reporters to visualize bone and cartilage signaling during development in vivo. *Front Endocrinol (Lausanne)*, *3*, 91. https://doi.org/10.3389/fendo.2012.00091

Hanson, S. R., Best, M. D., & Wong, C. H. (2004). Sulfatases: Structure, mechanism, biological activity, inhibition, and synthetic utility. *Angewandte Chemie-International Edition*, *43*(43), 5736-5763. https://doi.org/10.1002/anie.200300632

Hargest, T. E., Gay, C. V., Schraer, H., & Wasserman, A. J. (1985). Vertical distribution of elements in cells and matrix of epiphyseal growth plate cartilage determined by quantitative electron probe analysis. *J Histochem Cytochem*, *33*(4), 275-286. https://doi.org/10.1177/33.4.3980981

Harris, M. P., Rohner, N., Schwarz, H., Perathoner, S., Konstantinidis, P., & Nusslein-Volhard, C. (2008). Zebrafish eda and edar mutants reveal conserved and ancestral roles of ectodysplasin signaling in vertebrates. *PLoS Genet*, *4*(10), e1000206. https://doi.org/10.1371/journal.pgen.1000206

Heinegard, D., & Oldberg, A. (1989). Structure and Biology of Cartilage and Bone-Matrix Noncollagenous Macromolecules. *Faseb Journal*, *3*(9), 2042-2051. https://doi.org/DOI 10.1096/fasebj.3.9.2663581

Heubel, B. P., Bredesen, C. A., Schilling, T. F., & Le Pabic, P. (2021). Endochondral growth zone pattern and activity in the zebrafish pharyngeal skeleton. *Dev Dyn*, *250*(1), 74-87. https://doi.org/10.1002/dvdy.241

Holmes, R. S. (2016). Comparative and Evolutionary Studies of Vertebrate Arylsulfatase B, Arylsulfatase I and Arylsulfatase J Genes and Proteins: Evidence for an ARSB-like Sub-family. *Journal of Proteomics & Bioinformatics*, *9*(11), 1-8. https://doi.org/10.4172/jpb.1000418

Hopwood, J. J., & Ballabio, A. (2019). Multiple Sulfatase Deficiency and the Nature of the Sulfatase Family. In D. L. Valle, S. Antonarakis, A. Ballabio, A. L. Beaudet, & G. A. Mitchell (Eds.), *The Online Metabolic and Molecular Bases of Inherited Disease*. McGraw-Hill Education. ommbid.mhmedical.com/content.aspx?aid=1181465994

Howe, D. G., Bradford, Y. M., Conlin, T., Eagle, A. E., Fashena, D., Frazer, K., Knight, J., Mani, P., Martin, R., Moxon, S. A., Paddock, H., Pich, C., Ramachandran, S., Ruef, B. J., Ruzicka, L., Schaper, K., Shao, X., Singer, A., Sprunger, B., Van Slyke, C. E., & Westerfield, M. (2013). ZFIN, the Zebrafish Model Organism Database: increased support for mutants and transgenics. *Nucleic Acids Res*, *41*(Database issue), D854-860. https://doi.org/10.1093/nar/gks938

Howe, K., Clark, M. D., Torroja, C. F., Torrance, J., Berthelot, C., Muffato, M., Collins, J. E., Humphray, S., McLaren, K., Matthews, L., McLaren, S., Sealy, I., Caccamo, M., Churcher, C., Scott, C., Barrett, J. C., Koch, R., Rauch, G. J., White, S., Chow, W., Kilian, B., Quintais, L. T., Guerra-Assuncao, J. A., Zhou, Y., Gu, Y., Yen, J., Vogel, J. H., Eyre, T., Redmond, S., Banerjee, R., Chi, J., Fu, B., Langley, E., Maguire, S. F., Laird, G. K., Lloyd, D., Kenyon, E., Donaldson, S., Sehra, H., Almeida-King, J., Loveland, J., Trevanion, S., Jones, M., Quail, M., Willey, D., Hunt, A., Burton, J., Sims, S., McLay, K., Plumb, B., Davis, J., Clee, C., Oliver, K., Clark, R., Riddle, C., Elliot, D., Threadgold, G., Harden, G., Ware, D., Begum, S., Mortimore, B., Kerry, G., Heath, P., Phillimore, B., Tracey, A., Corby, N., Dunn, M., Johnson, C., Wood, J., Clark, S., Pelan, S., Griffiths, G., Smith, M., Glithero, R., Howden, P., Barker, N., Lloyd, C., Stevens, C., Harley, J., Holt, K., Panagiotidis, G., Lovell, J., Beasley, H., Henderson, C., Gordon, D., Auger, K., Wright, D., Collins, J., Raisen, C., Dyer, L., Leung, K., Robertson, L., Ambridge, K., Leongamornlert, D., McGuire, S., Gilderthorp, R., Griffiths, C., Manthravadi, D., Nichol, S., Barker, G., Whitehead, S., Kay, M., Brown, J., Murnane, C., Gray, E., Humphries, M., Sycamore, N., Barker, D., Saunders, D., Wallis, J., Babbage, A., Hammond, S., Mashreghi-Mohammadi, M., Barr, L., Martin, S., Wray, P., Ellington, A., Matthews, N., Ellwood, M., Woodmansey, R., Clark, G., Cooper, J., Tromans, A., Grafham, D., Skuce, C., Pandian, R., Andrews, R., Harrison, E., Kimberley, A., Garnett, J., Fosker, N., Hall, R., Garner, P., Kelly, D., Bird, C., Palmer, S., Gehring, I., Berger, A., Dooley, C. M., Ersan-Urun, Z., Eser, C., Geiger, H., Geisler, M., Karotki, L., Kirn, A., Konantz, J., Konantz, M., Oberlander, M., Rudolph-Geiger, S., Teucke, M., Lanz, C., Raddatz, G., Osoegawa, K., Zhu, B., Rapp, A., Widaa, S., Langford, C., Yang, F., Schuster, S. C., Carter, N. P., Harrow, J., Ning, Z., Herrero, J., Searle, S. M., Enright, A., Geisler, R., Plasterk, R. H., Lee, C., Westerfield, M., de Jong, P. J., Zon, L. I., Postlethwait, J. H., Nusslein-Volhard, C., Hubbard, T. J., Roest Crollius, H., Rogers, J., & Stemple, D. L. (2013). The zebrafish reference genome sequence and its relationship to the human genome. *Nature*, *496*(7446), 498-503. https://doi.org/10.1038/nature12111

Hruscha, A., Krawitz, P., Rechenberg, A., Heinrich, V., Hecht, J., Haass, C., & Schmid, B. (2013). Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish. *Development*, *140*(24), 4982-4987. https://doi.org/10.1242/dev.099085

Huysseune, A., Van der heyden, C., & Sire, J. Y. (1998). Early development of the zebrafish (Danio rerio) pharyngeal dentition (Teleostei, Cyprinidae). *Anat Embryol (Berl)*, *198*(4), 289-305. http://www.ncbi.nlm.nih.gov/pubmed/9764543

Hwang, W. Y., Fu, Y., Reyon, D., Maeder, M. L., Tsai, S. Q., Sander, J. D., Peterson, R. T., Yeh, J. R., & Joung, J. K. (2013). Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol*, *31*(3), 227-229. https://doi.org/10.1038/nbt.2501

Inada, M., Wang, Y., Byrne, M. H., Rahman, M. U., Miyaura, C., Lopez-Otin, C., & Krane, S. M. (2004). Critical roles for collagenase-3 (Mmp13) in development of growth plate cartilage and in

endochondral ossification. *Proc Natl Acad Sci U S A*, *101*(49), 17192-17197. https://doi.org/10.1073/pnas.0407788101

Ioannidis, J. P., Thomas, G., & Daly, M. J. (2009). Validating, augmenting and refining genome-wide association signals. *Nat Rev Genet*, *10*(5), 318-329. https://doi.org/10.1038/nrg2544

Izadifar, Z., Chang, T., Kulyk, W., Chen, X., & Eames, B. F. (2016). Analyzing Biological Performance of 3D-Printed, Cell-Impregnated Hybrid Constructs for Cartilage Tissue Engineering. *Tissue Eng Part C Methods*, *22*(3), 173-188. https://doi.org/10.1089/ten.TEC.2015.0307

Jiang, Z., Byers, S., Casal, M. L., & Smith, L. J. (2020). Failures of Endochondral Ossification in the Mucopolysaccharidoses. *Curr Osteoporos Rep*, *18*(6), 759-773. https://doi.org/10.1007/s11914-020-00626-y

Kabashi, E., Brustein, E., Champagne, N., & Drapeau, P. (2011). Zebrafish models for the functional genomics of neurogenetic disorders. *Biochim Biophys Acta*, *1812*(3), 335-345. https://doi.org/10.1016/j.bbadis.2010.09.011

Karsenty, G., & Wagner, E. F. (2002). Reaching a genetic and molecular understanding of skeletal development. *Dev Cell*, *2*(4), 389-406. https://doi.org/10.1016/s1534-5807(02)00157-0

Karve, T. M., & Cheema, A. K. (2011). Small changes huge impact: the role of protein posttranslational modifications in cellular homeostasis and disease. *J Amino Acids*, *2011*, 207691. https://doi.org/10.4061/2011/207691

Kawakami, K. (2007). Tol2: a versatile gene transfer vector in vertebrates. *Genome Biol*, *8 Suppl 1*, S7. https://doi.org/10.1186/gb-2007-8-s1-s7

Kearns, A. E., Vertel, B. M., & Schwartz, N. B. (1993). Topography of Glycosylation and Udp-Xylose Production. *Journal of Biological Chemistry*, *268*(15), 11097-11104. <Go to ISI>://WOS:A1993LD46600058

Kim, M., Bi, X., Horton, W. E., Spencer, R. G., & Camacho, N. P. (2005). Fourier transform infrared imaging spectroscopic analysis of tissue engineered cartilage: histologic and biochemical correlations. *J Biomed Opt*, *10*(3), 031105. https://doi.org/10.1117/1.1922329

Kim, Y. I., Lee, S., Jung, S. H., Kim, H. T., Choi, J. H., Lee, M. S., You, K. H., Yeo, S. Y., Yoo, K. W., Kwak, S., Lee, J. N., Park, R., Choe, S. K., & Kim, C. H. (2013). Establishment of a bone-specific col10a1:GFP transgenic zebrafish. *Mol Cells*, *36*(2), 145-150. https://doi.org/10.1007/s10059-013-0117-7

Kluppel, M., Wight, T. N., Chan, C., Hinek, A., & Wrana, J. L. (2005). Maintenance of chondroitin sulfation balance by chondroitin-4-sulfotransferase 1 is required for chondrocyte development and growth factor signaling during cartilage morphogenesis. *Development*, *132*(17), 3989-4003. https://doi.org/10.1242/dev.01948

Knopf, F., Hammond, C., Chekuru, A., Kurth, T., Hans, S., Weber, C. W., Mahatma, G., Fisher, S., Brand, M., Schulte-Merker, S., & Weidinger, G. (2011). Bone regenerates via dedifferentiation of osteoblasts in the zebrafish fin. *Dev Cell*, *20*(5), 713-724. https://doi.org/10.1016/j.devcel.2011.04.014

Knudson, C. B., & Knudson, W. (2001). Cartilage proteoglycans. *Semin Cell Dev Biol*, *12*(2), 69-78. https://doi.org/10.1006/scdb.2000.0243

Kok, F. O., Shin, M., Ni, C. W., Gupta, A., Grosse, A. S., van Impel, A., Kirchmaier, B. C., Peterson-Maduro, J., Kourkoulis, G., Male, I., DeSantis, D. F., Sheppard-Tindell, S., Ebarasi, L., Betsholtz, C., Schulte-Merker, S., Wolfe, S. A., & Lawson, N. D. (2015). Reverse genetic screening reveals poor correlation between morpholino-induced and mutant phenotypes in zebrafish. *Dev Cell*, *32*(1), 97-108. https://doi.org/10.1016/j.devcel.2014.11.018

Kong, R. Y., Kwan, K. M., Lau, E. T., Thomas, J. T., Boot-Handford, R. P., Grant, M. E., & Cheah, K. S. (1993). Intron-exon structure, alternative use of promoter and expression of the mouse collagen X gene, Col10a-1. *Eur J Biochem*, *213*(1), 99-111. http://www.ncbi.nlm.nih.gov/pubmed/8477738

Köster, R. W., & Sassen, W. A. (2015). A molecular toolbox for genetic manipulation of zebrafish. *Advances in Genomics and Genetics*, *5*, 151-163. https://doi.org/https://doi.org/10.2147/AGG.S57585

Kronenberg, H. M. (2003). Developmental regulation of the growth plate. *Nature*, *423*(6937), 332-336. https://doi.org/10.1038/nature01657

Kury, S., Besnard, T., Ebstein, F., Khan, T. N., Gambin, T., Douglas, J., Bacino, C. A., Sanders, S. J., Lehmann, A., Latypova, X., Khan, K., Pacault, M., Sacharow, S., Glaser, K., Bieth, E., Perrin-Sabourin, L., Jacquemont, M. L., Cho, M. T., Roeder, E., Denomme-Pichon, A. S., Monaghan, K. G., Yuan, B., Xia, F., Simon, S., Bonneau, D., Parent, P., Gilbert-Dussardier, B., Odent, S., Toutain, A., Pasquier, L., Barbouth, D., Shaw, C. A., Patel, A., Smith, J. L., Bi, W., Schmitt, S., Deb, W., Nizon, M., Mercier, S., Vincent, M., Rooryck, C., Malan, V., Briceno, I., Gomez, A., Nugent, K. M., Gibson, J. B., Cogne, B., Lupski, J. R., Stessman, H. A., Eichler, E. E., Retterer, K., Yang, Y., Redon, R., Katsanis, N., Rosenfeld, J. A., Kloetzel, P. M., Golzio, C., Bezieau, S., Stankiewicz, P., & Isidor, B. (2017). De Novo Disruption of the Proteasome Regulatory Subunit PSMD12 Causes a Syndromic Neurodevelopmental Disorder. *Am J Hum Genet*, *100*(2), 352-363. https://doi.org/10.1016/j.ajhg.2017.01.003

Kusche-Gullberg, M., & Kjellen, L. (2003). Sulfotransferases in glycosaminoglycan biosynthesis. *Curr Opin Struct Biol*, *13*(5), 605-611. http://www.ncbi.nlm.nih.gov/pubmed/14568616

Kwon, R. Y., Watson, C. J., & Karasik, D. (2019). Using zebrafish to study skeletal genomics. *Bone*, *126*, 37-50. https://doi.org/10.1016/j.bone.2019.02.009

Lai, J. P., Thompson, J. R., Sandhu, D. S., & Roberts, L. R. (2008). Heparin-degrading sulfatases in hepatocellular carcinoma: roles in pathogenesis and therapy targets. *Future Oncology*, *4*(6), 803-814. https://doi.org/10.2217/14796694.4.6.803

Law, S. H., & Sargent, T. D. (2014). The serine-threonine protein kinase PAK4 is dispensable in zebrafish: identification of a morpholino-generated pseudophenotype. *PLoS One*, *9*(6), e100268. https://doi.org/10.1371/journal.pone.0100268

Lawson, N. D. (2016). Reverse Genetics in Zebrafish: Mutants, Morphants, and Moving Forward. *Trends Cell Biol*, *26*(2), 77-79. https://doi.org/10.1016/j.tcb.2015.11.005

Lawson, N. D., & Weinstein, B. M. (2002). In Vivo Imaging of Embryonic Vascular Development Using Transgenic Zebrafish. *Dev Biol*, *248*(2), 307-318. https://doi.org/10.1006/dbio.2002.0711

Le Pabic, P., Dranow, D. B., Hoyle, D. J., & Schilling, T. F. (2022). Zebrafish endochondral growth zones as they relate to human bone size, shape and disease. *Front Endocrinol (Lausanne)*, *13*, 1060187. https://doi.org/10.3389/fendo.2022.1060187

Le Pabic, P., Ng, C., & Schilling, T. F. (2014). Fat-Dachsous signaling coordinates cartilage differentiation and polarity during craniofacial development. *PLoS Genet*, *10*(10), e1004726. https://doi.org/10.1371/journal.pgen.1004726

Leach, R. M., Jr., & Gay, C. V. (1987). Role of Epiphyseal Cartilage in Endochondral Bone Formation. *The Journal of Nutrition*, *117*(4), 784-790. https://doi.org/10.1093/jn/117.4.784

Leslie, E. J., Taub, M. A., Liu, H., Steinberg, K. M., Koboldt, D. C., Zhang, Q., Carlson, J. C., Hetmanski, J. B., Wang, H., Larson, D. E., Fulton, R. S., Kousa, Y. A., Fakhouri, W. D., Naji, A., Ruczinski, I., Begum, F., Parker, M. M., Busch, T., Standley, J., Rigdon, J., Hecht, J. T., Scott, A. F., Wehby, G. L., Christensen, K., Czeizel, A. E., Deleyiannis, F. W., Schutte, B. C., Wilson, R. K., Cornell, R. A., Lidral, A. C., Weinstock, G. M., Beaty, T. H., Marazita, M. L., & Murray, J. C. (2015). Identification of functional variants for cleft lip with or without cleft palate in or near PAX7, FGFR2, and NOG by targeted sequencing of GWAS loci. *Am J Hum Genet*, *96*(3), 397-411. https://doi.org/10.1016/j.ajhg.2015.01.004

Li, M., Li, C., & Guan, W. (2008). Evaluation of coverage variation of SNP chips for genome-wide association studies. *Eur J Hum Genet*, *16*(5), 635-643. https://doi.org/10.1038/sj.ejhg.5202007

Li, M., Zhao, L., Page-McCaw, P. S., & Chen, W. (2016). Zebrafish Genome Engineering Using the CRISPR-Cas9 System. *Trends Genet*, *32*(12), 815-827. https://doi.org/10.1016/j.tig.2016.10.005

Li, W., Li, F., Zhang, X., Lin, H. K., & Xu, C. (2021). Insights into the post-translational modification and its emerging role in shaping the tumor microenvironment. *Signal Transduct Target Ther*, *6*(1), 422. https://doi.org/10.1038/s41392-021-00825-8

Lieschke, G. J., & Currie, P. D. (2007). Animal models of human disease: zebrafish swim into view. *Nat Rev Genet*, *8*(5), 353-367. https://doi.org/10.1038/nrg2091

Lind, L. K., Stecksen-Blicks, C., Lejon, K., & Schmitt-Egenolf, M. (2006). EDAR mutation in autosomal dominant hypohidrotic ectodermal dysplasia in two Swedish families. *BMC Med Genet*, *7*, 80. https://doi.org/10.1186/1471-2350-7-80

Liu, L. Y., Fox, C. S., North, T. E., & Goessling, W. (2013). Functional validation of GWAS gene candidates for abnormal liver function during zebrafish liver development. *Dis Model Mech*, *6*(5), 1271-1278. https://doi.org/10.1242/dmm.011726

Liu, Y., Beyer, A., & Aebersold, R. (2016). On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell*, *165*(3), 535-550. https://doi.org/10.1016/j.cell.2016.03.014

Liu, Z., Chen, O., Wall, J. B. J., Zheng, M., Zhou, Y., Wang, L., Vaseghi, H. R., Qian, L., & Liu, J. (2017). Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. *Sci Rep*, *7*(1), 2193. https://doi.org/10.1038/s41598-017-02460-2

Long, F., Chung, U. I., Kronenberg, H., & McMahon, A. (2004). Ihh signaling is directly required for the osteoblast lineage in the endochondral skeleton. *Faseb Journal*, *18*(4), A398-A399. <Go to ISI>://WOS:000220470601920

Long, F., & Ornitz, D. M. (2013). Development of the endochondral skeleton. *Cold Spring Harb Perspect Biol*, *5*(1), a008334. https://doi.org/10.1101/cshperspect.a008334

Lorincz, P., & Juhasz, G. (2020). Autophagosome-Lysosome Fusion. *J Mol Biol*, *432*(8), 2462-2482. https://doi.org/10.1016/j.jmb.2019.10.028

Lu, H. H., & Thomopoulos, S. (2013). Functional attachment of soft tissues to bone: development, healing, and tissue engineering. *Annual Review of Biomedical Engineering, Vol 19*, *15*, 201-226. https://doi.org/10.1146/annurev-bioeng-071910-124656

Lu, P. F., Takai, K., Weaver, V. M., & Werb, Z. (2011). Extracellular Matrix Degradation and Remodeling in Development and Disease. *Cold Spring Harbor Perspectives in Biology*, *3*(12). https://doi.org/ARTNa00505810.1101/cshperspect.a005058

MacArthur, J., Bowler, E., Cerezo, M., Gil, L., Hall, P., Hastings, E., Junkins, H., McMahon, A., Milano, A., Morales, J., Pendlington, Z. M., Welter, D., Burdett, T., Hindorff, L., Flicek, P., Cunningham, F., & Parkinson, H. (2017). The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res*, *45*(D1), D896-D901. https://doi.org/10.1093/nar/gkw1133

Machado, R. G., & Eames, B. F. (2017). Using Zebrafish to Test the Genetic Basis of Human Craniofacial Diseases. *J Dent Res*, *96*(11), 1192-1199. https://doi.org/10.1177/0022034517722776

Mackay, E. W., Apschner, A., & Schulte-Merker, S. (2013). A bone to pick with zebrafish. *Bonekey Rep*, *2*, 445. https://doi.org/10.1038/bonekey.2013.179

Mackie, E. J., Ahmed, Y. A., Tatarczuch, L., Chen, K. S., & Mirams, M. (2008). Endochondral ossification: How cartilage is converted into bone in the developing skeleton. *International Journal of Biochemistry & Cell Biology*, *40*(1), 46-62. https://doi.org/10.1016/j.biocel.2007.06.009

Mackie, E. J., Tatarczuch, L., & Mirams, M. (2011). The skeleton: a multi-functional complex organ: the growth plate chondrocyte and endochondral ossification. *J Endocrinol*, *211*(2), 109-121. https://doi.org/10.1530/JOE-11-0048

Maddison, L. A., Li, M., & Chen, W. (2014). Conditional gene-trap mutagenesis in zebrafish. *Methods Mol Biol*, *1101*, 393-411. https://doi.org/10.1007/978-1-62703-721-1_19

Maier, T., Guell, M., & Serrano, L. (2009). Correlation of mRNA and protein in complex biological samples. *FEBS Lett*, *583*(24), 3966-3973. https://doi.org/10.1016/j.febslet.2009.10.036

Mangiavini, L., Merceron, C., & Schipani, E. (2016). Analysis of Mouse Growth Plate Development. *Curr Protoc Mouse Biol*, *6*(1), 67-130. https://doi.org/10.1002/9780470942390.mo150094

Marí-Beffa, M., Mesa-Román, A. B., & Duran, I. (2021). Zebrafish Models for Human Skeletal Disorders [Review]. *Frontiers in Genetics*, *12*. https://doi.org/10.3389/fgene.2021.675331

Mariotti, M., Carnovali, M., & Banfi, G. (2015). Danio rerio: the Janus of the bone from embryo to scale. *Clin Cases Miner Bone Metab*, *12*(2), 188-194. https://doi.org/10.11138/ccmbm/2015.12.2.188

Mathew, J., Jagadeesh, S. M., Bhat, M., Udhaya Kumar, S., Thiyagarajan, S., & Srinivasan, S. (2015). Mutations in ARSB in MPS VI patients in India. *Mol Genet Metab Rep*, *4*, 53-61. https://doi.org/10.1016/j.ymgmr.2015.06.002

McGurk, P. D., Lovely, C. B., & Eberhart, J. K. (2014). Analyzing craniofacial morphogenesis in zebrafish using 4D confocal microscopy. *J Vis Exp*(83), e51190. https://doi.org/10.3791/51190

Medeiros, D. M., & Crump, J. G. (2012). New perspectives on pharyngeal dorsoventral patterning in development and evolution of the vertebrate jaw. *Dev Biol*, *371*(2), 121-135. https://doi.org/10.1016/j.ydbio.2012.08.026

Meng, X., Noyes, M. B., Zhu, L. J., Lawson, N. D., & Wolfe, S. A. (2008). Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. *Nat Biotechnol*, *26*(6), 695-701. https://doi.org/10.1038/nbt1398

Mikami, T., & Kitagawa, H. (2013). Biosynthesis and function of chondroitin sulfate. *Biochim Biophys Acta*, *1830*(10), 4719-4733. https://doi.org/10.1016/j.bbagen.2013.06.006

Mis, E. K., Liem, K. E., Kong, Y., Schwartz, N. B., Domowicz, M., & Weatherbee, S. D. (2014). Forward genetics defines Xylt1 as a key, conserved regulator of early chondrocyte maturation and skeletal length. *Developmental Biology*, *385*(1), 67-82. https://doi.org/10.1016/j.ydbio.2013.10.014

Mitsunaga-Nakatsubo, K., Kusunoki, S., Kawakami, H., Akasaka, K., & Akimoto, Y. (2009). Cell-surface arylsulfatase A and B on sinusoidal endothelial cells, hepatocytes, and Kupffer cells in mammalian livers. *Med Mol Morphol*, *42*(2), 63-69. https://doi.org/10.1007/s00795-009-0447-x

Mizumoto, S., Yamada, S., & Sugahara, K. (2015). Molecular interactions between chondroitin-dermatan sulfate and growth factors/receptors/matrix proteins. *Current Opinion in Structural Biology*, *34*, 35-42. https://doi.org/10.1016/j.sbi.2015.06.004

Montano, A. M., Tomatsu, S., Gottesman, G. S., Smith, M., & Orii, T. (2007). International Morquio A Registry: clinical manifestation and natural course of Morquio A disease. *J Inherit Metab Dis*, *30*(2), 165-174. https://doi.org/10.1007/s10545-007-0529-7

Mork, L., & Crump, G. (2015). Zebrafish Craniofacial Development: A Window into Early Patterning. *Curr Top Dev Biol*, *115*, 235-269. https://doi.org/10.1016/bs.ctdb.2015.07.001

Moro, E., Vettori, A., Porazzi, P., Schiavone, M., Rampazzo, E., Casari, A., Ek, O., Facchinello, N., Astone, M., Zancan, I., Milanetto, M., Tiso, N., & Argenton, F. (2013). Generation and application of signaling pathway reporter lines in zebrafish. *Mol Genet Genomics*, *288*(5-6), 231-242. https://doi.org/10.1007/s00438-013-0750-z

Mougous, J. D., Green, R. E., Williams, S. J., Brenner, S. E., & Bertozzi, C. R. (2002). Sulfotransferases and sulfatases in mycobacteria. *Chem Biol*, *9*(7), 767-776. https://doi.org/10.1016/s1074-5521(02)00175-8

Muenzer, J. (2011). Overview of the mucopolysaccharidoses. *Rheumatology (Oxford)*, *50 Suppl 5*, v4-12. https://doi.org/10.1093/rheumatology/ker394

Nasevicius, A., & Ekker, S. C. (2000). Effective targeted gene 'knockdown' in zebrafish. *Nat Genet*, *26*(2), 216-220. https://doi.org/10.1038/79951

Neufeld, E. F., & Muenzer, J. (2019). The Mucopolysaccharidoses. In D. L. Valle, S. Antonarakis, A. Ballabio, A. L. Beaudet, & G. A. Mitchell (Eds.), *The Online Metabolic and Molecular Bases of Inherited Disease*. McGraw-Hill Education. ommbid.mhmedical.com/content.aspx?aid=1181463720

O'Brien, K. P., Westerlund, I., & Sonnhammer, E. L. (2004). OrthoDisease: a database of human disease orthologs. *Hum Mutat*, *24*(2), 112-119. https://doi.org/10.1002/humu.20068

Obaya, A. J. (2006). Molecular cloning and initial characterization of three novel human sulfatases. *Gene*, *372*, 110-117. https://doi.org/10.1016/j.gene.2005.12.023

Ohba, S. (2016). Hedgehog Signaling in Endochondral Ossification. *J Dev Biol*, *4*(2). https://doi.org/10.3390/jdb4020020

Orkin, R. W., Williams, B. R., Cranley, R. E., Poppke, D. C., & Brown, K. S. (1977). Defects in the cartilaginous growth plates of brachymorphic mice. *J Cell Biol*, *73*(2), 287-299. https://doi.org/10.1083/jcb.73.2.287

Ortega, N., Behonick, D. J., & Werb, Z. (2004). Matrix remodeling during endochondral ossification. *Trends Cell Biol*, *14*(2), 86-93. https://doi.org/10.1016/j.tcb.2003.12.003

Oshikawa, M., Usami, R., & Kato, S. (2009). Characterization of the arylsulfatase I (ARSI) gene preferentially expressed in the human retinal pigment epithelium cell line ARPE-19. *Mol Vis*, *15*, 482-494. https://www.ncbi.nlm.nih.gov/pubmed/19262745

Ota, S., Hisano, Y., Ikawa, Y., & Kawahara, A. (2014). Multiple genome modifications by the CRISPR/Cas9 system in zebrafish. *Genes Cells*, *19*(7), 555-564. https://doi.org/10.1111/gtc.12154

Otsuki, S., Hanson, S. R., Miyaki, S., Grogan, S. P., Kinoshita, M., Asahara, H., Wong, C. H., & Lotz, M. K. (2010). Extracellular sulfatases support cartilage homeostasis by regulating BMP and FGF signaling pathways. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(22), 10202-10207. https://doi.org/10.1073/pnas.0913897107 .

Otsuki, S., Taniguchi, N., Grogan, S. P., D'Lima, D., Kinoshita, M., & Lotz, M. (2008). Expression of novel extracellular sulfatases Sulf-1 and Sulf-2 in normal and osteoarthritic articular cartilage. *Arthritis Res Ther*, *10*(3), R61. https://doi.org/10.1186/ar2432

Payne, S. H. (2015). The utility of protein and mRNA correlation. *Trends Biochem Sci*, *40*(1), 1-3. https://doi.org/10.1016/j.tibs.2014.10.010

Pazzaglia, U. E., Reguzzoni, M., Casati, L., Minini, A., Salvi, A. G., & Sibilia, V. (2019). Long bone human anlage longitudinal and circumferential growth in the fetal period and comparison with the growth plate cartilage of the postnatal age. *Microsc Res Tech*, *82*(3), 190-198. https://doi.org/10.1002/jemt.23153

Pines, M., & Hurwitz, S. (1991). The role of the growth plate in longitudinal bone growth. *Poult Sci*, *70*(8), 1806-1814. https://doi.org/10.3382/ps.0701806

Ponten, F., Jirstrom, K., & Uhlen, M. (2008). The Human Protein Atlas - a tool for pathology. *Journal of Pathology*, *216*(4), 387-393. https://doi.org/10.1002/path.2440

Poole, A. R., 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*(2), 249-260. http://www.ncbi.nlm.nih.gov/pubmed/7037793

Postlethwait, J., Amores, A., Cresko, W., Singer, A., & Yan, Y. L. (2004). Subfunction partitioning, the teleost radiation and the annotation of the human genome. *Trends Genet*, *20*(10), 481-490. https://doi.org/10.1016/j.tig.2004.08.001

Postlethwait, J. H., Woods, I. G., Ngo-Hazelett, P., Yan, Y. L., Kelly, P. D., Chu, F., Huang, H., Hill-Force, A., & Talbot, W. S. (2000). Zebrafish comparative genomics and the origins of vertebrate chromosomes. *Genome Res*, *10*(12), 1890-1902. http://www.ncbi.nlm.nih.gov/pubmed/11116085

Power, R. C., Henry, A. G., Moosmann, J., Beckmann, F., Temming, H., Roberts, A., & Le Cabec, A. (2022). Synchrotron radiation-based phase-contrast microtomography of human dental calculus allows nondestructive analysis of inclusions: implications for archeological samples. *J Med Imaging (Bellingham)*, *9*(3), 031505. https://doi.org/10.1117/1.JMI.9.3.031505

Provot, S., & Schipani, E. (2005). Molecular mechanisms of endochondral bone development. *Biochem Biophys Res Commun*, *328*(3), 658-665. https://doi.org/10.1016/j.bbrc.2004.11.068

Prydz, K., & Dalen, K. T. (2000). Synthesis and sorting of proteoglycans. *J Cell Sci*, *113 Pt 2*, 193-205. https://doi.org/10.1242/jcs.113.2.193

Rabbani, B., Tekin, M., & Mahdieh, N. (2014). The promise of whole-exome sequencing in medical genetics. *J Hum Genet*, *59*(1), 5-15. https://doi.org/10.1038/jhg.2013.114

Rabenstein, D. L. (2002). Heparin and heparan sulfate: structure and function. *Natural Product Reports*, *19*(3), 312-331. https://doi.org/10.1039/b100916h

Rafferty, S. A., & Quinn, T. A. (2018). A beginner's guide to understanding and implementing the genetic modification of zebrafish. *Prog Biophys Mol Biol*, *138*, 3-19. https://doi.org/10.1016/j.pbiomolbio.2018.07.005

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nat Protoc*, *8*(11), 2281-2308. https://doi.org/10.1038/nprot.2013.143

Rattanasopha, S., Tongkobpetch, S., Srichomthong, C., Siriwan, P., Suphapeetiporn, K., & Shotelersuk, V. (2012). PDGFRa mutations in humans with isolated cleft palate. *Eur J Hum Genet*, *20*(10), 1058-1062. https://doi.org/10.1038/ejhg.2012.55

Ratzka, A., Mundlos, S., & Vortkamp, A. (2010). Expression patterns of sulfatase genes in the developing mouse embryo. *Dev Dyn*, *239*(6), 1779-1788. https://doi.org/10.1002/dvdy.22294

Reich, M., Shahbaz, M., Prajapati, D. H., Parmar, S., Hawkesford, M. J., & De Kok, L. J. (2016). Interactions of Sulfate with Other Nutrients As Revealed by H2S Fumigation of Chinese Cabbage. *Front Plant Sci*, *7*, 541. https://doi.org/10.3389/fpls.2016.00541

Richardson, L., Venkataraman, S., Stevenson, P., Yang, Y., Moss, J., Graham, L., Burton, N., Hill, B., Rao, J., Baldock, R. A., & Armit, C. (2014). EMAGE mouse embryo spatial gene expression database: 2014 update. *Nucleic Acids Res*, *42*(Database issue), D835-844. https://doi.org/10.1093/nar/gkt1155

Rieppo, L., Rieppo, J., Jurvelin, J. S., & Saarakkala, S. (2012). Fourier transform infrared spectroscopic imaging and multivariate regression for prediction of proteoglycan content of articular cartilage. *PLoS One*, *7*(2), e32344. https://doi.org/10.1371/journal.pone.0032344

Rivera-Colon, Y., Schutsky, E. K., Kita, A. Z., & Garman, S. C. (2012). The structure of human GALNS reveals the molecular basis for mucopolysaccharidosis IV A. *J Mol Biol*, *423*(5), 736-751. https://doi.org/10.1016/j.jmb.2012.08.020

Rizzo, R., Grandolfo, M., Godeas, C., Jones, K. W., & Vittur, F. (1995). Calcium, sulfur, and zinc distribution in normal and arthritic articular equine cartilage: a synchrotron radiation-induced X-ray emission (SRIXE) study. *J Exp Zool*, *273*(1), 82-86. https://doi.org/10.1002/jez.1402730111

Rosen, S. D., & Lemjabbar-Alaoui, H. (2010). Sulf-2: an extracellular modulator of cell signaling and a cancer target candidate. *Expert Opin Ther Targets*, *14*(9), 935-949. https://doi.org/10.1517/14728222.2010.504718

Rosenberg, M. I., Lynch, J. A., & Desplan, C. (2009). Heads and tails: evolution of antero-posterior patterning in insects. *Biochim Biophys Acta*, *1789*(4), 333-342. https://doi.org/10.1016/j.bbagrm.2008.09.007

Roughley, P. J., & Lee, E. R. (1994). Cartilage proteoglycans: structure and potential functions. *Microsc Res Tech*, *28*(5), 385-397. https://doi.org/10.1002/jemt.1070280505

Rubinsztein, D. C. (2006). The roles of intracellular protein-degradation pathways in neurodegeneration. *Nature*, *443*(7113), 780-786. https://doi.org/10.1038/nature05291

Saarakkala, S., & Julkunen, P. (2010). Specificity of Fourier Transform Infrared (FTIR) Microspectroscopy to Estimate Depth-Wise Proteoglycan Content in Normal and Osteoarthritic Human Articular Cartilage. *Cartilage*, *1*(4), 262-269. https://doi.org/10.1177/1947603510368689

Sakai, C., Ijaz, S., & Hoffman, E. J. (2018). Zebrafish Models of Neurodevelopmental Disorders: Past, Present, and Future. *Front Mol Neurosci*, *11*, 294. https://doi.org/10.3389/fnmol.2018.00294

Santagati, F., & Rijli, F. M. (2003). Cranial neural crest and the building of the vertebrate head. *Nat Rev Neurosci*, *4*(10), 806-818. https://doi.org/10.1038/nrn1221

Sardiello, M., Annunziata, I., Roma, G., & Ballabio, A. (2005). Sulfatases and sulfatase modifying factors: an exclusive and promiscuous relationship. *Hum Mol Genet*, *14*(21), 3203-3217. https://doi.org/10.1093/hmg/ddi351

Schlotawa, L., Adang, L. A., Radhakrishnan, K., & Ahrens-Nicklas, R. C. (2020). Multiple Sulfatase Deficiency: A Disease Comprising Mucopolysaccharidosis, Sphingolipidosis, and More Caused by a Defect in Posttranslational Modification. *Int J Mol Sci*, *21*(10). https://doi.org/10.3390/ijms21103448

Schwartz, N. B., & Domowicz, M. (2002). Chondrodysplasias due to proteoglycan defects. *Glycobiology*, *12*(4), 57r-68r. https://doi.org/DOI 10.1093/glycob/12.4.57R

Settembre, C., Annunziata, I., Spampanato, C., Zarcone, D., Cobellis, G., Nusco, E., Zito, E., Tacchetti, C., Cosma, M. P., & Ballabio, A. (2007). Systemic inflammation and neurodegeneration in a mouse model of multiple sulfatase deficiency. *Proc Natl Acad Sci U S A*, *104*(11), 4506-4511. https://doi.org/10.1073/pnas.0700382104

Settembre, C., Arteaga-Solis, E., McKee, M. D., de Pablo, R., Al Awqati, Q., Ballabio, A., & Karsenty, G. (2008). Proteoglycan desulfation determines the efficiency of chondrocyte autophagy and the extent of FGF signaling during endochondral ossification. *Genes Dev*, *22*(19), 2645-2650. https://doi.org/10.1101/gad.1711308

Shah, A. N., Davey, C. F., Whitebirch, A. C., Miller, A. C., & Moens, C. B. (2015). Rapid reverse genetic screening using CRISPR in zebrafish. *Nat Methods*, *12*(6), 535-540. https://doi.org/10.1038/nmeth.3360

Shaw, N. D., Brand, H., Kupchinsky, Z. A., Bengani, H., Plummer, L., Jones, T. I., Erdin, S., Williamson, K. A., Rainger, J., Stortchevoi, A., Samocha, K., Currall, B. B., Dunican, D. S., Collins, R. L., Willer, J. R., Lek, A., Lek, M., Nassan, M., Pereira, S., Kammin, T., Lucente, D., Silva, A., Seabra, C. M., Chiang, C., An, Y., Ansari, M., Rainger, J. K., Joss, S., Smith, J. C., Lippincott, M. F., Singh, S. S., Patel, N., Jing, J. W., Law, J. R., Ferraro, N., Verloes, A., Rauch, A., Steindl, K., Zweier, M., Scheer, I., Sato, D., Okamoto, N., Jacobsen, C., Tryggestad, J., Chernausek, S., Schimmenti, L. A., Brasseur, B., Cesaretti, C., Garcia-Ortiz, J. E., Buitrago, T. P., Silva, O. P., Hoffman, J. D., Muhlbauer, W., Ruprecht, K. W., Loeys, B. L., Shino, M., Kaindl, A. M., Cho, C. H., Morton, C. C., Meehan, R. R., van Heyningen, V., Liao, E. C., Balasubramanian, R., Hall, J. E., Seminara, S. B., Macarthur, D., Moore, S. A., Yoshiura, K. I., Gusella, J. F., Marsh, J. A., Graham, J. M., Jr., Lin, A. E., Katsanis, N., Jones, P. L., Crowley, W. F., Jr., Davis, E. E., FitzPatrick, D. R., & Talkowski, M. E. (2017). SMCHD1 mutations associated with a rare muscular dystrophy can also cause isolated arhinia and Bosma arhinia microphthalmia syndrome. *Nat Genet*, *49*(2), 238-248. https://doi.org/10.1038/ng.3743

Sheth, R., Marcon, L., Bastida, M. F., Junco, M., Quintana, L., Dahn, R., Kmita, M., Sharpe, J., & Ros, M. A. (2012). Hox Genes Regulate Digit Patterning by Controlling the Wavelength of a Turing-Type Mechanism. *Science*, *338*(6113), 1476-1480. https://doi.org/doi:10.1126/science.1226804

Simmons, D. (2008). *The Use of Animal Models in Studying Genetic Disease: Transgenesis and Induced Mutation*. Retrieved 22 June 2022 from https://www.nature.com/scitable/topicpage/the-use-of-animal-models-in-studying-855/

Simpson, M. A., Scheuerle, A., Hurst, J., Patton, M. A., Stewart, H., & Crosby, A. H. (2009). Mutations in FAM20C also identified in non-lethal osteosclerotic bone dysplasia. *Clin Genet*, *75*(3), 271-276. https://doi.org/10.1111/j.1399-0004.2008.01118.x

Singh, K., Menke, H., Andrew, M., Rau, C., Bijeljic, B., & Blunt, M. J. (2018). Time-resolved synchrotron X-ray micro-tomography datasets of drainage and imbibition in carbonate rocks. *Sci Data*, *5*, 180265. https://doi.org/10.1038/sdata.2018.265

St-Jacques, B., Hammerschmidt, M., & McMahon, A. P. (1999). Indian hedgehog signaling regulates proliferation and differentiation of chondrocytes and is essential for bone formation (vol 13, pg 2072, 1999). *Genes & Development*, *13*(19), 2617-2617. <Go to ISI>://WOS:000083381700015

Staines, K. A., Pollard, A. S., McGonnell, I. M., Farquharson, C., & Pitsillides, A. A. (2013). Cartilage to bone transitions in health and disease. *J Endocrinol*, *219*(1), R1-R12. https://doi.org/10.1530/JOE-13-0276

Stanier, P., & Pauws, E. (2012). Development of the lip and palate: FGF signalling. *Front Oral Biol*, *16*, 71-80. https://doi.org/10.1159/000337618

Staudt, N., Muller-Sienerth, N., Fane-Dremucheva, A., Yusaf, S. P., Millrine, D., & Wright, G. J. (2015). A panel of recombinant monoclonal antibodies against zebrafish neural receptors and secreted proteins suitable for wholemount immunostaining. *Biochem Biophys Res Commun*, *456*(1), 527-533. https://doi.org/10.1016/j.bbrc.2014.11.123

Stelzer, C., Brimmer, A., Hermanns, P., Zabel, B., & Dietz, U. H. (2007). Expression profile of Papss2 (3'-phosphoadenosine 5'-phosphosulfate synthase 2) during cartilage formation and skeletal development in the mouse embryo. *Dev Dyn*, *236*(5), 1313-1318. https://doi.org/10.1002/dvdy.21137

Stickens, D., Behonick, D. J., Ortega, N., Heyer, B., Hartenstein, B., Yu, Y., Fosang, A. J., Schorpp-Kistner, M., Angel, P., & Werb, Z. (2004). Altered endochondral bone development in matrix metalloproteinase 13-deficient mice. *Development*, *131*(23), 5883-5895. https://doi.org/10.1242/dev.01461

Streisinger, G., Walker, C., Dower, N., Knauber, D., & Singer, F. (1981). Production of clones of homozygous diploid zebra fish (Brachydanio rerio). *Nature*, *291*(5813), 293-296. http://www.ncbi.nlm.nih.gov/pubmed/7248006

Sugumaran, G., Katsman, M., & Silbert, J. E. (1992). Effects of Brefeldin-a on the Localization of Chondroitin Sulfate-Synthesizing Enzymes - Activities in Subfractions of the Golgi from Chick-Embryo Epiphyseal Cartilage. *Journal of Biological Chemistry*, *267*(13), 8802-8806. <Go to ISI>://WOS:A1992HR85400018

Swartz, M. E., Sheehan-Rooney, K., Dixon, M. J., & Eberhart, J. K. (2011). Examination of a palatogenic gene program in zebrafish. *Dev Dyn*, *240*(9), 2204-2220. https://doi.org/10.1002/dvdy.22713

T. Reinert, U. R., J. Vogt, T. Butz, A. Werner, W. Gründer. (2001). Spatially resolved elemental distributions in articular cartilage. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, *181*(1), 516 - 521. https://doi.org/https://doi.org/10.1016/S0168-583X(01)00481-5

Taher, L., McGaughey, D. M., Maragh, S., Aneas, I., Bessling, S. L., Miller, W., Nobrega, M. A., McCallion, A. S., & Ovcharenko, I. (2011). Genome-wide identification of conserved regulatory function in diverged sequences. *Genome Res*, *21*(7), 1139-1149. https://doi.org/10.1101/gr.119016.110

Tomatsu, S., Montano, A. M., Dung, V. C., Ohashi, A., Oikawa, H., Oguma, T., Orii, T., Barrera, L., & Sly, W. S. (2010). Enhancement of Drug Delivery: Enzyme-replacement Therapy for Murine Morquio A Syndrome. *Molecular Therapy*, *18*(6), 1094-1102. https://doi.org/10.1038/mt.2010.32

Tomatsu, S., Montano, A. M., Nishioka, T., Gutierrez, M. A., Pena, O. M., Tranda Firescu, G. G., Lopez, P., Yamaguchi, S., Noguchi, A., & Orii, T. (2005). Mutation and polymorphism spectrum of the

GALNS gene in mucopolysaccharidosis IVA (Morquio A). *Hum Mutat*, *26*(6), 500-512. https://doi.org/10.1002/humu.20257

Tonelli, F., Bek, J. W., Besio, R., De Clercq, A., Leoni, L., Salmon, P., Coucke, P. J., Willaert, A., & Forlino, A. (2020). Zebrafish: A Resourceful Vertebrate Model to Investigate Skeletal Disorders [Review]. *Frontiers in Endocrinology*, *11*. https://doi.org/10.3389/fendo.2020.00489

Uyama, T., Kitagawa, H., & Sugahara, K. (2007). Biosynthesis of glycosaminoglycans and proteoglycans. *J.P. Kamerling (Ed.), Comprehensive Glycoscience*, *3*, 79-104.

Valayannopoulos, V., Nicely, H., Harmatz, P., & Turbeville, S. (2010). Mucopolysaccharidosis VI. *Orphanet J Rare Dis*, *5*, 5. https://doi.org/10.1186/1750-1172-5-5

Van Assche, E., Van Puyvelde, S., Vanderleyden, J., & Steenackers, H. P. (2015). RNA-binding proteins involved in post-transcriptional regulation in bacteria. *Front Microbiol*, *6*, 141. https://doi.org/10.3389/fmicb.2015.00141

van der Kraan, P. M., Buma, P., van Kuppevelt, T., & van den Berg, W. B. (2002). Interaction of chondrocytes, extracellular matrix and growth factors: relevance for articular cartilage tissue engineering. *Osteoarthritis Cartilage*, *10*(8), 631-637. http://www.ncbi.nlm.nih.gov/pubmed/12479385

van Donkelaar, C. C., Janssen, X. J., & de Jong, A. M. (2007). Distinct developmental changes in the distribution of calcium, phosphorus and sulphur during fetal growth-plate development. *J Anat*, *210*(2), 186-194. https://doi.org/10.1111/j.1469-7580.2006.00680.x

Van Otterloo, E., Williams, T., & Artinger, K. B. (2016). The old and new face of craniofacial research: How animal models inform human craniofacial genetic and clinical data. *Dev Biol*, *415*(2), 171-187. https://doi.org/10.1016/j.ydbio.2016.01.017

Varshney, G. K., Pei, W., LaFave, M. C., Idol, J., Xu, L., Gallardo, V., Carrington, B., Bishop, K., Jones, M., Li, M., Harper, U., Huang, S. C., Prakash, A., Chen, W., Sood, R., Ledin, J., & Burgess, S. M. (2015). High-throughput gene targeting and phenotyping in zebrafish using CRISPR/Cas9. *Genome Res*, *25*(7), 1030-1042. https://doi.org/10.1101/gr.186379.114

Vauti, F., Stegemann, L. A., Vogele, V., & Koster, R. W. (2020). All-age whole mount in situ hybridization to reveal larval and juvenile expression patterns in zebrafish. *PLoS One*, *15*(8), e0237167. https://doi.org/10.1371/journal.pone.0237167

Vertel, B. M., Walters, L. M., Flay, N., Kearns, A. E., & Schwartz, N. B. (1993). Xylosylation Is an Endoplasmic-Reticulum to Golgi Event. *Journal of Biological Chemistry*, *268*(15), 11105-11112. <Go to ISI>://WOS:A1993LD46600059

Viviano, B. L., Paine-Saunders, S., Gasiunas, N., Gallagher, J., & Saunders, S. (2004). Domain-specific modification of heparan sulfate by Qsulf1 modulates the binding of the bone morphogenetic protein antagonist Noggin. *Journal of Biological Chemistry*, *279*(7), 5604-5611. https://doi.org/10.1074/jbc.M310691200

Wada, N., Javidan, Y., Nelson, S., Carney, T. J., Kelsh, R. N., & Schilling, T. F. (2005). Hedgehog signaling is required for cranial neural crest morphogenesis and chondrogenesis at the midline in the zebrafish skull. *Development*, *132*(17), 3977-3988. https://doi.org/10.1242/dev.01943

Wang, S., Ai, X., Freeman, S. D., Pownall, M. E., Lu, Q., Kessler, D. S., & Emerson, C. P., Jr. (2004). QSulf1, a heparan sulfate 6-O-endosulfatase, inhibits fibroblast growth factor signaling in mesoderm induction and angiogenesis. *Proc Natl Acad Sci U S A*, *101*(14), 4833-4838. https://doi.org/10.1073/pnas.0401028101

Wang, S. M., Sugahara, K., & Li, F. C. (2016). Chondroitin sulfate/dermatan sulfate sulfatases from mammals and bacteria. *Glycoconjugate Journal*, *33*(6), 841-851. https://doi.org/10.1007/s10719-016-9720-0

Wani, W. Y., Boyer-Guittaut, M., Dodson, M., Chatham, J., Darley-Usmar, V., & Zhang, J. (2015). Regulation of autophagy by protein post-translational modification. *Lab Invest*, *95*(1), 14-25. https://doi.org/10.1038/labinvest.2014.131

Watanabe, H., Kimata, K., Line, S., Strong, D., Gao, L. Y., Kozak, C. A., & Yamada, Y. (1994). Mouse Cartilage Matrix Deficiency (Cmd) Caused by a 7 Bp Deletion in the Aggrecan Gene. *Nature Genetics*, *7*(2), 154-157. https://doi.org/DOI 10.1038/ng0694-154

Waylen, L. N., Nim, H. T., Martelotto, L. G., & Ramialison, M. (2020). From whole-mount to single-cell spatial assessment of gene expression in 3D. *Communications Biology*, *3*(1), 602. https://doi.org/10.1038/s42003-020-01341-1

Weigele, J., & Franz-Odendaal, T. A. (2016). Functional bone histology of zebrafish reveals two types of endochondral ossification, different types of osteoblast clusters and a new bone type. *J Anat*, *229*(1), 92-103. https://doi.org/10.1111/joa.12480

White, A., & Wallis, G. (2001). Endochondral ossification: A delicate balance between growth and mineralisation. *Current Biology*, *11*(15), R589-R591. https://doi.org/Doi 10.1016/S0960-9822(01)00359-1

Woods, I. G., Wilson, C., Friedlander, B., Chang, P., Reyes, D. K., Nix, R., Kelly, P. D., Chu, F., Postlethwait, J. H., & Talbot, W. S. (2005). The zebrafish gene map defines ancestral vertebrate chromosomes. *Genome Res*, *15*(9), 1307-1314. https://doi.org/10.1101/gr.4134305

Woolfe, A., Goodson, M., Goode, D. K., Snell, P., McEwen, G. K., Vavouri, T., Smith, S. F., North, P., Callaway, H., Kelly, K., Walter, K., Abnizova, I., Gilks, W., Edwards, Y. J., Cooke, J. E., & Elgar, G. (2005). Highly conserved non-coding sequences are associated with vertebrate development. *PLoS Biol*, *3*(1), e7. https://doi.org/10.1371/journal.pbio.0030007

Wright, J. B., & Sanjana, N. E. (2016). CRISPR Screens to Discover Functional Noncoding Elements. *Trends Genet*, *32*(9), 526-529. https://doi.org/10.1016/j.tig.2016.06.004

Wuthier, R. E. (1969). A zonal analysis of inorganic and organic constituents of the epiphysis during endochondral calcification. *Calcified Tissue Research*, *4*(1), 20 - 38. https://doi.org/https://doi.org/10.1007/BF02279103

Yamamoto, A., Tagawa, Y., Yoshimori, T., Moriyama, Y., Masaki, R., & Tashiro, Y. (1998). Bafilomycin A1 prevents maturation of autophagic vacuoles by inhibiting fusion between autophagosomes and lysosomes in rat hepatoma cell line, H-4-II-E cells. *Cell Struct Funct*, *23*(1), 33-42. https://doi.org/10.1247/csf.23.33

Yang, J., Andre, P., Ye, L., & Yang, Y. Z. (2015). The Hedgehog signalling pathway in bone formation. *Int J Oral Sci*, *7*(2), 73-79. https://doi.org/10.1038/ijos.2015.14

Yao, Y., & Wang, Y. (2013). ATDC5: an excellent in vitro model cell line for skeletal development. *J Cell Biochem*, *114*(6), 1223-1229. https://doi.org/10.1002/jcb.24467

Yelick, P. C., & Schilling, T. F. (2002). Molecular Dissection of Craniofacial Development Using Zebrafish. *Critical Reviews in Oral Biology & Medicine*, *13*(4), 308-322. https://doi.org/10.1177/154411130201300402

Yin, L., Maddison, L. A., Li, M., Kara, N., LaFave, M. C., Varshney, G. K., Burgess, S. M., Patton, J. G., & Chen, W. (2015). Multiplex Conditional Mutagenesis Using Transgenic Expression of Cas9 and sgRNAs. *Genetics*, *200*(2), 431-441. https://doi.org/10.1534/genetics.115.176917

Yoshimori, T., Yamamoto, A., Moriyama, Y., Futai, M., & Tashiro, Y. (1991). Bafilomycin A1, a specific inhibitor of vacuolar-type H(+)-ATPase, inhibits acidification and protein degradation in lysosomes of cultured cells. *Journal of Biological Chemistry*, *266*(26), 17707-17712. https://www.ncbi.nlm.nih.gov/pubmed/1832676

Zare Mirakabad, H., Farsi, M., Malekzadeh Shafaroudi, S., Bagheri, A., Iranshahi, M., & Moshtaghi, N. (2019). Comparison the Effect of Ferutinin and 17beta-Estradiol on Bone Mineralization of

Developing Zebrafish (Danio rerio) Larvae. *Int J Mol Sci*, *20*(6). https://doi.org/10.3390/ijms20061507

Zare Mirakabad, H., & Khorramizadeh, M. R. (2022). Introduction to the potential of Ferula ovina in dental implant research due to estrogenic bioactive compounds and adhesive properties. *PLoS One*, *17*(1), e0262045. https://doi.org/10.1371/journal.pone.0262045

Zhu, S. M., Zhu, E. D., Provot, S., & Gori, F. (2010). Wdr5 Is Required for Chick Skeletal Development. *Journal of Bone and Mineral Research*, *25*(11), 2504-2514. https://doi.org/10.1002/jbmr.144

APPENDIX A

Methodology of preliminary results presented on Chapter 5

ATDC5 transfection

The constructs transfected into ATDC5 cells were previously generated in Eames lab by Tuanjie Cheng. The hARSI cDNA was obtained from human placenta (as described in item 3.3.13), and was subcloned into the pP2A-mCherry-N1 vector, which was obtained from Dorus Gadella (Addgene plasmid #84329). This vector enables bicistronic expression of the protein of interest and mCherry. GFP subcloned into the pP2A-mCherry-N1 vector was used as a control in transfection. hARSI was expressed under the control of a CMV enhancer (Figure 5.1 A). Transfection was performed using TransIT-LT1 Transfection Reagent (Mirus Bio, Madison, WI). After 48 hours, the 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. Micromass cultures were imaged at days 2, 4, 6, 8, and 10 by fluorescent microscopy, which showed that both GFP (used as a control) and mCherry remain expressed at these time points (Fig. 5.1 B). Cells were then harvested at day 10 and used for WB, as described in section 3.3.12. At this same time point, cells were examined for secretion of sulfated proteoglycans using Alcian blue staining as previously described (Izadifar et al. 2016).

Generating zebrafish mutants

<u>Guide RNAs (gRNAs)</u>

Based on the *arsia* and *arsib* sequences for *Danio Rerio* present in the NCBI database, gRNAs were designed to target 2 different regions of each of those genes. To

identify target sequences, the CHOPCHOP website (http://chopchop.cbu.uib.no/) was used. Selected gRNA sequences started with G or GG and had the lower off-target effects possible. Chosen sequences (20bp) were aligned to the zebrafish genome to ensure it matched only once. Instead of cloning the gRNA sequence into a vector before the *in vitro* transcription, we ordered two oligos that partially annealed to each other in a PCR reaction. The polymerase filed in the rest (free ends) during PCR, yielding a short double-stranded template with a T7 recognition site. One oligo was generic in that it is used in every template production reaction. The other was a sequence specific to the target sequence. The ordered oligos were as follows:

gRNA scaffold oligo (PAGE purified) from IDT:

5`-gatccgcaccgactcggtgccacttttcaagttgataacggactagccttattttaacttgctatttctagctctaaaac-3`

gene-specific oligos from IDT:

arsia1ex1:

5`- aattaatacgactcactata-GGATAAGCTGGCCTCGGAGG-gttttagagctagaaatagc-3`

arsia2ex1:

5`- aattaatacgactcactata- GAAGCTGCAGCGTTCAAGCC -gttttagagctagaaatagc-3`

arsib1ex1:

5`- aattaatacgactcactata-GGATCTCACCCAGGACTGGA -gttttagagctagaaatagc-3`

arsib2ex2:

5`- aattaatacgactcactata- GGATACGCCACGCACATGGT-gttttagagctagaaatagc-3`

Create *gRNA* DNA template

PCRs were performed with the scaffold oligo, plus one gene specific oligo at a time with Phusion polymerase, using the following conditions: 4 µl 5x Buffer, 4 µl dNTPs, 1 µl of gene specific oligo (10 µM), 1 µl of gRNA scaffold oligo (10 µM), 0.2 µl Phusion, 13.4 µl Water. Reactions proceeded for 98°C for 30 seconds, followed by 40 cycles of 98°C 10 seconds, 60°C for 10 seconds, 72°C 15 seconds, and one last cycle of 72ºC for 10 mins. PCR reactions were cleaned with column-based method, and product was analyzed on a 1% agarose gel for a single band of 125 bps. 4 µl of purified PCR products were used in a half-reaction of the MEGAscript T7 Kit for *in vitro* transcription overnight. DNase treatment was performed in all reactions with a column-based method.

Cas9

Competent *E. coli* were transformed with the *T3TS-nCas9n* plasmid (Addgene: plasmid 46757) and plated on LB agar plates with ampicillin (100 µg/ml). Colonies were collected and placed in tubes containing LB liquid media with ampicillin (100 µg/ml). The tubes were placed in a shaker at 37°C overnight. The next day, bacteria containing the plasmid with the DNA insert were purified using the QIAprep® Spin Miniprep. Plasmids were linearized with XbaI and purified using QiaPrep column (Qiagen). Purified samples were sent for DNA sequencing. Plasmids with the correct DNA sequence were used for *in vitro* transcription with the T3 mMESSAGE kit (Invitrogen) for 1-2hr, and the RNA was purified in a column (RNeasy Mini Kit from Qiagen).

Injection of *Cas9* mRNA and *gRNA* into zebrafish embryos

WT eggs were injected at a one-cell stage with an injection mixture containing 100 ng/µl *Cas9* mRNA, 50 ng/µl of gRNA, 0.0025% phenol red, and nuclease free water. The *gRNA/Cas9* mixture was injected into the yolk below the cell.

DNA extraction/genotyping

Injected fish were raised into adults, anesthetized in 4% Tricaine, and a small portion of the caudal fin was clipped. Fin clips were placed into individual PCR tubes together with 50 µl of alkaline lysis buffer (25 mM NaOH, 0.2 mM EDTA, pH 12). Tubes were placed into a thermocycler at 95℃ for 25 minutes and cooled down on ice for 15 minutes. Following the cooling step, 50µl of neutralizing buffer (40mM Tris-HCl pH 5) was added to each tube.

PCRs were performed with genomic DNA to check for a mutation created during CRISPR injections. PCR reactions in each well consisted of 2.5 µl 10x standard Taq buffer, 0.5 µl 10 mM dNTPs, 1 µl 10mM primers, 12.75 µl ddH20, 0.25 µl Thermofisher Taq polymerase, and 3 µl extracted DNA. The reaction was run for 40 cycles: 30 seconds at 95℃, 30 seconds at 56℃-60℃, and 40 seconds at 72℃. The PCR product of interest was run on a 1% agarose gel in 1xTAE buffer.

The band size of interest was extracted from the gel and purified using the QIAquick® Gel Extraction kit. The purified product was cloned into the pCR™4-TOPO™ vector and transformed into One Shot™ TOP10 Chemically Competent E. coli cells. The transformed product was grown overnight in 37°C on LB agar plates containing 100 µg/ml of Ampicillin. The next day, colonies were selected and transferred to plastic tubes containing 5 mL of LB media containing 100 µg/ml of ampicillin. The tubes were placed in a shaker at 37°C overnight. The next day, bacteria containing the plasmid with the DNA insert was purified using the QIAprep® Spin Miniprep. To ensure minipreps contained the DNA insert of interest, EcoR1 miniprep digestions were performed. 1 µl of the miniprep was placed in separate wells of a PCR plate, and then in each well the following reaction was made: 7.1 µl water, 0.1 µl 10x CutSmart® Buffer, 0.1 µl 100x BSA, and 0.8 µl HF EcoR1 enzyme. The reaction was run in an incubator at 37℃ overnight. The miniprep digestion product was visualized on a 1% agarose gel run with 1x TAE buffer. The minipreps containing the correct-sized insert were sent for sequencing with Eurofins. Sequencing results were analyzed by aligning the sequencing product with WT zebrafish DNA sequence from NCBI Genbank using Serial Cloner 2.6.

Zebrafish skeletal staining

Two potential mutants were found after cloning and sequence of the PCR product. Mutants were outcrossed with WT fish, and F1 fish were stained for Alcian blue and Alizarin red as described on item 4.3.5, and scored for the amounts of perichondral bone.

Zebrafish In situ hybridization (ISH)

ISH was performed on ZEBRAFISH CERATOHYALS (7UM frozen sections), to check for *arsia* and *arsib* endogenous expression. For that, embryos were fixed in 4% PFA at 4°C overnight, washed in 1xPBS, embedded in optimal cutting temperature (OCT) compound (Tissue Tek, Torrance, CA, United States), and immediately flash-frozen in liquid nitrogen-cooled isopentane. ISH proceeded as described in item 3.3.8 (*In situ* hybridization on frozen sections). The *col2a1a* zebrafish probe was previously produced in the lab, while probes for *arsia* and *arsib* were designed by me towards the 3'CDS region of the gene sequence. A T3 promoter (AATTAACCCTCACTAAAG) was inserted in the beginning of each sequence, and a T7 promoter reverse comp sequence (CTATAGTGAGTCGTATTA) was added by the end of each sequence. The oligos, which were ordered through Bio basic (https://www.biobasic.com), follow:

Zebrafish arsia 3'UTR probe

AATTAACCCTCACTAAAGCATTCCCATACACACACTCAGTGTTTTTCTTAAATGTAT
GATCATATTCTATTTGAAATAAACACGTTTGTTTTTGTGAAGGCTGATTAAAATGAAA
GAAAGCTGTCTTTATTGAGACTAGTGTAAATGTTGACTGCCCTCTAGCGGACACCT
GAGCTTTATTTGGATTTTGATTGGAGGACTGTACACGTGTAAATGTGTGTGTGTGTA
TGTGGGCTGTGTATGTGAAAGAGAGAGAGAGAGAGTGTGTGTGTGTGTGTGCATA
TGAGTGTTCTTGAATTGTATATTTTGTGTGTGTATATAATGTAAATGTTGACTGCCCT

CTAGTGGACACCTGAAGTTTATTTGTTATTCTGATTGGAGGACTGTACATGTGTAGG
TTCACGGTTCTCTTGTAAACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
GTGTGTGTGTTTAAAAAGACCGAAATATGTGTATGTATTTTAGATGTATTTTGTGCAT
GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTTAAATGTGTGTGTGTG
GTCTGTTTGCCTGCAGATACACTGGTCAGAGACCCTCTGGGCTGTGCTTTACAGTA
GCTGCTCCTCCAGTACACTATAGTGAGTCGTATTA

Zebrafish arsib CDS+3'UTR

AATTAACCCTCACTAAAGTACAGGCGGCCATTAGAGTTGGTGACTGGAAGCTCCTG
ACGGGTGATCCTGGGAATGGGGATTGGGTGCCGCCACATGTCCTTACACACTTCC
CCAGCAGCTGGTGGCATCTAGAGCGGGACATTGGAGAAAAGAGGAAATCCGTATG
GCTTTTCAATGTAACAGGTGACCCTTGTGAACGGCATGACCTTGCGGTGCACAGG
CCTGATGTTGTTAGGGAGCTGCTAGCACGGCTAGCCTTTCACAATCGCACCGCAAT
ACCTGTCCGCTACCCGCCTGATGATACCCGTGCCAACCCTAGTGCGAATGGAGGG
GCATGGAGACCTTGGGTTGGTGAAGATGACGAGGAGGAAAACTGGGATGGAGTTT
ATTATAAGAGAGGAAAGAATCGCAAAAGGAAGAAATGTAGGCTCTGCAAGCTGCAA
TCTTTTTTTAAGAAGTTTAACTTAAAGATCATGTCAAAACAAATATAGGGATCTTCTG
CTATTAGATATCACTGATAAGACTAAACCTAAGTGTTGACCATGTACAGCGGCTATA
GTGAGTCGTATTA

The cloned DNA fragments received from Bio basic were transformed into E.coli cells (OneShot, Invitrogen), selected for with ampicillin and cultured overnight in LB liquid medium. The plasmid DNA was extracted and purified to produce minipreps. Plasmids were then sequenced and compared to RNA sequences using the Basic Local Alignment Search Tool (BLAST) on NCBI. This comparison was conducted to verify that the cloned DNA sequences were accurate. Plasmids were linearized accordingly to produce the proper sense and anti-sense probes, transcribed using T7 or T3 RNA polymerases, and labelled with digoxigenin.

Histology of the developing chick humerus (HH32).

Safranin-O and Trichrome staining were performed on zebrafish cartilages as described on item 3.3.2

ISH of the developing chick humerus (HH32

ISH was performed on chick humeri at the HH32 stage (7 μm, frozen sections), in the same way described in item 3.3.8.

Immunohistochemistry for ARSI and COL10 in mouse and chick humerus sections

ICH for ARSI and COL10 in mouse and chick humerus were performed as described in item 3.3.3.