MULTIPLE CONSERVED ENHANCERS OF THE OSTEOBLAST MASTER TRANSCRIPTION FACTOR, *RUNX2,* INTEGRATE DIVERSE SIGNALING PATHWAYS TO DIRECT EXPRESSION TO DEVELOPING BONE

Christopher William Weber

A DISSERTATION

in

Cell and Molecular Biology

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2013

Supervisor of Dissertation

_____

Shannon Fisher, M.D, Ph.D.
Assistant Professor of Cell and Developmental Biology
Graduate Group Chairperson

_____

Daniel Kessler, Ph.D.
Associate Professor of Cell and Developmental Biology

Dissertation Committee

Douglas Epstein, Ph.D., Associate Professor of Genetics
Mary Mullins, Ph.D., Professor of Cell and Developmental Biology
Kurt Hankenson, M.S., D.V.M., Ph.D., Associate Professor of Equine Musculoskeletal Research
Eileen Shore, Ph.D., Cali and Weldon Research Professor in FOP

ACKNOWLEDGMENTS

Despite my peripatetic and inefficient path to actually sitting down and writing this thing, I've been surrounded constantly by an embarrassment of diversely sourced support and expertise that's kept my big dumb head directed towards this goal. I'm going to dispense with the prose in favor of bullet points now; please keep in mind that the choice is motivated by sloth and not style.

To my mother: Susan: You believe in a lot of weird shit, but you also believe in me. That means a lot, even if you still think crystals were sent from outer space to positively regulate your mood and that it's a good and useful exercise to pay a woman to mindmeld with your horses over the phone.

To my advisor, Shannon Fisher: Your support has been intellectual, material, patient, culinary and indispensible. Thank you. I don't know what I did to deserve it.

To my brothers, Brian and Andy: You are the funniest people I know, and you've never made enough fun of me to satisfy my need for self-flagellation. I know there's time.

To my thesis committee, Doug Epstein, Mary Mullins, Kurt Hankenson and Eileen Shore: Thank you for letting me fail in the privacy of your company, so that I could moderately succeed outside of it.

To other members of the Fisher lab, past and present, Thanks for showing me where we keep that. Also, I'm the one who left the scope on overnight and didn't sign in. **Gui**, I've been mispronouncing your name for four years. Many apologies.

To my graduate school colleagues and contemporaries, in particular Jessica Taaffe, Austin Thiel, and Dustin Hancks, and: Let's never do this again.

To the taxpayers of the Commonwealth of Pennsylvania and the United States of America: Thanks for subsidizing the last twelve years. I really wish I could promise you that it was money better spent than on a crumbling infrastructure, massive foreign debt, or the unfunded liabilities of Social Security and Medicare as the country undergoes a massive demographic shift. We'll see. I gotta finish school first.

"Personally, I liked the university. They gave us money and facilities; we didn't have to produce anything! You've never been out of college! You don't know what it's like out there! I've worked in the private sector. They expect results."

-Raymond Stantz, Ph.D

"Deformed, unfinish'd, sent before my time
Into this breathing world, scarce half made up"

---Richard III, Act 1, scene i

ABSTRACT

MULTIPLE CONSERVED ENHANCERS OF THE OSTEOBLAST MASTER TRANSCRIPTION FACTOR, *RUNX2,* INTEGRATE DIVERSE SIGNALING PATHWAYS TO DIRECT EXPRESSION TO DEVELOPING BONE

Christopher William Weber

Shannon Fisher

The vertebrate skeleton forms via two distinct modes of ossification, membranous and endochondral. Osteoblasts are also heterogeneous in embryonic origin; bone formed by either mode can be derived from neural crest cells or mesoderm. In contrast, all bone develops via a common genetic pathway regulated by the transcription factor Runx2. *Runx2* is required for bone formation, and haploinsufficiency in humans causes the skeletal syndrome cleidocranial dysplasia, demonstrating the importance of gene dosage. Despite the central role of *Runx2* in directing bone formation, little is understood about how its expression is regulated in development. We took an unbiased approach to identify direct regulatory inputs into *Runx2* transcription by identifying cis–regulatory elements associated with the human gene. We assayed conserved non-coding elements in a 1 Mb interval surrounding the gene for their ability to direct osteoblast expression in transgenic zebrafish. We identified three enhancers spaced out across the interval. Within each we identified conserved transcription factor binding sites required for their activity, and further showed

distinct and specific regulation of each. The enhancer in the last intron of *RUNX2* itself is positively regulated by the FGF signaling pathway, an enhancer in the last intron of the adjacent gene, *SUPT3H*, is regulated by canonical Wnt signaling, and a distant downstream enhancer requires a conserved *Dlx* binding site for its activity. While all of these pathways and factors have been previously implicated in bone formation, our results provide the first direct links to the common genetic pathway regulating osteogenesis, transcription of *Runx2*. These findings further illustrate the integration of multiple regulatory inputs at the level of transcription of a key developmental gene, and highlight the role of *Runx2* as the gatekeeper for changes in skeletal morphology achieved through alterations in gene expression.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

CHAPTER 1

Characteristics of the vertebrate skeleton

The presence of a mineralized endoskeleton is one of the common features of the vertebrate lineage[1]. In addition to its historically understood roles in support and as the sites of muscle attachment, the skeleton has more recently been understood to be the site of hematopoiesis[2], endocrine regulation of glucose metabolism[3], a reservoir for inorganic minerals[4] and critical in male reproductive function[5].

The vertebrate skeleton is chiefly composed of two tissue types: bone and cartilage.[a] Cartilage is the more evolutionarily primary of the two[6]. While not possessing a 'true' skeleton, the chordate amphioxus expresses orthologs of cartilage marker genes in the nascent notochord[7]. Cartilage is composed of chondrocytes suspended in a rigid matrix rich in collagen fibrillar proteins and acidic polysaccharides[8]. The most abundant of these proteins are type II collagen and aggrecan, whose negative charge accounts for the osmotic swelling of the tissue[9], resulting in the familiar rigid plasticity of the material. This property confers a biomechanical role in the fully realized skeleton, allowing articular surfaces of joints to tolerate compressive forces.

Conversely, bone is vascularized, has a higher metabolic activity and differs in its extracellular matrix (ECM) composition both in the content of secreted proteins, but also in the presence of inorganic calcium[10]. Unlike the

[a] Two other tissue types found exclusively in teeth are dentin and enamel, though they will not be discussed further in this document.

cartilage ECM, 90% of the total dry protein weight is type I collagen. Collagen I forms an extensively crosslinked fiber, around which calcium crystals in the form of spindles of hydroxyapatite are deposited, resulting in the characteristic rigidity of bone tissue[11]. Other well-characterized components include alkaline phosphatase, osteopontin, and osteocalcin[12].

Osteoblasts are the cells responsible for the deposition of this defining matrix. Correspondingly, they have a highly basophilic cytoplasm and extensive endoplasmic reticulum and Golgi apparatus[13] to produce substantial amounts of secreted protein. Following matrix deposition, osteoblasts either become lining cells or remain embedded in bone, the latter defined as osteocytes[14]. These cells account for 95% of mature bone tissue. Another bone cell type, osteoclasts, arises from the monocytic/macrophage lineage postnatally[15]. These cells have a resorbative role in bone homeostasis and therefore regulate bone mass density.

Embryonic origins of the vertebrate skeleton

Skeletogenesis describes the process by which mesenchymal stem cells (MSCs) differentiate into osteoblasts and chondrocytes in a defined program. MSCs are loose, multipotent cells with the capacity to differentiate into non-skeletal cell types such as adipocytes or myocytes[16]. Whether commitment to the skeletal lineage involves the existence of a bipotential skeletal precursor cell type, capable of adopting a bone or cartilage fate, is at issue in the literature[17]. Skeletal elements in the embryo forms via two distinct processes. Intramembrous ossification describes direct condensation of migrated MSCs and

subsequent transformation to bone[18]. This process is employed in the creation of the flat bones of the skull as well as fracture repair. Elsewhere, particularly in the long bones, endochondral ossification results in the calcification and invasion of a cartilaginous scaffold by osteoprogenitor cells[19]. A complementary heterogeneity is observed in the embroynic origin of MSCs, where neural crest, lateral plate, and somitic mesoderm all contribute to the developing skeleton[16].

Genetic origins of the vertebrate skeleton – *Runx2*

However, this diversity contrasts with the uniform genetic origin of skeletal tissues. Commitment to the osteoblast lineage requires the expression of the early marker gene and runt domain containing transcription factor *Runx2*[20]. The runt domain is a site of protein-protein interaction, as well as binding to the core sequence 5'- PyGPyGGTPy-3'[21]. *Runx2* $^{-/-}$ mice fail to generate any osteoblasts[22], and chondrocyte maturation and terminal differentiation are disturbed[23]. Additionally, haploinsufficency at the locus causes the skeletal disorder cleidocranial dysplasia, marked by delayed closure of the fontanelles of the skull, hypoplasticity of the clavical, and other features[24] (OMIM# 119600). *RUNX2* binds to and upregulates other osteoblast marker genes[25], which are also upregulated following forced expression of *Runx2* in non-skeletal tissues, including fibroblasts, C3H10T1/2 cells, primary myoblasts, and marrow stromal cells[25-27]. For these reasons, *Runx2* has been recognized as occupying an indispensable bottleneck position in the osteoblast fate switch and is often referred to the master regulator of osteoblast development[28].

The dosage of *Runx2* must be finely tuned in order to properly execute differentiation. Overexpression of *Runx2* in osteoblasts arrests bone development in a mouse model, resulting in an osteopenic phenotype[29]. Forced expression in chondrocytes produces precociously mature cells that produce osteoid tissue and bone marrow not present in orthologous structures in wild type animals[30]. Despite *Runx2*'s unquestioned indispensability early in fate commitment, the notion that *Runx2* might not have a role in mature osteoblasts has been proposed, as expression of a dominant negative form of the protein exclusively in mature osteoblasts does not affect transcription of the osteoblast marker osteocalcin, a gene that can be activated by forced expression of *Runx2* in non osteoblastic cells[25, 31].

Genetic origins of the vertebrate skeleton – *sp7/osx*

An answer to potential regulators of later osteoblast differentiation came with the identification of *Sp7/Osx* as a cDNA species specifically expressed in C2C12 cells undergoing osteoblastogenesis. *Sp7* codes for a zinc finger-containing transcription factor from the Kruppel-like factor family[32]. As with *Runx2*, inactivation of in mouse models yielded a skeleton devoid of osteoblasts; however mineralization did occur in bones formed by endochondral ossification, though the features of those tissues were more akin to a mineralized form of cartilage. Interestingly, *Runx2* expression levels were unaffected, indicating that *Sp7* is not upstream of *Runx2*. Further work located *Sp7* as a direct target of

Runx2 [b]. Sp7 is recruited to its own promoter in murine UMR106-01 osteosarcoma cells to the exclusion of other members of the sp transcription factor family in a manner that correlates with the expression of *sp7* [33] .

Sp7 is thought to function exclusively in later osteoblast differentiation and distinct from the activities of *Runx2* in cartilage. Among *Sp7*'s target genes [34] are the bone marker genes *Col1a1* [35] , *Bsp* [36] and *Ocn*. The regulation of *Col1a1* by *Sp7* is corroborated clinically by a report of a proband presenting with osteogenesis imperfecta and a frameshift mutation within the *SP7* coding region [37] . Finally, in contrast to *Runx2*, *Sp7* function appears to be critical for postnatal growth and maintenance of bone [38] .

Molecular signaling and the vertebrate skeleton – The BMP pathway

Bone morphogenetic proteins (BMPs) were initially identified on the basis of their ability to induce *de novo* bone and cartilage formation *in vivo* [39, 40] . Most BMPs are members of the transforming growth factor-β superfamily of proteins with important roles in both proper patterning and differentiation of the skeleton[c]. Canonically, signaling starts upon BMP ligand binding to heteromeric cell surface receptors composed of BMPR-I and BMPR-II receptors. This activated complex phosphorylates cytoplasmic SMAD proteins via a serine-threonine kinase domain. SMADs 1,5 and 8 bind to a co-SMAD upon phosphorylation and enter the nucleus to directly affect gene transcription via

[b] Allen, unpublished observation
[c] Notably, BMP-1 is a metalloproteinase.

chromatin binding. As will be discussed in future sections, the Smad proteins offer a context for crosstalk with other signaling pathways.

Understanding the role of BMP signaling in skeletogenesis is complicated by the presence of multiple components with distinct, yet overlapping, activities, as well the necessity of BMP signaling in early embryo patterning formation. BMP-2, --4,-6,and -7 are ligands with demonstrable osteogenic potential in vitro, yet genetic studies using conditional knockout alleles reveal more subtle and complementary roles. Both *Bmp2* and *Bmp4* activities are dispensable for the formation of the long bones, of the limbs, though deletion of the former results in an increase of fractures postnatally [41, 42] . Similarly, loss of *Bmpr2* [43] and *Bmp7* have no demonstrable effect on bone formation or fracture repair in the limbs [44] . However, a double knockout of *Bmp2* and *Bmp4* results in a severe impairment of osteogenesis, indicating a redundancy in these roles [45] .

Runx2 upregulation has been observed in in vitro systems following BMP stimulation [46, 47] , and consequently, *Runx2* is thought to be the principle mediator of downstream BMP actions [48] . However, there are also thought to be BMP signals capable of driving osteoblastogenesis independently of *Runx2*. Although BMP-2 administration is not capable of driving full differentiation of osteoblasts and chondroblasts in *Runx2*-deficient mouse calvarial cell lines, upregulation of alkaine phosphatase, osteocalcin and *sp7* is detectable [22, 49, 50] . BMP-2 treatment upregulates *sp7* expression in C2C12 cells independently of *Runx2* [51] . Also, preosteoblastic cell lines require autocrine BMP signaling for proper

differentiation, although they already express *Runx2*[52, 53] . Finally, activated SMAD proteins interact physically and functionally with RUNX2, suggesting a synergistic relationship to complement the *Runx2* dependent and independent BMP signaliing axes. SMAD1 and RUNX2 transcription factors complex to drive gene expression on target gene promoters[54] . An osteoblast specific deletion of *Smad1* causes an osteopenic phenotype[55] , and combined deletion of *Smad1/5/8* results in severe chondrodysplasia[56] .

Pretreatment with the ribosome inhibitor cyclohexamide prior to BMP-2 treatment blocks the induction of *Runx2*[57] and *sp7*[58] , indicating the need for the synthesis of an intermediate protein to complete the signaling axis. Among the direct targets of BMP signaling with known roles in skeletogenesis are homeodomain proteins. In particular, microarray experiments examining the transcriptional response to BMP-2 treatment in cultured C2C12 osteoprogenitor cells have identified members of the meshless(*Msx*), distalless(*Dlx*), and aristaless(*Alx*) transcription factor families as being immediately and transiently induced, prior to the commitment to osteogenesis evidenced by expression of *Runx2*[59-61] . Mutations associated with the *Msx1* and *Msx2* loci demonstrate consequences in skeletal patterning and differentiation. *Msx1*$^{-/-}$ mice exhibit craniofacial and tooth development abnormalities including a cleft palate phenotype[62] , while *Msx2*$^{-/-}$ mice possess delayed calvarial bone growth, defects in endochondrial ossification and chondrogenesis, as well as reduced expression of osteocalcin and *Runx2* [63] . Simultaneous deletion of both *Msx2* and *Msx1*

results in the complete absence of craniofacial bone[63, 64] A reversal of this dosage effect is evidenced in a human *MSX2* gain of function mutant with enhanced DNA binding, eliciting a premature fusion of the calvarial sutures and craniosynostosis[65] . Significantly, microduplications upstream of *MSX2* containing many conserved non-coding elements phenocopy CCD, suggesting a potentially rigid regulatory apparatus between BMP signaling and *Runx2* in vivo [66] .

In tetrapods, members of the *Dlx* gene family are grouped in binary clusters, facing each other via their 3' ends as a result of presumptive gene duplication events[67] . *Dlx1* and *Dlx4* have important roles in tooth development[68] and hematopoiesis[69] , respectively, but they have not been identified as expressed in osteoblasts. Although *Dlx3* inactivation results in embryonic lethality, it is expressed in osteoblastic lineage cells during endochondral ossification, and at its highest level in mature osteocalcin and *Runx2* expressing osteoblasts[70, 71] . A 4bp frameshift deletion in the human *DLX3* gene causes an autosomal dominant disease, tricho-dento-osseous syndrome (OMIM#600525), which is characterized by altered dermal bone formation in the skull as well as increased bone density[72] . Consistent with this observation, interaction between DLX3 and RUNX2 reduces the capacity of RUNX2 to direct transcription at the osteocalcin promoter in a cell culture context[73] .

Current opinion in the literature designates *Dlx5* as a critical regulator of BMP mediated osteogenesis[74] . Simultaneous knockout of the *Dlx5*/*6* cluster

results in gross skeletal abnormalities, including absence of the calvaria, maxilliary and mandibular bones, as well as a generalized ossification delay in the axial skeleton[75] . These anomalies are also seen in *Dlx5* $^{-/-}$ mice; curiously, no data on a *Dlx6* $^{-/-}$ phenotype has been published. *Dlx5* induction by BMPs has been observed in both the contexts of cell culture(MC3T3-E1 cells)[76] and in vivo development(chick skull development)[77] , where *Dlx5* expression is visible in proliferating suture mesenchyme not yet committed to an osteoblastic fate, suggesting a role in fate designation prior to *Runx2* induction. *Dlx5* induces *Runx2* in immature calvaria mesenchyme culture[78] , and *Dlx5* and *Runx2* have been shown to be recruited together at the stimulated promoters of induced osteoblast marker genes *Alp*[79] and *Ocn* [73] , so *Dlx5* appears to possess a duality of roles during osteoblast differentiation, both as a direct regulator of *Runx2* transcription, as well as a cooperative transcription factor at *Runx2* target genes.

Molecular signaling and the vertebrate skeleton – The Wnt pathway

Wnts are secreted, lipid-modified glycoproteins that activate cell surface receptor-mediated signal transduction pathways to regulate a variety of cellular activities, including cell fate determination, proliferation, migration, polarity, and gene expression[80] . In canonical, β-catenin-dependent WNT signaling, a Wnt ligand binds to binds to a Frizzled receptor and their co-receptors low-density lipoprotein receptor-related protein 5 (LRP5) or LRP6 to stabilize cytosolic β-catenin via inhibition of a ubiquitinating complex. β-catenin then enters the nucleus and stimulates the transcription of WNT target genes by interacting

with lymphoid enhancer-binding factor 1 (LEF1), T cell factor 1 (TCF1), TCF3 or TCF4. Non-canonical Wnt pathways not utilizing β-catenin include the noncanonical planar cell polarity, which also does not employ LRP5 or LRP6, and noncanonical Wnt/calcium pathway, which requires modulation of intracellular calcium ion levels. These pathways are further separated by their choice of ligand; canonical Wnt signaling uses WNT1, WNT3a, WNT8 or WNT10b, while noncanonical signaling relies on WNT4, WNT5a or WNT11.

Recognition of the Wnt pathway's involvement in bone biology began with a punctuation of discovery: in a single year, mutations causing severe alterations in bone density were identified in four groups of patients with bone mass disorders, pointing to the canonical branch of WNT signaling. Two of these mutations were detected in the LRP5 coreceptor necessary for Wnt signal propagation. Loss of function mutations in *LRP5* cause the autosomal recessive disorder osteoporosis-pseudoglioma syndrome (OMIM# 259770) [81] . Affected individuals have very low bone mass and are prone to developing fractures and deformation, though they lack any identifiable defects in collagen synthesis, anabolic and catabolic hormones, calcium homeostasis, endochondral growth, or bone turnover. A knockout mouse model confirmed the genotype-phenotype relationship, and provided insight into the bone mass deficit. *Lrp5*$^{-/-}$ mice have low bone mass compared to their wild type littermates, though this feature was only detectable postnatally [82] . Intriguingly, no aberrations in *Runx2* expression were detected in these mice.

The importance of the Wnt pathway in mediation of postnatal bone mass was further highlighted by the identification of gain of function mutations in LRP5 (G171V), causing an autosomal dominant high bone mass phenotype[83, 84]. Molecular investigations recognized the mutation as detrimental for the affinity of the protein for the extracellular Wnt signaling antagonists *DKK1*[85, 86] and *SOST*[86, 87]. *Sost* itself is a locus for mutations affecting bone mass. Premature termination mutations in *Sost*[88] cause sclerosteosis (OMIM #605740) whereas a 52 kb homozygous deletion downstream of the *SOST* gene is associated with van Buchem disease[89] both of whom are characterized by bone overgrowth.

Genetic analysis in the decade following these initial discoveries detailed the importance of many additional canonical Wnt components (β-catenin, Gsk-3ß, Axin2, and Dkk1; reviewed in[90]) in both osteoblast differentiation and postnatal bone mass density maintenance. Conditional deletion of β-catenin forces a chondrocytic fate on skeletal precursor cells[91, 92], a fate suppressed in these progenitors in response to ectopic activation of Wnt signaling[91, 93, 94]. However, finer dissection of this process reveals the stage of differentiation as a strong determinant of the response of a differentiating osteoblast to Wnt signaling. β-catenin stabilization in MSCs promotes proliferation at the expense of osteoblastic differentiation, while committed osteoblasts respond to the same stimulus by accelerating both growth and differentiation, at the expense of failure of terminal differentiation into mature osteoblasts. One possibility responsible for this state dependent response is the complex relationship

between Wnt signaling and the master regulatory transcription factors *Runx2* and s*p7*. While the Runx2 P1 promoter possesses a Wnt responsive element that recruits β-catenin and TCF/LEF transcription factors[95], cells lacking β-catenin, are, like *Lrp5*$^{-/-}$ mice, still capable of expressing *Runx2* in cells surrounding developing bone tissue. This suggests that other inputs in the *Runx2* regulatory apparatus are sufficient to induce the primary osteoblast differentiation genetic program in the absence of Wnt.

Consistent with their effects on mature bone, the regulatory relationship between *sp7* and the Wnt pathway appears to be reciprocal. While canonical Wnt signaling promotes both osteoblast differentiation and proliferation, *sp7* promotes differentiation of maturing osteoblasts, while inhibiting their proliferative potential. It appears that this is accomplished at least partially by an *sp7*-mediated inhibition of Wnt activity. *Sp7* appears to control the expression of the extracellular Wnt antagonist *Dkk1* by direct binding to its promoter, and its expression is indeed abolished in *sp7*-null embryonic calvarial cells. *Sp7* also inhibits β-catenin mediated transcription by direct interaction with the transcription factor *TCF1*. Therefore, it has been speculated that repeated downregulation of Wnt signaling is essential for balancing proliferative and cell fate priorities during osteoblastogenesis[96].

Molecular signaling and the vertebrate skeleton – The FGF pathway

Fibroblast growth factors (FGFs) are a family (22 members in both mouse and human)[97] of secreted growth factors with roles in diverse biological

processes that exert signaling activity by binding to tyrosine kinase fibroblast growth factor receptors (FGFRs), inducing intracellular pathways such as p38 MAPK, PLCγ, ERK1/2 or PI3K/AKT. Current thinking places the FGF signaling axis as a positive regulator of proliferation of progenitor cell populations and growth plate maturation during bone development [98] . In cell culture, FGF signaling increases proliferation of immature osteoblasts while simultaneously blocking differentiation [99, 100] .

The identities of and roles of specific Fgf ligands at discrete stages of skeletal development are poorly understood. *Fgf9* is expressed in early mesenchyme condensations prior to ossification, while *Fgf2*, *Fgf5*, *Fgf6*, and *Fgf7* are expressed in the mesenchyme surrounding the ossification. All Fgfs have been identified in the coronal suture of E17.5 embryos, save for *Fgf3* and *Fgf4*. While in vitro evidence has shown the capacity for Fgf ligands to stimulate osteoblast differentiation and or marker genes [101] , animal models have failed to provide striking evidence of the necessity of a given ligand for a skeletal process, though it is clear that excessive ligand disrupts proper development. The construction of an *Fgf2* knockout mouse provided an early illustration of this concept. *Fgf2*$^{-/-}$ mice are normal in apperance, but have lower bone mass density, concomitant with decrease thymidine incorporation in calvarial osteoblasts, suggesting an early proliferation defect behind the adult phenotype [102] . Consistent with these observations, overexpression of *Fgf2* in mice results in premature mineralization, achondroplasisa and shortening of the long bone [103]

Similar observations were made in experments using *Fgf9* knockout and transgenic mice [104, 105] .

As is not the case for the FGF ligands, there exists much human genetic evidence regarding the necessity of the FGFR genes in skeletal development. Of the four FGFRs, FGFRs1-3 are expressed in calvaria mesenchyme. Gain of function missense mutations in *FGFR1*, *FGFR2*, and *FGFR3* cause a spectrum of fourteen disorders, most of whom share a craniosynotosis or chondrodysplasia feature [106] . Despite the diversity in phenotypes resulting from those mutations, then, it makes sense to try to understand the common biology in these in these conditions. Craniosynotosis and chondrodysplasia differ fundamentally in the physiological process disrupted in their pathology. Craniosynostosis is a failure of the flat bones of the craniofactial skeleton to delay differentiation in progenitor cell populations, resulting in premature fusion of the sutures. However, chondrodysplasia is a defect in endochrondral ossification, often resulting in shortening of the long bones of the limbs and the axial skeleton in general. So despite affecting two distinct pathways to mature bone, upregulation of FGF signaling in developing skeletal tissue results in a common cell biology defect: premature differentiation of progenitor cells.

Genetic origins of the vertebrate skeleton – Signaling crosstalk

While understanding the functions of individual signaling pathways in bone development is a necessary effort towards a complete theory of skeletogenesis, these deconstructions, in isolation, lead to an impoverished view

of a highly integrated process. Therefore, it is necessary to consider how these pathways interact, either to enhance or police each other's activities. Therefore, I will discuss known connections of each of the systems discussed above.

Signaling crosstalk in the vertebrate skeleton - BMP and Wnt pathways

Studies investigating interactions between Wnt and BMP in osteoblast differentiation have identified both synergistic and epistatic relationships between components of these pathways. At a fundamental level, BMP-2 driven osteogenesis is dependent on the presence of β-catenin [107] . Many examples exist of synergistic activation of osteoblast marker genes by costimulation with BMP and Wnt ligands at early stages of osteoblastogenesis [107-109] . Several possible explanations involving intracellular mediators of these signals have been proposed. In *Xenopus* embryos and cos-7 cells, Wnt signaling extends the duration of a 'pulse' of BMP signaling by regulating SMAD1 activity via GSK-3β dependent phosphorylation [110] . Other researchers have described a mechanism involving the physical interaction of SMAD4 with TCF4 and the general co-activator protein p300 [111] .

However, at later stages in bone biology, they may have distinctly antagonistic roles. Where continued Wnt signaling is crucial for maintaining sufficent levels of bone mass density, BMP signaling at this stage actually acts in an catabolic manner. Deletion of *Noggin*, which codes for an extracellular inhibitor of BMP ligands, led to decreased bone mineral density (BMD) and bone formation in mice [112] . The extracellular Wnt inhibitors *Dkk1* and *Sost* are

downstream of BMP signaling [113] , and through their upregulation, Wnt signaling is attenuated. Genetic evidence for this interaction was observed in an osteoblast-specific *Bmpr1a* knockout mouse, which had a high bone mass phenotype concomitant with upregulated Wnt signaling [113] . BMP inhibition of Wnt occurs in uncommitted bone marrow cells via sequestration of the GSK-3β inhibitor/Wnt activator DSH by SMAD1 [114] .

Another level of BMP and Wnt integration to discuss is the combined cis regulation of both component and target genes. Chip-seq data from erythoid cells indicate that many active enhancers in these cells recruit both SMAD1 and TCF7L2 [115] . The promoters of *Dlx5* and *Msx2*, which are routinely and essentially upregulated in response to BMP signaling, respond synergistically to BMP and Wnt activation. Unsurprisingly, SMAD1, TCF4 and β-catenin are recruited to these promoters following dual stimulation of these pathways [108] .

Signaling crosstalk in the vertebrate skeleton - BMP and FGF pathways

Unlike the complicated relationship between BMP and Wnt signaling, the association between BMP and FGF signaling has been described as largely cooperative. Similarly to its relationship with Wnt, many examples exist where BMP signaling is in part dependent on the presence of active FGF signaling to achieve full osteogenic effect. Mice null for *Fgf2* have decreased *Bmp2* expression [116] , while FGF-2 and FGF-9 increase expression of *Bmp2* in calvarial osteoblasts. Additionally, these ligands inhibit the expression of *noggin*, an extracellular BMP inhibitor normally upregulated in response to BMP signaling [117] . FGF

mediated suppression of *noggin* is also observed in vivo in the coronal dura mater during suture development. *Noggin* maintains the patency of flat bone sutures in the skull, so it is possible that some of the craniosynostosis phenotype arising from gain of function FGFR mutations is due, in part, to coordinately misregulated BMP signaling [116, 118] . FGF also upregulates BMP signaling beyond the context of increasing ligand-receptor association; *Fgf2*$^{-/-}$ osteoblasts have impaired colocalization of phosphorylated SMADs and RUNX2 in response to BMP-2 signaling, though the reason for this deficit is unclear [116, 119] . Finally, FGF-2 and BMP-2 have a synergistic effect on fracture healing: FGF-2 has a critical function at early stage while BMP-2 promotes mineralization at later stage [120] .

Signaling crosstalk in the vertebrate skeleton – Wnt and FGF pathways

Wnt and FGF signaling have opposing effects during osteoblast differentiation [121] . The convergence of Wnt and FGF signaling in skeletogenesis occurs primarily by the suppression of Wnt signaling by FGF signaling. Multiple mechanisms have been described underlying this process. At a fundamental level, the expression of components of the canonical Wnt pathway requires FGF signaling. mRNA expression of *Wnt10b*, *Lrp6*, and β-catenin are significantly downregulated in bone marrow stromal cells from *Fgf2*$^{-/-}$ mice [122] . Exogenous application of Fgf2 ligand to these cells rescues both the osteogenesis defects while increasing β-catenin stabilization and nuclear localization. Comparative microarray analysis of osteoblasts derived from patients with gain of function

FGFR2 mutations identified the transcription factor SOX2 as dramatically (15 to 121 fold) upregulated compared to wildtype cells[99] . Coimmunoprecipitation experiments demonstrated that SOX2 associates with β-catenin in osteoblasts and can repress activity of a reporter plasmid drive by TCF/LEF binding sites.

Wnt and FGF signaling interactions have also been studied genetically in the context of skull suture formation. Tellingly, deletion of the gene encoding the Wnt negative regulator, *Axin2*, resulted in a phenotype similar to that observed in craniosynostosis in FGFR gain of function mutations[123] . Upregulation of Wnt signaling in *Axin2*-deficient mice was confirmed by increased nuclear accumulation of β-catenin. In concert, the proportion of FGFR positive cells at the suture was significantly reduced[124] . Further altering the FGF/WNT balance by generating $Axin2^{-/-}$, $Fgfr1^{+/-}$ mice produce sutures with ectopic cartilage formation[125] . Cells at the front of these sutures had upregulated BMP signalling, as evidenced by increased SMAD phosphorylation. A complex mechanism in suture mesenchyme has been proposed, where Wnt signaling expands the population of skeletal precusors, while stimulating BMP signaling to counteract FGF signaling. In the presence of relatively high levels of FGF signaling, BMP signaling promoters osteoblastogenesis in the microenvironment, while reduced FGF signaling results in the effect of BMP to signaling to promote a chondrocytic fate.

Zebrafish as a model to study vertebrate skeletogenesis

Zebrafish enjoy a burgeoning status as a more tractable alternative to the standard skeletal biology models of mouse and chick[126]. Development of both the craniofacial[127] and axial[128] skeletons has been well characterized. Zebrafish produce the same skeletal cell types as higher vertebrates, albeit in simpler patterns[129]. Additionally, gene expression events in skeletal elements are orthologous to those observed in higher vertebrates[130]. Specifically for the purposes of this work, both zebrafish runx2 orthologs, *runx2a* and *runx2b* are expressed in nascent skeletal elements[131]. Moreover, the appearance of the zebrafish skeleton is rapid; the first bony structure, the cleithrum[d], is visible within 72 HPF[132], though expression of bone marker genes in the anlagen begins at approximately 36 HPF[133]. Potential bone specific deficiencies of the system, such as the lack of osteocytes or hematopoietic activity in the bone marrow, are not hindrances for exploring early development[134].

Study aims

Despite the identification of skeletogenesis specific roles of the signaling pathways discussed above and others[e], a coherent narrative of the genetics and cell biology underlying this process still eludes the field. Because of its singular

[d] Most of the imaging in this document will focus on two bones as proxies for the expression in the rest of the skeleton. The cleithrum is a bone of mesodermal origin, and is roughly analogous to the shoulder girdle in mammals. Conversely, the opercle is derived from neural crest, and adopts a fan shaped morphology to lend structural support to the gill flap.

[e] These include Notch, Indian Hedgehog, calcineurin, retinoic acid, p38 MAPK, among others.

role in both marking and initiating early stages of osteoblastogenesis, understanding the regulation of *Runx2* itself is a potentially fruitful approach to understanding the biology of this process. Therefore, the rest of this document will be committed to describing the cis regulatory architecture responsible for regulating both the presence of the RUNX2 transcription factor itself in putative skeletal cells, but also the modulation of gene dosage that is so critical for proper execution of this process. Chapter 2 describes the results of a conservation based screen for conserved non-coding elements associated with the human *RUNX2* locus capable of directing expression to bone. Chapter 3 relates a series of functional studies on individual elements, identifying upstream regulators with previously confirmed roles in skeletogenesis. Finally, Chapter 4 integrates the results from Chapters 2 & 3 for a summary, discussion of the implications of the work, and suggestions for future avenues of experimental inquiry informed by this effort.

CHAPTER 2

A SCREEN FOR RUNX2 ASSOCIATED ENHANCERS IDENTIFIES THREE CONSERVED NON-CODING ELEMENTS CAPABLE OF DIRECTING EXPRESSION TO OSTEOBLASTS IN VIVO.

CHAPTER 2

Introduction

The gene encoding the transcription factor RUNX2 was identified as the underlying cause of the human skeletal syndrome cleidocranial dysplasia (CCD) [24, 126] , resulting from haploinsufficiency. RUNX2 regulates expression of downstream genes important for osteoblast function, and its forced expression can upregulate those target genes [25] . Mutation of the mouse gene demonstrated the requirement for *Runx2* in bone formation throughout the skeleton, and its continued expression is also required for normal bone homeostasis [22, 127] .

The years since its identification have yielded a detailed understanding of the pathway downstream of *Runx2* leading to differentiated osteoblasts, with identification of many genes whose transcription is directly regulated by *Runx2*. Comparatively, almost nothing is known about the transcriptional regulation of *Runx2* itself. This is a critical question, since initiation of *Runx2* expression in development determines when and where bones will form, and its ongoing expression is important for proper maintenance of bone throughout life. Numerous signaling pathways have been implicated in its induction, but none has been shown to directly regulate *Runx2* transcription *in vivo* [128] .

Direct regulation of a gene is accomplished by the binding of diffusible trans regulatory factors, either directly or to other trans factors, to cis-regulatory elements (CREs) [129] . CREs are regions of genomic DNA with some role in activating, maintaining, or repressing transcription of an mRNA product.

Elements containing or immediately adjacent to the one of the transcriptional start sites of a gene are generally classified as promoters. They mark the site of recruitment of RNA polymerase and melting of the DNA strand and consequently possess an innate directionality. Conversely, enhancers can positively regulate gene transcription without regard to DNA strand orientation, and may be located at either a great distance from the transcriptional start site or, potentially on other chromosomes altogether [130, 131] , though this is not known to be a common phenomenon in vertebrate genomes. Additionally, exons of neighboring genes can also function as enhancers [132, 133] . Other forms of cis regulatory elements include locus control regions and silencers, capable of preventing gene activation and insulators, which form boundaries to prevent the spread of a repressive heterochromeric chromatin environment through the association of the CTCF protein.

Runx2 is somewhat noteworthy in that it possesses two distinct promoters. The proximal P2 promoter regulates the type I isoform, while the distal P1 promoter (*Runx2* P1) regulates the type II isoform. The two proteins share the same functional domains and are similarly capable of transactivating target genes [134] . The P2 promoter is active at a basal level in a broad number of cells and tissues, including the thymus, cartilage, periosteum, and suture tissue of the calvarium [135-137] whereas the P1 promoter is active in hypertrophic chondrocytes and mature osteoblasts [25, 138] . Although *in vivo*, transcription from both promoters are necessary for fine-tuning *Runx2* expression [139] , they are

incapable of directing proper expression of a reporter transgene by themselves [140, 141], indicating the existence and necessity of a more elaborate cis regulatory architecture. Also, while most characterized CCD mutations affect the *RUNX2* coding sequence, some cases have been associated with translocations of distal regions [142, 143] or have no identified coding sequence [66] mutations, suggesting the presence of critical regulatory sequences whose mutation or disruption severely impairs gene expression.

Therefore, I hypothesized that there must exist additional enhancer elements necessary to direct *Runx2* expression [144]. Methodologies for identifying functional CREs are an area of ongoing inquiry in the literature, each possessing relative strengths and inherent limitations. In a developmental context, the current 'gold standard' experiment for confirming regulatory potential of a DNA region is the deletion of that element in the germline or a relevant integrated BAC *in vivo* and confirmation of a phenotypic of transgene expression change. The chief advantage of the approach is the opportunity to observe an element in its native regulatory environment in a variety of tissue types. Unfortunately, isolating individual elements using germline modifications in mice is costly and time consuming. Additionally, functional redundancy between elements may mask or buffer the consequences of element loss.

Other methods rely on the flood of bioinformatic data that has been made available as the result of numerous genome sequencing projects as well as massively parallel sequencing technologies [145]. Current estimates place

approximately 5% of the human genome under negative purifying selection [146, 147] , where evolutionary forces conserve sequence against the caprice of mutation. However, only 1.5% of the genome appears to have exonic coding potential. The difference between these proportions—popularly referred to as the 'dark matter' of the genome—is due to the relative difficulty of annotating aspects of genes without protein coding potential. Conservation based methods assume conserved non-coding elements possess functional importance whose ablation or alteration would have fitness consequences. While conservation is a generally reliable and accepted criterion for identifying candidate CREs associated with a given gene, there has not yet been an agreement in the field regarding the algorithm or parameters that are best suited for identifying CREs amongst diverse biological contexts [129] . Notably, deletion of many 'ultraconserved' elements in mice resulted in no observable phenotype [148] . Additionally, while conservation might be a good approach to discern sites of input for relatively ancient signaling connections, it is less useful at identifying newly arisen CREs, which are likely to be of the greatest interest from a evolutionary perspective.

Marrying chromatin immunoprecipitation with genome wide interrogation techniques such as microarrays and next generation sequencing permits the mapping of locations of both known transcription factors as well as histone modifications associated with regulatory activity to the genome in a variety of cell types and environmental situations. In particular, enrichment of methylation of 4th lysine of the N-terminus of the histone H3 has been repeatedly

shown to be a powerful technique for locating promoters (H3K4me3) and enhancers (H3K4me1/2) in the absence of detectable sequence conservation [149, 150] . However, these approaches are generally poorly suited for study of enhancers involved in development, because of cell number requirements of precious embryonic materials as well as their necessarily static nature.

With these experimental approaches and limitations in mind, I decided rely on moderate conservation amongst vertebrates to identify CREs associated with the human *RUNX2* locus. A list of candidate CREs was generated by interrogating the locus using PhastCons [151] , an algorithm which relies on multiple vertebrate genomes to both identify conserved elements, as well as a quantitative measure of the evidence of that conservation, permitting a ranking and prioritization of element testing. Because location is a poor predictor of regulatory function, and enhancers can exist at great distances from their associated genes, I examined sequences in an interval of >1 Mb containing human *RUNX2*.

Finally, the qualities of the systems used to actually confirm regulatory activity merits discussion. Especially in developmental contexts, *in vivo* systems are preferable to cell based assays because of the ability to simultaneously observe the full range of likely dynamic regulatory activity in all tissues. Traditionally, mice have been used for this purpose, but for the reasons discussed above, these are a difficult choice for studying embryonic expression, as it requires sacrificing a transgenic animal at each time point of interest.

Additionally, because many transgenes show position dependent effects resulting in ectopic activation in cells that do not actually reflect biological reality, it is necessary to examine a number of embryos to ensure that uniform and robust pattern of expression can be elucidated. In contrast, zebrafish produce hundreds of optically accessible embryos in single clutch. Finally, a highly efficient transgenesis methodology based on the Tol2 transposase permits the construction of potentially hundreds of mosaic transgenic fish in a single morning [152] .

Specifically to skeletal biology, zebrafish are increasingly being employed as an alternative to the standard models of mouse and chick [153] . In addition to the reasons of tractability listed above, they produce the same skeletal cell types as higher vertebrates, albeit in simpler patterns. The appearance of the zebrafish skeleton is rapid; the first bone, the cleithrum[f], is visible within 72 HPF [154] , though expression of bone marker genes in the anlagen begins at approximately 36 HPF [155] . Potential bone specific deficiencies of the system, such as the lack of osteocytes or hematopoietic activity in the bone marrow, are not hindrances for exploring early development [156] .

In this chapter, I describe a screen for *RUNX2*-associated CREs in the human genome. Candidates were selected on the basis of moderate conservation

[f] Most of the imaging in this document will focus on two bones as proxies for the expression in the rest of the skeleton. The cleithrum is a bone of mesodermal origin, and is roughly analogous to the shoulder girdle in mammals. Conversely, the opercle is derived from neural crest, and adopts a fan shaped morphology to lend structural support to the gill flap.

amongst vertebrates and tested for *in vivo* activity in a zebrafish system. Commonalities and differences in expression domains are also noted.

Materials and Methods

Ethics statement. All animal work was conducted according to national and international guidelines, and with the knowledge and approval of the Institutional Animal Care and Use Committee of the University of Pennsylvania.

Identification of conserved non-coding elements associated with RUNX2. To identify sequences for *in vivo* analysis, the candidate locus (hg18; chr6: 44904448-45974166) was interrogated for conserved non-coding elements, as defined by the PhastCons Vertebrate Conserved Elements track during January, 2010[151] . Elements overlapping known RefSeq exons were excluded from further analysis. The top 50 scoring elements (LOD >454) were amplified by PCR (Table 1) from human genomic DNA with LA Taq polymerase (Takara), and cloned into the Tol2 transposon containing vector p*att*P-Tol2-EGFP as previously described [152] .

Transgenesis and expression analyses. Fish were cared for following standard protocols. Each construct for analysis was injected as previously described in at least two separate experiments, and mosaic expression of eGFP analyzed in a minimum of 150 embryos[157] . Embryos were screened from 1- 5 DPF using a Zeiss V12 Stereomicroscope, and imaged with AxioVision 4.5 software. For those constructs regulating a consistent expression pattern, embryos were raised to adulthood and their progeny examined for expression after germline

transmission. All constructs were examined in at least three independent transgenic lines for consistent expression. The *+210RUNX2:*mCherry construct was made in a version of the same vector with mCherry coding sequence instead of eGFP. *sp7:*mCherry fish were made by injection of a modified BAC containing medaka *sp7* regulatory sequences, a kind gift from Christoph Winkler[158] .

Confocal imaging. Embryos were anesthetized in Tricaine and mounted in MatTek glass bottom culture dishes in 1% low melt SeaPlaque agarose. Images were acquired on an Olympus IX 81 microscope equipped with a Yokogawa CSU 10 scan head combined with a Hamamatsu EMCCD camera (model C9100-13, Bridgewater, NJ). Hardware was controlled by Slidebook version 5.0 (Intelligent Imaging Innovations). Diode lasers for excitation (488 nm for eGFP and 561 nm for mCherry) were housed in a Spectral Applied Research launch (Richmond Hill, Ontario). Confocal image stacks were processed with ImageJ (http://rsbweb.nih.gov/ij).

In situ hybridizations. Whole mount *in situ* hybridizations were performed as described[159] , with the following modifications: 0.05% CHAPS detergent was added to the pre-hybridization and hybridization solutions to prevent embryo clumping, and the concentration of NBT was reduced 10-fold in the staining solution to permit overnight development with low background. Stained embryos were dehydrated by successive methanol washes, cleared in methyl salicylate, and mounted in Permount medium (Fisher; SP15-100) between bridged coverslips. Microscopy was performed on an Olympus BX51 with Nomarski

optics. Images were acquired using Spot Basic version 4.6 (Diagnostic Instruments). Further adjustments to white balance and contrast were performed with Adobe Photoshop.

Quantitative PCR. RNA was extracted from whole zebrafish embryos as previously described [160] . cDNA was synthesized using the cDNA high capacity transcription kit (Invitrogen). qPCR was performed in 20 µl reactions using ABI Sybr Green Master mix, and 250µM primer concentration. Samples were amplified using an ABI StepOnePlus real time PCR system.

Identification of cis regulatory elements associated with the zebrafish orthologs runx2a *and* runx2b. The program WPH finder was downloaded from http://rana.lbl.gov/downloads/wph.tar.gz. The three characterized human enhancers (-460*RUNX2*, +210*RUNX2* and +542*RUNX2*) were used as substrates to build word profiles based on the occurrence of 8-mers. These were individually used scan the loci containing the zebrafish *Runx2* orthologs *runx2a* (chr17: 5385672-5740215;Zv8) and *runx2b* (chr20: 44206838-44359224;Zv8) using 250 base pair windows offset by 100 base pairs. Repetitive sequences were removed with RepeatMasker prior to scanning. An arbitrary cut off of Z>5 determined which candidate elements to progress to functional testing.

Results

A screen for RUNX2 associated enhancers identifies three conserved non-coding elements capable of directing expression to osteoblasts in vivo.

As potential cis–regulatory elements for *RUNX2*, we selected the 50 most conserved non–coding sequences within the ~1Mb interval bounded by the 3' UTR of the overlapping gene *SUPT3H* and the 3' UTR of the downstream gene *CLIC5* in the human genome(Figure 1). We grouped proximate sequences into larger amplicons, resulting in 36 constructs (Table 1). Each sequence was tested for *in vivo* enhancer activity, through its ability to direct tissue–specific eGFP expression in zebrafish. We initially screened by examining approximately 150 mosaic injected embryos for fluorescence from 1-6 days post fertilization (DPF); constructs positive in the initial screen were passed through the germline for detailed characterization.

In total, I identified three enhancers capable of directing reporter gene expression to osteoblasts: a distant upstream enhancer located in the last intron of *SUPT3* (*-460RUNX2*)[g], a downstream enhancer in the intergenic space between *RUNX2* and *CLIC5* (*+542RUNX2*), and one in the last intron of *RUNX2* itself (*+210RUNX2*).

Comparison of transgene activity by confocal microscopy and in situ hybridization reveals distinct, yet redundant expression patterns.

While all three enhancers direct expression to osteoblasts, they do not have identical activities. Prior to formation of the first bones, *+210RUNX2* is transiently active in the branchial arches, as evidenced by GFP expression in

[g] The nomenclature used here and subsequently in this document to identify specific enhancer elements is the distance from the transcriptional start site of a gene in kilobases, relative to the directionality of the open reading frame.

live embryos at 2DPF(Figure 2.2j). An earlier and less enduring expression in the branchial arches is directed by +542*RUNX2* between 24-48 HPF (data not shown). +542*RUNX2*:eGFP expression is first detectable in the cleithrum anlagen between 28-32 HPF, while this activity is relatively delayed in embryos carrying the other two transgenes (Figure 2.3a,d). Referring to the confocal data at 3DPF shows that, again, while all three transgenes drive expression in cleithrum in osteoblasts, the expression is much more pronounced in the +542*RUNX2*:eGFP (Figure 2.2f) and -460*RUNX2*:eGFP lines (Figure 2.2o). Furthermore, +542*RUNX2* mediated expression in the cleithrum at this stage is uniform along the dorsal ventral gradient, while in -460*RUNX2*, expression is relatively punctuated at the dorsal and ventral extremes of the bone.

All three enhancers are active in cells of the opercle (Figure 2.4), a neural–crest derived bone[161] that forms by membranous ossification lateral to the branchial arches. However, *+210RUNX2* directs expression to the osteoprogenitors surrounding the edges of the bone (Figure 2.4j,m,p), while the activity of the other two enhancers is confined to cells within the bone itself. Similarly to the expression differences observed in the cleithrum, +542*RUNX2* directs expression uniformly throughout the bone (Figure 2.4s,v,y), while +460*RUNX2* expression is enhanced in the strut and fan structures of the opercle (Figure 2.4 a,d,g)

*+154*runx2a*, the zebrafish ortholog of +210RUNX2, is conserved at the levels of sequence and function.*

While identifying conserved enhancers between distantly related species is not always a straightforward effort, alignment of vertebrate genomes using MultiZ[162] in the UCSC genome browser revealed apparent linear alignment delement, +154*runx2a,* directs expression the branchial arches at 3DPF (Figure 2.5 a) as well as to bony elements such as the opercle and branchiostegal rays at 5DPF (Figure 2.5 b,e).

*Only +542*RUNX2 *directs expression to the developing vertebrae.*

Later in zebrafish development (10-21 DPF), the vertebrae form from migrating sclerotome cells that surround the spinal cord[163] . Though at 5DPF, all three enhancers direct expression to all visible skeletal structures, the ability to direct expression to the vertebral arches at 14DPF is limited to +542*RUNX2* (Figure 2.6c). Some expression by -460*RUNX2*:eGFP is visible in the centra of the vertebrae. (Figure 2.6a). +210*RUNX2* failed to direct expression to any aspect of the vertebral column (Figure2.6b).

Chromatin features of identified enhancers.

As discussed in the introduction, enhancer associated chromatin marks are an often-employed method to identify loci with regulatory activity in a cell or tissue type being studied. There exists an unpublished, publicly available ChIP-Seq data set from a normal human osteoblast cell line as part of the ENCODE project[h]. Enrichment for regulatory element associated marks (H3K4me1,

[h] According the documentation accompanying the data set, the cell line is normal human osteoblasts (NHOst) from Lonza (#CC-2538). The ChIP-Seq data was produced by a collaboration of Bradley Bernstein and Greg Crawford.

H3K4me2, H3K27Ac) was examined as well as recruitment of the enhancer associated coactivator p300 (Figure 2.7). +542*RUNX2* and +210*RUNX2* were significantly enriched for p300, as well as H3K4me1 and H3K4me2, and were identified as peaks in the data set as determined by a hidden Markov model algorithm. Curiously, no enrichment for any of the studied histone modifications or proteins was observed in -460*RUNX2*.

The three RUNX2*-associated enhancers contain enough information content to efficiently identify other skeletal specific enhancers in a non-conservation based approach.*

With the limitations of conservation based approaches to discovering CREs in mind, I wanted to explore whether other methodologies might be efficient at identifying enhancers associated with the zebrafish *RUNX2* orthologs, *runx2a* and *runx2b*, for which there are few alignable genomes available to detect conservation. One such strategy involves utilizing the information content of known enhancers as a basis for predicted new ones from untested sequence data. WPH Finder is such an algorithm that has been successfully used to recognize enhancer elements not identifiable by linear conservation between Drosophila species[164] . It trains itself by counting the occurrence of specific eight letter DNA 'words' (which likely correspond to transcription factor binding sites, or other features conferring regulatory activity upon an element), forming a profile of these word, and then testing windows of candidate sequences for overrepresentation relative to it.

Using this methodology, the loci of the zebrafish *Runx2* orthologs, *runx2a* and *runx2b* were scored for word profile similarity to the each of the three known human enhancers (Tables 2.2. 2.3). Eighteen candidates were tested from the *runx2a* locus, and ten were examined from that of *runx2b*. In contrast to the relatively low rate of skeletal enhancer activity identified in conserved candidates from the human locus (3/38 = 7.8%), 38.9% (7/18) of amplicons tested from the *runx2a*[i] locus and 50% (4/10) of those from *runx2b* showed some pattern of skeletal expression (Figure 2.8a,b). Notably, almost all of the positive elements (11/12) fall within the coding region of either one of the *Runx2* orthologs, or *supt3h*, whose syntenic relationship with *Runx2* is ancient. Representative views of expression patterns of these elements show expression in both subsets of osteoblasts as well as cartilaginous structures (Figure2.8c-h).

[i] These numbers exclude +154*runx2a*, which had already been shown to regulate a skeletal expression pattern.

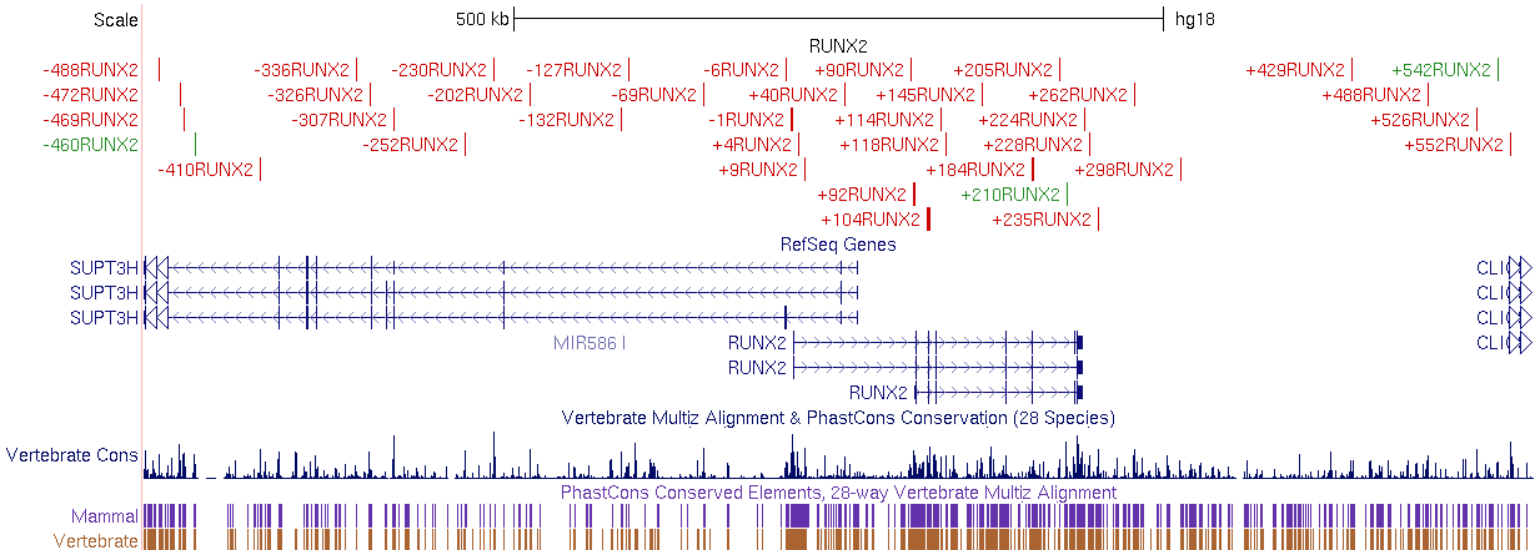
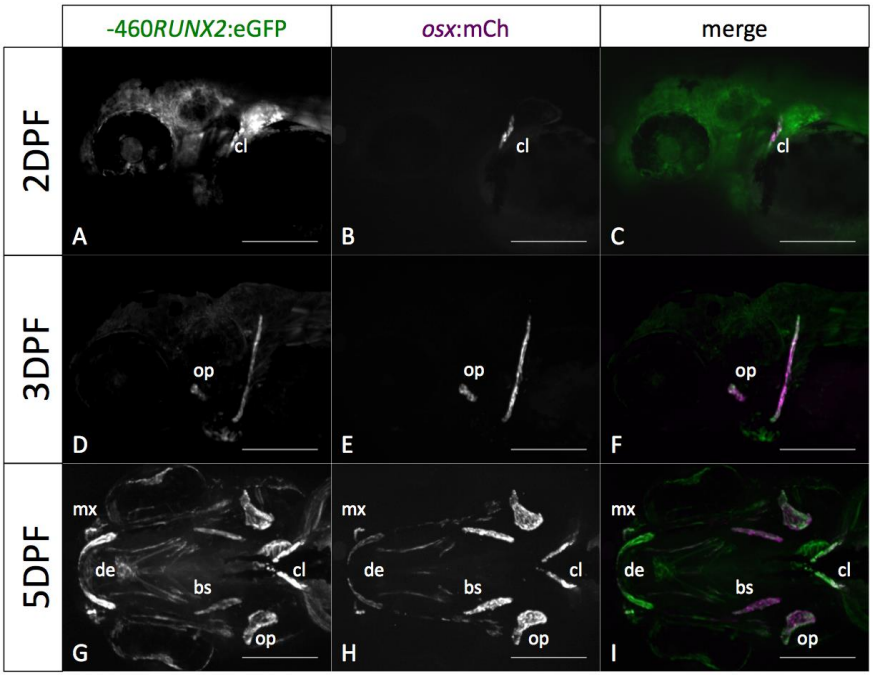
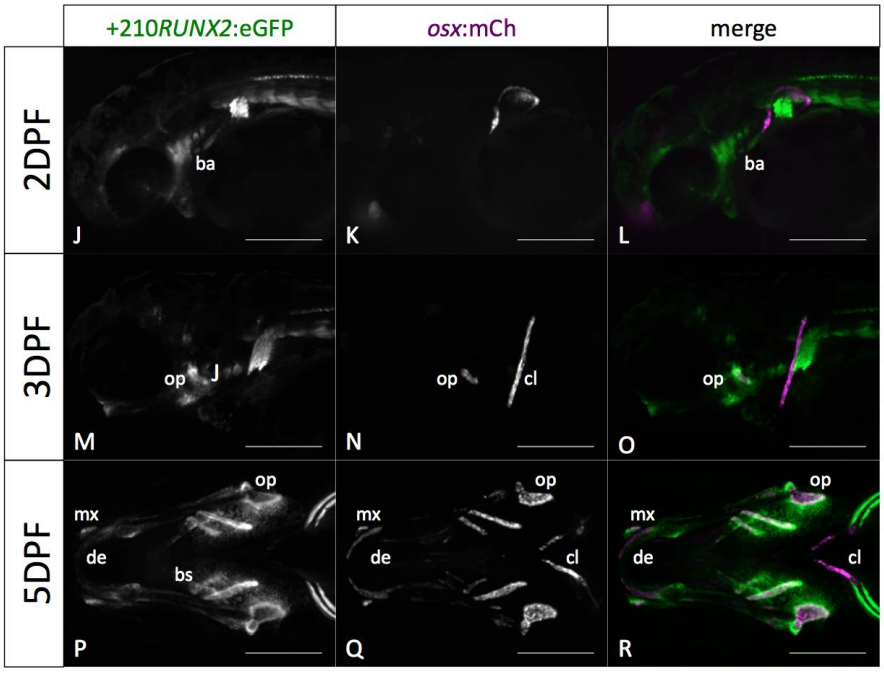

Scale
500 kb
hg18
RUNX2
-488RUNX2
-472RUNX2
-469RUNX2
-460RUNX2
-410RUNX2
-336RUNX2
-326RUNX2
-307RUNX2
-252RUNX2
-230RUNX2
-202RUNX2
-132RUNX2
-127RUNX2
-69RUNX2
-6RUNX2
-1RUNX2
+4RUNX2
+9RUNX2
+80RUNX2
+40RUNX2
+92RUNX2
+104RUNX2
+205RUNX2
+145RUNX2
+114RUNX2
+118RUNX2
+184RUNX2
+210RUNX2
+262RUNX2
+224RUNX2
+228RUNX2
+298RUNX2
+235RUNX2
+429RUNX2
+488RUNX2
+526RUNX2
+542RUNX2
+552RUNX2
RefSeq Genes
SUPT3H
SUPT3H
SUPT3H
MIR586 I
RUNX2
RUNX2
RUNX2
CLI
CLI
CLI
CLI
Vertebrate Multiz Alignment & PhastCons Conservation (28 Species)
Vertebrate Cons
PhastCons Conserved Elements, 28-way Vertebrate Multiz Alignment
Mammal
Vertebrate

Figure 2.1. – Broad distribution of osteoblast specific enhancers at the *RUNX2* locus. The *RUNX2* locus (chr6: 44,904,448-45,971,166, hg18) as represented in the UCSC Genome Browser is shown, with the conserved non–coding elements tested for enhancer assay indicated at top. Elements testing positive for osteoblast expression *in vivo* are shown in green, while those with no activity in skeletal tissues are indicated in red. Tracks displaying all conserved elements as defined by PhastCons amongst vertebrates and mammals are displayed at the bottom of the figure to visualize relative conservation across the genomic region.


-460RUNX2:eGFP
osx:mCh
merge
2DPF
3DPF
5DPF
cl
op
mx
de
bs
A
B
C
D
E
F
G
H
I


+210RUNX2:eGFP
osx:mCh
merge
2DPF
3DPF
5DPF
ba
op
J
cl
mx
de
bs
K
L
M
N
O
P
Q
R

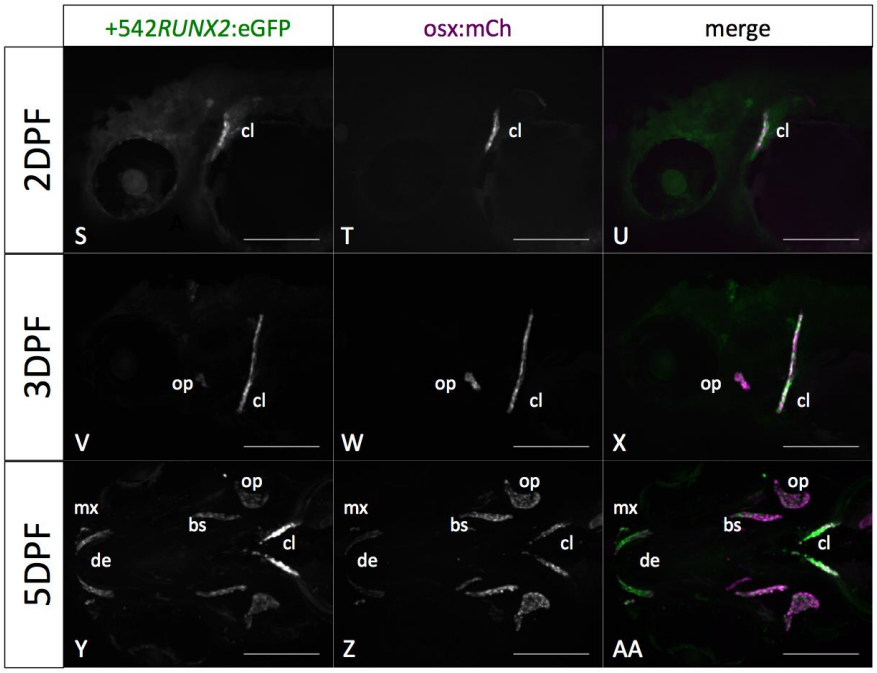


Figure 2.2 -- Expression patterns of three human *RUNX2* associated enhancers. Lateral views of doubly transgenic zebrafish embryos for *sp7*:mCh and -460 *RUNX2*:eGFP (A-I), +210*RUNX2*:eGFP (J-R), and +542*RUNX2* (Q-AA) were imaged at 2DPF and 3DPF. In 2DPF embryos, coexpression of *sp7*:mCh and -460*RUNX2*:eGFP (A) and +542*RUNX2* :eGFP (S) in the developing cleithrum, but not in +210*RUNX2* (J). Conversely, branchial arch expression in +210*RUNX2*:eGFP is apparent at 2DPF. All three *RUNX2* transgenes express in the opercle anlage at 3DPF with *sp7:*mCh (D,M,V). Ventral views imaged at 5DPF demonstrate concomitant RUNX2 transgene expression in later ossifications (G,P,Y).ba, branchial arches; bs, brachiostegal ray; cl, cleithrum; de, dentary mx, maxilla; op, opercle; Scale bar = 50 mm

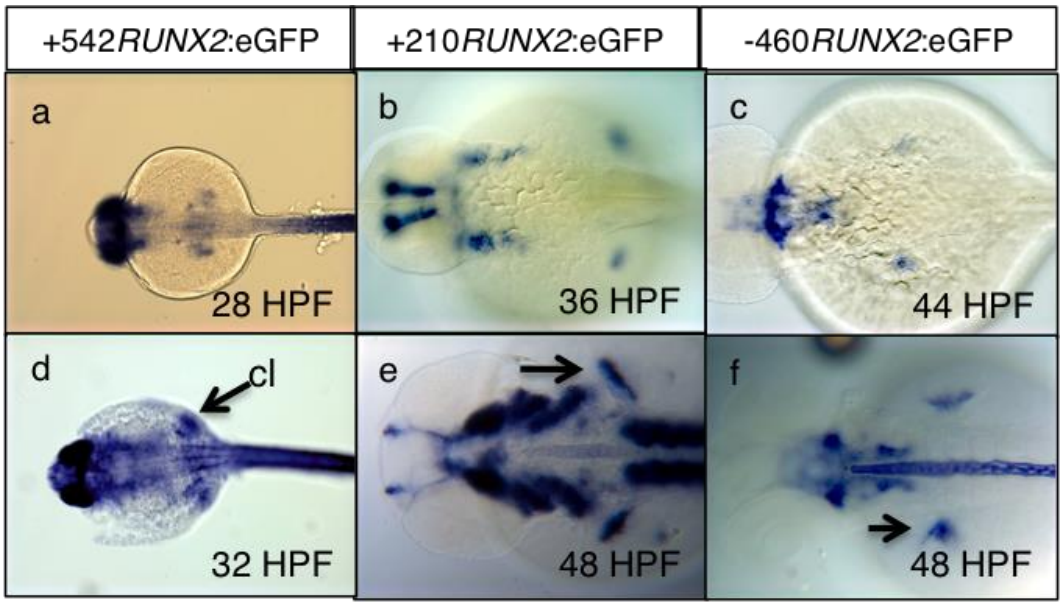
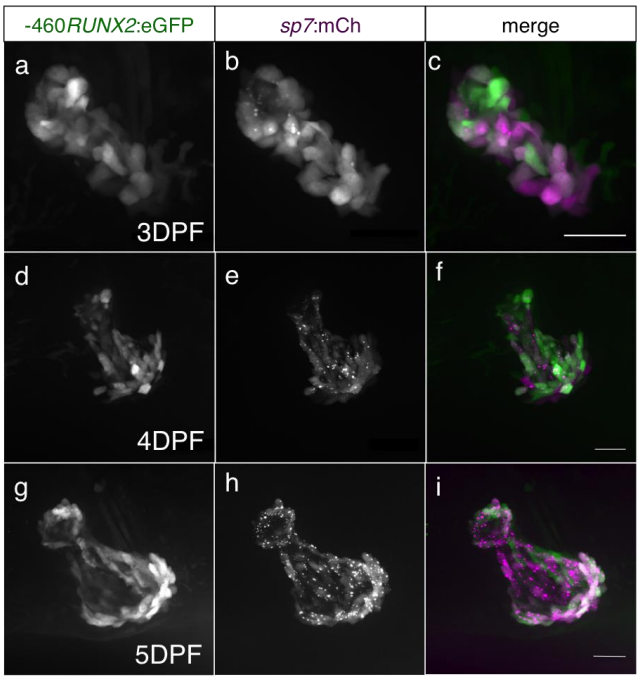
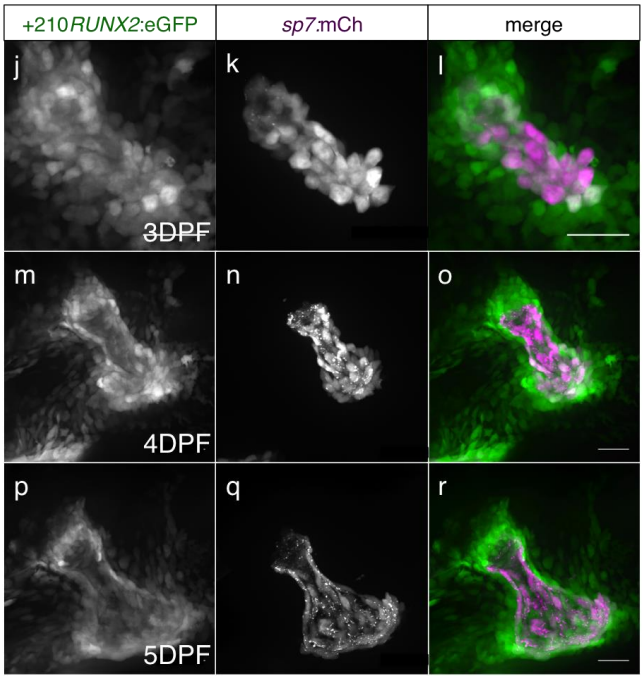


Figure 2.3 -- Differential expression initiation times in the cleithrum anlagen. Embryos from each transgenic line were fixed at 4 hour intervals and analyzed by in situ hybridization with an eGFP probe to determine onset of expressing in the cleithrum. Dorsal views of representative embryos of +542RUNX2:eGFP (a,d), +210*RUNX2*:eGFP (b,e) , and +542*RUNX2*:eGFP (c,f).


-460RUNX2:eGFP
sp7:mCh
merge
a
b
c
3DPF
d
e
f
4DPF
g
h
i
5DPF


+210RUNX2:eGFP
sp7:mCh
merge
j
k
l
3DPF
m
n
o
4DPF
p
q
r
5DPF

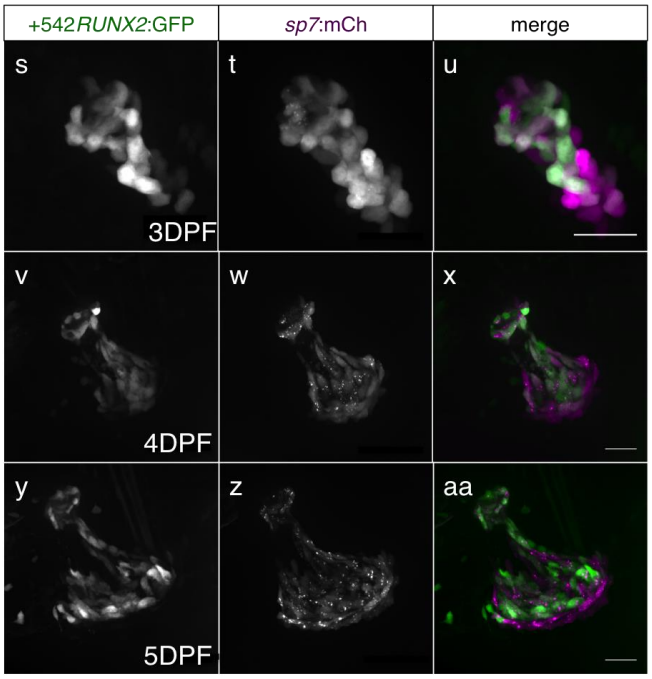


Figure 2.4 – Distinct expression domains in the developing opercle.

Lateral views of the opercle in doubly transgenic zebrafish embryos for *sp7*:mCh and (a-i), -460 *RUNX2*:eGFP (j-r) +210*RUNX2*:eGFP, and (q-aa) +542*RUNX2* were imaged at 3DPF, 4DPF and 5DPF; Scale bar = 10 mm

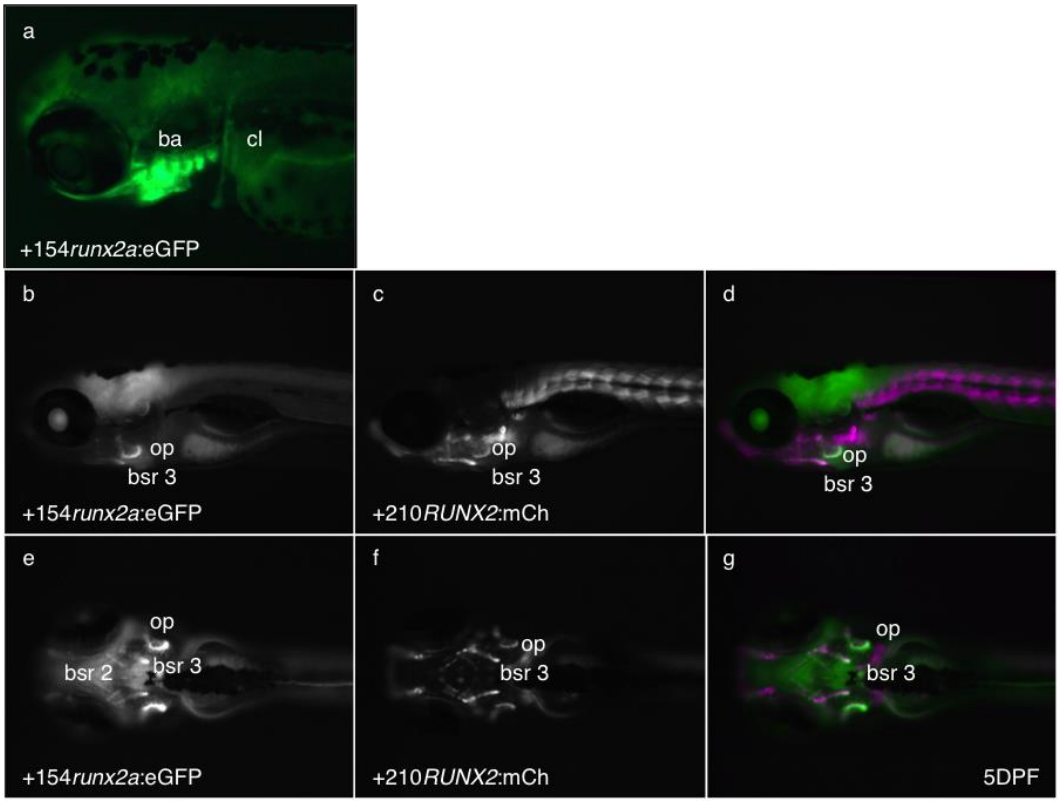


Figure 2.5 – +154*runx2a*, a conserved ortholog of +210*RUNX2,* directs expression to the branchial arches and osteoblasts. Lateral (a-d) and ventral (e-g) of doubly transgenic zebrafish for +154*runx2a*:eGFP and +210*RUNX2*:mCh at 3DPF(a) and 5DPF (b-g). ba, branchial arches; bsr2, brachiostegal ray 2 bsr3, brachiostegal ray 3; op, opercle;

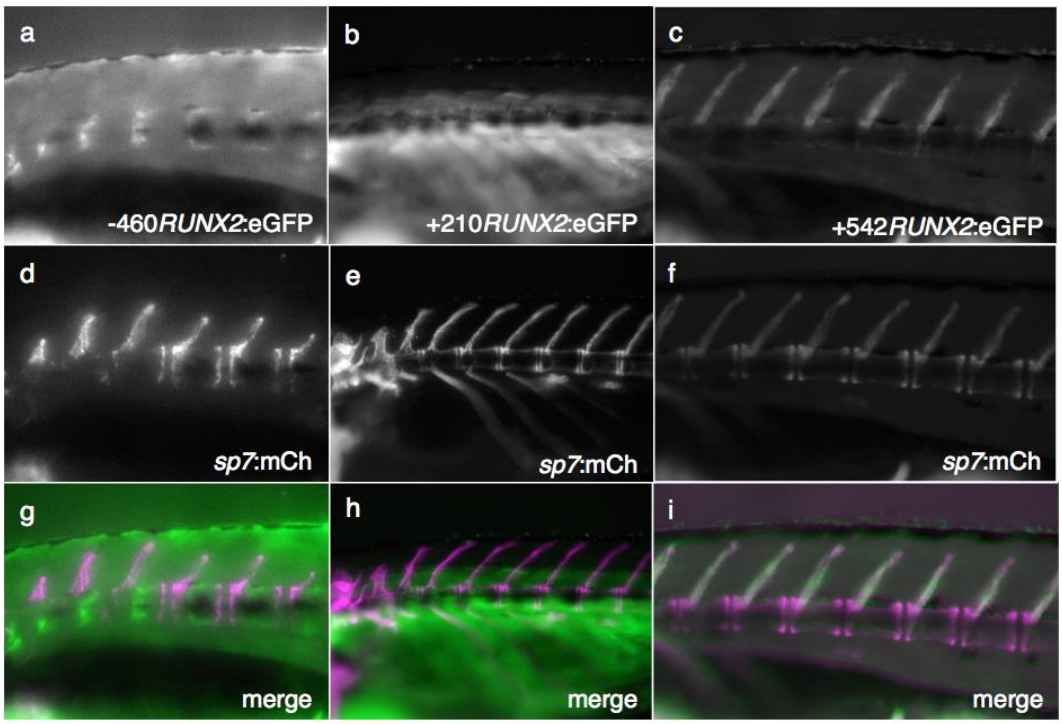


Figure 2.6 -- +542RUNX2 directs expression to the developing vertebral arches. Lateral views of the developing vertebrate column of doubly transgenic zebrafish for the indicated *RUNX2*:eGFP transgene and *sp7*:mCh at 14 DPF. -460*RUNX2*:eGFP expression (a) is directed to the anterior edge of the presumptive vertebrate. +210*RUNX2*:eGFP (b) does not express in any portion of the anatomy and +542*RUNX2*:eGFP (c) directs expression to the vertebral arches. (d-f) demonstrate *sp7*:mCh mediated expression; (g-i) shows a merge of these two pattern,

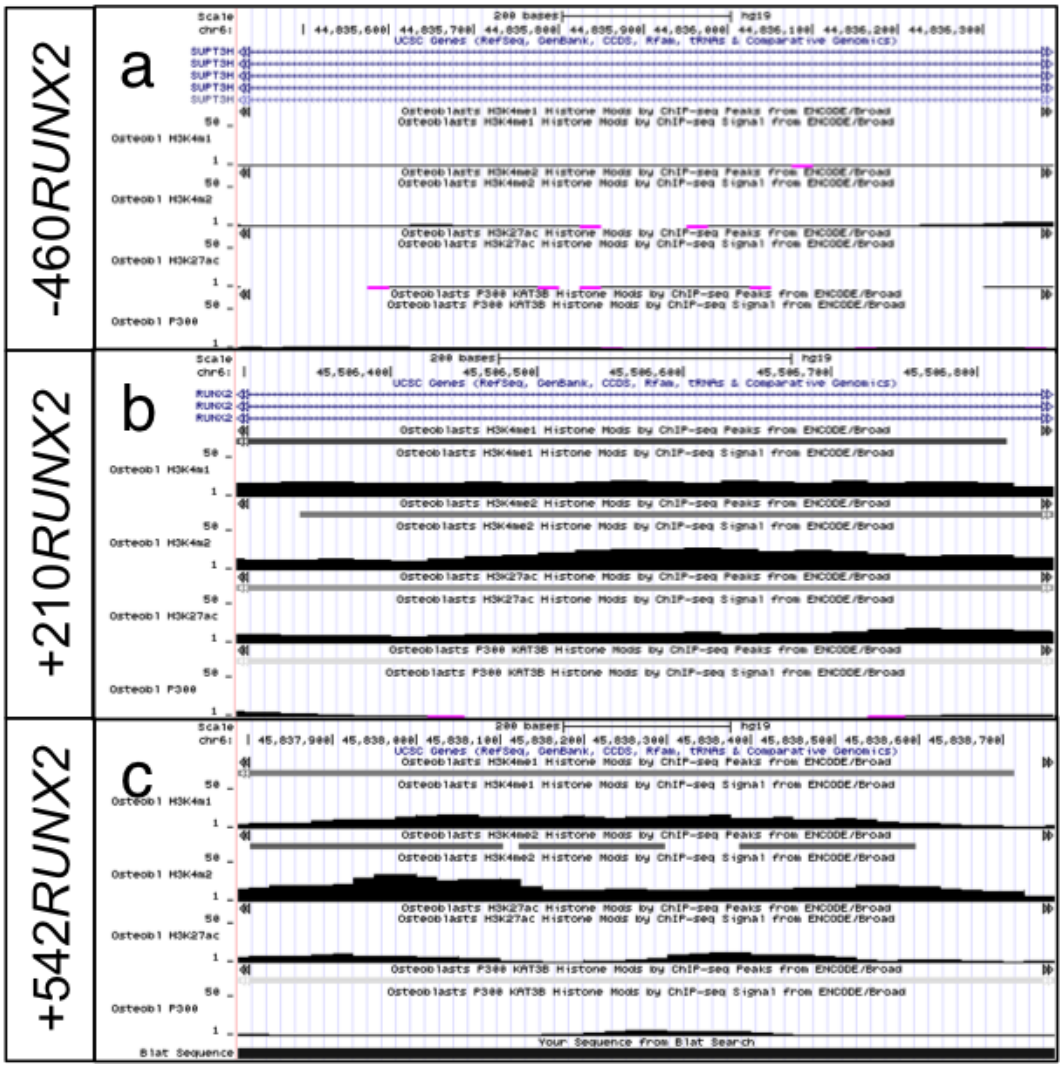

-460RUNX2
+210RUNX2
+542RUNX2
a
b
c
Scale
chr6:
200 bases
hg19
UCSC Genes (RefSeq, GenBank, CCDS, Rfam, tRNAs & Comparative Genomics)
SUPT3H
RUNX2
Osteobl H3K4me1
Osteobl H3K4me2
Osteobl H3K27ac
Osteobl P300
Osteoblasts H3K4me1 Histone Mods by ChIP-seq Peaks from ENCODE/Broad
Osteoblasts H3K4me1 Histone Mods by ChIP-seq Signal from ENCODE/Broad
Osteoblasts H3K4me2 Histone Mods by ChIP-seq Peaks from ENCODE/Broad
Osteoblasts H3K4me2 Histone Mods by ChIP-seq Signal from ENCODE/Broad
Osteoblasts H3K27ac Histone Mods by ChIP-seq Peaks from ENCODE/Broad
Osteoblasts H3K27ac Histone Mods by ChIP-seq Signal from ENCODE/Broad
Osteoblasts P300 KAT3B Histone Mods by ChIP-seq Peaks from ENCODE/Broad
Osteoblasts P300 KAT3B Histone Mods by ChIP-seq Signal from ENCODE/Broad
Blat Sequence
Your Sequence from Blat Search

Figure 2.7 – Analysis of chromatin environment at three RUNX2 associated enhancers in a normal human osteoblast cell line. The human skeletal enhancers -460*RUNX2* (a), +210*RUNX2* (b), +542*RUNX2* (c) are shown as represented in the UCSC genome browser along with tracks indicating both the enrichment of and presence of peak of the following histone modification/proteins in normal human osteoblast cell lines as defined by ChIP-Seq.

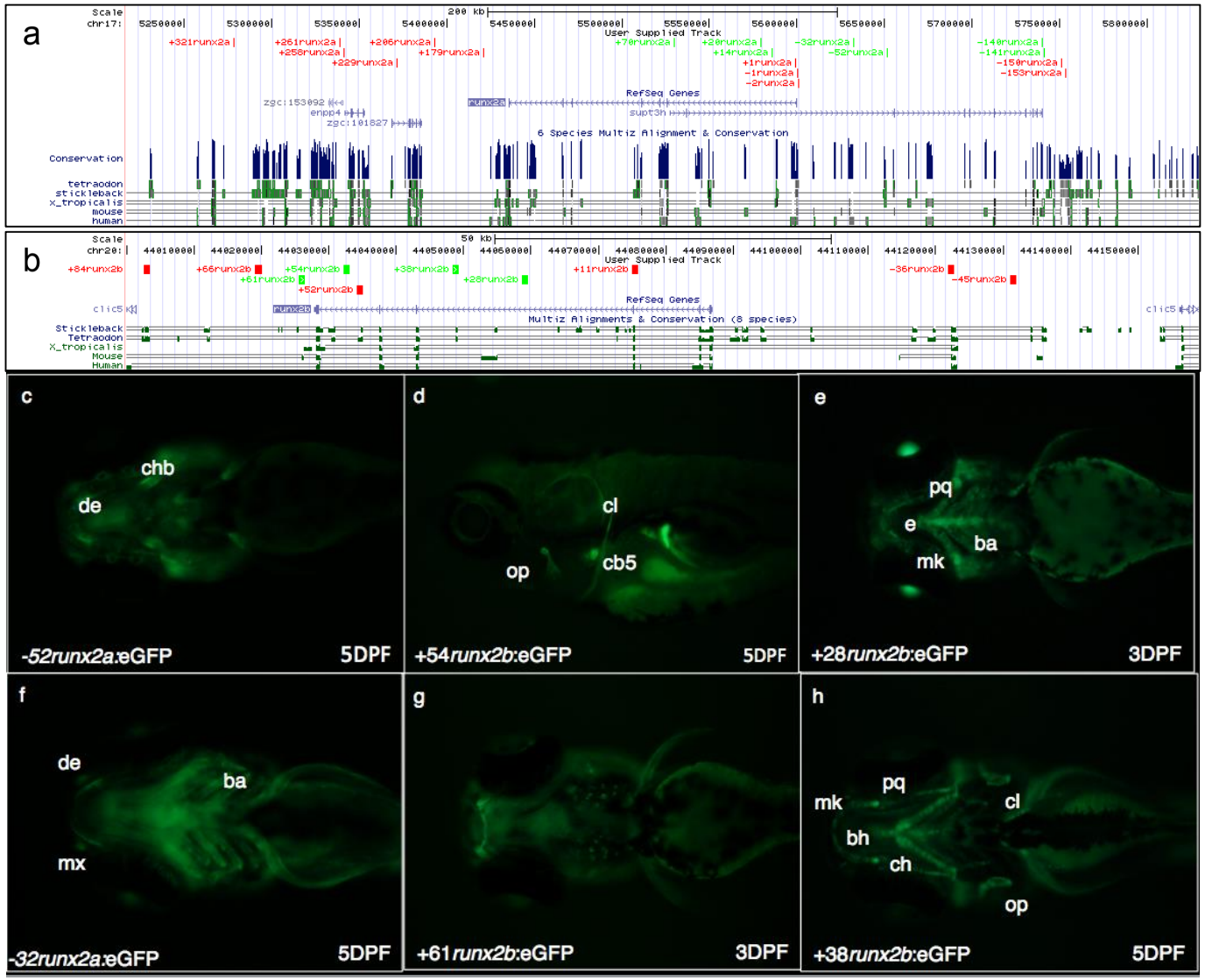

a
Scale
chr17:
200 kb
User Supplied Track
+321runx2a|
+261runx2a|
+258runx2a|
+229runx2a|
+206runx2a|
+179runx2a|
+70runx2a|
+28runx2a|
+14runx2a|
+1runx2a|
-1runx2a|
-2runx2a|
-32runx2a|
-52runx2a|
-140runx2a|
-141runx2a|
-150runx2a|
-153runx2a|
RefSeq Genes
zgc:153092
enpp4
zgc:101827
runx2a
supt3h
6 Species Multiz Alignment & Conservation
Conservation
tetraodon
stickleback
X_tropicalis
mouse
human
b
Scale
chr20:
50 kb
User Supplied Track
+84runx2b
+66runx2b
+61runx2b
+54runx2b
+52runx2b
+38runx2b
+28runx2b
+11runx2b
-36runx2b
-45runx2b
RefSeq Genes
clic5
runx2b
Multiz Alignments & Conservation (8 species)
Stickleback
Tetraodon
X_tropicalis
Mouse
Human
c
chb
de
-52runx2a:eGFP
5DPF
d
cl
op
cb5
+54runx2b:eGFP
5DPF
e
pq
e
ba
mk
+28runx2b:eGFP
3DPF
f
de
ba
mx
-32runx2a:eGFP
5DPF
g
+61runx2b:eGFP
3DPF
h
mk
pq
cl
bh
ch
op
+38runx2b:eGFP
5DPF

Figure 2.8 – Screen for skeletal enhancers associated with the zebrafish *Runx2* orthologs, *runx2a* and *runx2b*. The *runx2a*(a; chr17: 5385672-5740215;Zv8) and *runx2b*(b; chr20: 44206838-44359224;Zv8) loci(chr6: 44,904,448-45,971,166, hg18) as represented in the UCSC Genome Browser is shown, with the elements tested for enhancer assay indicated at top. Elements testing positive for skeletal expression *in vivo* are shown in green, while those with no activity in skeletal tissues are indicated in red. Tracks displaying all conserved elements as defined by 6 way MultiZ alignment displayed at the bottom of the figure to visualize relative conservation across the genomic region. Representative images of transgenic fish carrying these elements illustrate the diversity of skeletal expression observed. -52*runx2a*:eGFP (c) expresses in a subset of ossifying structures at 5DPF; +54*runx2b*:eGFP (d) expression is exclusive to osteoblasts; +28*runx2b*:eGFP (e) expression is found in the cartilaginous elements of the neurocranium; -32*runx2a*:eGFP (f) expressed in the branchial arches as well as the dentary and maxilla forming the mouth; and +38*runx2b*:eGFP (h) expresses in the pharyngeal skeleton and dermal bones. ba, branchial arches; bh, basihyal; bs, brachiostegal ray; cb5, ceratobranchial 5; ch, ceratohyal; chb, ceratohyal bone; cl, cleithrum; de, dentary; e, ethmoid; mk, Meckel's cartilage; mx, maxilla; op, opercle; pq, palatoquadrate

Element	Location	Size	LOD	F primer	R primer
-488	44808206-8579	373	1105	agcccagcaccagtttttga	agaaggcaaaggggaagcag
-472	44824048-4376	328	780	tgcgctcacattctttttctg	tccattatggcgctttgtgtt
-469	44827066-7407	341	677	gcacctggagaaggacctga	cagttcagccttgggacagc
-460	**44835694-6092**	**398**	**692**	**gcaaaaaccacaacccctca**	**aggcagaaatggcaccttca**
-410	44886294-6648	354	1190	agccacctggttttctccaa	cgccccctgaaaaagaaact
-336	44959617-9931	314	454	gctcctcagcatggatgcac	tcgtgtccactgttggtgatg
-326	44970259-0571	312	741	gtaccctgctgacccccagt	ttgcatttgagtgccttcca
-307	44988634-9028	394	1126	ctttgcgcatcaaaaacaga	gcccatattcttctcaacactg
-252	45043906-4200	294	625	gcaagttccaggcagaagaa	ggctaatgattaaaggcgtca
-230	45065752-6659	907	1983	acagaagcaaaggcagagga	aacttggcccaaaatgtgac
-202	45093924-4357	433	855	gtgggcttcagcaaaacctg	gcatcgtctcttcccccttt
-132	45163927-4407	480	2070	ttctgatcccgggcaagtta	tccatgttgtggctttttacg
-127	45169299-9615	316	600	tgcaacttgccaacttaaaatgc	tgcctgattttgagagctggat
-69	45227036-7415	379	1240	ggcttcaaagtatcccttccaa	gcaatgaacctgtcatttccaa
-6	45290403-1852	1449	973; 670	tgcctatgcatcaagaaacaga	ccctccctccctctccctat
-1	45294619-6531	1912	896; 888; 3526	ctgcatgttaaatgaacgctttt	ttcatttttcaattcttaatgcaaca
+4	45299716-0238	522	697; 668	tgtggaaaaagatctggctggt	gactgtttcggtgaatccttcat
+9	45305220-5485	265	758	aagccaaccccgaacatagg	ccagttcagcagaatcttgtttga
+40	45335673-5997	324	892	aaccccagcctccagaaagt	ttctgaggaaaggcgtgcat
+90	45386404-6653	249	520	gggagtggtgggtgggatta	cctcctcccaggaagtccat
+92	45388499-0699	2200	674	gagcagtgcatggaggaagg	ctccccacccctaggacaag
+104	45399557-1838	2281	1546	ccatcatctgaatccatattcaaaaa	tgcaagttccactgatgctct
+114	45409604-0034	430	1238	ttttcccccctgcagatgac	ttttccatcctctgcacaacaa
+118	45413657-3853	196	688	tccgaagtgcccaaggaata	gccatggcacaaacaaaataga
+145	45441266-1580	314	734	cccacctgccaattgtgtct	gacaaaatcatttggcgcaga
+184	45479980-1419	1439	726	gacctggctgcccctctatc	cgcatcagccttcgttatcc
+205	45501079-1314	235	516	agatgtggccctgcatctgt	ggtatgccttctcgcactgg
+210	**45506418-6731**	**313**	**1302**	**aatttggctccatgttttgg**	**ggcaggcagtagatgtgtga**
+224	45520471-0722	251	511	aggcagaggctgagggactt	ttgtcctccacttccccaga
+228	45524043-4304	261	698	gtcattcctgcctcccagtg	gcagcctggtcactgtctga
+235	45531043-1392	349	786	cctccttgcgggctactctt	ttcaaaagcccttgccgata
+262	45558534-8861	327	1283	tgacaccaggctgcaagtgt	ggggatgactgcacatggat
+298	45593889-4243	354	648	ttgccctcatcacatcaacg	atgacttgggcgccatgtag
+429	45725477-5774	297	1021	agctcaagcatggcaagagg	cctcaaccccaagacagacg
+488	45784423-4811	388	911	ccccctctttctttcattttcc	cccccgtcactctctctcct
+526	45822026-2328	302	629	gggcaatgcatttgactggt	acagctcagcatttggcaca
+542	**45838035-8397**	**362**	**773**	**acagccgccattttactcca**	**cacagagagagacagagaggagaa**
+552	45848037-8491	454	635; 643	caacacccagggagaagtcg	cccagctgagctttgtccttt

Table 2.1. Sequences at *RUNX2* locus tested for enhancer activity. Each element is named for the distance in kilobases from the transcription start site of *RUNX2*; the location of each conserved region is given as coordinates in hg19, chr6. Where multiple conserved sequences were grouped into one amplicon, the LOD scores for each are listed separately. The primer sequences are those used to amplify the elements from genomic DNA for testing. The three elements in bold are those with activity in osteoblasts.

Name	Coordinates(Zv8)	F primer	R primer	Product size (bp)	Similarity
+321runx2a	chr17:5,278,105-5,278,353	ACAGATGGCGGTGAAGCAGT	GCACGTCCGAATGTCAACAA	1061	-460
+261runx2a	chr17:5,338,455-5,338,703	TGTGAAGCTGCTTTGACACAATC	CGGTGTCCTGTTGCTCAGTG	870	+542
+258runx2a	chr17:5,341,255-5,341,503	CGAGCGTCCATTCACAAACA	AGGAAACACTGAAGGACAAAATGC	852	-460
+229runx2a	chr17:5,371,055-5,371,303	CCCAACGTCGGCTTACGATA	TTGCAGTGAGATTGCGTTGG	892	+542
+206runx2a	chr17:5,392,750-5,393,159	AGCACCACCACCAACTGGAT	GGAGCAGCTGAAGAGGCTTG	938	-460,+210
+179runx2a	chr17:5,420,855-5,421,103	CTAGCGCCATTGCTGGTTTC	ATCAGATTCCATGCGGTTCG	991	-460
+70runx2a	**chr17:5,529,955-5,530,203**	**AGCAGTTACGCTTTTGATGGAG**	**ACATATTTGGCGCTCGCAGT**	**1017**	**-460**
+20runx2a	**chr17:5,579,405-5,579,653**	**TCTGCTGGCCCATAAGAAAAA**	**TCTTGGAGCAACTGGCAAGC**	**882**	**-460**
+14runx2a	**chr17:5,585,655-5,585,903**	**TTCAATGGACTTTGATTCAGCTT**	**CAAGCAGTGACTGACAAATGAAATG**	**852**	**+542**
+1runx2a	chr17:5,599,105-5,599,353	CGGCAGTGATGACAAAACCA	TCACCACGACCTGCAGAAGA	895	-460
-1runx2a	chr17:5,600,005-5,600,253	TCCTAAAGCGGGAGCACAAA	TGACCCCGAAACAGGAGAGA	865	-460
-2runx2a	chr17:5,600,505-5,600,753	TGTCAGTGGTCCTGCGTTGT	TGACTGAAGGCAGTCGACCA	884	+542
-32runx2a	**chr17:5,632,105-5,632,353**	**GGCAACCACAAATTGAAAACC**	**CCGCAACCATACGGGACTAA**	**1455**	**+542**
-52runx2a	**chr17:5,651,555-5,651,803**	**CTCCCTTCATGGTGGCTTCA**	**CCTCAAACCAGGGCACTAAGAC**	**949**	**-460**
-140runx2a	**chr17:5,740,055-5,740,303**	**AGCGATAGAGCCGAGACGTG**	**ACTGAAGCTGCGTCCCAAAA**	**2456**	**+542**

-141runx2a	**chr17:5,741,305-5,741,553**	**CGAACAGACAGATGAATAAAAAGACAA**	**GCAACCCATCTCTGGGAAAC**	**1262**	**+542**
-150runx2a	chr17:5,750,755-5,751,003	TGGGCTGGAGACCAAGAAAA	TGCATTGCACATAGGGGACA	1192	+210
-153runx2a	chr17:5,753,255-5,753,503	TCCCGAAGATCTGGCAAATG	GGCCTGGATGCATCATTTTT	863	+460

Table 2.2. Sequences at *the runx2a* locus tested for enhancer activity. Each element is named for the distance in kilobases from the transcription start site of *runx2a*. The primer sequences are those used to amplify the elements from genomic DNA for testing. The seven elements in bold are those with skeletal activity. Similarity indicates the human *RUNX2* associated enhancer to which WPHFinder identified word profile similarity.

Name	Coordinates(Zv8)	F primer	R primer	Product size (bp)
-45*runx2b*	chr20:44,131,047-44,131,967	TGCGCTCTTTCCAGCAATTT	TGGGCACGCTATGATGTGAC	921
-36*runx2b*	**chr20:44,121,758-44,122,750**	**GTGAAGGCTCCGCTCACCT**	**GGCAAAGACATACCAGGCAAA**	**993**
+11*runx2b*	chr20:44,075,029-44,075,935	GCGGGGCATGTCAGATTCTA	CGACAGAGAGGTGAGCGTGA	907
+28*runx2b*	**chr20:44,058,608-44,059,525**	**CCAAAACAATGAACGGCAGA**	**CGGCAGCCAGAAGAGAGAAC**	**918**
+38*runx2b*	**chr20:44,121,758 - 44,122,750**	**ACCATCCGACAAGCTGATCC**	**TGGAAATCAATGGGGCAAAA**	**909**
+52*runx2b*	chr20:44,034,157-44,035,096	TTGAAGCGGGGTTCATTTTG	GCCCGAACATGAAGGTAACTC	940
+54*runx2b*	**chr20:44,032,209-44,033,092**	**TGTGCTCACCTTTAAGTGGTTCA**	**GGGAGAGAGCCCTGAGCATA**	**884**
+61*runx2b*	**chr20:44,025,533-44,026,461**	**TGGTTCTTAGATGGCAATGAGC**	**CCGTCTCGATTCCTTCAATCC**	**930**
+66*runx2b*	chr20:44,019,065-44,020,088	GGGAAACATCCATACATAAAAAGTGTT	GAAACACACACTCAATCACACTCA	1000
+84*runx2b*	chr20:44,002,595-44,003,488	TGCTTTAATTTATCATCCTTTTGCAG	CCTAACGTGGCGAAAAGGCTA	895

Table 2.3. Sequences at *the runx2b* locus tested for enhancer activity. Each element is named for the distance in kilobases from the transcription start site of *runx2b*. The primer sequences are those used to amplify the elements from genomic DNA for testing. The five elements in bold are those with skeletal activity.

CHAPTER 3

FUNCTIONAL CHARACTERIZATION OF *RUNX2* ASSOCIATED ENHANCERS.

CHAPTER 3

Introduction

Having identified three *RUNX2* associated enhancers from the human genome capable of directing expression to osteoblasts in Chapter 2, I attempted to characterize upstream regulators for each CRE. Fortunately, the transgenic fish lines I created to describe expression patterns also serve as a useful platform to study the underlying biology of their reporter constructs.

Materials and Methods

Site directed mutagenesis. Predicted, conserved transcription factor binding sites in *+210RUNX2* were identified via linear alignment to its zebrafish ortholog, +154*runx2a,* using Clustal W (http://www.ebi.ac.uk/Tools/msa/clustalw2/). Binding sites were individually mutated via PCR. Briefly, each site was converted into a unique restriction site via two parallel PCR reactions with primer pairs that introduced the restriction site and attB sequence flanking each amplicon[j]. A subsequent digestion and ligation step produced a mutagenized enhancer competent for Gateway recombination. For *-460RUNX2mutTCF* and *+542RUNX2mutDLX*, the mutated enhancer sequences were synthesized (GeneWiz), and cloned directly into p*att*P-Tol2-egfp as described in Chapter 2.

Drug treatments. To screen for responsiveness to candidate signaling pathways, embryos transgenic for each enhancer construct were treated from 48

[j] Mutagenized +210RUNX2 constructs were constructed by Gina Mahatma.

to 72 HPF with inhibitors of FGF signaling (SU5402; 10μM) BMP signaling (dorsomorphin; 30μM), retinoic acid (DEAB; 50μM), notch (DAPT; 20μM), hedgehog (cyclopamine; 50μM) and calcineurin (FK506; 3μM) signaling and screened for changes in GFP expression. Additionally, *+210RUNX2:egfp* embryos were treated from 28-32 HPF, or *+210RUNX2:mCh* embryos from 100-104 HPF, with SU5402 at 10 μM, and *-460RUNX2:mCh* transgenic embryos were treated from 52-56 HPF with GSK3β inhibitor XV at 5μM, before being harvested for *in situ* hybridization or Q-PCR analysis. For all treatments, drugs were dissolved in DMSO to make stock solutions, which were diluted into embryo medium; additional DMSO was added to equalize concentration for all treatments.

Heat shock treatments. Embryos doubly transgenic for *-460RUNX2:mCh* and *hsp70:dkk1*[165] were immersed in pre–warmed embryo medium at 37°C for 30 minutes at 52 HPF. Following heat shock, embryos were transferred to fresh embryo medium at 28.5°C and incubated for 3.5 hours before harvesting for analysis. For *+542RUNX2:egfp,* embryos doubly transgenic with either *hsp70:bmp2b*[166] or *hsp70:chd*[167] were similarly heat shocked at 48 HPF, and harvested at 56 HPF.

DNA sequence alignments. Orthologs of human sequences were identified by BLAT. Sequences were downloaded from the UCSC genome browser, curated into FASTA files and aligned using Clustal X (http://www.clustal.org)

Confocal imaging, *in situ* hybridizations, zebrafish transgenesis and quantitative RT-PCR all performed as described in Chapter 2.

Results – Characterization of +210*RUNX2*

*Identification of conserved predicted transcription factor binding sites in +210*RUNX2

The search for sequence features potentially critical for the ability of an enhancer to direct skeletal expression was facilitated greatly in the example of +210*RUNX2*. Unlike -460*RUNX2* and +542*RUNX2*, *+210RUNX2* is deeply conserved, with orthologous sequences alignable from mammals to teleosts (Fig. 3.1). The site of conserved sequences themselves was similarly preserved; all are located in the final intron of either *Runx2* or a putatively orthologous gene in more poorly annotated genomes. There are several conserved predicted transcription factor binding sites, including two adjacent inverted binding sites for Ets-related factors (containing a characteristic 5'-GGA(A/T)-3' core), a binding site for proteins containing a POU DNA binding domain, and one for *RUNX2* itself.

*Functional testing of conserved predicted transcription factor binding sites in +210*RUNX2

In order to test what, if any, function these deeply conserved sequences had with regard to the function of the enhancer, these potentially critical residues were individually and specifically ablated in new transgenic constructs (Figure 3.1) Single insert transgenic lines were constructed as in Chapter 2, and these were crossed onto fish carrying the wild type +210*RUNX2* sequence driving the expression of mCherry (+210*RUNX2*:mCh). Although only representative

microscopic photography is demonstrated below, at least three lines of each transgene have been constructed to confirm the changes in enhancer activity.

*The +210*RUNX2 RUNX2 *binding site mediates bone expression.*

Removal of the RUNX2 binding site produced an enhancer that failed to direct expression to bone at any time during the first five days post fertilization. (Figure 3.2 a-c). However, this altered enhancer was still capable of directing early expression to the branchial arches (Figure 3.2d-f).

*The +210*RUNX2 *Ets binding sites mediate branchial arch expression.*

Without the conserved Ets binding sites, +210*RUNX2* is still competent to direct expression to bony tissues, though this does appear to be less robust compared to the activity of the wild type enhancer (Figure 3.3 a-c). Possible position integration effects on the autonomy of the transgene currently confound confirming this quantitatively. More strikingly, however, this altered enhancer failed to direct expression to the branchial arches at 3DPF (Figure 3.3 d-f).

*The +210*RUNX2 *POU binding site has no confirmable effect on transgene activity.*

Eliminating the POU binding site in +210*RUNX2* did not compromise its ability to direct expression to the domains of either the branchial arches or osteoblasts (Figure 3.4a,d). eGFP expression driven by +210*RUNX2*mutPOU did appear to be more intense than that typically driven by the wildtype enhancer, suggesting that the conserved sequence might actually have a role in attenuating expression.

*+210*RUNX2 *directed expression is a direct target of FGF signaling.*

Because of the presence and functionality of two inverted Ets binding sites in +210*RUNX2*, I sought to understand what signaling pathways might be mediating gene regulation through those conserved sites. Ets transcription factors are often found to be downstream of the FGF signaling pathway [168-170] , which, in turn, has a well-appreciated role in skeletogenesis and *Runx2* regulation[k]. The small molecule inhibitor SU-5402 is a potent and selective vascular endothelial growth factor receptor (VEGFR) and fibroblast growth factor receptor (FGFR) inhibitor [171] . An initial treatment of 10uM from 24-48 HPF reduced +210*RUNX2*:eGFP mediated fluorescence in the branchial arches (Fig3.5a,b) To confirm this downregulation was a direct result of modulating FGF signaling, I treated *+210RUNX2:*eGFP transgenic embryos with the Fgf inhibitor SU5402 from 28-32 HPF. *In situ* hybridization showed specific reduction of *egfp* expression in the branchial arches (Fig. 3.5c-f), and Q-PCR confirmed a quantitative reduction in transcript levels (Fig. 3.5g). While the ETS binding sites are not absolutely required for the later activity of *+210RUNX2* in differentiated osteoblasts, activity of the enhancer was quantitatively decreased by pharmacological inhibition of FGF signaling from 100-104 HPF (Fig. 3.5h), demonstrating continued regulation of enhancer activity by the FGF pathway during osteoblast differentiation.

Results – Characterization of +542*RUNX2*

[k] Reviewed in Chapter 1

The *-460RUNX2* and *+542RUNX2* enhancers are less deeply conserved, complicating prediction of transcription factor binding sites. To provide evidence for specific regulatory inputs, I pharmacologically altered activity of candidate signaling pathways in transgenic embryos. Changes in BMP, FGF, retinoic acid, notch, hedgehog, calcineurin, and canonical Wnt signaling had no effect on *+542RUNX2* activity (data not shown).

A conserved subelement of +542RUNX2 is sufficient to direct osteoblast expression.

To better localize the essential components of +542*RUNX2*, I created transgenic lines containing the most conserved cores of the element (Figure 3.6a; Table3.1) driving eGFP expression. The more conserved of the two (MC1; Phastcons LOD = 773) directed expression to osteoblasts in a similar manner to the entire element (Figure 3.6b,c). However, bone expression was notably less robust than that driven by +542*RUNX2*. +542*RUNX2*MC2:eGFP (Phastcons = 334) expresses in the basihyal and ceratohyal cartilages (Figure 3.6d,e), components of the pharyngeal skeleton, but this does not comprise part of the expression pattern normally dictated by the intact wild type element.

*A conserved DLX binding site is necessary for +542*RUNX2 *activity.*

Alignment of the *+542RUNX2* enhancer with the orthologous sequence from chicken revealed several conserved predicted binding sites (Figure 3.7a). Initially, I hypothesized that SATB2 might be directly regulating +542*RUNX2* because of the generalized delay in bone formation observed in *Satb2*$^{-/-}$ mice [172] .

However, ablation of this sequence failed to curtail osteoblast expression directed by the enhancer (Figure 3.7b,c). Subsequently, I mutated the core of the *Dlx* binding site and observed no eGFP expression in 500 embryos injected with the resulting construct, compared to the readily observable mosaic expression regulated by the wild–type sequence (Figure 3.7d,e). Transgenic embryos from +542*RUNX2*mutDLX:eGFP founders showed no eGFP expression, despite evidence of germline transmission of the transgene (Figure 3.7h) confirming requirement of the *Dlx* binding site for enhancer activity.

Results – Characterization of -460*RUNX2*

A drug screen identifies GSK-3β as an inhibitor of -460RUNX2 mediated expression.

Treatment *-460RUNX2:egfp* fish with a small molecule inhibitor of GSK3β from 48-72 HPF resulted in a broad upregulation of eGFP expression (Figure 3.8a,b). To confirm that this effect is a direct effect of modulating the Wnt pathway, a narrower window of treatment from 52-56 HPF (which is relevant to the initiation of -460RUNX2 mediated expression in the cleithrum; Figure 2.3) demonstrated a similar pattern of upregulation (Figure 3.8c,d). GSK3β is an inhibitory component of the canonical Wnt signaling pathway, but can also function in other signaling pathways. To confirm Wnt regulation of *-460RUNX2*, we generated embryos doubly transgenic for *-460RUNX2:*mCh and the Wnt inhibitory protein Dkk-1 under control of the hsp70 promoter (*hs:Dkk1GFP*)[165] . A brief heat shock substantially reduced expression of mCherry (Figure 3.8e,f).

-460*RUNX2* requires two conserved TCF/LEF binding sites to direct expression.

Canonically, the endpoint of Wnt mediated signaling is the recruitment of members of the TCF/LEF family of transcription factors to cognate DNA binding sites[173] . There are two predicted TCF/Lef1 binding sites in the *-460RUNX2* sequence, conserved among mammals (Figure 3.9a,b), so I created a transgene in which these sites had been mutated (-*460RUNX2*mutTCF:eGFP). We observed no eGFP expression in >500 injected embryos (Figure 3.9d). -460*RUNX2*mutTCFLEF:eGFP founders showed no eGFP expression, despite evidence of germline transmission of the transgene (Figure 3.9e,f) confirming requirement of the TCF/LEF binding sites for enhancer activity.

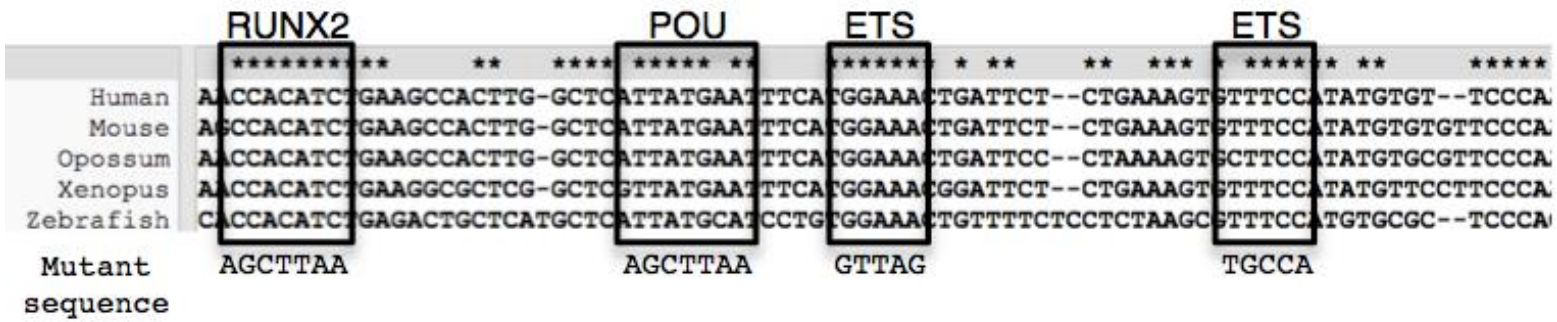

RUNX2
POU
ETS
ETS
Human
Mouse
Opossum
Xenopus
Zebrafish
Mutant sequence
AGCTTAA
AGCTTAA
GTTAG
TGCCA

Figure 3.1 -- Deep linear conservation between +210*RUNX2* and other vertebrate orthologs. Alignment to orthologous sequences from other vertebrates reveals conservation of predicted transcription factor binding sites, including RUNX2 itself, a binding site for transcription factor containing a POU domain and a pair of inverted sites for the Ets family of transcription factors.

5DPF
3DPF
a
b
c
d
e
f
mutRUNX2
+210RUNX2
merge
cl
op
cb5
ba

Figure 3.2 – The Runx2 binding site mediates +210*RUNX2* directed bone expression. (a-c) Lateral views of a doubly transgenic +210*RUNX2*mutRUNX2:eGFP; +210*RUNX2*:mCh zebrafish show no expression in any bony tissue at 5DPF driven by +210*RUNX2*mutRUNX2, while this activity is intact in +210*RUNX2*:mCh. (d-f) Ventral views of a doubly transgenic +210*RUNX2*mutRUNX2:eGFP; +210*RUNX2*:mCh zebrafish show both transgenes expressing in the branchial arches at 3 DPF. ba, branchial arches; cb5, ceratobranchial 5; cl, cleithrum; op, opercle

5DPF
3DPF
a
cl
cb5
op
mutETS
b
cl
cb5
op
+210RUNX2
c
cl
cb5
op
d
op
cl
mutETS
e
op
ba
cl
+210RUNX2
f
op
cl

Figure 3.3 – The ETS binding sites mediates branchial arch expression. (a-c) Lateral views of a doubly transgenic +210*RUNX2*mutETS:eGFP; +210*RUNX2*:mCh zebrafish show attenuated expression in bony tissue at 5DPF driven by +210*RUNX2*mutETS, while this activity is intact in +210*RUNX2*:mCh. (d-f) Dorsal views of a doubly transgenic +210*RUNX2*mutETS:eGFP; +210*RUNX2*:mCh zebrafish show failure of +210*RUNX2*mutETS, to express in branchial arches at 3DPF, while this activity is intact in +210*RUNX2*:mCh. ba, branchial arches; cb5, ceratobranchial 5; cl, cleithrum; op, opercle

5DPF
3DPF
a
cl
cb5
op
mutPOU
b
cl
cb5
op
+210RUNX2
c
cl
cb5
op
d
op
ba
mutPOU
e
op
ba
+210RUNX2
f
op
ba

Figure 3.4 – The POU binding site is not essential for +210*RUNX2* activity during embryogenesis. (a-c) Lateral views of a doubly transgenic +210*RUNX2*mutPOU:eGFP; +210*RUNX2*:mCh zebrafish demonstrate coexpression in bony tissue at 5DPF by both transgenes. (d-f) Dorsal views of a doubly transgenic +210*RUNX2*mutPOU:eGFP; +210*RUNX2*:mCh zebrafish show coexpression in branchial arches at 3DPF by both transgenes. ba, branchial arches; cb5, ceratobranchial 5; cl, cleithrum; op, opercle

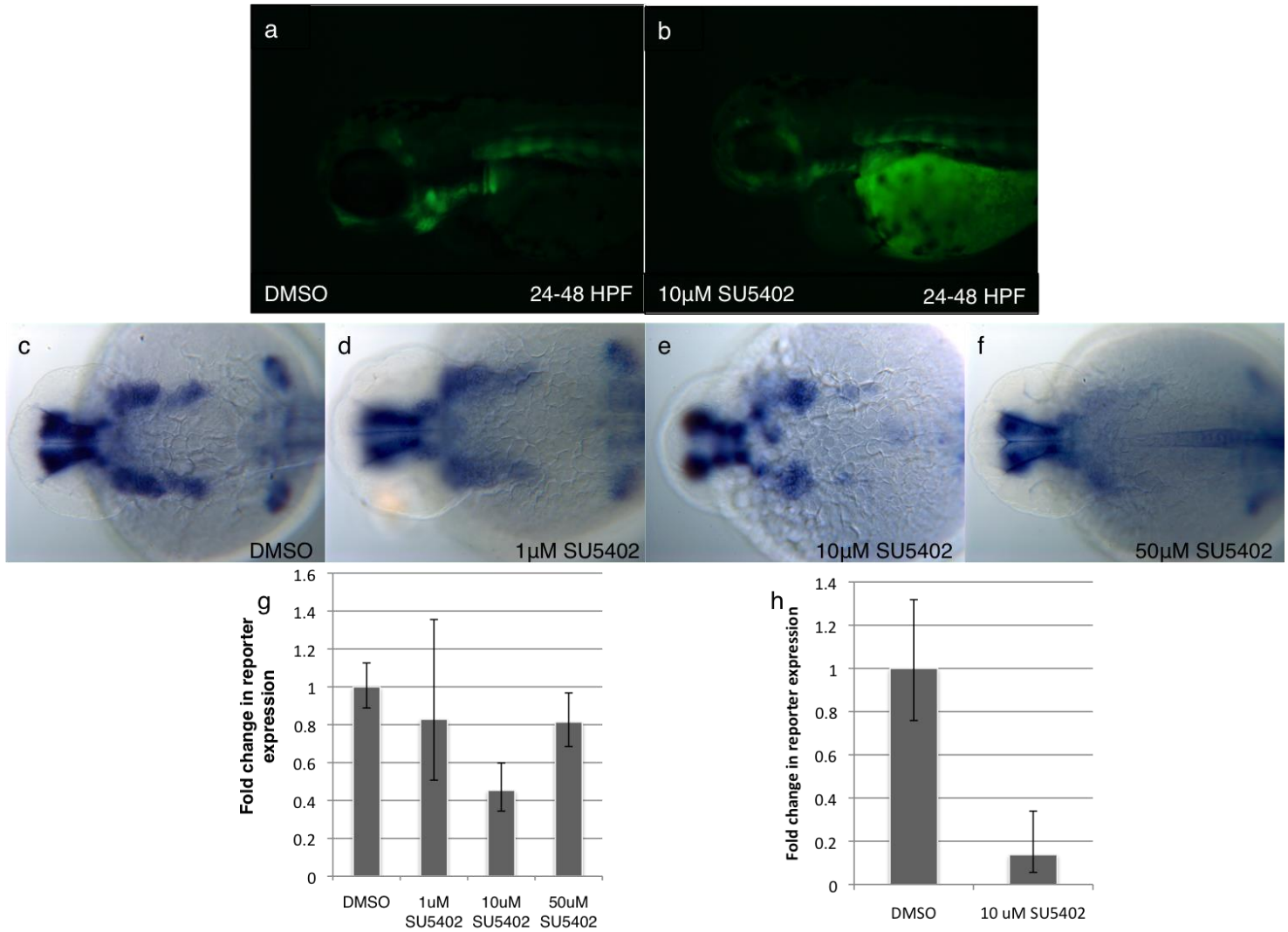

a
DMSO
24-48 HPF
b
10µM SU5402
24-48 HPF
c
DMSO
d
1µM SU5402
e
10µM SU5402
f
50µM SU5402
g
Fold change in reporter expression
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
DMSO
1uM SU5402
10uM SU5402
50uM SU5402
h
Fold change in reporter expression
1.4
1.2
1
0.8
0.6
0.4
0.2
0
DMSO
10 uM SU5402

Figure 3.5 – +210*RUNX2* is regulated by the FGF signaling pathway. a,b) Treatment of *+210RUNX2:egfp* transgenics with the FGF inhibitor SU5042 from 24-48 HPF resulted in loss of transgene expression in the branchial arches. c-f) As show shown by *in situ* hybridization for *egfp*, a narrow window of treatment from 28-32 HPF this loss of expression is specific and direct. g,h) Q-PCR confirmed a quantitative decrease in reporter gene expression following treatment for the same interval, and similarly following a later treatment from 100-104 HPF. Views in a and b are lateral, and in c-f, dorsal, with anterior to the left.

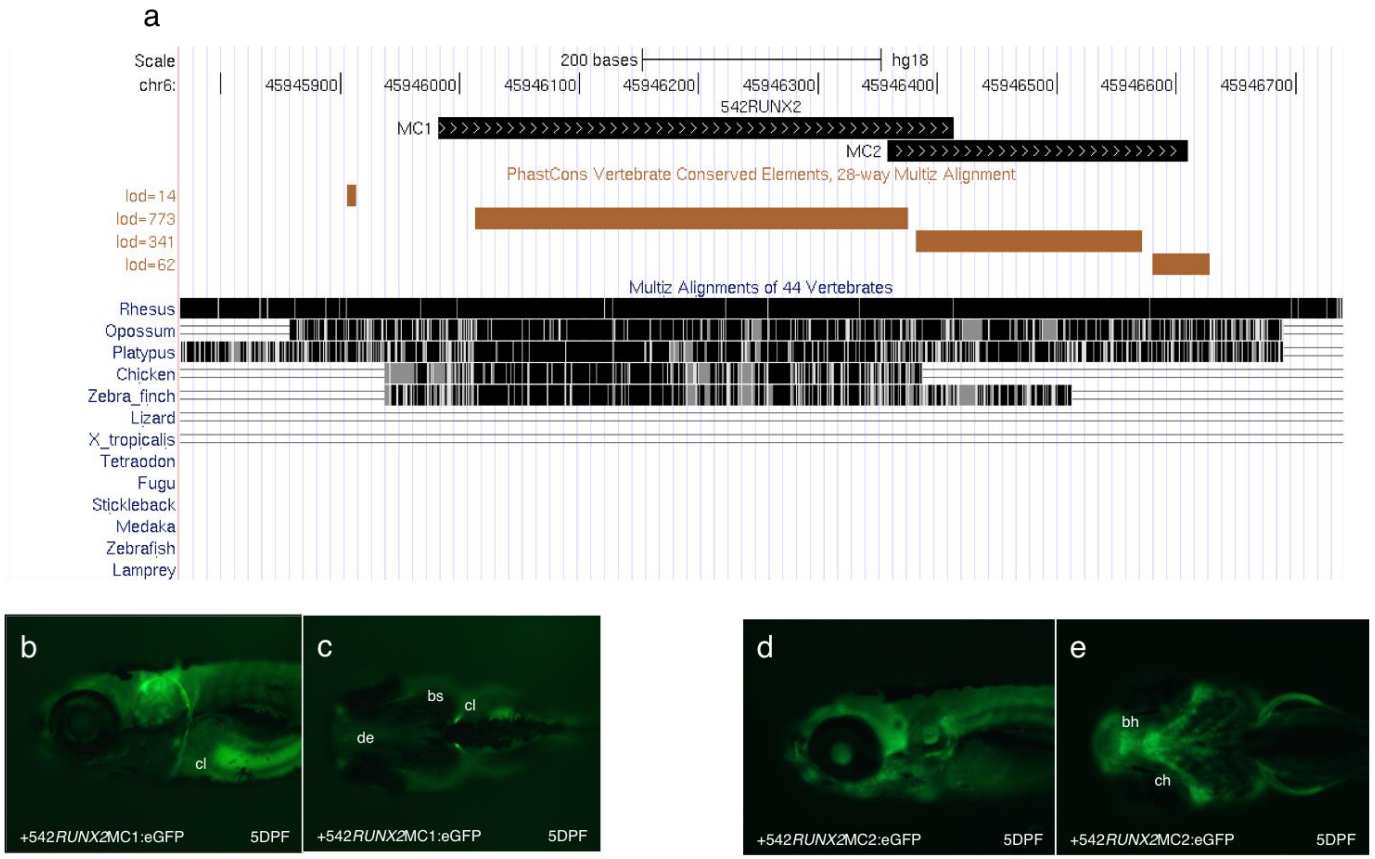

a
Scale
chr6:
200 bases
hg18
45945900
45946000
45946100
45946200
45946300
45946400
45946500
45946600
45946700
542RUNX2
MC1
MC2
PhastCons Vertebrate Conserved Elements, 28-way Multiz Alignment
lod=14
lod=773
lod=341
lod=62
Multiz Alignments of 44 Vertebrates
Rhesus
Opossum
Platypus
Chicken
Zebra_finch
Lizard
X_tropicalis
Tetraodon
Fugu
Stickleback
Medaka
Zebrafish
Lamprey
b
cl
+542RUNX2MC1:eGFP
5DPF
c
bs
cl
de
+542RUNX2MC1:eGFP
5DPF
d
+542RUNX2MC2:eGFP
5DPF
e
bh
ch
+542RUNX2MC2:eGFP
5DPF

Figure 3.6 – Subcloning of +542RUNX2 localizes osteoblast activity to a 433 bp fragment. a) Genome browser view of the original +542RUNX2 element with tracks indicating the location of the amplicons for subcloning (black) and Phastcon elements upon which those amplicons were designed (brown). b,d) Lateral views of 5DPF transgenic embryos carrying +542*RUNX2*MC1:eGFP (b) or (d) +542*RUNX2*MC2:eGFP. c,e) Ventral views of 5DPF transgenic embryos carrying +542*RUNX2*MC1:eGFP (c) or (e) +542*RUNX2*MC2:eGFP. ba, branchial arches; bh, basihyal; bs, brachiostegal ray; ch, ceratohyal; cl, cleithrum; de, dentary; op, opercle;

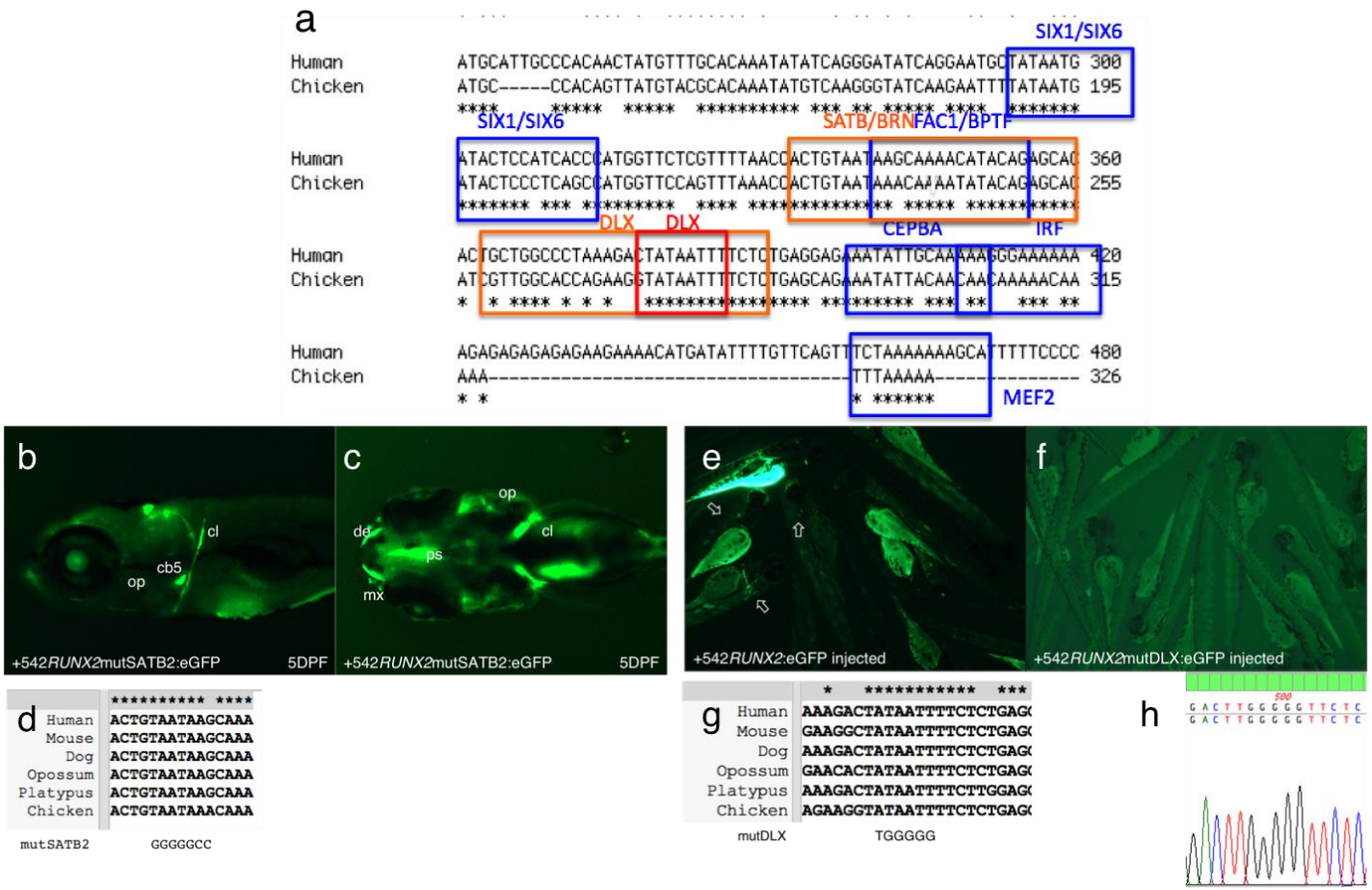

a
SIX1/SIX6
Human ATGCATTGCCCACAACTATGTTTGCACAAATATATCAGGGATATCAGGAATGCTATAATG 300
Chicken ATGC-----CCACAGTTATGTACGCACAAATATGTCAAGGGTATCAAGAATTTTATAATG 195
SIX1/SIX6
SATB/BRN
FAC1/BPTF
Human ATACTCCATCACCCATGGTTCTCGTTTTAACCACTGTAATAAGCAAAACATACAGAGCAC 360
Chicken ATACTCCCTCAGCCATGGTTCCAGTTTAAACCACTGTAATAAACAAAATATACAGAGCAC 255
DLX
DLX
CEPBA
IRF
Human ACTGCTGGCCCTAAAGACTATAATTTTCTCTGAGGAGAAATATTGCAAAAAGGGAAAAAA 420
Chicken ATCGTTGGCACCAGAAGGTATAATTTTCTCTGAGCAGAAATATTACAACAACAAAAACAA 315
Human AGAGAGAGAGAGAAGAAAACATGATATTTTGTTCAGTTTCTAAAAAAAGCATTTTTCCCC 480
Chicken AAA-----TTTAAAAA---------- 326
MEF2
b
cl
cb5
op
+542RUNX2mutSATB2:eGFP 5DPF
c
op
de
cl
ps
mx
+542RUNX2mutSATB2:eGFP 5DPF
e
+542RUNX2:eGFP injected
f
+542RUNX2mutDLX:eGFP injected
d
Human ACTGTAATAAGCAAA
Mouse ACTGTAATAAGCAAA
Dog ACTGTAATAAGCAAA
Opossum ACTGTAATAAGCAAA
Platypus ACTGTAATAAGCAAA
Chicken ACTGTAATAAACAAA
mutSATB2 GGGGGCC
g
Human AAAGACTATAATTTTCTCTGAG
Mouse GAAGGCTATAATTTTCTCTGAG
Dog AAAGACTATAATTTTCTCTGAG
Opossum GAACACTATAATTTTCTCTGAG
Platypus AAAGACTATAATTTTCTTGGAG
Chicken AGAAGGTATAATTTTCTCTGAG
mutDLX TGGGGG
h

Figure 3.7 – Identification and functional testing of conserved transcription factor binding sites in +542*RUNX2*A) Alignment of +542*RUNX2* and +312*RUNX2*(gg) along with transcription factor binding sites identified in both by Transfac (blue), Genomatix (orange), or Uniprobe (red) b,c) Representative views of eGFP expression pattern in +542RUNX2mutSATB2:eGFP at 5DPF. d) Alignment of a predicted SATB2 binding site to other vertebrate orthologs. Sequence of mutagenized construct is shown below. e) Following injection of *+542RUNX2:egfp*, mosaic expression in bones, including the cleithrum of two different embryos (arrows) is readily apparent. f) In contrast, >500 embryos injected with +542*RUNX2*mut*Dlx*:eGFP showed no mosaic expression. g) Alignment of a predicted *Dlx* binding site to other vertebrate orthologs. Sequence of mutagenized construct is shown below. h) Presence of +542*RUNX2*mutDLX:eGFP in non-expressing progeny of injected founder was confirmed by PCR and sequencing of transgene. cb5; ceratobranchial 5; cl, cleithrum; de, dentary; mx, maxilla; op, opercle.

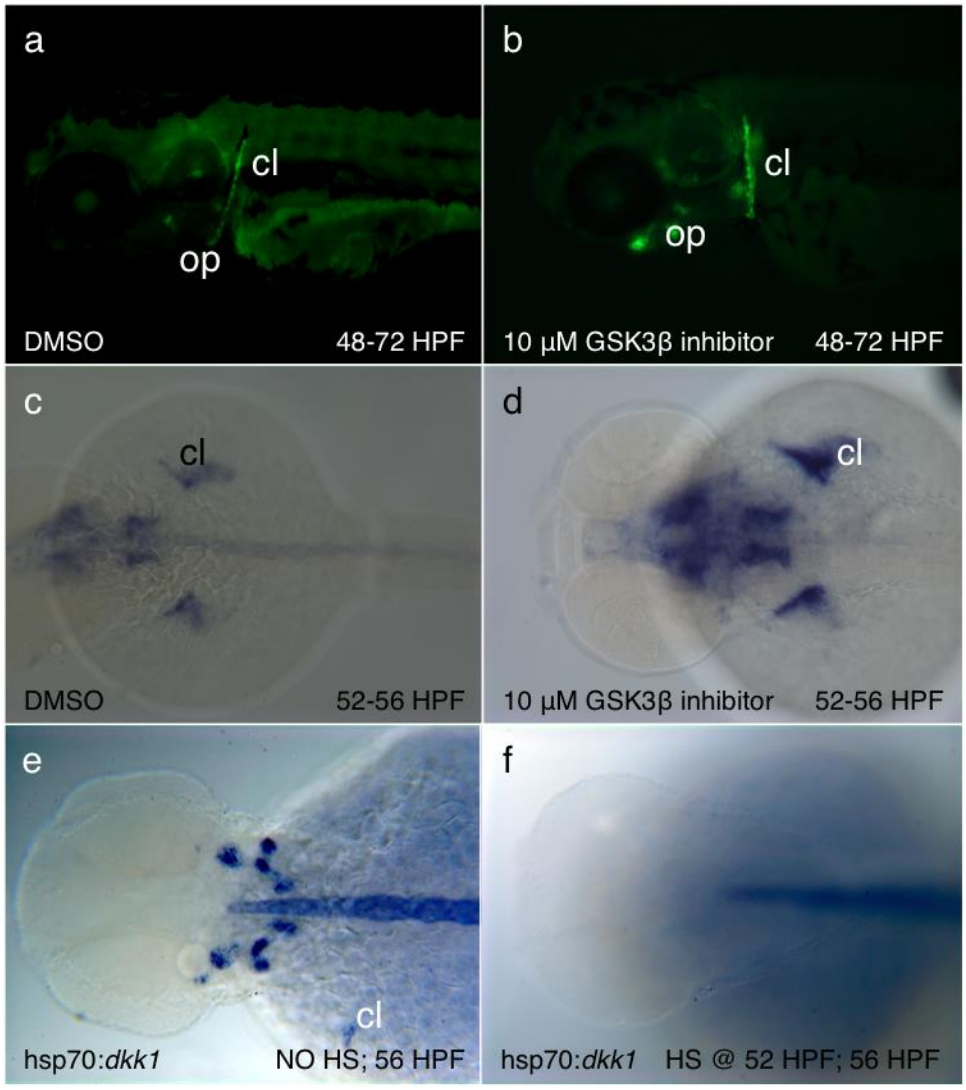

a
cl
op
DMSO
48-72 HPF
b
cl
op
10 µM GSK3β inhibitor
48-72 HPF
c
cl
DMSO
52-56 HPF
d
cl
10 µM GSK3β inhibitor
52-56 HPF
e
cl
hsp70:dkk1
NO HS; 56 HPF
f
hsp70:dkk1
HS @ 52 HPF; 56 HPF

Figure 3.8 – -460*RUNX2* mediated expression is responsive to Wnt mediated signaling. Compared to control (a), activity of *-460RUNX2* is increased by treatment from 48-72 HPF with GSK3b inhibitor XV. (b) A narrower window of treatment (52-56HPF) shows rapid upregulation (d) relative to control (c). -460*RUNX2* mediated expression (e) is abolished by ectopic expression of the Wnt inhibitor *dkk1* by heat shock (f). cl, cleithrum; op, opercle

a

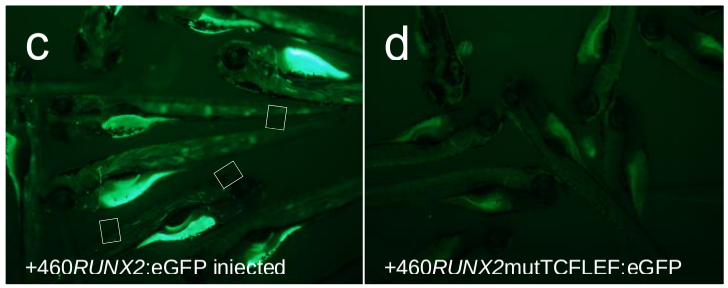
```
                 * ******
Human            AGAACTTTGATTTTC
Chimp            AGAACTTTGATTTTC
Orangutan        AGAACTTTGATTTTC
Rhesus           AGAACTTTGATTTTC
Marmoset         AGAACTTTGATTTTC
Mouse            AGAACTTTGATTTTC
Rat              AGAACTTTGATTTTC
Kangaroo_rat     GGAACTTTGATTTTC
Rabbit           GGAACTTTGATTTTC
Cow              GTAACTTTGATGTCC
Armadillo        GGAACTTTGATTTCC
Shrew            GGAACTTTGATTTCC
Cat              GGAACTTTGATTTCC
Dog              GGAACTTTGATTCTC
Microbat         GGAACTTTGATTTCC
Elephant         GGAACTTTGATTTCC
Megabat          GAAACTTTGAATTTT
Sloth            GGAACTTTGA-TTTT
Horse            TGAACTTTGATTTTC
Guinea_pig       GGAACTTTGATTTCC
Squirrel         GGAACTTTGATTTCC
Bushbaby         GGAACTTTGATTTTT
Platypus         GGATCTTTGATTTTT
Hedgehog         AAAACTTTGATTTTT
Opossum          GGATCTTTGATTCTT
Rock_hyrax       GGAACTTTGATTTTT
```

mutTCF

b

```
                 *******
Human            TGCTTTGAAGT
Chimp            TGCTTTGAAGT
Orangutan        GGCTTTGAAGT
Marmoset         TGCTTTGAAGT
Tarsier          TGCTTTGAAGT
Mouse_lemur      TGCTTTGAAGT
Guinea_pig       TGCTTTGAAGT
Kangaroo_rat     TGCTTTGAAGT
Mouse            TGCTTTGAAGT
Sloth            TGCTTTGAAGT
Rat              TGCTTTGAAGT
Squirrel         TGCTTTGAAGT
Rock_hyrax       TGCTTTGAAGT
Rabbit           TGCTTTGAAGT
Pika             TGCTTTGAAGT
Dolphin          TGCTTTGAAGT
Megabat          TGCTTTGAAGT
Cat              TGCTTTGAAGT
Dog              TGCTTTGAAGT
Cow              TGCTTTGAAGT
Shrew            TGCTTTGAAGT
Hedgehog         TTCTTTGAAGT
Opossum          TGCTTTGAAGT
Platypus         TGCTTTGAACC
```

mutTCF



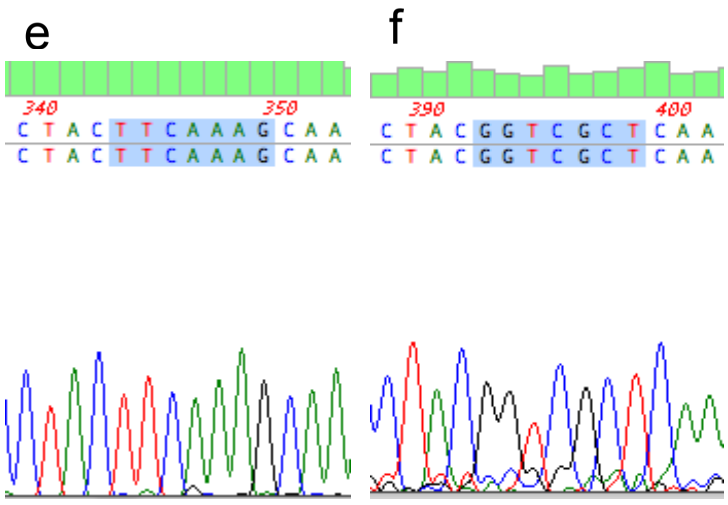


Figure 3.9 – -460*RUNX2* regulatory competency requires two conserved TCF/LEF binding sites. a,b) Alignment of vertebrate genomes to conserved TCF/LEF binding sites (a) at chr6:44,835,639-44,835,645(hg19), and b) chr6:44,836,005-44,836,012(hg19) in -460*RUNX2*. Residues that have been changed in the mutagenized transgene are indicated below the alignment. c) Following injection of -460*RUNX2:*eGFP, mosaic expression in bone is readily apparent at 5 DPF, seen in the cleithra of multiple injected embryos (arrows). d) In contrast, following injection of -*460RUNX2*mutTCFLEF:eGFP, no expression was seen in >500 embryos. PCR and sequencing of progeny from injected fish confirmed the presence of the transgene with ablated TCF/LEF binding sites (f), compared to sequencing of the transgenics with wildtype sequence (e).

Name	F primer	R primer	Coordinates (hg18)	Size (bp)
+542*RUNX2* MC 1	AGACAACACGGGCTCATCGT	CCCCAAGGGTCTCTGGATTT	chr6:45,945,982-45,946,414	433
+542*RUNX2* MC 2	CTGGGATGGCCAGAGAGAGG	TGGCTTCGATATGCCTCTAGTGTA	chr6:45,946,358-45,946,610	253

Table 3.1. Sequences of +542*RUNX2* tested for enhancer activity. Both elements are wholly located within +542RUNX2. The primer sequences are those used to amplify the elements from genomic DNA for testing.

CHAPTER 4

DISCUSSION, SUMMARY AND FUTURE DIRECTIONS

CHAPTER 4

Summary of human RUNX2 *associated enhancers*

Runx2 is the common denominator in osteoblast development throughout the skeleton, and its levels of expression are critical. As a crucial step in understanding the regulation of the gene, and subsequently the skeleton itself, I have identified distant osteoblast specific enhancers associated with *RUNX2* and characterized signaling pathways acting on them. Despite their common feature of directing osteoblast expression, they are strikingly diverse. They are widely spaced across the locus (Figure 2.1), have no obvious sequence similarity to each other, and are conserved across species to varying degrees. While they all direct expression to osteoblasts, they do so with differing spatiotemporal dynamics. The cleithrum is the first bone to ossify in the zebrafish skeleton and does so intramembraneously. Using it as a proxy for the relative timing of expression onset yields a sequence of +542*RUNX2* → +210*RUNX2* → -460*RUNX2* (Figure 2.3). Whether this is consistent across all bony structures is unclear; it appears that the ability to drive expression to the vertebral arches is exclusively a property of +542*RUNX2* (Figure 2.6).

In addition to expression at the resolution of individual bones, subpopulations of osteoblasts express the three transgenes differentially. As is evidenced by study of opercle development in these transgenic lines (Figure 2.4) these enhancers direct expression to different cells within that structure. +542*RUNX2:*eGFP expression is uniform throughout the opercle and cleithrum

(Figure 2.3). Combining these observations with its unique expression in the vertebral arches and early expression in the cleithrum, it is intriguing to speculate that +542*RUNX2*-mediated expression is the most 'fundamental' of the three characterized enhancers and tied primarily to osteoblast identity itself. The expression pattern directed by +210*RUNX2* with respect to the developing opercle is similar to that of the *fli1*:eGFP transgenic line, which labels all neural-crest derived mesenchyme[161] and may indicate cells that have recently become RUNX2+. Finally, -460*RUNX2* expression is relatively strongest in the strut and the leading edge in the fan structure of the bone. These cells are also *sp7*:mCh positive, indicating their likely active deposition of bone ECM components. The opercle fan structure expands via a banding pattern[161] , so -460*RUNX2* positive cells may be those that have committed to remaining in a specific iteration of that process.

Finally, these three CREs do not fit the definition of redundant 'shadow' enhancers[174] that reinforce a response to a single inductive event and ensure transcriptional robustness to environmental variability. Rather, they appear to integrate inputs from different signaling pathways to induce or maintain *Runx2* expression (Figure 4.1). This observation parallels and complements the diversity of signaling inputs capable of accelerating *Runx2* expression and osteogenic differentiation in the literature.

Summary of human RUNX2 *associated enhancer activity -- +210*RUNX2 *directs expression to osteoblasts separably through FGF signaling and Runx2 autoregulation.*

That +210*RUNX2* shares orthology both in sequence and function with an element similarly placed in the last intron of the zebrafish *Runx2* ortholog *runx2a* (+154*runx2a*;Figure 2.5) suggests its role in regulating *Runx2* activity is ancient, and consequential. It also possesses a modularity competent to respond to FGF signaling to direct expression to the branchial arches (essentially) or osteoblasts (qualitatively), while also possessing a conserved binding site for RUNX2 required to direct bone expression. The involvement for FGF signaling during osteoblast differentiation generally has been discussed in Chapter 1, so it is not surprising that +210RUNX2 directs expression to the calvarial sutures that are so sensitive to that signaling axis[l].

+210*RUNX2* also presumably functions as a site for positive autoregulation of *Runx2* activity. In diverse biological systems, the existence of a positive feedback loop is an essential step in the creation of switches with an all-or-none 'digital' output characteristic[175] . And where better to place a switch incapable of nuance than at a gatekeeper gene whose expression above threshold is sufficient to completely alter cell fate? Whether the +210*Runx2* response to RUNX2 is a required step in the commitment of MSCs or chondrocytes to an osteoblastic fate is unclear. There may be other unannotated positive and negative feedback loops

[l] Kague, E. Unpublished observation

involving recruitment of RUNX2 to target CREs, to alter cell fate kinetics. Analysis of the rat and mouse *Runx2* P1 promoters indicated that RUNX2 binding to the 5' UTR coding region of *Runx2* was capable of *suppressing* transcription *in vitro*[176] . However, a reasonable hypothesis is that +210*RUNX2* functions to 'lock in' a cell to an osteoblastic fate commitment, due to its inability to direct bone-specific expression without a conserved binding site.

Summary of human RUNX2 *associated enhancers -- +542*RUNX2 *directs expression to early osteoblasts*.

Dissection of the activity of +542RUNX2 was focused on the two most biologically likely conserved direct upstream regulators. SATB2 is a nuclear matrix attachment protein that also functions as a transcription factor. *Satb2*$^{-/-}$ mice have generalized osteoblast differentiation delays as well as craniofacial patterning defects. Deletion of the predicted SATB2 binding site from +542*RUNX2* did not affect the ability of the enhancer to direct expression to osteoblasts. Although, SATB2 binds the promoters of and upregulates bone marker genes, it is also hypothesized to act as a negative regulator of *Hoxa2* expression during osteoblast differentiation. Whether it might be executing a similar role with regards to regulation of +542*Runx2* is unclear with respect to current experimental evidence.

Members of the *Dlx* family are dynamically expressed during osteoblast maturation, suggesting roles in different aspects of this process[177] . During skull formation in chick, *Dlx5* is expressed in osteoblast progenitors, specifically in

response to BMP but not FGF signaling, and its expression activates *Runx2* and osteogenic differentiation in uncommitted embryonic calvarial mesenchyme[78]. Zebrafish *dlx5a* is expressed in the cleithrum at least as early as the long pec stage (~42 HPF)[178], consistent both with early expression of *runx2b* and early activity of *+542RUNX2*. Coexpression of *bmp2a*[178] and *bmp2b* early in the cleithrum is also consistent with a BMP->DLX->RUNX2 signaling axis in these cells. However, +542RUNX2 did not demonstrate response to perturbation by induction of *bmp2b* or *chd* via heat shock (data not shown), confounding the impulse to arrive at such a conclusion. While the ablated binding site was identified by multiple algorithms as one similar to others capable of recruiting DLX proteins, transcription factor binding site profiles are famously degenerate, and so it is reasonable that other homeodomain containing proteins could be signaling through +542*RUNX2*.

Summary of human RUNX2 *associated enhancers -460*RUNX2 *potentially links Wnt signaling, Runx2 regulation and variation in common skeletal phenotypes and diseases.*

Although experimental and clinical data indicate that gross aberrations in *Runx2* expression cause skeletal disorders, smaller individual variations in *Runx2* dosage might be responsible for differences in variation of non-pathologic skeletal phenotypes or susceptibility to disease. A cluster of SNPs associated with skeletal conditions (bone mass density (BMD) and osteoarthritis (OA), and height in three different populations) by genome wide association studies cluster

around the Wnt responsive enhancer -460*RUNX2* (Figure 4.3; Table4.1) No other SNPs associated with skeletal phenotypes are located in the remainder of the *RUNX2* locus. Wnt signaling has been well implicated in affecting BMD, although the precise mechanism of that effect is not clear in the literature. Some evidence suggests the effect of Wnt signaling on bone mass is indirect, mediated by serotonin secretion by neuroendocrine cells of the gut [179] , although this has been disputed [180] . The presented data strongly support a direct role for Wnt signaling in osteoblasts, acting via transcriptional regulation of *Runx2*.

The location of *-460RUNX2* suggests that variations in the enhancer itself alter the risk of low BMD and OA and influence height through changes in *RUNX2* expression. Interestingly, in addition to its positive association with BMD, the canonical Wnt pathway has been implicated in increased osteoarthritis risk [181, 182] , as has increased *RUNX2* expression [183, 184] Therefore, sequence variants in the population may affect either the basal activity of -460*RUNX2* or its responsiveness to Wnt signaling, accounting for the genetic associations with both of these skeletal phenotypes. An intriguing possibility is that two alleles at a single location could lead either to increased enhancer activity and increased arthritis risk, or decreased enhancer activity and increased risk of osteoporosis.

Runx2 expression modulation as a source of evolutionary skeletal diversity.

Runx2 protein activity is positively correlated to facial length in carnivores, especially domesticated dogs [185] . This relationship is not generally true among

placental mammals, suggesting that other changes, such as in gene expression levels, are more likely to correlate with intra– or inter-specific variation[186] . The sensitivity of normal skeletal development to precise levels of Runx2 has led to the suggestion that alterations in Runx2 activity provide a mechanism for skeletal evolution, acting as a 'tuning knob' to either accelerate or delay osteoblast differentiation during development[186] . Following assembly of the Neanderthal genome sequence, the *RUNX2* locus was identified as one of the regions with the strongest evidence of positive selection in the evolution of modern humans[187] . Specifically, the 3' end of *RUNX2*, encompassing +*210RUNX2,* shows a deficit of derived alleles in modern humans (Figure 4.3a). No fixed differences in the *RUNX2* coding sequence are present between Neanderthal and modern humans, suggesting that the positive selection has acted on changes in regulatory sequences. Comparing the human +210*RUNX2* sequence to other primates identifies three derived, human specific SNPs that could potentially link this element to the evolution of the human skeleton (Figure 4.3b). Interestingly, many of the differences between the skeletons of Neanderthal and modern humans—clavicular morphology, frontal bossing of the skull—are similar to the differences observed in cleidocranial dysplasia, which is caused by a *Runx2* gene dosage defect[127, 188, 189] .

Future directions

Future investigation of the functional consequences of specific sequence alterations in the *RUNX2* enhancers will shed light on the role of its regulation in development, evolution, and disease. An obvious question resulting from the screen in Chapter 2 is the thoroughness of it. While it is likely impossible to ever know the cis regulatory architecture of single gene in a complex eukaryotic genome has been exhaustively annotated, some potential experiments in other model systems present themselves to address this question. However, this will require divesting ourselves of the zebrafish model. While it has been shown to be an effective system for evaluating the regulatory potential of discrete elements in the human genome, the existence of two *runx2* genes as well as the current size limitations of BAC-mediated transgenesis make the fish a poor choice to study the intact human *Runx2* locus. To try to get a broader locus-wide view of *Runx2* enhancer dynamics during development, we must turn turn to a system with a more similar *Runx2* structure, namely the mouse.

Although not discussed in this document, it is not difficult to obtain a population of cells uniformly positive for the *Runx2* transgene from early embryos. This process necessitates enzymatic digestion followed by flow-cytometry sorting to derive an enriched population of transgene expressing osteoblasts at a relatively discrete stage of development. Sorting based on multiple colors/transgenes can further refine this staging. Creating one or more mouse transgenic lines using either the +542*RUNX2* or -460*RUNX2* elements

would, assuming expression patterns are similar in the mouse, permit exploration for bone specific enhancers in early osteoblasts. This could be done by looking for enrichment for histone modifications associated with regulatory sequences, or for regions directly associated with RUNX2 itself. Circular chromosome conformation capture (4C) allows to us to ask questions specifically about the dynamics of the Runx2 locus itself. Using one or both Runx2 promoters as 'bait', comparing the physical interaction of distal elements with the Runx2 TSS would presumably yield a list of currently unknown cis regulators, as well as informing how they and the currently known Runx2 enhancers function dynamically through osteoblast development[189].

Additionally, as the number of and knowledge regarding individual RUNX2 cis regulators grows, the genetic tools they offer might be applicable to studying bone biology in other contexts than embryonic differentiation. +210*RUNX2* transgenic fish have already been used in a to study bone regeneration post-amputation[190] as well as suture development in the skull vault[m]. Certainly, how bones heal post-fracture is a robust area of research[191], and the ability to visualize *Runx2* expression during in vivo assays would presumably augment them. Finally, the transgenic lines could be incorporate in a high throughput screen against a pharmaceutical library[192], enhancing drug discovery for *Runx2* expression mediators that may aid in therapies for common skeletal disorders such as OA and osteoporosis.

[m] E. Kague, unpublished

Conclusion

Taken together, the data presented in this document provide evidence for direct regulation of *Runx2* transcription by biologically important signaling pathways and transcription factors through three independent enhancers. This complex regulatory landscape has allowed the fine–tuning of expression of this critical developmental gene through alterations in enhancer activities. Furthermore, I hypothesize that these alterations have been selected for in evolution, and help account for differences in skeletal morphology among species. This data also supports the model that variations in a distal enhancer of *RUNX2* account for genetic associations in the region with height, BMD, and increased OA risk.

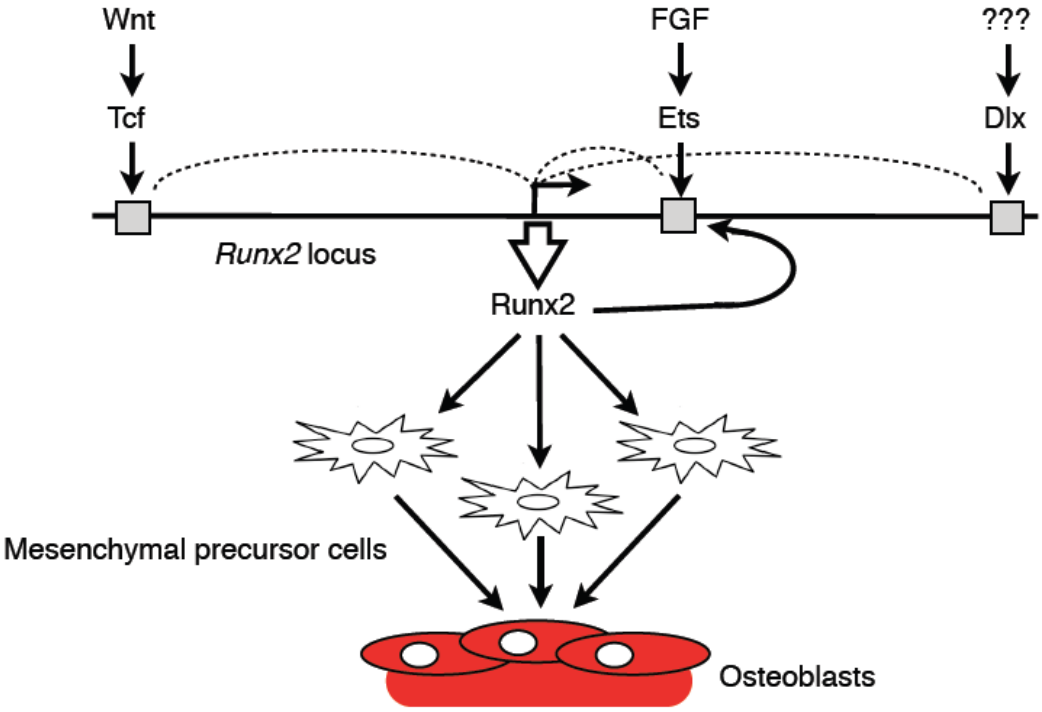

Wnt
FGF
???
Tcf
Ets
Dlx
Runx2 locus
Runx2
Mesenchymal precursor cells
Osteoblasts

Figure 4.1 -- Model for integration of multiple signaling inputs at the *Runx2* locus. Three identified enhancers at the *Runx2* locus are regulated by different upstream factors, and each is capable of interacting independently with the transcriptional start site (dotted lines) to activate gene transcription. Once transcription is activated via one or more external signals, it is stabilized by Runx2 auto-regulation through the intronic enhancer. Downstream, expression of Runx2 in mesenchymal precursor cells of diverse embryologic origins leads to activation of genes necessary for development of osteoblasts.

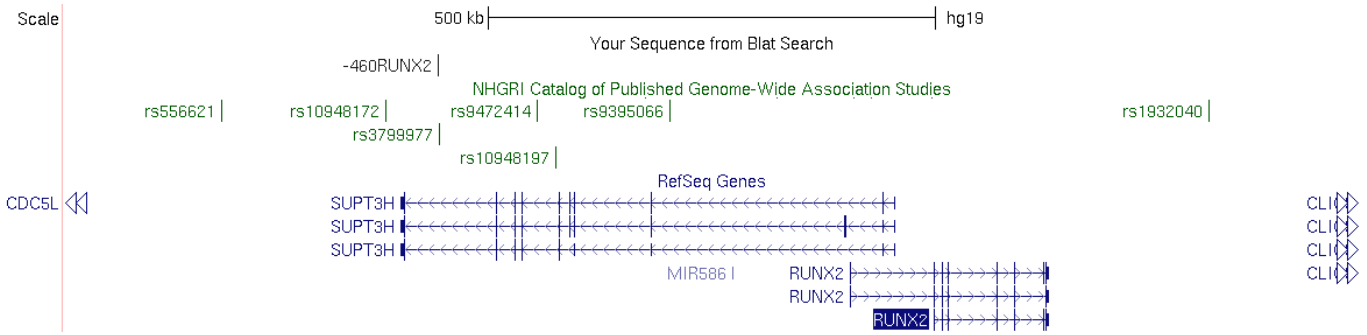
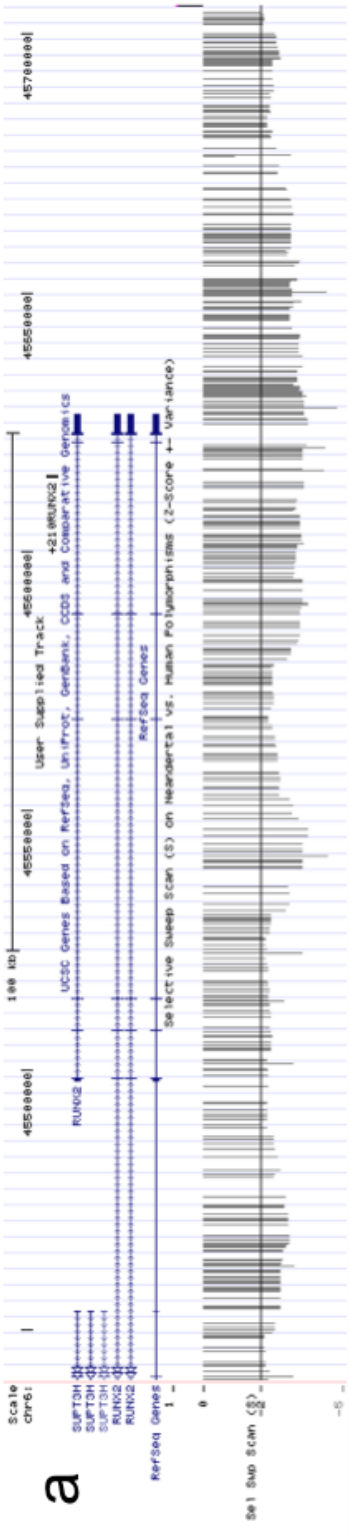

Scale
500 kb
hg19
Your Sequence from Blat Search
-460RUNX2
NHGRI Catalog of Published Genome-Wide Association Studies
rs556621
rs10948172
rs9472414
rs9395066
rs1932040
rs3799977
rs10948197
RefSeq Genes
CDC5L
SUPT3H
SUPT3H
SUPT3H
MIR586 I
RUNX2
RUNX2
RUNX2
CLI
CLI
CLI
CLI

Figure 4.2 – SNPs associated with skeletal phenotypes and disorders cluster near the Wnt responsive enhancer -460*RUNX2*. A genome browser view of the human *RUNX2* locus interrogated for regulatory activity shows the location of all SNPs associated with a human phenotype by genome wide association studies. SNPs associated with skeletal phenotypes are listed in Table 4.1; rs1932040 is associated with attention deficit hyperactivity disorder.


a
Scale
chr6:
100 kb
45540000
45560000
45580000
45600000
45620000
45640000
UCSC Genes Based on RefSeq, UniProt, GenBank, CCDS and Comparative Genomics
User Supplied Track
+21@RUNX2
RefSeq Genes
Selective Sweep Scan (S) on Neanderthal vs. Human Polymorphisms (Z-Score +/- Variance)
RUNX2
SUPT3H
RefSeq Genes
Sel Swp Scan (S)

b


+210runx2 -AATTTGGCTCCATGTTTTGGTAGCCATAAGTCATTAGTGGTTTTTAATTTTTCTGAAAC
Chimp AATTGTGGCTCCATGTTTTGGTAGCCATAAGTCATTAGTGGTTTTTAATTTTTCTGAAAC
Orang AATTGTGGCTCCATGTTTTGGTAGCCATAAGTCATTAGTGGTTTTTAATTTTTCTGAAAC
Gibbon AATTGTGGCTCCATGTTTTGGTAGCCATAAGTCATTAGTGGTTTTTAATTTTTCTGAAAC
Rhesus --------CTCCATGTTTTGGTAGCCATAAGTCATTAGTGGTTTTTAATTTTTCTGAAAC
+210runx2 TGAAATAAGGTCTGAAAGGGTGGATTTGTCCCTTTTTTCCTCCTTT----------TT
Chimp TGAAATAAGGTCTGAAAGGGTGGATTTGTCCCTTTTTTCCTCCTTT----------TT
Orang TGAAATAAGGTCTGAAAGGGTGGATTTGTCCCTTTTTTCCTCCTTT----------TT
Gibbon TGAAATAAGGTCTGAAAGGGTGGATTTGTCCCTTTTTTCCTCCTTT----------TT
Rhesus TGAAATAAGGTCTGAAAGGGTGGATTTGTCCCTTTTTTTCCTCCTTTTTTCCTTTTTTT
+210runx2 TCCTTTTT--TTTTTTGCTGGCTAATCGGAGGAACGCAACCACATCTGAAGCCACTTGG
Chimp TCCTTTA---TTTTTTGCTGGCTAATCGGAGGAACGCAACCACATCTGAAGCCACTTGG
Orang TCCTTTTTTTTTTTTTTGCTGGCTAATCGAAGGAACGCAACCACATCTGAAGCCACTTGG
Gibbon TCCTTTT---TTTTTTGCTGGCTAATCGGAGGAACGCAACCACATCTGAAGCCACTTGG
Rhesus TTTTTTTTTTTTTTTTTGCTGGCTAATCGGAGGAACGCAACCACATCTGAAGCCACTTGG
+210runx2 CTCATTATGAATTTCATGGAAACTGATTCTCTGAAAGTGTTTCCATATGTGTTCCCAATA
Chimp CTCATTATGAATTTCATGGAAACTGATTCTCTGAAAGTGTTTCCATATGTGTTCCCAATA
Orang CTCATTATGAATTTCATGGAAACTGATTCTCTGAAAGTGTTTCCATATGTGTTCCCAATA
Gibbon CTCATTATGAATTTCATGGAAACTGATTCTCTGAAAGTGTTTCCATATGTGTTCCCAATA
Rhesus CTCATTATGAATTTCATGGAAACTGATTCTCTGAAAGTGTTTCCATATGTGTTCCCAATA
+210runx2 TATCCCATCCTCTCTAAATGGGAGCACAGACGTGTGCATGGGCTGGAAATGGCTTCAGGA
Chimp TATCCCATCCTCTCTAAATGGGAGCACAGACGTGTGCATGGGCTGGAAATGGCTTCAGGA
Orang TATCCCATCCTCTCTAAATGGGAGCACAGACGTGTGCATGGGCTGGAAATGGCTTCAGGA
Gibbon TATCCCATCCTCTCTAAATGGGAGCACAGACGTGTGCATGGGCTGGAAATGGCTTCAGGA
Rhesus TATCCCATCCTCTCTAAATGGGAGCACAGACGTGTGCATGGGCTGGAAATGGCTTCAGGA
+210runx2 TGTGACAGGGCAGTCATATGACTGCCACGGCACCGTGGGTGGCCCCTTCTGAAAGACACG
Chimp TGTGACAGGGCAGTCATATGACCGCCACGGCACCGTGGGTGGCCCCTTCTGAAAGACACG
Orang TGTGACAGGGCAGTCATATGACCGCCACGGCACTGTGGGTGGCCCCTTCTGAAAGACACG
Gibbon TGTGACAGGGCAGTCACATGACTGCCACGGCACCGTGGGTGGCCCCTTCTGAAAGACACG
Rhesus TGTGACAGGGCAGTCATATGACTGCCACGGCACCGTGGGTGGCCCCTTCTGAAAGACACG
+210runx2 GTGTCCTGGGTGCTGGCCTCCTTCCTGCAGCGGATCTGGTTTCAAAGACTATCCACCCTT
Chimp GTGTCCTGGGTGCTGGCCTCCTTCCTGCAGCGGATCTGGTTTCAAAGACTATCCACCCTT
Orang GTGTCCTGGGTGCTGGCCTCCTTCCTGCAGCAGATCTGGTTTCAAAGACTATCCACCCTT
Gibbon GTGTCCTGGGTGCTGGCCTCCTTCCTGCAGCGGATCTGGTTTCAAAGACTATCCACCCTT
Rhesus GTGTCCTGGGTGCTGGCCTCCTTCCTGCAGCGGATCTGGTTTCAAAGACTATCCACCCTT
+210runx2 CCTAGGATTTAAAAAGTTCTTAATCAAGGCCTAGCAGGAAGATCTCTCTCATTTTTCCTC
Chimp CCTAGGATTTAAAAAGTTCTTAATCAAGGCCTAGCAGGAAGATCTCTCTCATTTTTCCTC
Orang CCTAGGATTTAAAAAGTTCTTAATCAAGGCCTAGCAGGAAGATCTCTCTCATTTTTCCTC
Gibbon CCTAGGATTTAAAAAGTTCTTAATCAAGGCCTAGCAGGAAGATCTCTCTCATTTTTCCTC
Rhesus CCTAGGATTTAAAAAGTTCTTAATCAAGGCCTAGCAGGAAGATCTCTCTCATTTTTCCTC
+210runx2 CCTTTTCTCTCCCCCTCCATCTTTGGCTCGTTCCTTCCCTTACTCTCCCTCCCTGTATGC
Chimp CCTTTTCTCTCCCCCTCCATCTTTGGCTCGTTCCTTCCCTTACTCTCCCTCCCTGTATGC
Orang CCTTTTCTCTCCCCCTCCATCTTTGGCTCGTTCCTTCCCTTACTCTCCCTCCCTGTATGC
Gibbon CCTTTTCTCTCCCCCTCCATCTTTGGCTCGTTCCTTCCCTTACTCTCCCTCCCTGTATGC
Rhesus CCTTTTCTCTCCCCCTCCATCTTTGGCTCCTTCCTTCCCTTGCTCTCCCTCCCTGTATGC
+210runx2 ATACACACTCACTCACACATCTACTGCCTGCC
Chimp ATACACACTCACTCACACATCTACTGCCTGCC
Orang ATACACACTCACTCACACATCTACTGCCTGCC
Gibbon ATACACACTCACTCACACATCTACTGCCTGCC
Rhesus ATACACACTCACTCACACTTCTACTGCCTGCC

Figure 4.3 – Recent positive selection in the human lineage near +210*RUNX2*. (a) Genome browser view of the human *RUNX2* locus. Signals of positive selection based on scoring of individual SNPs. A negative score indicates more derived alleles in modern humans relative to Neanderthals and is evidence of positive selection. SNP scan data obtained from [187]. (b) Three derived SNPs (indicated by red rectangles) in +210*RUNX2* are candidate alleles for a recent selective sweep in the human lineage. All three are derived with respect to the ancestral primate state and have not been observed to be polymorphic in human populations.

SNP	Phenotype	MAF	Dist from -460RUNX2 (kb)
rs556621	Atherosclerotic Stroke	0.3	-241
rs11755164	Bone mass density	0.4	-196
rs10948172	Osteoarthritis	0.29	58
rs9472414	Adult Height (European)	0.22	111
rs10948197	Adult Height (Korean)	0.34	132
rs9395066	Adult Height (DECODE)	0.48	260

Table 4.1 – SNPs associated with human skeletal phenotypes in the human RUNX2 locus. Minor allele frequency is given in the studied population. Distance from the Wnt responsive enhancer -460*RUNX2* is given in kilobases. A negative distance indicates distance 5' to the element, while a positive one denotes 3' separation.

References

1. Lefebvre, V. & Bhattaram, P. Vertebrate Skeletogenesis. *Curr. Top. Dev. Biol.* **90C**, 291-317 (2010).

2. Zhang, J. *et al*. Identification of the haematopoietic stem cell niche and control of the niche size. *Nature* **425**, 836 <last_page> 841 (2003).

3. Ferron, M. *et al*. Insulin Signaling in Osteoblasts Integrates Bone Remodeling and Energy Metabolism. *Cell* **142**, 296-308 (2010).

4. Copp, D. H. & Shim, S. S. The homeostatic function of bone as a mineral reservoir. *Oral Surgery, Oral Medicine, Oral Pathology* **16**, 738-744 (1963).

5. Oury, F. *et al*. Endocrine Regulation of Male Fertility by the Skeleton. *Cell* (2011).

6. Zhang, G., Eames, B. F. & Cohn, M. J. in *Current Topics in Developmental Biology* 15-42 (Academic Press.

7. - Meulemans, D. & - Bronner-Fraser, M. - Insights from Amphioxus into the Evolution of Vertebrate Cartilage. - *PLoS ONE*, - e787.

8. Person, P. & Mathews, M. B. Endoskeletal Cartilage in a Marine polychaete, Eudistylia polymorpha. *Biol. Bull.* **132**, 244-252 (1967).

9. Wilson, R., Belluoccio, D. & Bateman, J. F. Proteomic analysis of cartilage proteins. *Methods* **45**, 22-31 (2008).

10. Long, F. Building strong bones: molecular regulation of the osteoblast lineage. *Nat. Rev. Mol. Cell Biol.* **13**, 27-38 (2011).

11. Lammi, M. J., Häyrinen, J. & Mahonen, A. Proteomic analysis of cartilage- and bone-associated samples. *Electrophoresis* **27**, 2687-2701 (2006).

12. Weinreb, M., Shinar, D. & Rodan, G. A. Different pattern of alkaline phosphatase, osteopontin, and osteocalcin expression in developing rat bone visualized by in situ hybridization. *Journal of Bone and Mineral Research* **5**, 831-842 (1990).

13. Doty, S. B. in *Enzyme histochemistry of bone and cartilage cells* 37 (Fischer, Stuttgart, 1976).

14. Noble, B. S. The osteocyte lineage. *Arch. Biochem. Biophys.* **473**, 106-111 (2008).

15. Teitelbaum, S. L. Bone Resorption by Osteoclasts. *Science* **289**, 1504-1508 (2000).

16. Uccelli, A., Moretta, L. & Pistoia, V. Mesenchymal stem cells in health and disease. *Nat. Rev. Immunol.* **8**, 726-736 (2008).

17. Hall, B. K. in *Bones and cartilage: developmental and evolutionary skeletal biology* 760 (San Diego, Calif. ; Elsevier Academic Press, c2005.

18. Franz-Odendaal, T. A. Induction and patterning of intramembranous bone. *Front. Biosci.* **16**, 2734-2746 (2011).

19. Long, F. & Ornitz, D. M. Development of the Endochondral Skeleton. *Cold Spring Harbor Perspectives in Biology* **5** (2013).

20. Kania, M. A., Bonner, A. S., Duffy, J. B. & Gergen, J. P. The Drosophila segmentation gene runt encodes a novel nuclear regulatory protein that is also expressed in the developing nervous system. *Genes Dev.* **4**, 1701-1713 (1990).

21. Kitayner, M., Rozenberg, H., Rabinovich, D. & Shakked, Z. Structures of the DNA-binding site of Runt-domain transcription regulators. *Acta Crystallogr. D Biol. Crystallogr.* **61**, 236-246 (2005).

22. Komori, T. *et al*. Targeted Disruption of Cbfa1 Results in a Complete Lack of Bone Formation owing to Maturational Arrest of Osteoblasts. *Cell* **89**, 755 <last_page> 764 (1997).

23. Inada, M. *et al*. Maturational disturbance of chondrocytes in Cbfa1-deficient mice. *Developmental Dynamics* **214**, 279-290 (1999).

24. Mundlos, S. *et al*. Mutations involving the transcription factor CBFA1 cause cleidocranial dysplasia. *Cell* **89**, 773-779 (1997).

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

26. Gersbach, C. A., Le Doux, J. M., Guldberg, R. E. & García, A. J. Inducible regulation of Runx2-stimulated osteogenesis. *Gene Ther.* (2006).

27. Zheng, H., Guo, Z., Ma, Q., Jia, H. & Dang, G. Cbfa1/osf2 transduced bone marrow stromal cells facilitate bone formation in vitro and in vivo. *Calcif. Tissue Int.* **74**, 194-203 (2004).

28. Stein, G. S. *et al*. Runx2 control of organization, assembly and activity of the regulatory machinery for skeletal gene expression. *Oncogene* **23**, 4315-4329 (2004).

29. Liu, W. *et al*. Overexpression of Cbfa1 in osteoblasts inhibits osteoblast maturation and causes osteopenia with multiple fractures. *J. Cell Biol.* **155**, 157-166 (2001).

30. Ueta, C. *et al*. Skeletal Malformations Caused by Overexpression of Cbfa1 or Its Dominant Negative Form in Chondrocytes. *The Journal of Cell Biology* **153**, 87-100 (2001).

31. Maruyama, Z. *et al*. Runx2 determines bone maturity and turnover rate in postnatal bone development and is involved in bone loss in estrogen deficiency. *Dev. Dyn.* **236**, 1876-1890 (2007).

32. Nakashima, K. *et al*. The Novel Zinc Finger-Containing Transcription Factor Osterix Is Required for Osteoblast Differentiation and Bone Formation. *Cell* **108**, 17-29 (2002).

33. Barbuto, R. & Mitchell, J. Regulation of the osterix (Osx, Sp7) promoter by osterix and its inhibition by parathyroid hormone. *Journal of Molecular Endocrinology* **51**, 99-108 (2013).

34. Zhang, C. *et al*. Inhibition of Wnt signaling by the osteoblast-specific transcription factor Osterix. *Proceedings of the National Academy of Sciences* **105**, 6936 <last_page> 6941 (2008).

35. Ortuño, M. J., Susperregui, A. R. G., Artigas, N., Rosa, J. L. & Ventura, F. Osterix induces Col1a1 gene expression through binding to Sp1 sites in the bone enhancer and proximal promoter regions. *Bone* **52**, 548 <last_page> 556 (2013).

36. Okamura, H., Yoshida, K., Yang, D. & Haneji, T. Protein phosphatase 2A C? regulates osteoblast differentiation and the expressions of bone sialoprotein and osteocalcin via osterix transcription factor. *J. Cell. Physiol.* **228**, 1031-1037 (2013).

37. Lapunzina, P. *et al*. Identification of a Frameshift Mutation in Osterix in a Patient with Recessive Osteogenesis Imperfecta. *The American Journal of Human Genetics* **87**, 110-114 (2010).

38. Zhou, X. *et al*. Multiple functions of Osterix are required for bone growth and homeostasis in postnatal mice. *Proceedings of the National Academy of Sciences* **107**, 12919-12924 (2010).

39. Wozney, J. M. *et al*. Novel regulators of bone formation: molecular clones and activities. *Science* **242**, 1528-1534 (1988).

40. Urist, M. R. Bone: formation by autoinduction. *Science* **150**, 893-899 (1965).

41. Tsuji, K. *et al*. BMP4 is dispensable for skeletogenesis and fracture-healing in the limb. *J. Bone Joint Surg. Am.* **90 Suppl 1**, 14-18 (2008).

42. Tsuji, K. *et al*. BMP2 activity, although dispensable for bone formation, is required for the initiation of fracture healing. *Nat. Genet.* **38**, 1424-1429 (2006).

43. Gamer, L. W. *et al*. BMPR-II is dispensable for formation of the limb skeleton. *Genesis* **49**, 719-724 (2011).

44. Tsuji, K. *et al*. Conditional deletion of BMP7 from the limb skeleton does not affect bone formation or fracture repair. *J. Orthop. Res.* **28**, 384-389 (2010).

45. Bandyopadhyay, A. *et al*. Genetic analysis of the roles of BMP2, BMP4, and BMP7 in limb patterning and skeletogenesis. *PLoS Genet.* **2**, e216 (2006).

46. Lee, K. S. *et al*. Runx2 is a common target of transforming growth factor beta1 and bone morphogenetic protein 2, and cooperation between Runx2 and Smad5 induces osteoblast-specific gene expression in the pluripotent mesenchymal precursor cell line C2C12. *Mol. Cell. Biol.* **20**, 8783-8792 (2000).

47. Lee, M. H. *et al*. Transient upregulation of CBFA1 in response to bone morphogenetic protein-2 and transforming growth factor beta1 in C2C12 myogenic cells coincides with suppression of the myogenic phenotype but is not sufficient for osteoblast differentiation. *J. Cell. Biochem.* **73**, 114-125 (1999).

48. Gersbach, C. A., Byers, B. A., Pavlath, G. K. & Garcia, A. J. Runx2/Cbfa1 stimulates transdifferentiation of primary skeletal myoblasts into a mineralizing osteoblastic phenotype. *Exp. Cell Res.* **300**, 406-417 (2004).

49. Liu, T. *et al*. BMP-2 promotes differentiation of osteoblasts and chondroblasts in Runx2-deficient cell lines. *J. Cell. Physiol.* **211**, 728-735 (2007).

50. Kobayashi, H., Gao, Y., Ueta, C., Yamaguchi, A. & Komori, T. Multilineage differentiation of Cbfa1-deficient calvarial cells in vitro. *Biochem. Biophys. Res. Commun.* **273**, 630-636 (2000).

51. Lee, M., Kwon, T., Park, H., Wozney, J. M. & Ryoo, H. BMP-2-induced Osterix expression is mediated by Dlx5 but is independent of Runx2. *Biochem. Biophys. Res. Commun.* **309**, 689 <last_page> 694 (2003).

52. Abe, E. *et al*. Essential requirement of BMPs-2/4 for both osteoblast and osteoclast formation in murine bone marrow cultures from adult mice: antagonism by noggin. *J. Bone Miner. Res.* **15**, 663-673 (2000).

53. Xiao, G. *et al*. Bone morphogenetic proteins, extracellular matrix, and mitogen-activated protein kinase signaling pathways are required for osteoblast-specific gene expression and differentiation in MC3T3-E1 cells. *J. Bone Miner. Res.* **17**, 101-110 (2002).

54. Jonason, J. H., Xiao, G., Zhang, M., Xing, L. & Chen, D. Post-translational Regulation of Runx2 in Bone and Cartilage. *J. Dent. Res.* **88**, 693-703 (2009).

55. Wang, M. *et al*. Smad1 plays an essential role in bone development and postnatal bone formation. *Osteoarthritis and Cartilage* **19**, 751-762 (2011).

56. Retting, K. N., Song, B., Yoon, B. S. & Lyons, K. M. BMP canonical Smad signaling through Smad1 and Smad5 is required for endochondral bone formation. *Development* **136**, 1093-1104 (2009).

57. Lee, M. *et al*. BMP-2-induced Runx2 Expression Is Mediated by Dlx5, and TGF-ß1 Opposes the BMP-2-induced Osteoblast Differentiation by Suppression of Dlx5 Expression. *Journal of Biological Chemistry* **278**, 34387-34394 (2003).

58. Lee, M., Kwon, T., Park, H., Wozney, J. M. & Ryoo, H. BMP-2-induced Osterix expression is mediated by Dlx5 but is independent of Runx2. *Biochem. Biophys. Res. Commun.* **309**, 689-694 (2003).

59. de Jong, D. S. *et al*. Identification of Novel Regulators Associated With Early-Phase Osteoblast Differentiation. *Journal of Bone and Mineral Research* **19**, 947-958 (2004).

60. Balint, E. *et al*. Phenotype discovery by gene expression profiling: Mapping of biological processes linked to BMP-2-mediated osteoblast differentiation. *J. Cell. Biochem.* **89**, 401-426 (2003).

61. Harris, S. E., Guo, D., Harris, M. A., Krishnaswamy, A. & Lichtler, A. Transcriptional regulation of BMP-2 activated genes in osteoblasts using gene expression microarray analysis: role of Dlx2 and Dlx5 transcription factors. *Front. Biosci.* **8**, s1249-65 (2003).

62. Satokata, I. & Maas, R. Msx1 deficient mice exhibit cleft palate and abnormalities of craniofacial and tooth development. *Nat. Genet.* **6**, 348-356 (1994).

63. Satokata, I. *et al*. Msx2 deficiency in mice causes pleiotropic defects in bone growth and ectodermal organ formation. *Nat. Genet.* **24**, 391-395 (2000).

64. Han, J. *et al*. Concerted action of Msx1 and Msx2 in regulating cranial neural crest cell differentiation during frontal bone development. *Mech. Dev.* **124**, 729-745 (2007).

65. Jabs, E. W. *et al*. A mutation in the homeodomain of the human MSX2 gene in a family affected with autosomal dominant craniosynostosis. *Cell* **75**, 443-450 (1993).

66. Ott, C. E. *et al*. Microduplications upstream of MSX2 are associated with a phenocopy of cleidocranial dysplasia. *J. Med. Genet.* **49**, 437-441 (2012).

67. Panganiban, G. & Rubenstein, J. L. R. Developmental functions of the Distal-less/Dlx homeobox genes. *Development* **129**, 4371-4386 (2002).

68. Thomas, B. L. *et al*. Role of Dlx-1 and Dlx-2 genes in patterning of the murine dentition. *Development* **124**, 4811-4818 (1997).

69. Chase, M. B. *et al*. BP1, a Homeodomain-Containing Isoform of DLX4, Represses the ß-Globin Gene. *Molecular and Cellular Biology* **22**, 2505-2514 (2002).

70. Ghoul-Mazgar, S. *et al*. Expression pattern of Dlx3 during cell differentiation in mineralized tissues. *Bone* **37**, 799-809 (2005).

71. Morasso, M. I., Grinberg, A., Robinson, G., Sargent, T. D. & Mahon, K. A. Placental failure in mice lacking the homeobox gene Dlx3. *Proceedings of the National Academy of Sciences* **96**, 162-167 (1999).

72. Haldeman, R. J. *et al*. Increased bone density associated with DLX3 mutation in the tricho-dento-osseous syndrome. *Bone* **35**, 988-997 (2004).

73. Hassan, M. Q. *et al*. Dlx3 Transcriptional Regulation of Osteoblast Differentiation: Temporal Recruitment of Msx2, Dlx3, and Dlx5 Homeodomain

Proteins to Chromatin of the Osteocalcin Gene. *Molecular and Cellular Biology* **24**, 9248-9261 (2004).

74. Samee, N. *et al*. Dlx5, a Positive Regulator of Osteoblastogenesis, is Essential for Osteoblast-Osteoclast Coupling. *The American Journal of Pathology* **173**, 773-780 (2008).

75. Robledo, R. F., Rajan, L., Li, X. & Lufkin, T. The Dlx5 and Dlx6 homeobox genes are essential for craniofacial, axial, and appendicular skeletal development. *Genes & Development* **16**, 1089-1101 (2002).

76. Miyama, K. *et al*. A BMP-Inducible Gene, Dlx5, Regulates Osteoblast Differentiation and Mesoderm Induction. *Dev. Biol.* **208**, 123-133 (1999).

77. Holleville, N., Quilhac, A., Bontoux, M. & Monsoro-Burq, A. é. BMP signals regulate Dlx5 during early avian skull development. *Dev. Biol.* **257**, 177-189 (2003).

78. Holleville, N., Matéos, S., Bontoux, M., Bollerot, K. & Monsoro-Burq, A. Dlx5 drives Runx2 expression and osteogenic differentiation in developing cranial suture mesenchyme. *Dev. Biol.* **304**, 860-874 (2007).

79. Kim, Y., Lee, M., Wozney, J. M., Cho, J. & Ryoo, H. Bone Morphogenetic Protein-2-induced Alkaline Phosphatase Expression Is Stimulated by Dlx5 and Repressed by Msx2. *Journal of Biological Chemistry* **279**, 50773-50780 (2004).

80. Klaus, A. & Birchmeier, W. Wnt signalling and its impact on development and cancer. *Nature Reviews Cancer* **8**, 387 <last_page> 398 (2008).

81. Gong, Y. *et al*. LDL Receptor-Related Protein 5 (LRP5) Affects Bone Accrual and Eye Development. *Cell* **107**, 513-523 (2001).

82. Kato, M. *et al*. Cbfa1-independent decrease in osteoblast proliferation, osteopenia, and persistent embryonic eye vascularization in mice deficient in Lrp5, a Wnt coreceptor. *The Journal of Cell Biology* **157**, 303-314 (2002).

83. Boyden, L. M. *et al*. High Bone Density Due to a Mutation in LDL-Receptor–Related Protein 5. *N. Engl. J. Med.* **346**, 1513-1521 (2002).

84. Little, R. D. *et al*. A Mutation in the LDL Receptor–Related Protein 5 Gene Results in the Autosomal Dominant High–Bone-Mass Trait. *The American Journal of Human Genetics* **70**, 11-19 (2002).

85. Ai, M., Holmen, S. L., Van Hul, W., Williams, B. O. & Warman, M. L. Reduced affinity to and inhibition by DKK1 form a common mechanism by which high bone mass-associated missense mutations in LRP5 affect canonical Wnt signaling. *Mol. Cell. Biol.* **25**, 4946-4955 (2005).

86. Balemans, W. *et al*. The binding between sclerostin and LRP5 is altered by DKK1 and by high-bone mass LRP5 mutations. *Calcif. Tissue Int.* **82**, 445-453 (2008).

87. Ellies, D. L. *et al*. Bone density ligand, Sclerostin, directly interacts with LRP5 but not LRP5G171V to modulate Wnt activity. *J. Bone Miner. Res.* **21**, 1738-1749 (2006).

88. Balemans, W. *et al*. Increased bone density in sclerosteosis is due to the deficiency of a novel secreted protein (SOST). *Hum. Mol. Genet.* **10**, 537-543 (2001).

89. Balemans, W. *et al*. Identification of a 52 kb deletion downstream of the SOST gene in patients with van Buchem disease. *J. Med. Genet.* **39**, 91-97 (2002).

90. Regard, J. B., Zhong, Z., Williams, B. O. & Yang, Y. Wnt Signaling in Bone Development and Disease: Making Stronger Bone with Wnts. *Cold Spring Harbor Perspectives in Biology* **4** (2012).

91. Day, T. F., Guo, X., Garrett-Beal, L. & Yang, Y. Wnt/ß-Catenin Signaling in Mesenchymal Progenitors Controls Osteoblast and Chondrocyte Differentiation during Vertebrate Skeletogenesis. *Developmental Cell* **8**, 739-750 (2005).

92. Hill, T. P., Später, D., Taketo, M. M., Birchmeier, W. & Hartmann, C. Canonical Wnt/ß-Catenin Signaling Prevents Osteoblasts from Differentiating into Chondrocytes. *Developmental Cell* **8**, 727-738 (2005).

93. Rudnicki, J. A. & Brown, A. M. C. Inhibition of Chondrogenesis byWntGene Expressionin Vivoandin Vitro. *Dev. Biol.* **185**, 104-118 (1997).

94. Reinhold, M. I., Kapadia, R. M., Liao, Z. & Naski, M. C. The Wnt-inducible Transcription Factor Twist1 Inhibits Chondrogenesis. *Journal of Biological Chemistry* **281**, 1381-1388 (2006).

95. Dong, Y., Soung, D. Y., Schwarz, E. M., O'Keefe, R. J. & Drissi, H. Wnt induction of chondrocyte hypertrophy through the Runx2 transcription factor. *J. Cell. Physiol.* **208**, 77-86 (2006).

96. Marcellini, S., Henriquez, J. P. & Bertin, A. Control of osteogenesis by the canonical Wnt and BMP pathways in vivo. *Bioessays* **34**, 953-962 (2012).

97. Ornitz, D. M. & Itoh, N. Fibroblast growth factors. *Genome Biol.* **2**, REVIEWS3005 (2001).

98. Ornitz, D. M. FGF signaling in the developing endochondral skeleton. *Cytokine Growth Factor Rev.* **16**, 205-213 (2005).

99. Mansukhani, A., Ambrosetti, D., Holmes, G., Cornivelli, L. & Basilico, C. Sox2 induction by FGF and FGFR2 activating mutations inhibits Wnt signaling and osteoblast differentiation. *The Journal of Cell Biology* **168**, 1065-1076 (2005).

100. Debiais, F., Hott, M., Graulet, A. M. & Marie, P. J. The Effects of Fibroblast Growth Factor-2 on Human Neonatal Calvaria Osteoblastic Cells Are Differentiation Stage Specific. *Journal of Bone and Mineral Research* **13**, 645-654 (1998).

101. Woei Ng, K. *et al*. Osteogenic differentiation of murine embryonic stem cells is mediated by fibroblast growth factor receptors. *Stem Cells Dev.* **16**, 305-318 (2007).

102. Montero, A. *et al*. Disruption of the fibroblast growth factor-2 gene results in decreased bone mass and bone formation. *J. Clin. Invest.* **105**, 1085 <last_page> 1093 (2000).

103. Coffin, J. D. *et al*. Abnormal bone growth and selective translational regulation in basic fibroblast growth factor (FGF-2) transgenic mice. *Molecular Biology of the Cell* **6**, 1861-1873 (1995).

104. Hung, I. H., Yu, K., Lavine, K. J. & Ornitz, D. M. FGF9 regulates early hypertrophic chondrocyte differentiation and skeletal vascularization in the developing stylopod. *Dev. Biol.* **307**, 300-313 (2007).

105. Garofalo, S. *et al*. Skeletal Dysplasia and Defective Chondrocyte Differentiation by Targeted Overexpression of Fibroblast Growth Factor 9 in Transgenic Mice. *Journal of Bone and Mineral Research* **14**, 1909-1915 (1999).

106. Wilkie, A. O. M. Bad bones, absent smell, selfish testes: The pleiotropic consequences of human FGF receptor mutations. *Cytokine Growth Factor Rev.* **16**, 187-203 (2005).

107. Mbalaviele, G. *et al*. ?-Catenin and BMP-2 synergize to promote osteoblast differentiation and new bone formation. *J. Cell. Biochem.* **94**, 403-418 (2005).

108. Rodríguez-Carballo, E. *et al*. Conserved regulatory motifs in osteogenic gene promoters integrate cooperative effects of canonical Wnt and BMP pathways. *Journal of Bone and Mineral Research* **26**, 718-729 (2011).

109. Fukuda, T. *et al*. Canonical Wnts and BMPs cooperatively induce osteoblastic differentiation through a GSK3beta-dependent and beta-catenin-independent mechanism. *Differentiation* **80**, 46-52 (2010).

110. Fuentealba, L. C. *et al*. Integrating Patterning Signals: Wnt/GSK3 Regulates the Duration of the BMP/Smad1 Signal. *Cell* **131**, 980-993 (2007).

111. Hirota, M. *et al*. Smad2 functions as a co-activator of canonical Wnt/beta-catenin signaling pathway independent of Smad4 through histone acetyltransferase activity of p300. *Cell. Signal.* **20**, 1632-1641 (2008).

112. Kamiya, N. The role of BMPs in bone anabolism and their potential targets SOST and DKK1. *Curr. Mol. Pharmacol.* **5**, 153-163 (2012).

113. Kamiya, N. *et al*. Wnt inhibitors Dkk1 and Sost are downstream targets of BMP signaling through the type IA receptor (BMPRIA) in osteoblasts. *Journal of Bone and Mineral Research* **25**, 200-210 (2010).

114. Liu, Z., Tang, Y., Qiu, T., Cao, X. & Clemens, T. L. A Dishevelled-1/Smad1 Interaction Couples WNT and Bone Morphogenetic Protein Signaling Pathways in Uncommitted Bone Marrow Stromal Cells. *Journal of Biological Chemistry* **281**, 17156-17163 (2006).

115. Trompouki, E. *et al*. Lineage Regulators Direct BMP and Wnt Pathways to Cell-Specific Programs during Differentiation and Regeneration. *Cell* **147**, 577-589 (2011).

116. Naganawa, T. *et al*. Reduced expression and function of bone morphogenetic protein-2 in bones of Fgf2 null mice. *J. Cell. Biochem.* **103**, 1975-1988 (2008).

117. Fakhry, A. *et al*. Effects of FGF-2/-9 in calvarial bone cell cultures: differentiation stage-dependent mitogenic effect, inverse regulation of BMP-2 and noggin, and enhancement of osteogenic potential. *Bone* **36**, 254-266 (2005).

118. Warren, S. M., Brunet, L. J., Harland, R. M., Economides, A. N. & Longaker, M. T. The BMP antagonist noggin regulates cranial suture fusion. *Nature* **422**, 625-629 (2003).

119. Agas, D., Sabbieti, M. G., Marchetti, L., Xiao, L. & Hurley, M. M. FGF-2 enhances Runx-2/Smads nuclear localization in BMP-2 canonical signaling in osteoblasts. *J. Cell. Physiol.*, n/a-n/a (2013).

120. Hughes-Fulford, M. & Li, C. F. The role of FGF-2 and BMP-2 in regulation of gene induction, cell proliferation and mineralization. *J. Orthop. Surg. Res.* **6**, 8-799X-6-8 (2011).

121. Raucci, A., Bellosta, P., Grassi, R., Basilico, C. & Mansukhani, A. Osteoblast proliferation or differentiation is regulated by relative strengths of opposing signaling pathways. *J. Cell. Physiol.* **215**, 442-451 (2008).

122. Fei, Y., Xiao, L., Doetschman, T., Coffin, D. J. & Hurley, M. M. Fibroblast Growth Factor 2 Stimulation of Osteoblast Differentiation and Bone Formation Is Mediated by Modulation of the Wnt Signaling Pathway. *Journal of Biological Chemistry* **286**, 40575-40583 (2011).

123. Yu, H. I. *et al*. The role of Axin2 in calvarial morphogenesis and craniosynostosis. *Development* **132**, 1995-2005 (2005).

124. Liu, B., Yu, H. I. & Hsu, W. Craniosynostosis caused by Axin2 deficiency is mediated through distinct functions of ß-catenin in proliferation and differentiation. *Dev. Biol.* **301**, 298-308 (2007).

125. Maruyama, T., Mirando, A. J., Deng, C. & Hsu, W. The Balance of WNT and FGF Signaling Influences Mesenchymal Stem Cell Fate During Skeletal Development. *Sci. Signal.* **3**, ra40 (2010).

126. Spoorendonk, K. M., Hammond, C. L., Huitema, L. F. A., Vanoevelen, J. & Schulte-Merker, S. Zebrafish as a unique model system in bone research: the power of genetics and in vivo imaging. *J. Appl. Ichthyol.* **26**, 219-224 (2010).

127. Eames, B. F. *et al.* FishFace: interactive atlas of zebrafish craniofacial development at cellular resolution. *BMC Developmental Biology* **13**, 23 (2013).

128. Bird, N. C. & Mabee, P. M. Developmental morphology of the axial skeleton of the zebrafish, Danio rerio (Ostariophysi: Cyprinidae). *Developmental Dynamics* **228**, 337-357 (2003).

129. Yelick, P. C. & Schilling, T. F. Molecular Dissection of Craniofacial Development Using Zebrafish. *Critical Reviews in Oral Biology & Medicine* **13**, 308-322 (2002).

130. Li, N., Felber, K., Elks, P., Croucher, P. & Roehl, H. H. Tracking gene expression during zebrafish osteoblast differentiation. *Developmental Dynamics* **238**, 459-466 (2009).

131. Flores, M. V. *et al.* Duplicate zebrafish runx2 orthologues are expressed in developing skeletal elements. *Gene Expression Patterns* **4**, 573-581 (2004).

132. Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. & Schilling, T. F. Stages of embryonic development of the zebrafish. *Developmental Dynamics* **203**, 253-310 (1995).

133. Li, N., Felber, K., Elks, P., Croucher, P. & Roehl, H. H. Tracking gene expression during zebrafish osteoblast differentiation. *Developmental Dynamics* **238**, 459-466 (2009).

134. Witten, P. E. & Huysseune, A. A comparative view on mechanisms and functions of skeletal remodelling in teleost fish, with special emphasis on osteoclasts and their function. *Biological Reviews* **84**, 315-346 (2009).

135. Lee, B. *et al.* Missense mutations abolishing DNA binding of the osteoblast-specific transcription factor OSF2/CBFA1 in cleidocranial dysplasia. *Nat. Genet.* **16**, 307 <last_page> 310 (1997).

136. Otto, F. *et al.* Cbfa1, a candidate gene for cleidocranial dysplasia syndrome, is essential for osteoblast differentiation and bone development. *Cell* **89**, 765-771 (1997).

137. Long, F. Building strong bones: molecular regulation of the osteoblast lineage. *Nat. Rev. Mol. Cell Biol.* **13**, 27-38 (2011).

138. Hardison, R. C. & Taylor, J. Genomic approaches towards finding cis-regulatory modules in animals. *Nature Reviews Genetics* **13**, 469-483 (2012).

139. Muller, H. P. & Schaffner, W. Transcriptional enhancers can act in trans. *Trends Genet.* **6**, 300-304 (1990).

140. Bateman, J. R., Johnson, J. E. & Locke, M. N. Comparing enhancer action in cis and in trans. *Genetics* **191**, 1143-1155 (2012).

141. Birnbaum, R. Y. *et al*. Coding exons function as tissue-specific enhancers of nearby genes. *Genome Res.* **22**, 1059-1068 (2012).

142. Schubert, M., Ritter, D. I., Dong, Z., Guo, S. & Chuang, J. H. Transcriptional Enhancers in Protein-Coding Exons of Vertebrate Developmental Genes. *PLoS ONE* **7**, e35202 (2012).

143. Harada, H. *et al*. Cbfa1 Isoforms Exert Functional Differences in Osteoblast Differentiation. *Journal of Biological Chemistry* **274**, 6972-6978 (1999).

144. Ogawa, E. *et al*. PEBP2/PEA2 represents a family of transcription factors homologous to the products of the Drosophila runt gene and the human AML1 gene. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 6859-6863 (1993).

145. Banerjee, C. *et al*. Differential Regulation of the Two Principal Runx2/Cbfa1 N-Terminal Isoforms in Response to Bone Morphogenetic Protein-2 during Development of the Osteoblast Phenotype. *Endocrinology* **142**, 4026-4039 (2001).

146. Park, M. H. *et al*. Differential expression patterns of Runx2 isoforms in cranial suture morphogenesis. *J. Bone Miner. Res.* **16**, 885-892 (2001).

147. Enomoto, H. *et al*. Cbfa1 Is a Positive Regulatory Factor in Chondrocyte Maturation. *Journal of Biological Chemistry* **275**, 8695-8702 (2000).

148. Liu, J. C. *et al*. Runx2 protein expression utilizes the Runx2 P1 promoter to establish osteoprogenitor cell number for normal bone formation. *J. Biol. Chem.* **286**, 30057-30070 (2011).

149. Xiao, Z. S., Liu, S., Hinson, T. K. & Quarles, L. D. Characterization of the upstream mouse Cbfa1/Runx2 promoter*. *J. Cell. Biochem.* **82**, 647-659 (2001).

150. Tamiya, H. *et al.* Analysis of the Runx2 promoter in osseous and non-osseous cells and identification of HIF2A as a potent transcription activator. *Gene* **416**, 53 <last_page> 60 (2008).

151. Fernandez, B. A., Siegel-Bartelt, J., Herbrick, J. A., Teshima, I. & Scherer, S. W. Holoprosencephaly and cleidocranial dysplasia in a patient due to two position-effect mutations: case report and review of the literature. *Clin. Genet.* **68**, 349-359 (2005).

152. Purandare, S. M. *et al.* De novo three-way chromosome translocation 46,XY,t(4;6;21)(p16;p21.1;q21) in a male with cleidocranial dysplasia. *Am. J. Med. Genet. A.* **146A**, 453-458 (2008).

153. Dickmeis, T. & Muller, F. The identification and functional characterisation of conserved regulatory elements in developmental genes. *Brief Funct. Genomic Proteomic* **3**, 332-350 (2005).

154. Brown, C. T. Computational approaches to finding and analyzing cis-regulatory elements. *Methods Cell Biol.* **87**, 337-365 (2008).

155. Mouse Genome Sequencing Consortium *et al.* Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520-562 (2002).

156. Miller, W., Makova, K. D., Nekrutenko, A. & Hardison, R. C. Comparative genomics. *Annu. Rev. Genomics Hum. Genet.* **5**, 15-56 (2004).

157. Ahituv, N. *et al*. Deletion of ultraconserved elements yields viable mice. *PLoS Biol.* **5**, e234 (2007).

158. Barski, A. *et al*. High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823-837 (2007).

159. Roh, T. Y., Wei, G., Farrell, C. M. & Zhao, K. Genome-wide prediction of conserved and nonconserved enhancers by histone acetylation patterns. *Genome Res.* **17**, 74-81 (2007).

160. Siepel, A. *et al*. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* **15**, 1034-1050 (2005).

161. Fisher, S. *et al*. Evaluating the biological relevance of putative enhancers using Tol2 transposon-mediated transgenesis in zebrafish. *Nat. Protoc.* **1**, 1297-1305 (2006).

162. Kague, E., Weber, C. & Fisher, S. Mosaic zebrafish transgenesis for evaluating enhancer sequences. *J. Vis. Exp.* **(41). pii: 1722. doi**, 10.3791/1722 (2010).

163. Renn, J. & Winkler, C. Osterix-mCherry transgenic medaka for in vivo imaging of bone formation. *Developmental Dynamics* **238**, 241-248 (2009).

164. Thisse, C. & Thisse, B. High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nat. Protoc.* **3**, 59-69 (2008).

165. Lan, C. C., Tang, R., Un San Leong, I. & Love, D. R. Quantitative real-time RT-PCR (qRT-PCR) of zebrafish transcripts: optimization of RNA extraction,

quality control considerations, and data analysis. *Cold Spring Harb Protoc.* **2009**, pdb.prot5314 (2009).

166. Kimmel, C. B. Modes of Developmental Outgrowth and Shaping of a Craniofacial Bone in Zebrafish. *ONE Alerts* (2010).

167. Blanchette, M. *et al*. Aligning Multiple Genomic Sequences With the Threaded Blockset Aligner. *Genome Research* **14**, 708-715 (2004).

168. Willems, B. *et al*. Conditional ablation of osteoblasts in medaka. *Dev. Biol.* **364**, 128-137 (2012).

169. Leung, G. & Eisen, M. B. Identifying cis-regulatory sequences by word profile similarity. *PLoS One* **4**, e6901 (2009).

170. Stoick-Cooper, C. L. *et al*. Distinct Wnt signaling pathways have opposing roles in appendage regeneration. *Development* **134**, 479-489 (2007).

171. Chocron, S., Verhoeven, M. C., Rentzsch, F., Hammerschmidt, M. & Bakkers, J. Zebrafish Bmp4 regulates left–right asymmetry at two distinct developmental time points. *Dev. Biol.* **305**, 577-588 (2007).

172. Hashiguchi, M. & Mullins, M. C. Anteroposterior and dorsoventral patterning are coordinated by an identical patterning clock. *Development* **140**, 1970-1980 (2013).

173. Zhang, Z., Verheyden, J. M., Hassell, J. A. & Sun, X. FGF-Regulated Etv Genes Are Essential for Repressing Shh Expression in Mouse Limb Buds. *Developmental Cell* **16**, 607 <last_page> 613 (2009).

174. Münchberg, S. R. & Steinbeisser, H. The Xenopus Ets transcription factor XER81 is a target of the FGF signaling pathway. *Mech. Dev.* **80**, 53-65 (1999).

175. Miya, T. & Nishida, H. An Ets transcription factor, HrEts, is target of FGF signaling and involved in induction of notochord, mesenchyme, and brain in ascidian embryos. *Dev. Biol.* **261**, 25-38 (2003).

176. Sun, L. *et al*. Design, Synthesis, and Evaluations of Substituted 3-[(3- or 4-Carboxyethylpyrrol-2-yl)methylidenyl]indolin-2-ones as Inhibitors of VEGF, FGF, and PDGF Receptor Tyrosine Kinases. *J. Med. Chem.* **42**, 5120 <last_page> 5130 (1999).

177. Dobreva, G. *et al*. SATB2 is a multifunctional determinant of craniofacial patterning and osteoblast differentiation. *Cell* **125**, 971-986 (2006).

178. Westendorf, J. J., Kahler, R. A. & Schroeder, T. M. Wnt signaling in osteoblasts and bone diseases. *Gene* **341**, 19-39 (2004).

179. Perry, M. W., Boettiger, A. N., Bothma, J. P. & Levine, M. Shadow Enhancers Foster Robustness of Drosophila Gastrulation. *Current Biology* **20**, 1562-1567 (2010).

180. Brandman, O., Ferrell, J. E.,Jr, Li, R. & Meyer, T. Interlinked fast and slow positive feedback loops drive reliable cell decisions. *Science* **310**, 496-498 (2005).

181. Drissi, H. *et al*. Transcriptional autoregulation of the bone related CBFA1/RUNX2 gene. *J. Cell. Physiol.* **184**, 341-350 (2000).

182. Verreijdt, L. *et al*. Expression of the dlx gene family during formation of the cranial bones in the zebrafish (Danio rerio): Differential involvement in the visceral skeleton and braincase. *Developmental Dynamics* **235**, 1371-1389 (2006).

183. Thisse, B. *et al*. Spatial and temporal expression of the zebrafish genome by large-scale in situ hybridization screening. *Methods Cell Biol.* **77**, 505-519 (2004).

184. Yadav, V. K. *et al*. Lrp5 Controls Bone Formation by Inhibiting Serotonin Synthesis in the Duodenum. *Cell* **135**, 825-837 (2008).

185. Cui, Y. *et al*. Lrp5 functions in bone to regulate bone mass. *Nat. Med.* **17**, 684-691 (2011).

186. Leijten, J. C. H. *et al*. Gremlin 1, Frizzled-related protein, and Dkk-1 are key regulators of human articular cartilage homeostasis. *Arthritis & Rheumatism* **64**, 3302-3312 (2012).

187. Pasold, J. *et al*. Reduced expression of Sfrp1 during chondrogenesis and in articular chondrocytes correlates with osteoarthritis in STR/ort mice. *Exp. Cell Res.* **319**, 649-659 (2013).

188. Kamekura, S. *et al*. Contribution of runt-related transcription factor 2 to the pathogenesis of osteoarthritis in mice after induction of knee joint instability. *Arthritis & Rheumatism* **54**, 2462-2470 (2006).

189. Orfanidou, T., Iliopoulos, D., Malizos, K. N. & Tsezou, A. Involvement of SOX-9 and FGF-23 in RUNX-2 regulation in osteoarthritic chondrocytes. *J. Cell. Mol. Med.* **13**, 3186-3194 (2009).

190. Sears, K. E., Goswami, A., Flynn, J. J. & Niswander, L. A. The correlated evolution of Runx2 tandem repeats, transcriptional activity, and facial length in Carnivora. *Evol. Dev.* **9**, 555 <last_page> 565 (2007).

191. Pointer, M. A. *et al*. RUNX2 tandem repeats and the evolution of facial length in placental mammals. *BMC Evol. Biol.* **12**, 103-2148-12-103 (2012).

192. Green, R. E. *et al*. A draft sequence of the Neandertal genome. *Science* **328**, 710-722 (2010).

193. Gunz, P. *et al*. A uniquely modern human pattern of endocranial development. Insights from a new cranial reconstruction of the Neandertal newborn from Mezmaiskaya. *J. Hum. Evol.* **62**, 300-313 (2012).

194. Voisin, J. The Omo I hominin clavicle: Archaic or modern? *J. Hum. Evol.* **55**, 438-443 (2008).

195. Knopf, F. *et al*. Bone Regenerates via Dedifferentiation of Osteoblasts in the Zebrafish Fin. *Developmental Cell* **20**, 713-724 (2011).

196. Dimitriou, R., Tsiridis, E. & Giannoudis, P. V. Current concepts of molecular aspects of bone healing. *Injury* **36**, 1392-1404 (2005).

197. Baker, M. Screening: the age of fishes. *Nature Methods* **8**, 47 <last_page> 51 (2011).