ELUCIDATION OF THE GENE REGULATORY NETWORK CONTROLLING EMBRYONIC SKELETAL DEVELOPMENT

CHAN HSIAO YUN

(B.Sc (Merit), NUS)

A THESIS SUBMITTED FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF BIOLOGICAL SCIENCES

NATIONAL UNIVERSITY OF SINGAPORE

2011

ACKNOWLEDGEMENTS

My heartfelt thanks goes out to many who have supported me one way or another throughout this enriching Ph.D journey.

To my supervisor, Dr. Thomas Lufkin, for having first encouraged me to embark on this journey and for his patient guidance and confidence in my work.

To the postdocs in the lab, Sook Peng, Patricia and Sumantra, for being so willing to share with me their experience and knowledge.

To Jiewei, Gerry and Val, for taking care of my little friends in the BRC and to Petra for all the microinjections.

To other members of the Lufkin Lab, past and present, Siva, Songjie, Jean, Xing xing, Siewlan, Serene, Joel, Clara, Mathia, Max, for scientific discussions and suggestions, technical help and most of all for making the past 6 years in the lab such an enjoyable experience.

To Michelle, Keefe and Thye Seng, for their friendly help with the FACSAria Sorter.

Special thanks to Siva who always went that extra mile to help and for making the paper writing experience more pleasant than it would have been. To Nirmala, for the fun scientific talks during lunches and encouragement during that last lap.

To my husband, Kenn, for love, patience, constant encouragement and affirmation to sustain me through difficult and frustrating times in research.

To my family and Kenn's family, for making concessions to take care of Reuel so that I can fully concentrate on completing this.

And most of all to God, for seeing me through this journey.

TABLE OF CONTENTS

SUMMARY

In the last two decades, much progress has been made to understand skeletal biology. A growing number of transcription factors, co-factors, chromatin modifiers and signalling molecules have been discovered to contribute to the precise control of gene regulation that determines osteo-chondrogenic lineage specification and bone formation. Not only have these factors been identified, increasing number of genes are being placed in genetic pathways that control skeletogenesis. Since these diverse molecules work together in a highly interconnected network to tightly regulate gene expression, skeletal disorders such as cleidocranial dysplasias can be more effectively addressed if the molecular mechanisms behind bone development can be mapped onto graphic gene regulatory diagrams.

Researchers in the field have identified several master regulators in the lineage restriction of multipotent mesenchymal cells into chondrocytes and osteoblasts and of these are Sox9 and Runx2. Sox9 is the key factor driving chondrogenesis while Runx2 is the master regulator of osteoblast differentiation. Runx3 is another factor found to possibly play redundant roles with Runx2 in chondrocyte maturation. My study focuses on elucidating the gene regulatory network governing skeletogenesis that is centred on Runx2 and Runx3.

Currently, there are many known regulators and target genes of Runx2. However, the regulatory information is limited and it is believed that there are many more undiscovered factors controlled by Runx2. The expression profiling of Runx3 in the context of skeletal formation has also not been undertaken. By coupling mouse transgenic techniques with high throughput genomic studies such as gene expression profiling and chromatin immunoprecipitation followed by sequencing

(ChIP-Seq) for both Runx2 and Runx3, gene expression profiles for Runx2 and Runx3 using an enriched pool of *Runx2*- and/or *Runx3*-expressing cells isolated from fluorescing mouse embryos were generated. Additionally, Runx2-specific binding sites were mapped using HA_3-tagged mouse embryos and anti-HA antibody to identify Runx2 direct targets to complement the expression profiling data. This study is the first to reveal the vast number of factors controlled by both Runx2 and Runx3 cooperatively, antagonistically and uniquely which were partly validated by published data and by RNA section in-situ hybridization. As a result, a preliminary gene regulatory network centering Runx2 and Runx3 is established.

Finally, this work serves as a proof of principle that the strategy employed is feasible and can be built upon as a common platform to fabricate gene regulatory networks governing any developmental systems studied.

LIST OF TABLES AND FIGURES

CHAPTER 1 – INTRODUCTION

In any one cell of an organism, the traffic of events occurring is tightly regulated to ensure its proper development and survival. Though every cell contains identical genetic material, the genes are differentially expressed in each cell or group of cells at specific time-points in development. The dynamics of gene expression that occur within a group of cells as they change fate can be controlled at various levels – transcription, translation and stability of mRNA and/or protein. It involves multiple genes, transcription factors, miRNAs and even long non-coding RNAs. A complex organism's final form is thus the result of temporal and spatial control of gene expression through the interaction of various factors within complex regulatory networks.

In order to capture these dynamic events onto a visual network, many have tried to decipher the 'regulatory code' encoded within the genetic code, elucidating regulatory networks that govern cell fate specifications from pluripotency to tissue-specific states. In our case, we are trying to map the regulatory relationships that dictate how pluripotent mesenchymal stem cells differentiate into chondrocytes and osteoblasts during skeletal development onto a graphic gene regulatory network diagram. This enables visualization of the regulatory dynamics during cell fate commitments to gain a better understanding of skeletal development. Through that, any alterations in the network leading to skeletal diseases may be identified and developed as targets for bone regenerative therapy and tissue engineering.

1.1 Gene Regulatory Networks

The concept of using gene regulatory networks (GRNs) to represent regulatory interactions governing the development and well-being of organisms emerged more

than 40 years ago (Britten and Davidson, 1969; Kauffman, 1969). It has since taken on a systems-level approach to depict complex interactions between key regulatory genes, typically transcription factors (TFs), and the genes they regulate, known as target genes, across biological systems and species. So far, the most extensive developmental GRN built is of the endomesoderm development in the sea urchin (Davidson et al., 2002). Ongoing efforts are made with increasing accuracy to assemble comprehensive GRNs of other systems and species such as vulva development in the *Caenorhabditis elegans* (Inoue et al., 2005; Ririe et al., 2008), endomesoderm development in the *Xenopus* (Koide et al., 2005; Loose and Patient, 2004), and a variety of developmental systems in the *Drosophila* (Furlong, 2004; Silver et al., 2005; Stathopoulos and Levine, 2005).

Gene regulatory networks are usually depicted as diagrams containing nodes, which represent genes, and directional edges connecting the regulator to the target genes. The edges represent the functional links between the two connected genes and an edge that ends with an arrowhead means the target gene is activated by the TF and likewise one which ends with a short perpendicular line represents inhibitory action by the TF on the target gene.

There are two main components in a GRN. The first is a causal link between the functional state or abundance of the TF and the target gene's expression. The second is the relevant site on the genome which the TF binds to control the spatio-temporal expression of the target gene. This non-coding DNA sequences bound by the TF is often known as the "cis-regulatory module" (CRM) of the target gene (Levine and Davidson, 2005). The former is often elucidated by genetic perturbations of the TF of interest in an organism coupled with global expression profiling while the latter is unravelled by either genome-wide chromatin immunoprecipitation studies or bioinformatics *in silico* predictions of binding sites followed by electrophoretic mobility

shift assay (EMSA) verification experiments. As embryonic development is a dynamic process which presents a real challenge to constructing developmental GRNs, one also needs to consider a third aspect in the GRN – the spatio-temporal expression of the TF of interest (Wilczynski and Furlong, 2010) – in order to build a more accurate and comprehensive GRN that represents developmental events in a particular organ.

1.2 The Skeleton and Its Development

1.2.1 Function of the Skeleton

The skeleton is a pivotal organ in all vertebrates. Primarily, it serves as a scaffold to support the body and protect vital visceral organs as well as to allow sophisticated movements. However, this complex and multifunctional organ has roles that are not only limited to structural ones. It is also a home to haematopoiesis, a reserve of essential minerals (mainly calcium and phosphorus) for homeostatic functions and a detoxification site where toxins such as heavy metals are adsorbed (Lefebvre and Bhattaram, 2010). Furthermore, it has a more influential role as an endocrine organ, regulating phosphate metabolism via the release of fibroblast growth factor 23 (Fgf23) which acts on kidneys to reduce phosphate re-absorption and more recently found to contribute to blood sugar regulation and fat deposition through the secretion of a hormone called osteocalcin (Jensen et al., 2010; Lee and Karsenty, 2008).

1.2.2 Cellular Structure of the Skeleton

The strong and hard yet lightweight skeletal frame comprises two types of tissue – the cartilage and the bone – and is composed of three types of cells: chondrocytes in the cartilage and osteoblasts and osteoclasts in the bone. While chondrocytes and osteoblasts are derived from mesenchymal stem cells, osteoclasts have their origins traced back to the monocyte-macrophage cell lineage and have bone resorption functions (Bar-Shavit, 2007). The chondrocytes are embedded in an extracellular

cartilaginous matrix that is rich in proteoglycans, mainly aggrecans, and type II collagen whereas the bone matrix is abundant in type I collagen (a major constituent of osteoid), osteocalcin (also known as bone gamma-carboxyglutamic acid-containing protein, BGLAP) and bone sialoprotein. These bone matrix proteins aid the precipitation of the bone mineral hydroxyapatite which is the main component that confers hardness to the bone (Lefebvre and Bhattaram, 2010).

1.2.3 Formation of the Skeleton

Skeletogenesis can be described as two major phases. The first begins with migrations of multipotent ectoderm-derived (the neural crest) and mesoderm-derived mesenchymal cells to the sites where future bones form. The former give rise to craniofacial skeletal structures and the latter are divided into two parts. The lateral plate mesoderm gives rise to some craniofacial skeletal elements, the appendicular skeleton (limbs) and part of the axial skeleton (sternum) while the paraxial mesoderm gives rise to the remaining axial skeleton (vertebrae and ribs) via somitogenesis and sclerotome formation. Upon resting at these pre-skeletal sites, the mesenchymal cells condense to form skeletal templates of osteochondroprogenitor cells in the shape of the eventual skeletal elements. This process of migration and condensation is governed by numerous factors expressed in a spatial- and temporal-specific manner and is commonly known as skeletal patterning. Thereafter, the next phase, bone formation, commences (Karsenty et al., 2009; Lefebvre and Bhattaram, 2010).

1.2.3.1 Endochondral Ossification

During embryonic development, bone formation occurs by two distinct processes – endochondral ossification and intramembranous ossification. The former occurs in majority of the bones in the body especially the long bones of the axial and appendicular skeleton. It entails a two-step process where the condensed

mesenchymal cells in the centre first differentiate into early proliferating chondrocytes which produce copious amounts of type II collagen (Garofalo et al., 1993) and aggrecans – the foundational proteins of the extracellular matrix (ECM), to form a cartilage anlage. A thin layer of periphery cells, however, do not differentiate into chondrocytes. Instead, they remain as type I collagen-producing mesenchymal cells, flatten and aggregate to form a structure called the perichondrium that envelops the skeletal template. Upon the right cues, the chondrocytes at the core of the cartilage anlage stop proliferating and elongate to transiently form prehypertrophic chondrocytes before maturing fully into hypertrophic chondrocytes. The transient prehypertrophic chondrocytes continue to produce type II collagen as well as secrete a signalling molecule, Indian hedgehog (Ihh), which induces the adjacent perichondrial cells to differentiate into osteoblasts forming the periosteum (St-Jacques et al., 1999). The hypertrophic chondrocytes, on the other hand, exclusively synthesize type X collagen which provides an extracellular environment that is conducive for mineralization of the ECM. The hypertrophic chondrocytes further advance into terminal hypertrophic chondrocytes. The second step of endochondral ossification initiates with the apoptosis of the terminal hypertrophic chondrocytes and the vascular endothelial growth factor (VEGF)-induced capillary invasion of the mineralized cartilage template, concomitantly transporting osteoblast progenitor cells from the periosteum to replace the dying chondrocytes (Gerber et al., 1999; Zelzer et al., 2004). These osteoblasts along with osteoclasts secrete matrix metalloproteinases (Mmp9, 13, 14) to degrade the cartilage matrix (Stickens et al., 2004; Vu et al., 1998). Simultaneously, the osteoblasts secrete bone sialoprotein and type I collagen which contributes to the bone matrix in addition to the type X collagen-rich mineralized ECM previously laid down by the hypertrophic chondrocytes. This forms the bone shaft in the middle that is flanked by organized layers of chondrocytes at different maturation stages known as growth plates. Proliferation and hypertrophy of chondrocytes at the distal part of the growth plates

promote bone lengthening and subsequent chondrocyte apoptosis, vascularization and calcification events at the proximal ends complete the endochondral ossification with the progressive replacement of the cartilage by bone (Karsenty and Wagner, 2002; Kronenberg, 2003; Lefebvre and Bhattaram, 2010).

1.2.3.2 Intramembranous Ossification

A minority of the skeletal tissue, such as the flat bones of the skull and parts of the mandible and clavicle, are formed via the more straightforward process termed intramembranous ossification. In the course of intramembranous ossification, the condensed mesenchymal cells skip the cartilage intermediary and differentiate directly into osteoblasts progenitors. As these osteoprogenitors mature into osteoblasts, they secrete osteoid, a non-mineralized matrix made up of ninety percent type I collagen, osteocalcin and chondroitin sulphate to build up the initial bone scaffold. The osteoblasts subsequently express alkaline phosphatase (*Alp*) and bone sialoprotein (*Bsp*) to facilitate the mineralization of the osteoid matrix. Eventually, the osteoblasts get embedded in the calcified bone and differentiate into osteocytes (Hartmann, 2009; Jensen et al., 2010; Karsenty et al., 2009).

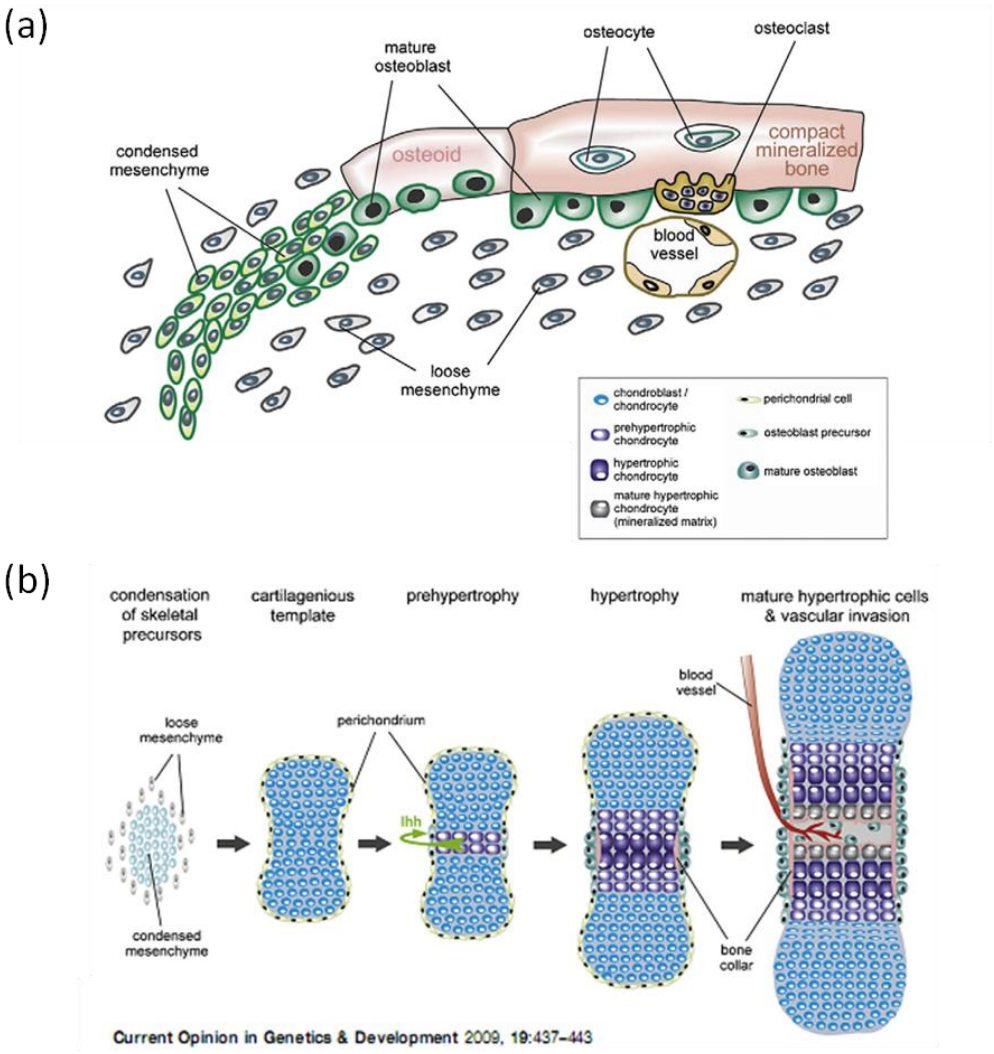


Figure 1. Bone Formation

(a) Intramembranous ossification occurs in the flat bones of the skull and the clavicle. Condensed mesenchymal cells differentiate directly into osteoblast progenitors and mature into osteoid-producing osteoblasts. (b) Endochondral ossification takes place in long bones. Condensed mesenchyme first differentiates into an intermediate cartilaginous template and matures into terminal hypertrophic chondrocytes before osteoblasts from the perichondrium invade the apoptotic chondrocytes via vascular invasion. Figure from (Hartmann, 2009).

1.3 Transcriptional Control of Skeletal Formation

Mesenchymal stem cells (MSCs) are the common precursors to a variety of tissue cell types: adipocytes (adipose), myocytes (muscle), chondrocytes (cartilage) and osteocytes (bone). Several significant transcription factors have been identified as critical regulators that govern each lineage commitment of the multipotent MSCs. For example, (1) Pparγ2 promotes adipogenesis; (2) MyoD expression is vital for myogenesis; (3) Sox9 is the key factor directing MSC down the chondrogenic path, and (4) Runx2 is the master regulator of bone formation. These factors were discovered from human diseases and mouse knockout studies and the severe relevant phenotype resulting from their absence earned them their key title in the respective tissue development pathways. As the thesis is focused on skeletal development, I will only expand the discussion on the last two points in the following segments.

Bone and cartilage tissue are different in their composition, structure and molecular regulation but their developments are highly interconnected and tightly coordinated. The discoveries of genetic mutations in several human skeletal diseases have sparked mouse genetic manipulation and recent genome wide expression studies that propagated our understanding of the transcriptional events that control skeletogenesis over the last two decades.

1.3.1 Chondrogenesis

One of the factors being manifested early in skeletal precursors and being critical for initiation of chondrogenesis is Sox9. It is a transcription factor containing a high mobility group (HMG)-box DNA-binding domain which bears homology to that of the mammalian sex-determining region Y factor, Sry. *Sox9* was first discovered when its inactivating mutation was found to cause the severe cartilage disease, Campomelic

dysplasia, and was later shown to regulate expression of cartilage proteins such as type II collagen and aggrecan (Bell et al., 1997; Wagner et al., 1994). Formation of the sclerotome, the initial pre-skeletal element, requires the paired-box members, Pax1 and Pax9 and the homeobox family members Nkx3.1 and Nkx3.2 (aka Bapx1) (Herbrand et al., 2002; Peters et al., 1999) in addition to Sox9 for prechondrogenic condensation of the mesenchymal cells and the subsequent chondrogenesis of the vertebrae. Both pairs of transcription factors Pax1 and Pax9, and Nkx3.1 and Nkx3.2 (Bapx1) play synergistic and redundant roles respectively in axial skeleton development (Herbrand et al., 2002; Peters et al., 1999). Pax1 and Pax9 activate *Bapx1* during Sox9-mediated chondrogenesis of the sclerotome (Rodrigo et al., 2003). Furthermore, the transactivating function of Sox9 is enhanced when bound by two other Sox family members that do not have the transactivation domains, Sox5 and Sox6. All three Sox proteins form the well-known Sox trio complex essential for proper differentiation of early proliferating chondrocyte but the exact mechanism of the Sox trio complex in chondrocyte differentiation is still not well understood (Akiyama et al., 2002; Lefebvre et al., 1998). Hypoxia inducible factor-1 (Hif1α) is another factor that supports chondrocyte survival through the up-regulation of Vegf to promote vascularisation of the developing bone (Schipani et al., 2001).

The core factors driving chondrocyte maturation into prehypertrophic and hypertrophic chondrocytes are runt-domain TFs Runx2 and Runx3 evident from the complete blockage of chondrocyte hypertrophy in the *Runx2*$^{-/-}$*Runx3*$^{-/-}$ mouse (Yoshida et al., 2004). Additional co-factors such as Grg5 and Hdac4 physically interact with Runx2 to enhance and inhibit its activity respectively during chondrocyte hypertrophy (Vega et al., 2004; Wang et al., 2004). Dlx5 and Dlx6 are reported to work synergistically with Runx2 to positively regulate chondrocyte hypertrophy (Chin et al., 2007; Roca et al., 2005). Additionally, basic helix-loop-helix TF Mef2c as well

as AP1 family member, Fra2 are required for chondrocyte maturation (Arnold et al., 2007; Karreth et al., 2004).

The mechanisms of terminal maturation are not well studied and so far only the basic leucine zipper protein cMaf is implicated in this process. cMaf is expressed in late hypertrophic chondrocytes and mice deficient in this gene have impaired terminal maturation of hypertrophic chondrocytes (MacLean et al., 2003).

1.3.2 Osteoblastogenesis

Runx2 is the crucial factor for the initial commitment of perichondrial cells and condensed mesenchymal anlagen of the intramembranous bones to osteoblast lineage cells (Komori et al., 1997; Otto et al., 1997). In endochondral ossification, the link between chondrocyte maturation and osteoblast differentiation hinges on Ihh signalling. While Runx2 regulates *Ihh* in the prehypertrophic chondrocytes, Ihh induces *Runx2* expression in the adjacent perichondrium (Karsenty, 2001). *Runx2* expression, however, is not sufficient for osteoblast differentiation reflected by the ectopic maturation of chondrocytes without any defects in osteoblast differentiation in transgenic mice constitutively expressing *Runx2* (Takeda et al., 2001). Further commitment of the *Runx2*-expressing osteoblast progenitor cells to fully committed osteoblasts in both endochondral and intramembranous bones requires a Krüppel-like zinc finger domain-containing transcription factor Sp7 (aka Osterix) (Nakashima et al., 2002). The activity of Osterix is enhanced through interaction with nuclear factor of activated T cells (Nfatc1) transcription factor (Koga et al., 2005).

Osteoblast maturation involves activating transcription factor 4 (ATF4), Sp-family of Krüppel-like zinc finger protein Sp3 and Fos-related antigen Fra1 (Eferl et al., 2004; Gollner et al., 2001; Yang and Karsenty, 2004). Atf4 determines the function of

osteoblast and mice deficient in Atf4 have delayed bone formation while Atf4 activity in early osteoblasts is inhibited by co-dimerizing with a nuclear leucine zipper protein, Fiat (Yang et al., 2004; Yu et al., 2009). The transactivation ability of Atf4 is increased through phosphorylation by Rsk2 kinase and through association with Satb2 (Dobreva et al., 2006; Yang et al., 2004). The exact mechanism of Sp3 in ossification is not known except that an *in vitro* study demonstrated that it binds to the promoter of receptor activator of nuclear factor kappa B ligand, RANKL, and regulates its basal promoter activity in osteoblasts (Liu et al., 2005b). RANKL is required for osteoclastogenesis hence Sp3 might be involved in osteoclast activity. Lastly, Fra1 up-regulates bone matrix proteins such as osteocalcin, col1a2 and matrix Gla protein (Mgp). Ectopic expression of Fra1 leads to increased bone mass in the transgenic mice while $Fra1^{\Delta/\Delta}$ mice lacking functional Fra1 developed osteopenia (Eferl et al., 2004).

The skeleton is the core organ in the vertebrate body and tremendous progress has been made in the past few decades to increase our knowledge in its development. Many crucial factors influencing chondrogenesis or osteoblastogenesis that have been linked to human skeletal diseases have been discovered and these diverse factors often work together in a highly organized and cooperative network to propagate bone formation. Some light has also been shed on the molecular mechanisms behind the cell type specification and maintenance. However, the spatio-temporal expression and activity of many more factors and the complex regulatory mechanisms among the factors cooperating to bring about the complex but finely tuned events during bone formation is still a huge piece of information missing. Evidently, there is still a lot more to uncover in order to unravel the transcriptional network controlling embryonic skeletogenesis.

1.4 The Runx Family of Proteins

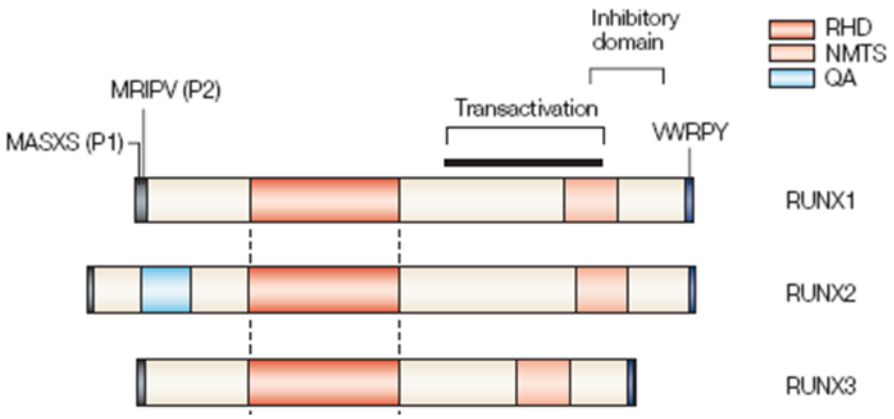


Figure 2. The Structure of Runx Proteins

Figure from (Blyth et al., 2005). Runx proteins share a highly conserved runt domain and are transcribed from two distinct promoters P1 and P2. RHD, runt homology domain; NMTS, nuclear-matrix-targeting signal; QA, Glutamine-alanine repeat domain of Runx2.

The *Runx* family of genes encodes for transcription factors that contain the characteristic DNA-binding runt domain which derived its name from the *Drosophila* pair-rule gene, *runt,* owing to the high degree of homology between the two sequences (Gergen and Butler, 1988). This highly conserved 128-amino-acid runt motif found proximal to the N-terminus has functions in (1) DNA binding, recognizing a canonical DNA motif TGPyGGTPy (where Py refers to pyrimidine) (Crute et al., 1996), (2) protein-protein interactions (Kagoshima et al., 1993) and (3) nuclear import that is in addition to the conserved nuclear matrix-targeting signal (NMTS) in the C-terminus (Kanno et al., 1998). On the other hand, the proline, serine and threonine (PST)-rich C-terminal portion of Runx proteins is responsible for transcriptional activation and repression, and regulation of DNA-binding affinity of the runt domain (Gu et al., 2000; Kim et al., 1999). The final five conserved amino acids VWRPY serve as an interaction motif mainly to recruit the Groucho/TLE family of

transcriptional co-repressors. The Runx proteins, consisting of only the α-subunit, position at DNA binding sites as heterodimers with a non DNA-binding β-subunit protein, Cbfβ. This cofactor association at the runt domain enhances the DNA-binding ability of the Runx proteins (Cohen, 2001; Liu et al., 2006). Another defining feature of the *Runx* gene is the presence of two alternative promoters which give rise to multiple protein isoforms with either of the two distinct N-terminal start sequences: (1) MASNS from the distal P1 promoter generating what is known as type II isoform proteins and (2) MRIPV from the proximal P2 promoter yielding type I isoform proteins (Bangsow et al., 2001; Levanon et al., 2001; Park et al., 2001). Alternative splicing further augments the diversity of *Runx* gene products.

The stability and activity of Runx proteins are modulated at the post-translational level via phosphorylation and acetylation with the former modification conferring activating functions to the protein. The Runx proteins can also engage co-activators or co-repressors at enhancer or promoter regions of target genes to initiate or inhibit a cascade of transcriptional events during development. Therefore, Runx proteins serve as transcriptional coordinators in a complex regulatory network directly and indirectly governing the expression of a myriad of genes pertaining to a broad spectrum of cellular and molecular functions.

There are three members in the mammalian *Runx* gene family commonly known as *Runx1*, *Runx2* and *Runx3* [refer to (van Wijnen et al., 2004) for the alternative nomenclature]. Despite the extent of semblance in their sequence and structure, each of the three *Runx* genes has a distinct primary biological role in mammalian development elucidated from the overt phenotypes manifested in the respective knockout mice. Runx1 is essential in definitive haematopoiesis as *Runx1*^{null} mice die early at E12.5 from blood development failure in the fetal liver (Okuda et al., 1996).

Runx2 is crucial for osteoblast differentiation in bone development as *Runx2*^{null} mice die at birth with a non-ossified skeleton (Komori et al., 1997; Otto et al., 1997). Runx3 has a primary role in neurogenesis apparent from the display of severe limb ataxia in *Runx3*^{null} mice due to impaired development of proprioceptive neurons in the dorsal root ganglia (DRG) (Inoue et al., 2002; Levanon et al., 2002). In another independent study (Li et al., 2002), the *Runx3*^{null} mice died of starvation shortly after birth as a result of hyperplasia in the gastric mucosa. This led to the implication of Runx3 as a tumour suppressor in gastric cancer cells. Evidently, each of the *Runx* genes is a critical lineage determinant of blood, skeletal, neuronal cells and gut development respectively.

Although the individual *Runx* genes have distinct primary roles, it is no surprise that they also have overlapping functions since all the members of the Runx family can bind to the same nucleotide sequence. Particularly, there are two highly conserved, in sequence and location, Runx binding sites at the 5' UTR of all three *Runx* genes (Drissi et al., 2000) which hint at the possible cross regulation amongst the Runx proteins. Furthermore, their expression domains often coincide at specific tissues (in either the same cells or different cells of the same tissue) and developmental stages, with some variations in their expression levels. This suggests that they function synergistically, redundantly or even complementarily to specify a population of cells into a particular type, with one usually taking on as the major player and the other a supportive or redundant role. For example, (1) *Runx1* and *Runx3* are both expressed in the developing DRG and mature T-cells but are expressed in different groups of sensory neurons in the DRG and at different stages of T-cell development (Kramer et al., 2006; Zhong et al., 2006). Moreover, the requirement of Runx1 in the lymphoid system always takes precedence over Runx3 (de Bruijn and Speck, 2004). A study by (Brady and Farrell, 2009) probing into the mutually exclusive expression of *Runx1* in mature resting B-cells and *Runx3* in proliferating B-cells revealed that the former is

repressed by the latter in proliferating B-cells. This is one prototype of the *Runx* genes working in complementary to define a subpopulation of cells. (2) *Runx1*, *Runx2* and *Runx3* expressions are found at variable intensity, extent and tempo in skeletal tissues (Levanon et al., 2001; Stein et al., 2004). (a) In a particular study (Yoshida et al., 2004), *Runx2* and *Runx3* expressions were noted to coincide in the prehypertrophic and hypertrophic chondrocytes. *Runx2*$^{-/-}$*Runx3*$^{-/-}$ mice were also observed to have a more severe skeletal phenotype than *Runx2*$^{-/-}$ or *Runx3*$^{-/-}$ mice. Notably, the *Runx2*$^{-/-}$*Runx3*$^{-/-}$ mice were devoid of large columnar cells typical of hypertrophic chondrocytes in the limbs at E18.5. These observations suggest that Runx2 and Runx3 play compensatory roles in chondrocyte maturation during endochondral ossification. However, Runx2 plays a bigger role in advancing chondrocyte maturation than Runx3 as chondrocyte maturation was more impeded in *Runx2*$^{-/-}$ mice than in *Runx3*$^{-/-}$ mice (Yoshida et al., 2004). (b) Recently, it was reported that Runx1 and Runx2 work together acting via *Sox5* and *Sox6* during sternal development as demonstrated by the absence of a sternum in the mouse when both Runx1 (conditionally knocked out by *Prx1-Cre* transgene) and Runx2 were abrogated (Kimura et al., 2010). This loss was not observed in either the *Prx1*-conditional *Runx1*$^{-/-}$ or *Runx2*$^{-/-}$ mice.

While early mouse knockout experiments delineated the major roles of the individual Runx members (Ducy et al., 1997; Inoue et al., 2002; Komori et al., 1997; Li et al., 2002; Okuda et al., 1996; Otto et al., 1997), subsequent studies gradually uncovered more merging and cooperative functions of the Runx proteins (de Bruijn and Speck, 2004; Kimura et al., 2010; Stein et al., 2004; Yoshida et al., 2004). However, the mechanisms of these overlapping and compensatory roles have not been fully elucidated and the scope of functional redundancy among the Runx proteins remains to be discovered, hence, part of my PhD research will attempt to deal with that.

Apart from cell lineage determination during foetal development, *Runx* genes are also implicated in human cancer development because RUNX-CBFβ complexes regulate scores of genes pertinent to cell-fate decisions which are compromised in cancer cells; decisions to promote apoptosis or proliferation and differentiation or self-renewal. Most genes are easily classified as tumour suppressors or oncogenes but both conflicting functions have been associated with Runx genes in cancer cells. This dual function is not surprising since RUNX–CBFβ complexes can activate as well as repress the expression of vital regulators of cell proliferation and differentiation. Hence, deregulation of the normally balanced roles of RUNX is associated with acute myeloid leukemia (Runx1) (Song et al., 1999), osteosarcoma (Martin et al., 2011) and metastatic bone disease of breast and prostate cancer (Runx2) (Pratap et al., 2011) and gastric cancer (Runx3) (Li et al., 2002). Therefore, a more comprehensive knowledge of the factors and pathways regulated by the RUNX-CBFβ heterodimeric complex in normal conditions is essential to progress applications that will benefit cancer diagnosis and treatments.

The *Runx* genes are largely conserved in sequence and structure between the *Homo sapiens* and the *Mus musculus* and human diseases caused by *RUNX* mutations are also recapitulated in the mouse when the respective mouse *Runx* genes are deleted. Hence, the elucidation of the GRN governed by each Runx transcription factors during mouse development has the potential to contribute to developing applications in human gene therapy and drug development for the treatment of human developmental diseases and cancer.

In the next two major sub-chapters, I will give a more detailed introduction of Runx2 and Runx3 as I have decided to focus on only these two Runx members in my research into elucidating the GRN controlling embryonic skeletal development.

1.4.1 RUNX2

The mouse *Runx2* gene is found on chromosome 17 and the translated Runx2 protein contains a short exclusive region of glutamine/alanin (QA) repeats at the N-terminus that is absent in its co-orthologs. This stretch of QA repeats regulates its transactivation and heterodimerization activity (Thirunavukkarasu et al., 1998). Like the other *Runx* members, it has two distinct promoters, P1 and P2 (Bangsow et al., 2001; Miyoshi et al., 1995; Xiao et al., 1998), which give rise to two major transcript variants. In addition to the two unique promoters, alternative splicing of the eight protein-coding exons generates at least nine protein isoforms, two of which are dominant in skeletal tissues: type II isoform starting with MASNS that is derived from the osteoblast-specific P1 promoter (Geoffroy et al., 2002; Ogawa et al., 2000; Stewart et al., 1997; Xiao et al., 1999; Xiao et al., 2001) and type I isoform beginning with MRIPV that is expressed from the chondrocyte-specific P2 promoter (Takeda et al., 2001). The two proteins, which differ only in 19 amino acids at the N-terminus, otherwise bear the same functional domains, interact with similar co-factors and are capable of transactivating similar target genes *in vitro* (Banerjee et al., 2001; Harada et al., 1999; Javed et al., 2001). However, their expression sites are not identical. Type I *Runx2* expression is more widespread. It is predominantly expressed in the perichondrium and periosteum but low levels of expression can be found in a variety of tissues including the thymus, proliferating chondrocytes of the cartilage and suture tissue of the calvarium. Type II *Runx2* is more restricted to the bone tissues and is intensely expressed in the hypertrophic chondrocytes and mature osteoblasts but not in the perichondrium (Banerjee et al., 1997; Choi et al., 2002; Enomoto et al., 2000; Park et al., 2001; Ueta et al., 2001). The distinct spatiotemporal expressions of the two protein isoforms can be attributed to the different cis-regulatory elements situated at the P1 and P2 promoter regions which allow different factors to discriminately regulate them (Gaur et al., 2005; Hassan et al., 2006; Xiao et al., 2008). The equal

functional capabilities of the two isoforms together with their differentially regulated expression at various sites may contribute to a more specific yet robust control of skeletal development. As I will not be investigating the differential roles of the two Runx2 isoforms, I will refer to both isoforms as Runx2.

1.4.1.1 Spatiotemporal Expression of Runx2 in the Developing Skeleton

Runx2 expression is first detected in the developing mouse embryo at embryonic day 10.5 (E10.5) at the mesenchyme anterior of the forelimb bud which prefigures the shoulder bones (Stricker et al., 2002). At E11.5, *Runx2* is strongly expressed in the first and second brachial arches, which will eventually give rise to the maxilla and mandible, and the condensed mesenchyme of the ulna (Stricker et al., 2002). At E12.5, *Runx2* expression can be detected in most of the developing skull, axial and appendicular skeleton including the condensed mesenchymal cells of the vertebral perichondrium, the chondrocytes of the Meckel's cartilage and the prechondrogenic mesenchyme of the limbs (Kaufman et al., 1992). At this stage, the osteochondro-progenitor cells have the potential to either differentiate into chondrocytes or osteoblasts (Ducy et al., 1997). Ossification begins at E13.0 and *Runx2* expression progresses to the digits of the limbs by E13.5 (Stricker et al., 2002). The first signs of osteoblasts can be detected at E14.5 (Kaufman et al., 1992) and *Runx2* is expressed transiently in prehypertrophic and hypertrophic chondrocytes (Takeda et al., 2001) as well as osteoblasts at that stage. *Runx2* expression is abundant in the frontal, nasal, baso-occipital, baso-sphenoid and hyoid bones and the mandible but is absent in the chondrocytes of the Meckel's cartilage of the E16.0 mouse embryo (Ducy et al., 1997; Otto et al., 1997). By E16.5, *Runx2* expression in the maturing chondrocytes is minimal or absent while its expression remains high in the perichondrium and the osteoblast progenitors (Bialek et al., 2004; Ducy et al., 1997; Hinoi et al., 2006; Karsenty, 2008).

1.4.1.2 Runx2 Phenotype, Function and Associated Skeletal Diseases

Runx2 was identified as the earliest master driver of osteoblast differentiation in both intramembranous and endochondral ossification through genetic analyses of human skeletal dysplasias and studies of genetically modified mouse models (Ducy 1997, Komori 1997, 2002, Otto 1997). In those mouse studies, the *Runx2*$^{+/-}$ mice appeared normal but on closer examination revealed a core defect in intramembranous ossification characterized by hypoplastic clavicles and delayed fusion of the cranial fontanelles. These abnormalities reflected some of the symptoms in the autosomal dominant human skeletal disorder, cleidocranial dysplasia (CCD) (Lee et al., 1997; Mundlos et al., 1997; Otto et al., 1997). The *Runx2*$^{-/-}$ mice died from respiratory failure shortly after birth owing to the inability to respire from a non-osseous rib cage. The mutant mice were clearly smaller with shorter limbs and a foreshortened snout and were virtually devoid of a mineralised skeleton. Histological analysis of all the bones showed absence of osteoblasts and a lack of bone-associated marrow precursors while chondrocytes were still present. This demonstrated that Runx2 is essential for osteoblast differentiation and vascularisation of the bone and has no positive regulatory functions in chondrocyte differentiation and proliferation. Although *Runx2* gene deletion has an impact on both intramembranous and endochondral ossification, the former appears more sensitive to Runx2 deficiency.

Apart from the osteoblast-deficient phenotype, there was also a lack of hypertrophic chondrocytes in some but not all skeletal tissues of the *Runx2*$^{-/-}$ mice denoted by a drastic reduction in *type X collagen* expression levels (Inada et al., 1999; Kim et al., 1999). Skeletal elements such as the tibia, fibula, radius and ulna, where hypertrophic chondrocytes expressing *type X collagen* were still detected, however, did not express proteins characteristic of terminal hypertrophic chondrocytes such as secreted phosphoprotein 1 (Spp1) and bone sialoprotein (Bsp) (Inada et al., 1999; Kim et al., 1999). In another study, over-expression of *Runx2* under the control of

chondrocyte-specific *type II collagen* promoter augmented *type X collagen* expression and boosted chondrocyte maturation in mice whereas the expression of dominant-negative *Runx2* under the same promoter reduced *type X collagen* expression in chondrocytes and inhibited chondrocyte maturation (Ueta et al., 2001). These studies indicate that Runx2 is involved in chondrocyte maturation.

Similar studies in the osteoblast lineage yielded results that were contrary to expectations. Transgenic mice over-expressing *Runx2* under the control of osteoblast-specific *type I collagen* promoter exhibited low bone mass (osteopenia) with multiple fractures owing to restricted mineralization and heightened osteoclastogenic activity which resulted in high bone resorption (Geoffroy et al., 2002; Kanatani et al., 2006; Liu et al., 2001). The osteoblasts in these mice were immature and the number of terminally differentiated osteoblasts and osteocytes detected were greatly reduced. Conversely, transgenic mice expressing dominant-negative *Runx2* under the same promoter yielded higher bone mass and were able to prevent ovariactomy-induced bone loss in the mice (Maruyama et al., 2007).

Taken together, Runx2 has dual functions – osteoblast differentiation and chondrocyte maturation. While an adequate level of Runx2 is required for osteoblast differentiation and maintenance, excessive Runx2 activity in the osteoblast progenitors increases osteoblast-directed osteoclastogenesis disproportionately. Therefore, a tight control of *Runx2* expression level is of paramount importance to proper skeletal formation and maintenance. Notably, Runx2 exerts opposing control on chondrocyte maturation, its secondary role. Its early expression (E12.5) in the proliferating chondrocytes propagates Ihh-dependent chondrocyte hypertrophy (Yoshida et al., 2004). Conversely, its constitutive expression in the perichondrium, coupled with its diminishing expression in the chondrocytes, at later stages impedes chondrocyte hypertrophy via the up-regulation of *Fgf18* (Hinoi et al., 2006). In this

manner, Runx2 prepares the skeletal elements for subsequent endochondral ossification events and thus prevents ectopic bone formation and dwarfism. Hence, in addition to a balanced level of *Runx2* expression, a precise regulation of its spatio-temporal expression is vital for proper bone formation. Furthermore, the restricted effect of *Runx2* deletion on chondrocyte maturation is probably attributed to the compensatory actions of Runx3 since genetic studies have shown that Runx3 is also important for chondrocyte hypertrophy and mice lacking both Runx2 and Runx3 have no hypertrophic chondrocytes or type X collagen expression (Yoshida et al., 2004).

1.4.1.3 Upstream, Downstream and Co-regulators of Runx2

At the molecular level, *Runx2* expression is regulated by a number of different molecules at different stages of skeletogenesis. At the onset of *Runx2* expression, Twist proteins repress *Runx2* at the post-transcriptional level to prevent premature osteoblast differentiation which accounted for the early expression of *Runx2* preceding the appearance of osteoblasts by 4 days (Bialek et al., 2004). Hdac3 and Hdac7 associate physically with Runx2 and act as co-repressors to negatively regulate osteoblastogenesis (Jensen et al., 2008; Westendorf, 2006). Stat1 inhibits the nuclear translocation of Runx2 while Shn3 mediates the degradation of Runx2 thereby inhibiting the activity of Runx2 in osteoblast differentiation (Jones et al., 2006; Kim et al., 2003). At the transcriptional level, Sox8 strongly reduces the expression of *Runx2* again to avert precocious osteoblast differentiation (Schmidt et al., 2005) while *Bapx1*, upon up-regulation by Sox9, directly represses *Runx2* in cells that are destined for chondrocytic lineage (Lengner et al., 2005; Yamashita et al., 2009). On the contrary, Bapx1 was shown to positively regulate *Runx2* expression exclusively in osteochondrogenic progenitor cells prefiguring the vertebral column (Tribioli and Lufkin, 1999). Other transcription factors that likely regulate *Runx2* are Dlx5 and Msx2 which activate the expression of *Runx2* thus promoting osteoblast differentiation, however, Msx2 takes on an opposite role during osteoblast maturation

decreasing the activity of Runx2 and inhibiting the maturation of osteoblasts (Holleville et al., 2007; Lee et al., 2005; Shirakabe et al., 2001).

Subsequent to the expression of *Runx2* in osteoprogenitors, the osteoblast lineage is reinforced by the expression of *Sp7*, also known as Osterix (Nakashima et al., 2002). The regulation of *Osterix* in osteoblasts is both Runx2-dependent and Runx2-independent via Msx2 (Matsubara et al., 2008) or the recently discovered Osterix master, Xbp1 (Tohmonda et al., 2011). In osteoblasts, Runx2 directly up-regulates bone matrix protein genes such as *Col1a1*, *Col1a2*, *Spp1* (aka *Opn*, *Osteopontin*), *Ibsp*, *Bglap* (aka *Ocn*, *Osteocalcin*) and *Fn1* (*Fibronectin1*) (Banerjee et al., 1997; Ducy et al., 1997; Harada et al., 1999; Jimenez et al., 1999; Kern et al., 2001; Lee et al., 2000; Sato et al., 1998). The regulation of *osteocalcin* expression by Runx2 is enhanced by association with Satb2 and Atf4/Creb2 and inhibited by Dlx3 and Msx2 (Dobreva et al., 2006; Hassan et al., 2004; Xiao et al., 2005). However, in an *in vivo* study by Maruyama *et al*, 2007, it was observed that osteocalcin expression was not affected in a dominant-negative *Runx2* transgenic mouse model. Hence, there is a need to verify the *in vitro* interactions at the osteocalcin promoter using an *in vivo* model (Hartmann, 2009).

Apart from osteoblastogenesis, Runx2 also promotes chondrocyte maturation and this is indirectly inhibited by the Sox trio (Sox9, Sox5 and Sox6) through the reduction of *Runx2* transcript levels (Saito et al., 2007). Grg5, a Groucho-family member, interacts with Runx2 as a positive co-factor while Dlx5 and Dlx6 interact cooperatively with Runx2 to enhance chondrocyte hypertrophy (Chin et al., 2007; Roca et al., 2005; Wang et al., 2004). Hdac4 associates with Runx2 and inhibits both its activity and its transcription to hinder chondrocyte advancement (Vega et al., 2004). Shox2 and Mef2c, on the other hand, are upstream positive regulators of

Runx2 to progress chondrocyte maturation (Arnold et al., 2007; Cobb et al., 2006). Runx2 positively regulates the expression of *Col10a1* in hypertrophic chondrocytes (Drissi et al., 2003; Enomoto et al., 2000; Higashikawa et al., 2009; Zheng et al., 2003) and *Spp1*, *Ibsp* and *Mmp13* in terminal chondrocytes (Hess et al., 2001; Jimenez et al., 1999; Porte et al., 1999; Selvamurugan et al., 2000). Runx2 works synergistically with Ets1 to directly regulate *Spp1* (Sato et al., 1998). Runx2 also interacts with other activating transcription factors such as Smads, C/EBP, Rb and with the transcriptional repressor Tle (Gutierrez et al., 2002; Javed et al., 2000; McCarthy et al., 2000; Thomas et al., 2001; Zhang et al., 2000). In addition, Runx2 up-regulates the expression of *Fgf18* in the perichondrium which encodes for a secreted molecule that activates Fgfr3 to inhibit chondrocyte hypertrophy (Hinoi et al., 2006).

Chondrocyte maturation and osteoblast differentiation are linked by Ihh signalling. Ihh secretion by prehypertrophic chondrocytes enhances the expression of *Runx2* in the perichondrium to promote osteoblast differentiation via the up-regulation of *Gli2* expression (Shimoyama et al., 2007) and the inhibition of the repressor form of *Gli3* which hampers DNA-binding ability of Runx2 (Ohba et al., 2008). In hypertrophic chondrocytes, Runx2 was found to directly up-regulate *Ihh* expression (Yoshida et al., 2004) which in turn induces chondrocyte proliferation and inhibits chondrocyte maturation through the induction of parathyroid hormone related peptide, *Pthrp*. Pthrp inhibits *Runx2* expression through the PKA signalling pathway (Iwamoto et al., 2003; Li et al., 2004b; St-Jacques et al., 1999; Vortkamp et al., 1996) thus forming a negative feedback loop to prevent premature chondrocyte hypertrophy. Hence, while Ihh was found to stimulate osteoblast differentiation in the perichondrium by directly up-regulating *Runx2* expression, Runx2 was found to directly induce *Ihh* in the growth plate to promote chondrocyte proliferation and delay chondrocyte maturation via a negative feedback.

Runx2 is transcriptionally activated by several other signalling molecules such as bone morphogenetic proteins (BMPs), fibroblast growth factors (FGFs), and retinoic acid (RA), and inhibited by 1,25(OH)2D3 (vitamin D_3) and tumor necrosis factor-α (TNF-α) (D'Souza et al., 1999; Drissi et al., 2002; Ducy et al., 1997; Gilbert et al., 2002; Jimenez et al., 2001). Another pathway linked to Runx2 is Wnt signalling. It was first implicated in osteoblastogenesis when *Wnt5a*, *Wnt5b* and *Wnt4* were found to be expressed in the perichondrium, subpopulation of the prehypertrophic chondrocytes and cells in the joint of the chicken limb bone respectively (Hartmann and Tabin, 2000). It was also later demonstrated that the canonical Wnt/β-catenin signalling was required to push skeletal precursors towards osteoblast differentiation while the lack of β-catenin resulted in chondrocyte formation instead (Hill et al., 2005). In one study, Runx2 was found to be a direct target of the canonical Wnt/β-catenin signalling pathway and that the promoter of *Runx2* was directly stimulated by Tcf/Lef complex to promote osteogenesis (Gaur et al., 2005). In another study, Runx2 was demonstrated to form a ternary complex with Tcf/Lef at the promoter of Fgf18 to induce bone formation (Reinhold and Naski, 2007). Combining the two findings, *Runx2* may well be transcriptionally activated by the canonical Wnt/β-catenin signalling pathway and subsequently work synergistically with it via Fgf18 to trigger bone formation.

Lastly, six *Runx* binding motifs have been found in its promoter and 5'UTR sequences which suggests that Runx2 auto-regulates itself. There are opposing data demonstrating that Runx2 positively (Ducy et al., 1999) as well as negatively regulates (Drissi et al., 2000) itself during osteoblastogenesis. Perhaps, Runx2 autoregulates its variant isoforms uniquely depending on the cellular context and the

stage of skeletal development and this remains to be examined in an *in vivo* model. It is also unclear if the Runx proteins cross regulate each other.

1.4.2 RUNX3

Mammalian *Runx3*, the smallest member of the *Runx* family with only 6 exons and spanning a small stretch of 67kb, is located on the human and mouse chromosomes 1p36.1 and 4 respectively (Avraham et al., 1995; Bae et al., 1995; Calabi et al., 1995; Levanon and Groner, 2004). Among the *Runx* members, it has the highest occurrence of ancient mammalian-wide interspersed repeats (MIRs) which led to the speculation that *Runx3* is the most primitive in the mammalian *Runx* family (Bangsow et al., 2001). With no exceptions, *Runx3* is translated from two distinct promoters, P1 (MASNS) and P2 (MRIPV). Similar to *Runx2*, the two promoter regions harbour a different repertoire of DNA binding motifs which enable differential regulation of the two major Runx3 isoforms (Bangsow et al., 2001). For instance, the P1 promoter contains more T- and B-cell-specific transcription factor binding sites such as Ikaros, Ets, CREB/ATF and an E-box while the P2 promoter, being more GC-rich, contains binding sites for Sp1 and Egr-1 (Leiden and Thompson, 1994; O'Riordan and Grosschedl, 2000).

1.4.2.1 Spatiotemporal Expression of Runx3 in the Developing Skeleton

Runx3 is abundantly expressed in the haematopoietic system particularly in the spleen, thymus and blood of an adult mouse (Bangsow et al., 2001; Le et al., 1999; Levanon et al., 1996; Levanon et al., 1994). The first expressions of *Runx3* are detected at E10.5 in the haematopoietic precursors in the liver, cranial trigeminal ganglia and dorsal root ganglia of a developing mouse embryo (Levanon et al., 2001; Stricker et al., 2002). At E11.5, *Runx3* transcripts start to appear in the mesenchymal condensations of the developing mouse limb. Subsequently, its expression progresses to the perichondrium of the metacarpals, the chondrocytic condensations

in the distal phalanges of the mouse digits as well as the vertebrae by E13.5 (Stricker et al., 2002; Yoshida et al., 2004). By E15.5, *Runx3* is manifested in most cartilaginous elements of the appendicular and axial skeleton such as the scapulae, limbs, ribs, pelvic bones and the vertebrae (Yoshida et al., 2004). *Runx3* expression first coincides with that of *Runx2* in the limb condensation at E11.5 and in the digits, ribs and vertebrae by E13.5. Expression of *Runx3* is found primarily in the prehypertrophic and hypertrophic chondrocytes and to a smaller extent in the perichondrium of the developing mouse endochondral bone (Levanon et al., 2001; Stricker et al., 2002; Yoshida et al., 2004).

1.4.2.2 Runx3 Phenotype, Function and Associated Skeletal Diseases

Runx3 is involved in both neuronal (Inoue et al., 2002; Levanon et al., 2002) and T-cell development (Taniuchi et al., 2002; Woolf et al., 2003) however the severe limb ataxia phenotype of *Runx3* knockout mice due to the failure of proprioceptive neurons to develop in the dorsal root ganglia suggest that Runx3 plays a primary role in neurogenesis. Furthermore, Runx3 was implicated as a tumour suppressor in gastric cancer in a separate *Runx3* mouse knockout study where it was demonstrated that the loss of functional Runx3 resulted in excessive growth of the gastric endothelial cells causing the mice to die of starvation shortly after birth (Li et al., 2002).

In addition to gastric cancer, *Runx3* is often found deleted or inactivated in cancers of the colon, lung, bladder, bile duct, pancreas, liver, prostate and breast (Goel et al., 2004; Ito et al., 2008; Kang et al., 2004; Kim et al., 2005; Lau et al., 2006; Li et al., 2004a; Wada et al., 2004; Xiao and Liu, 2004). Besides the involvement in neurogenesis, thymopoiesis and cancers, Runx3 also cooperates with Runx2 to contribute to chondrocyte maturation evident from the complete lack of distinguishable hypertrophic chondrocytes or the expression of the hypertrophic

chondrocyte marker, *Col10a1*, in the skeleton of a *Runx2*$^{-/-}$*Runx3*$^{-/-}$ mouse embryo (Soung do et al., 2007; Yoshida et al., 2004).

1.4.2.3 Transcriptional Regulation by Runx3

Of the three *Runx* genes, *Runx3* is the least studied especially in the context of skeletogenesis. In an *in vitro* study (Otto et al., 2003), Runx3 was found to positively regulate *Col10a1* and *Runx2* transcripts. Runx3 was also observed to regulate *Ihh* indirectly in hypertrophic chondrocytes (Yoshida et al., 2004). In gastric cancers, Runx3 associates with Smads to activate *p21* expression and up-regulates *Bim*, an apoptotic gene, to propagate Tgfβ-induced apoptosis and tumour suppression (Chi et al., 2005; Yano et al., 2006). There was no previous expression profiling or binding sites data of Runx3. The mechanisms underlying cooperative regulation among the Runx family members during chondrocyte and osteoblast development are yet to be determined.

1.5 Research Aims, Strategy and Significance

The main research objective of our laboratory is to elucidate the complex spatial-temporal interactions, amongst regulatory proteins and between transcription factors and the genetic material, that occur during the commitment of multipotent mesenchymal stem cells (MSCs) towards the osteo-chondrogenic lineage in a developing mammalian embryo. By coupling conventional mouse genetics with high-throughput genomic technologies such as microarray and chromatin immunoprecipitation-sequencing (ChIP-Seq), we pursue the grand goal of building gene regulatory networks (GRNs) that govern embryonic skeletal development.

In our research, we mainly employ the use of mice as they are ideal mammalian models owing to their (1) small size, (2) short generation time, (3) prolific breeding

and (4) the ease of handling them. Furthermore, the availability of mouse embryonic stem cells (ESCs) and well-established gene-targeting protocols in the mouse make them an excellent species for genetic studies. And in our case, human skeletal diseases are often recapitulated in the relevant gene knockout mice; insights and knowledge gleaned from mouse studies are thus applicable and beneficial in developing drugs and genetic therapies for human skeletal disorders.

From the literature review in this introduction chapter, it is noted that there has been much progress in our understanding of the molecular events that take place during embryonic bone formation and already an extensive number of transcription factors, signalling molecules and hormones have been identified to be implicated in the process. However, there is still a huge chunk of information missing to date such as the regulatory relationships among these identified factors and the yet numerous more undiscovered factors involved in skeletal lineage specification. Therefore, my specific aim is to unravel a part of the gene regulatory network controlling embryonic skeletogenesis by focusing on two *Runx* family genes – *Runx2* and *Runx3*. The former is a key factor driving osteoblastogenesis while both have functions in chondrocyte maturation. I chose to work with only these two *Runx* genes because they play a bigger role in skeletogenesis.

The strategy I have undertaken to elucidate a portion of the GRN is two-prong. The first is gene expression profiling of Runx2 and Runx3 (comparing knockout against wild-type profiles) to uncover their individual downstream targets. The second is ChIP-Seq to differentiate the direct targets from the indirect targets that are identified through the microarray analyses. Ultimately, by means of bioinformatics to make sense of the massive amount of complex microarray and ChIP-Seq data generated, I seek to answer these questions:

1. What other undiscovered factors are under the influence of Runx2 and Runx3 during skeletogenesis?
2. Which of those are the primary targets of Runx2 and Runx3?
3. What is the molecular mechanism behind the possible compensatory roles of Runx2 and Runx3?

Finally, if we can map out these regulatory relationships onto graphic skeletal GRN diagrams, it will enable us to visualize the regulatory dynamics that can be interrogated, not only to gain insight into bone development but also to understand human skeletal diseases as a consequence of network perturbations and in turn allow identification of targets for therapeutic intervention, regenerative therapy and tissue engineering.

CHAPTER 2 – MATERIALS AND METHODS

2.1 BAC Modification and Subcloning

Mouse bacterial artificial chromosomes (BACs) containing a partial or the entire genomic locus of *Runx2* and *Runx3* (purchased from BACPAC Resource Centre, CHORI) were used for modification using the Gene Bridges Quick and Easy BAC Modification kit via the Red/ET recombineering technology according to the manufacturer's protocol. Appropriate 50 bp homology arms were added to both ends of the insertion cassette (*EGFP-loxP-PGKgb2-Neo-loxP*, *F2A-EGFP-Frt-PGKgb2-Neo-Frt* or *HA_3-TGA-loxP-PGKgb2-Neo-loxP*) via PCR. DH10B *E.coli* strain hosting the appropriate BAC was first transformed with *pRed/ET* and selected with chloramphenicol (12.5 µg/mL) and tetracycline (3 µg/mL) antibiotics before a second transformation of the insertion cassette flanked with 50 bp homology arms into the *pRed/ET*-containing BAC clones for homologous recombination (Fig. 3). The positive recombinants were selected with chloramphenicol and kanamycin antibiotics and screened by colony PCR.

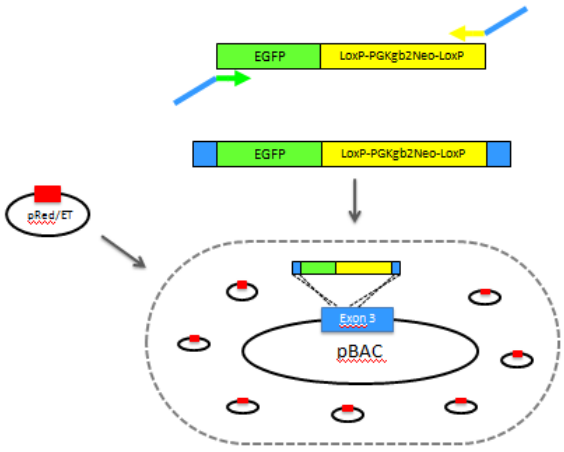


Figure 3. BAC Modification by Recombineering Technology

Upon successful BAC modification, the region-of-interest was amplified by PCR and verified by double-stranded sequencing to ensure that there were no errors in the modification. Mutation-free modified BAC clones were subcloned into a minimal vector using the Gene Bridges Quick and Easy BAC Subcloning kit. Grabbing arms (50 bp) complementary to the regions to be subcloned were attached to the minimal vector by PCR. Similarly, the *pRed/ET* was first transformed into the modified BAC host followed by a second transformation of the minimal vector flanked by 50 bp homology arms into the *pRed/ET*-containing modified BAC clones for subcloning by homologous recombineering (Fig. 4). The positive recombinants were selected with kanamycin (20 µg/mL) and ampicilin (100 µg/mL) antibiotics and screened by colony PCR. The subcloned plasmids were designed to contain the modified region flanked by short and long homology arms with a combined length of at least 10 kb. Subcloned plasmids were verified again by double-stranded sequencing to ensure no mutations had occurred within the modified region during subcloning. Subsequently, these subclones were linearized with *AcI* or *XmnI* restriction enzymes which cut only within the minimal vector and electroporated into the mouse ESCs for gene-targeting.

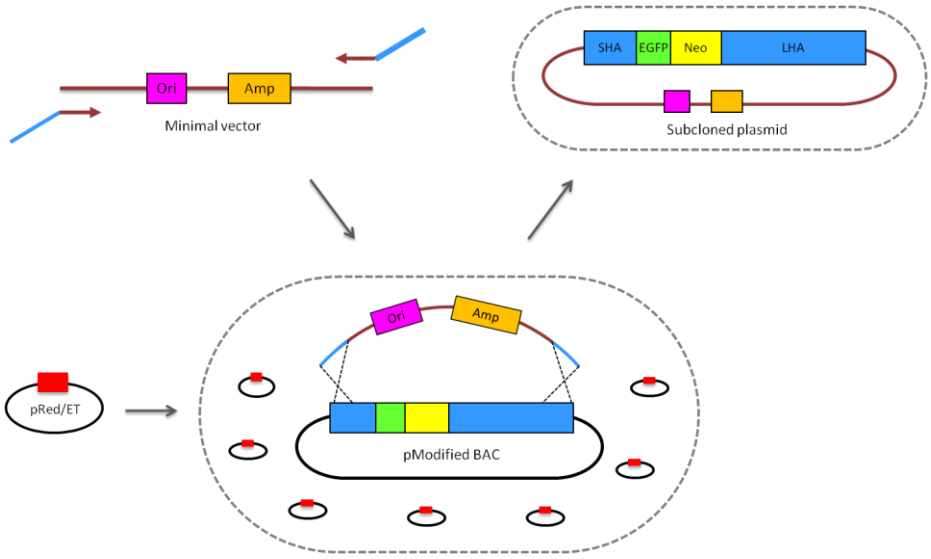


Figure 4. BAC Subcloning by Recombineering Technology

Legend: SHA, short homology arm; LHA, long homology arm.

2.2 Homologous Recombination in Mouse ES Cells

2.2.1 ES Cell Culture

The hybrid (C57BL/6 x 129) mouse ES cell line V6.4 was used for all gene targeting experiments as these hybrid ESCs were shown to remain robust after extended periods of *in-vitro* culture and gene-targeting procedures, producing a good number of viable cloned mice derived entirely from these cells (Eggan et al., 2001). Culture dishes were gelatinized with 0.1% gelatine and cells were grown on a feeder layer of irradiated mouse primary embryonic fibroblasts (PEF) and maintained in ES media (ESM) containing DMEM supplemented with 15% heat-inactivated ES grade fetal bovine serum (FBS) (Gibco, Invitrogen), 0.1 mM β-mercaptoethanol, 4 mM L-glutamine, 1 mM sodium pyruvate, 40 µg/mL gentamicin, and 500 U/mL LIF (ESGRO, Chemicon) at 37°C in 5% CO_2. Media was refreshed daily, and cells were trypsinized with 0.05% trypsin (Gibco, Invitrogen) every 3-4 days.

2.2.2 Electroporation of ES Cells

A day prior to electroporation, V6.4 ES cells were passaged. Cells were again trypsinized just before electroporation and counted on a haemocytometer. 10 x 10^6 cells were resuspended in 400 µL of ESM, mixed with 10-20 µg of linearized targeting vector, then placed in an electroporation cuvette (Biorad, 1 mm gap) and pulsed at 125 µFarads, 0.4 kVolts. A no vector control was also electroporated in parallel. After 5 minutes at room temperature, the cells were divided into 6 DR4 (multi-drug resistant) feeder-layered 10 cm culture dishes and allowed to recover overnight before G418 (150-400 µg/mL) antibiotics selection for successful recombinants.

2.2.3 ES Cell Colony Picking

Putative positive recombinants were selected for with G418 antibiotics over 8-10 days. No ES cell colonies survived in the control plates. Surviving colonies ranging in the hundreds from the vector-electroporated plates were picked into 96-well plates containing 0.05% trypsin – one colony per well. Selection media was first replaced with warm PBS before each colony was picked under the light microscope in the laminar flow hood with a 20 µL pipette set at 15 µL. Upon picking a row of colonies, they were broken up immediately with a multi-channel pipette and divided into 2 sets of 96-well PEF-layered plates. After a week of culture, one of the duplicates was expanded to 24-well plates for genomic DNA extraction and subsequent screening by PCR and Southern blot while the other was further replicated and frozen 'in-situ' as 96-well plate duplicates.

2.2.4 ES Cell Freezing

Freezing media comprising 70% DMEM, 20% FBS and 10% DMSO was prepared fresh and sterile filtered each time. When one well of any plate was 80% confluent, the entire plate was added freezing media (100 or 200 µL/well in 96- or 24-well plate), sealed with cling wrap and aluminium foil, placed in a Styrofoam box and slowly frozen in a -80°C freezer. Putative positive cells were thawed from plates, expanded, trypsinized, resuspended in freezing media (500 µL/vial) and frozen in cryo-vials in a special freezing container (Invitrogen; 1°C drop per minute) at -80°C. Cells, whether frozen in plates or in vials, were stored long term at -150°C or in liquid nitrogen tanks.

2.3 ES Cell Clone Screening and Genotyping

2.3.1 Genomic DNA Extraction

- **ES Cells and Mouse Tail Tips**

ES cells and mouse tail tips were subjected to Proteinase K (Sigma; final concentration: cells 0.2 mg/mL, tails 0.5 mg/mL) digestion overnight at 37^0C and 57^0C with agitation respectively. For high quality and yield, genomic DNA was extracted using MaXtract High Density tubes (Qiagen; Cat.#129046) by adding an equal volume of phenol: chloroform: isoamyl alcohol (pH7.9) to the digested material and vortexed for 30 seconds before spinning down at maximum speed for 5 minutes to separate the phases. The DNA-containing aqueous layer was decanted into a fresh 1.5 ml tube and DNA was precipitated with twice the volume of 100% ethanol and washed with 70% ethanol. DNA pellet was air-dried and re-dissolved in sterile water or elution buffer (Qiagen).

PK Digestion Buffer (PKDB):

50 mM	TRIS, pH 7.0-8.0
5 mM	EDTA
1%	SDS
0.2 M	NaCl

2.3.2 Southern Blotting

- **DIG Probe Design and Synthesis**

Restriction sites that would differentiate modified alleles from the wild-type based on their restricted lengths were identified and any significant repeats found in 40kb of genomic sequences flanking the modified region were masked using the RepeatMasker (www.repeatmasker.org). Probes spanning from 400 bp to 1 kb in size were designed to sit within the restriction sites but either outside of the homology arms termed as ‘external probes’ or within the modified region including the homology arms termed as ‘internal probes’. Short probes between repetitive

sequences were combined for enhanced DNA band detection. The external probes distinguished ES cell clones that have undergone homologous recombination at the targeted locus from those that have simply assumed random genomic integration. The internal probes further investigated these clones if random integrations on top of homologous recombination have occurred. DNA hybridization probes were PCR-labelled with non-radioactive digoxygenin (DIG) by incorporating DIG-dUTP into the nucleotide sequence using the PCR DIG Probe Synthesis kit (Roche, Cat #1636090). These PCR-labelled products are either column purified with Zymo DNA Clean & Concentrator kit and eluted in sterile water or used directly in the hybridization buffer. Probe quality and yield were assessed by gel electrophoresis and stored at -20°C until use.

- **DNA Transfer**

Genomic DNA (10-15µg) was digested for 8-16 hours with appropriate restriction enzymes and ran on 0.75% TAE agarose gels without ethidium bromide at a constant voltage ranging from 20-30 V for 12-24 hours or until DNA bands are well separated. Gels were stained with ethidium bromide, imaged under UV light for DNA quality, depurinated with 0.25 M HCl (necessary for DNA bands > 5 kb) for 10 minutes then denatured in 0.5 M NaOH with gentle agitation for 2 x 30 minutes. The gel was subsequently laid on a DNA transfer assembly consisting of a positively charged nylon membrane (Roche, Cat #1417240) resting on sheets of Whatman 3MM filter paper and a stack of C-fold towels as illustrated below (Fig. 5). DNA was transferred in 0.5 M NaOH onto the membrane by downward capillary movement. The gel was drizzled with transfer buffer 3-4 times every hour before the entire assembly is finally sealed with cling wrap and left to transfer overnight.

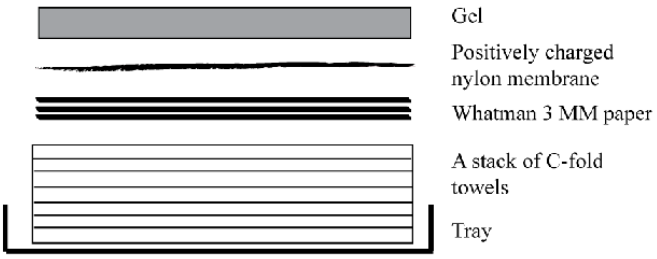


Figure 5. Southern Blot DNA Transfer Assembly

- **Hybridisation and Washing**

The membrane was washed briefly in 5X SSC and placed in a roller bottle containing pre-warmed DIG Easy Hyb buffer (Roche, Cat #1603558; 10 mL/100 cm^2) for pre-hybridization at 42°C for 1-4 hours to block non-specific sites. Thereafter, DIG-labelled probes were denatured at 99°C for 5 minutes, quick-chilled on ice before added to fresh pre-hybridization solution at 20-50 ng/mL probe concentration and the membrane was incubated in this hybridization buffer overnight at 42°C in a rotating oven. After overnight hybridization, the hybridization buffer was stored at -20°C and re-used over several hybridisations upon warming at 68°C for 10 minutes and the membrane was washed twice at room temperature for 5 minutes (2X SSC, 0.1% SDS), then twice for 15 minutes at a higher temperature (60-68°C) in a pre-heated higher stringency wash buffer (0.5X SSC, 0.1% SDS) followed by a brief rinse in 1X MABT.

- **Blocking and Detection**

Blocking of the washed membrane was done with 1% Blocking buffer (Roche Blocking Reagent, Catalog #11096176001; 1g/100mL of 1X MAB) for 30 minutes and then incubated with alkaline-phosphatase (AP)-conjugated anti-DIG antibody (Roche, Catalog #11093274910) at a 1:10,000 dilution in 1% Blocking buffer for another 30 minutes at room temperature with gentle shaking. The antibody was washed off with

1X MABT for 15 minutes twice at room temperature. The blot was rinsed briefly in detection buffer and CSPD chemiluminescent AP substrate was added (1mL/100 cm^2). After incubation at room temperature for 5 minutes, the blot was warmed at 37°C for 10 minutes to enhance enzymatic activity of the substrate before signal detection on an X-ray film. The film generally has to be exposed for 3-6 hours before a reasonable signal is detected.

- **Stripping and Re-probing**

Membranes can be stripped by washing with 0.2 M NaCl /0.1%SDS for 2X15mins after a brief rinse in double-distilled water. Next, the membrane is washed with 2X SSC prior to pre-hybridization and hybridisation with another probe.

Maleic Acid Buffer (10X MAB)
0.2 M Maleic Acid
0.3 M NaCl
Adjust to pH7.5 with NaOH
Filter sterilize

1X MABT
1X MAB
0.3% Tween 20

Detection Buffer
0.1 M Tris-HCl, pH 9.5
0.1 M NaCl

2.4 Generation of Transgenic Mice

2.4.1 Microinjection of ES Cells

Successfully targeted ES cell clones were thawed out, cultured for 3-4 days and passaged a day prior to microinjection. On the morning of microinjection, cells were trypsinized, washed with PBS and resuspended in M2 medium. Approximately 8-10 ES cells were microinjected into each of the 2 to 8-cell stage embryos (Kraus et al, 2010) harvested from C57BL/6J mice and re-implanted into the oviducts of CD-1 pseudopregnant mice (6-10 microinjected embryos per oviduct) on the same day. This method of microinjection generated many high-percentage, germline transmitting chimeras.

2.4.2 Breeding and Genotyping of Transgenic Mice

Heterozygous transgenic mice were obtained by crossing male chimeras to female C57BL/6J mice and homozygous mice, if viable, were obtained by intermating heterozygous mice. Mice were weaned and ear-tagged 3 weeks after birth. The genotype of the mice were determined by PCR, with primers flanking the modified region, on genomic DNA extracted from the tail tips cut during ear-tagging.

2.5 Fluorescence – Activated Cell Sorting (FACS)

2.5.1 Dissociation of Mouse Skeletal Tissue into Single Cells

Mouse embryos expressing enhanced green fluorescent protein (EGFP) under the control of endogenous *Runx2* and/or *Runx3* promoters were harvested and critically staged based on M.H. Kaufmann, Theiler morphological criteria. Skeletal tissues that showed green fluorescence i.e. the limbs, ribs, vertebral column, maxilla and mandible were dissected out from either E13.5 or E14.5 mouse embryos. Yolk sacs of each embryo were retained to verify the genotype of the embryos. Dissociation buffer, composed of Leibovitz's L-15 Medium (Invitrogen, Catalog #21083027),

collagenase I and II (Sigma, 150 U/mL), DNase (Sigma, 50 U/mL) and 0.05% trypsin, was freshly made and sterile filtered (Millipore, 0.2 µm filter disc) each time right after the embryos were removed from the uterine horns. Dissected mouse tissues were placed in a single well of a 24-well plate containing 1 mL of dissociation buffer and broken down by mechanically pipetting up and down initially with a wide bore then a narrow bore 1000 µL pipette tip. An equal volume of 20% FBS in Leibovitz medium (sterile filtered) was added to stop the dissociation activity. The cell suspension was then filtered through a 100 µm strainer basket and again through a 40 µm strainer basket into a 50 mL Falcon tube to remove small bits of tissue. The first filtrate was placed on ice while the dissociation and filtration steps were repeated several times with the bits of tissue left on the strainer baskets after each filtration until there was no more tissue bits visible on the strainer baskets. These subsequent filtrates were combined to give the second filtrate. The second and subsequent dissociation steps incorporated a 5-minute incubation at 37°C before the mechanical pipetting step to enhance the breaking down of the tougher tissues such as cartilage and bone retained by the strainer baskets. Prior to centrifugation, 10 µL of cell suspension was pipetted into a haemocytometer for cell counting and Trypan Blue was used to assay for cell viability. Single cells in filtrates 1 and 2 were then pelleted at 2000 rpm for 5 minutes in a 4°C centrifuge and resuspended in 5% FBS/ 50% Accumax^{TM}/ 2.5 µM EDTA (sterile filtered) to give a desired concentration of 3 million cells/ mL for cell sorting. A wild-type CD-1 embryo at the same embyonic stage was dissociated in parallel and used as a control for gating the GFP detection threshold. Cells were sorted into 1.5 mL centrifuge tubes containing 500 µL of 20% FBS. Sorted cells were spun down in a 4°C tabletop centrifuge at 1,400 rcf and 20% FBS was replaced with Trizol (Invitrogen Cat#15596-018; 100,000 cells /mL Trizol) and stored at -80°C.

2.6 Microarray

2.6.1 RNA Extraction

Total RNA extraction was carried out with the hybrid method coupling the use of Trizol and the Qiagen RNeasy Micro kit. Sorted cells stored frozen in Trizol were thawed, homogenized by repetitive pipetting and incubated at room temperature for 5 minutes. The lysed samples were transfered to 2 mL MaXtract High Density tubes (Qiagen; Cat. #129046). Chloroform (0.2 mL per mL of Trizol) was added to the samples and shaken vigorously by hand for 15 seconds before further incubation at room termperature for 3 minutes. Samples were centrifugated at maximum speed for 15 minutes at 4°C. The aqueous layers at the top were transferred to RNase-free microcentrifuge tubes and precipitated with an equal volume (about 60% of the initial sample in Trizol) of 70% ethanol (made fresh). Well-mixed samples were loaded onto the RNeasy MinElute columns (Qiagen) each seated in a 2 mL collection tube and spun at > 10,000 rpm at room temperature for 30 seconds. The flowthroughs were discarded and the columns were washed with 350 µL wash buffer RW1. On-column DNase treatment was done using 80 µL of 1U/µL DNaseI in buffer for 15 minutes at room temperature after which the columns were washed with 350 µL buffer RW1, 500 µL RPE and 500 µL 80% ethanol (made fresh). The columns were spun dry at full speed for 5 minutes in a new collection tube. Purified RNA samples were eluted in 14 µL of RNase-free water into fresh 1.5 mL centrifuge tubes by a full speed spin for a minute. Elution was accompanied by a 3-minute incubation each time prior to centrifugation. A second elution was performed by returning the eluate to the column for incubation and centrifugation to recover more yield. RNA samples were quantified using Quanti-iT™ Ribogreen® RNA Reagent kit according to the manufacturer's protocol. The integrity of the RNA samples were assessed with the Agilent RNA Pico Chip (#5067-1513) and Agilent 2100 Bioanalyzer software according to the manufacturer's protocol. RNA was stored at -80°C until further use.

2.6.2 RNA Amplification and Biotin Labeling

Isolated RNA was amplified using the NuGEN Ovation™ RNA Amplification V2 kit and biotin-labeled with NuGEN Encore™ BiotinIL Module kit according to the manufacturer's protocol. In brief, the first cDNA strand was synthesized from 1 ng of total RNA samples using a mix of reverse transcriptase and SPIA™ DNA/RNA chimeric primers complementary to the 5' region of the polyA sequence. The second cDNA strand was synthesized with DNA polymerase, SPIA™ DNA/RNA chimeric primers and RNaseH to fragment the chimeric primer to create a priming site for the DNA polymerase. As a result, double-stranded cDNA with unique DNA/RNA heteroduplex at one end was synthesized. Finally, linear isothermal DNA amplification of the double-stranded cDNA was carried out using a master mix of SPIA™ DNA/RNA chimeric primers, DNA polymerase and RNaseH. The amplified cDNA was purified using the Zymo Research DNA Clean and Concentrator™-25 kit (Cat. D4005) according to the manufacturer's protocol prior to biotin-labeling.

Three micrograms of purified and amplified cDNA was brought to a final volume of 25 µL with nuclease-free water. Five microliters of UNG enzyme (1U/µL) in 5 µL of UNG buffer (10 mM $K_2HPO_4.3H_2O$, 4 mM $MgCL_2$) was added to the cDNA samples, incubated at 50°C for 30 minutes in a thermal cycler and placed on ice. Next, 5 µL of ARP [N-aminooxyacetly)-N'-(D-biotinoyl)hydrazine, trifluoroacetic acid salt] solution (Molecular Probes, 11.3 mg/mL ARP in 22.4 mM phosphate buffer) and 5 µL of labeling buffer (0.952 M acetic acid, 28 mM $MgCl_2$) were added and incubated at 50°C for another 60 minutes. Biotin-labeled cDNA was purified with Zymo Research DNA Clean and Concentrator™-5 and kept at -20°C before hybridization onto the BeadChip arrays.

2.6.3 Hybridization on Illumina Mouse WG-6 BeadChip

Reagents from Illumina MouseWG-6 Expression BeadChip kit were used for the microarray. The hybridization oven and the water bath were first preheated to 48°C and 58°C respectively and 1.5 μg of each biotin-labeled cDNA sample was resuspended in 10 μL of nuclease-free water and incubated at room temperature for 10 minutes. Twenty microliters of GEX-HYB (prewarmed to 58°C and cooled to dissolve any precipitation) was added to each cDNA sample and heated at 65°C for 5 minutes. Meanwhile, 200 μL of GEX-HCB was added to the humidifying buffer reservoirs of the hybridization chambers. The warmed cDNA was briefly vortexed, centrifugated at full speed for a minute and cooled to room temperature before all 30 μL of biotin-labeled cDNA was loaded onto the array. There were at least 4 biological replicates per genotype assessed and the biological replicates of the same genotype were randomized on the BeadChip arrays. The loaded BeadChips were placed horizontally and sealed within the hybridization chamber. The whole chamber was placed in the hybridization oven and incubated for 18 hours at 48°C. The hybridization temperature was reduced to compensate the less stable cDNA/DNA pairs compared to the cRNA/DNA pairs. The 1X High-Temp wash buffer (500 mL) was prewarmed overnight at 55°C in a waterbath-insert-heatblock.

Coverseals of the hybridized BeadChips were removed in Wash E1BC solutions (3 mL of E1BC in 1 L of RNase-free water) the next day and incubated in the 1X High Temp wash buffer at 58°C for 10 minutes. Every incubation and wash required the BeadChips to be completely submerged in solution. Thereafter, the BeadChips were washed in E1BC solution for 5 minutes, 100% ethanol for 10 minutes and back in E1BC solution for 2 minutes. Next, the BeadChips were blocked with Block E1 buffer (prewarmed to room temperature; 4 mL per chip) for 10 minutes in a provided wash tray on a rocker at medium speed. BeadChips were transferred to new wash trays

containing 2 µL of Streptavidin-Cy3 in 2 mL of Block E1 buffer, covered and left to incubate in the dark on the rocker for another 10 minutes. A final 5-minute wash was done with the Wash E1BC solution before the BeadChips, placed in a chip rack, were spun dry at 275 rcf for 4 minutes. BeadChips were scanned with the Illumina® BeadArray Reader on the same day.

2.6.4 Gene Expression Analysis using GeneSpring GX 11.0

Raw image data was interpreted with the Illumina® BeadStudio software and exported as GeneSpring compatible gene expression output data with background subtraction and no normalization. The sample probe profiles in text file format were imported into GeneSpring GX 11.0 for gene expression analysis. Entities with detection p-value > 0.8 were flagged as "Present", those with detection p-value < 0.6 were flagged as "Absent" and anything in between were flagged as "Marginal". Negative raw values were shifted to a minimum threshold of 1.0 and a percentile shift to 75% with a baseline to median normalization algorithm was applied. Averaging among biological replicates was not applied. All entities were first filtered by flags retaining entities with "Present" or "Marginal" flags in at least 1 out of the total number of samples. The filtered entities were further filtered by retaining entities with raw expression data that fell between 20-100th percentile in at least 1 out of the total number of samples. Statistical analysis such as the ANOVA for a three-way comparison of *Runx2*$^{+/+}$ vs. *Runx2*$^{+/-}$ vs. *Runx2*$^{-/-}$ and the unpaired Student's t-test for pair-wise comparisons of *Runx3*$^{+/-}$ vs. *Runx3*$^{-/-}$, *Runx2*$^{-/-}$*Runx3*$^{-/-}$ vs. *Runx2*$^{-/-}$ and *Runx2*$^{-/-}$*Runx3*$^{-/-}$ vs. *Runx3*$^{-/-}$ were employed and asymptotic p-values were subjected to the Benjamini Hochberg False Discovery Rate (B-H FDR) multiple testing correction to identify significant differentially expressed entities. Significance was defined as corrected p-value < 0.05 and fold change > 1.5.

2.7 Chromatin Immunoprecipitation – Sequencing (ChIP-Seq)

2.7.1 Tissue Harvesting and Cross-linking

E13.5 mouse embryos were harvested from the mouse uterine horns and critically staged based on M.H. Kaufmann, Theiler morphological criteria. Embryonic tissues of interest were dissected from E13.5 mouse embryos on ice into chilled Leibovitz medium. Tissues were spun at 1,100 g for 5 minutes at 4˚C and resuspended in 10 mL of fresh Leibovitz medium. Resuspended tissues were homogenized on ice with a 15 mL Douncer. Minced tissues were pelleted at 1,100 g for 5 minutes at 4˚C, weighed and resuspended in 10X the tissue pellet volume of room termperature 1X PBS. Cross-linking buffer (11% Formaldehyde) was added at one tenth the volume of PBS and tissues were cross-linked for 10 minutes on a nutator. Cross-linking was quenched for 5 minutes with 2.5 M Glycine added at one-tenth of the volume. Tissues were spun down at 1,100 g for 5 minutes at 4˚C, resuspended in 10 mL of chilled 1X PBS and homogenized on ice again with the Douncer until there were no more visible clumps. Cross-linked cells were washed with chilled PBS, pelleted, weighed, snap frozen in liquid nitrogen and stored at -80˚C.

11% Formaldehyde Solution (Cross-linking Buffer):

50 mM	Hepes-KOH (pH 7.5)
100 mM	NaCl
1 mM	EDTA (pH 8.0)
0.5 mM	EGTA (pH 8.0)
11%	Formaldehyde

2.7.2 Binding of Antibodies to Magnetic Beads

Magnetic Dynabeads Protein G (Invitrogen; Cat. 100.04D) (100 μL per IP) were placed in 1.5 mL centrifuge tubes, washed 3 times with 1 mL of blocking solution (0.5% BSA) using a magnetic stand and incubated with 5 μg of rabbit IgG or 10 μg of rabbit anti-HA antibody overnight or for at least 6 hours in 250 μL blocking solution at 4˚C on a 360° rotating platform. The bead-conjugated antibodies were washed with

the same blocking solution the next day and resuspended in 100 μL of blocking solution to be added to the sonicated chromatin. IgG bound to magnetic beads were used for preclearing the chromatin.

2.7.3 Cell Lysis, Sonication and Chromatin Immunoprecipitation

Frozen cell pellets were thawed, resuspended in lysis buffer 1 (LB1) at 10 times the volume/weight of the pellet, nutated for 10 minutes at 4˚C, spun down at 1,350 g for 5 minutes at 4˚C and resuspended in lysis buffer 2 (LB2) at the same volume as LB1. The cell lysates in LB2 were nutated at room temperature for 10 minutes and spun down at 1,350 g for 5 minutes at 4˚C to pellet the nuclei. The nuclei pellets were weighed and the nuclear lysis buffer 3 (LB3) was added at 10 times the volume/weight of the pellet. The nuclear lysates (2 - 3 mL) were transferred to 15 mL bacterial culture tubes containing 1 mL of glass beads (BioSpec Products; #11079105) and kept on ice prior to sonication.

Lysis Buffer 1 (LB1):

50 mM	Hepes-KOH (pH 7.5)
140 mM	NaCl
1 mM	EDTA (pH 8.0)
10%	Glycerol
0.5%	Igepal CA360
0.25%	Triton X-100

Lysis Buffer 2 (LB2):

10 mM	Tris-HCl (pH 8.0)
200 mM	NaCl
1 mM	EDTA (pH 8.0)
0.5 mM	EGTA (pH 8.0)

Lysis Buffer 3 (LB3) (made with chilled water):

10 mM	Tris-HCl (pH 8.0)
100 mM	NaCl
1 mM	EDTA (pH 8.0)
0.5 mM	EGTA (pH 8.0)
0.5 mM	Na-Deoxcholate
0.5%	Sodium Dodecyl Sulfate (SDS)

* Protease inhibitor (Roche) was added to each lysis buffer

Sonication was performed with the Branson Digital Sonifier® in the cold room. Chromatin sonication was previously optimized at 12 minutes (15-cycles of "ON" and "OFF" sonication for 20 seconds each time at 40% amplitude and 30 seconds rest in between) to shear the chromatin to a recommended size range of 100-500bp using E13.5 CD-1 wild-type embryonic tissues (Fig. 6). The sonicator probe was placed into the nuclear lysate in LB3 until it was just touching the glass beads. Sonication was done using the above optimized parameters in a beaker of ice water to prevent overheating of the nuclear lysates. The sonicated chromatin was transferred to a 2 mL centrifuge tube and 1% Triton X-100 was added. The sonicated chromatin was centrifugated to remove any debris from sonication and quantitated with the Nanodrop spectrophotometer (Thermo Scientific). The average concentration range obtained was 1.0 – 1.4 μg/μL. Two milligrams of sheared chromatin were used with 10 μg of antibody for each immunoprecipitation (IP). The volume of nuclear lysates containing 2 mg of chromatin was brought to a total volume of 1.8 mL with 1% Triton X-100 in LB3. The sheared chromatin was pre-cleared with bead-conjugated IgG for an hour at 4˚C on a 360° rotating platform before overnight chromatin immunoprecipitation with the anti-HA antibody (Abcam 9110) at 4˚C on a 360° rotating platform. Preclearing and IP were done in 2 mL centrifuge tubes sealed with parafilm.

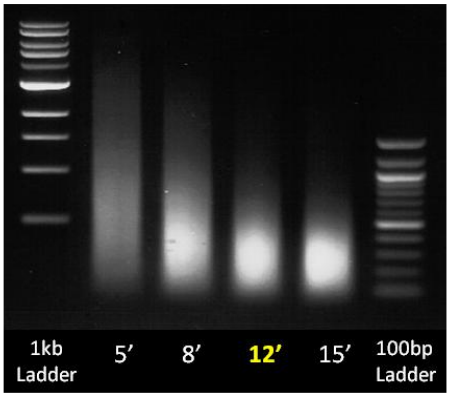


Figure 6. Optimizing Chromatin Sonication

Chromatin sonication was optimized using tissues collected from E13.5 CD-1 wild-type embryos. The recommended sheared chromatin size range of 100-500 bp was attained with 12 minutes of "ON" and "OFF" sonication for 20secs each time at 40% amplitude with 30 secs rest in between each cycle using the Branson Digital Sonifier®.

2.7.4 Wash, Elution and Reverse Cross-link

The immunoprecipitated chromatin was washed with 1 mL of chilled 1X RIPA buffer (Upstate; Cat# 20-188) added with protease inhibitor the next day. Each wash included a 5-minute 360° rotation at 4°C before recovering the immunoprecipitated chromatin with a magnetic stand and discarding the wash buffer. A total of 6 washes were performed. Thereafter, the immunoprecipitated chromatin was washed once with 1 mL of TE buffer (10 mM Tris, pH 8.0; 1mM EDTA; 50 mM NaCl) and the immunocomplex was eluted off the beads at 65°C in 210 μL of Elution buffer (50 mM Tris, pH 8.0; 10 mM EDTA; 1% SDS) for 30 minutes with shaking at 1,400 rpm in a heat block. The magnetic beads were removed with the magnetic stand and the supernatants were transferred to a clean 1.5 mL centrifuge tube for reverse cross-linking by incubating at 65°C overnight.

2.7.5 ChIP DNA Clean Up

The reverse crosslinked samples were transferred to MaXtract High Density tubes (Qiagen; Cat.#129046), added with 200 μL TE buffer and incubated with 0.2 μg/mL RNaseA at 37°C for 2 hours to remove excessive RNA that might interfere with the subsequent DNA purification. Proteins were removed next with 0.2 μg/mL of Proteinase K at 55°C for another 2 hours. An equal amount of phenol-chloroform-isoamyl alcohol (Ambion; Cat# AM9732) was added to each sample and vortexed for 30 seconds. The phases were separated by centrifugation at full speed for 3 minutes and the aqueous layer was transferred to a new 1.5 mL tube containing 30 μg of glycogen in 16 μL of 5M NaCl. The samples were precipitated with 800 μL of absolute ethanol and incubation at -80°C for 30 minutes after which the DNA was pelleted at full speed for 10 minutes at 4°C. The DNA was then washed with 80% ethanol (made fresh), air-dried and resuspended in 30 μL of 10 mM Tris-HCL, pH

8.0. The DNA was quantified with the Nanodrop spectrophotometer and stored at -80°C until further use.

2.7.6 ChIP-Seq DNA Library Prep

The purified ChIP DNA was end-repaired, added with 'A' bases at the 3' ends, ligated with universal adaptors, amplified and size selected (200-400 bp) using the NEBNext® ChIP-Seq Sample Prep Reagent kit (#E6200S) according to the manufacturer's protocol. The DNA libraries were eventually sequenced with the Illumina Solexa Sequencer.

2.8 Western Blotting

ChIP samples eluted from the magnetic beads were added with 4X loading buffer and 10X sample reducing agents (Invitrogen; Cat# NP0004), boiled at 80°C for 10 minutes to denature the proteins and ran on a 10-well NuPAGE® 4%-12% Bis-Tris precast gel (Invitrogen) at 120 V for 2 hours or until the blue front ran out. A polyvinylidene difluoride (PVDF) membrane (Bio-Rad Immun-Blot™) and 2 extra thick filter papers (Biorad) were soaked in 0.2% methanol in Tris-Glycine transfer buffer (1st Base #BUF-2020-10X1L) just before the gel run ended. The Western gel was taken out of the covers, placed on the PVDF membrane, sandwiched between the soaked filter papers and placed on the transfer plate of the Bio-Rad Trans-Blot® Semi-Dry Transfer Cell. The protein bands were transferred onto the membrane at 25 V for 30 minutes. Next, the membrane was blocked with 3% skim milk (BD Difco) in TBST (1X TBS and 0.05% Tween20; Bio-Rad) for 1 hour at room temperature, rinsed briefly with TBST and probed with HRP-conjugated anti-HA goat primary antibody (Bethyl; Cat# A190-107P; 1:20,000) diluted in 3% BSA (Sigma; A3059) for 1 hour at room temperature and washed with TBST 4 x 15 minutes. Protein bands were detected with Supersignal™ West Pico Enhanced Chemiluminiscent (ECL)

Reagents (Thermo Scientific) and exposed onto Amersham Hyperfilm™ X-ray films (GE Healthcare; #28906844).

2.9 Embryo Processing for Histology

Mouse embryos were fixed in 4% paraformaldehyde (PFA) in PBS overnight at 4°C, washed in PBS for 5 minutes and dehydrated through 50% and 70% ethanol/PBS for 20 minutes each the next day. Dehydrated mouse embryos were processed in an automated tissue processor machine (Leica TP 1020), embedded in paraffin and sectioned at 10 microns with the microtome (Leica RM 2165).

2.10 Section In-Situ Hybridization (SISH)

- **RNA Probe Synthesis**

cDNA clones purchased from Open BioSystems were used as templates for in vitro transcription synthesis of antisense DIG-labelled RNA probes using the DIG RNA labelling kit (Roche, Cat.#11 175 025 910) according to the manufacturer's instructions. In vitro transcription reaction was stopped with 0.2 M EDTA. Small nucleotides were removed with the RNA Spin Columns (Roche; Cat# 11274015001) and the DIG-labelled RNA probes were precipitated with 5 µL of 3M NaOAc, pH 5.5 and 2.5 X100% ethanol with incubations at -20°C for 30 minutes. RNA probes were pelleted at full speed for 10 minutes at 4°C, washed twice with 70% ethanol, air-dried and resuspended in DEPC water. Probe concentrations were measured by Nanodrop spectrophotometer. Probes were run on 1% DEPC-treated TAE gel to check for integrity.

- **Section Pre-treatment, Pre-hybridization and Hybridization**

Mouse embryo sections on polysine-coated histology slides were de-waxed in Histo-Clear™ (a non-toxic substitute for xylene), rehydrated through 100%, 90%, 70% and 30% ethanol for 5 minutes each, post-fixed in 4% PFA for 20 minutes, digested with 10 μg/mL Proteinase K in 0.1 M Tris, pH 7.5 for 10 minutes and post-fixed a second time with 4% PFA for 10 minutes. Sections were then pre-hybridized at 67°C for 2-3 hours and hybridized overnight at the same temperature with antisense DIG-labeled RNA probes at 600-1200 ng/mL concentration in pre-hybrization solution.

Pre-hybridization Solution:

50%	Formamide (Roche; Cat#11814310001)
5X	SSC (1st Base)
1X	Denhardt's (5 g Ficoll, 5 g BSA and 5 g Polyvinylpyrrolidine in 500 mL DEPC water)
0.1%	Tween20
0.1 mg/mL	Yeast tRNA (Ambion; Cat#AM7118)
0.05 mg/mL	Heparin

- **Post-hybridization Washes and Probing with anti-DIG Antibody**

The hybridized mouse sections were washed the next day with pre-warmed Solution I at 67°C for 3x30 minutes, TNT for 3x5 minutes, TNT:Solution II (1:1) for 5 minutes, pre-warmed Solution II at 63°C for 3x30 minutes and MABT for 3x5 minutes before blocking with 2% blocking solution (Roche Blocking Reagent dissolved in MAB) for 2-3 hours. Thereafter, the sections were probed with alkaline phosphatase (AP)-conjugated anti-DIG antibody diluted 2000-fold in blocking solution overnight at 4°C.

Solution I:	
50%	Formamide (Roche)
5X	SSC (1st Base)
1%	SDS
TNT:	
10 mM	Tris-HCl, pH 7.5
0.5 M	NaCl
0.1%	Tween20
Solution II:	
50%	Formamide (Roche)
2X	SSC (1st Base)
0.2%	SDS

- **Post-Antibody Washes and Colour Development**

Mouse sections probed with anti-DIG antibody were briefly washed with MABT for 3x10 minutes before the hourly washes with MABT for 3-4 hours. Sections were conditioned in NTMT for 3x10 minutes before NBT/BCIP (Roche; Cat#11681451001) substrate (200 μL diluted in 10 mL of 0.1 M Tris-HCl, pH9.5 and 0.1 M NaCl) was added to the sections for colour development. Hybridization cover slips were placed over the sections and the slides were left at 4°C in the dark for slow colour development over several days to reduce the background levels. Cover slips were removed in PBS when colour development was complete and washed in PBS for 2x5 minutes prior to mounting with glycerol-gelatin and glass cover slips.

NTMT:	
100 mM	Tris-HCl, pH 9.5
100 mM	NaCl
50 mM	$MgCl_2$
0.1%	Tween20

All solutions used for SISH were kept RNase-free by using Diethyl Pyrocarbonate (DEPC)-treated PBS or water in the solution preparations and the washes. All SISH sections were imaged with Zeiss Axio Imager Z1.

2.11 Immunohistochemistry (IHC)

Mouse sections on slides were baked on a slide warmer at 54°C for 10 minutes to remove any trapped moisture. Sections were de-waxed in Histo-Clear™ for 20 minutes and rehydrated through 100%, 95%, 90%, 70% ethanol each for 5 minutes prior to steaming in 0.01 M sodium citrate buffer, pH 6.0, in a water bath steamer at 121°C for 15 minutes to retrieve the antigens on the proteins. The slides were left in the buffer to cool for 2-3 hours. The slides were then washed with PBS for 3x5 minutes, incubated with 0.6% hydrogen peroxide for 20 minutes in the dark, washed with 0.2% Tween20 in PBS for 2x5 minutes and blocked with horse serum provided by the Vectastain® ABC kit (Vector Laboratories) for 30 minutes and another 30 minutes with blocking solution (2% BSA, 5% sheep serum in PBS) at room temperature before overnight incubation with the primary antibody (rabbit anti-Runx2: Abcam 23981, 10 μg/mL; rabbit anti-Runx3: Abcam 68938, 1:700; rabbit anti-HA: Acris Antibodies, AP09230PU-N, 1:400) diluted in blocking buffer at 4°C.

Incubation with the primary antibody was continued at room temperature for an hour the next day. The slides were washed with 0.1% Tween20 in PBS for 3x5 minutes and incubated with either biotin-conjugated or HRP-conjugated secondary anti-rabbit antibody diluted in PBS for 2 hours at room temperature. The secondary antibody was washed off with 0.1% Tween20 for 3x5 mins, PBS for 10 minutes, TBS for 10 minutes prior to colour development using BD Biosciences Pharmingen™ DAB substrate kit. Reaction was stopped with water when colour development reached the desired levels. Further washing was done under running water for 5 minutes and

slides were mounted with glycerol-gelatin. For biotin-conjugated secondary antibody, an extra 1-hour incubation step with the Avidin DH: Biotinylated Horseradish Peroxidase H complex in PBS before the 10-minute PBS wash was necessary. The complex-forming reagents A and B from the Vectastain® ABC kit were previously incubated at 1:1:100 (Reagent A: Reagent B: PBS) proportions for an hour at 4°C prior to application on the sections. The Vectastain® ABC kit gave less background and better signals when used with the anti-Runx2 and anti-Runx3 primary antibodies. All IHC sections were imaged with Zeiss Axio Imager Z1.

CHAPTER 3 – RESULTS AND DISCUSSIONS

3.1 Gene Expression Profiling of Runx2 and Runx3

3.1.1 Generation of Wild-type and Knockout Fluorescing Mice

In this post-genomic era, the use of increasingly affordable high-throughput technologies such as microarray and RNA-sequencing has greatly propelled us forward in the field of functional genomics. The possibility of analysing global transcriptome patterns during events of differentiation (embryonic development), transformation (cancer) or comparing differences between two or more gene expression profiles (normal vs. gene mutation) permits discovery of novel links between genes and functions.

For any causal link between a transcription factor and its targets within a network to be established there must be some form of loss of function experiments. More importantly, in order to draw an accurate relationship between the investigated factor and the downstream targets interpreted from the gene expression studies, sufficient cellular resolution of the gene activity is vital. Therefore, my initial step was to generate knockout gene-targeting constructs of *Runx2* and *Runx3* using BACs harboring either a relevant segment or the entire *Runx2* or *Runx3* gene for homologous recombination in mouse ESCs (v6.4). By replacing a small portion (30 amino acids) of *Runx2* and *Runx3* with enhanced green fluorescence protein (EGFP) whereby the EGFP is driven by the endogenous *Runx2* or *Runx3* promoter, cells with disrupted *Runx2* or *Runx3* gene activity can be isolated and enriched for gene expression profiling by fluorescence-activated cell sorting (FACS).

BACs containing the 5' region of the genomic locus of *Runx2* (RP24-217E4) and the entire genomic locus of *Runx3* (RP24-118B14) were used for BAC modification using the Gene Bridges Quick and Easy BAC Modification kit. The *EGFP-loxP-PGKgb2-*

Neo-loxP insertion cassette was inserted in frame six amino acids after the second ATG at the P2 promoter of *Runx2* or *Runx3* in exon 3 or exon 2 respectively, deleting 90bp of DNA sequences simultaneously (Figs. 7, 8 & 9). This enabled both major isoforms of Runx2 and Runx3 to be disrupted by the insertion.

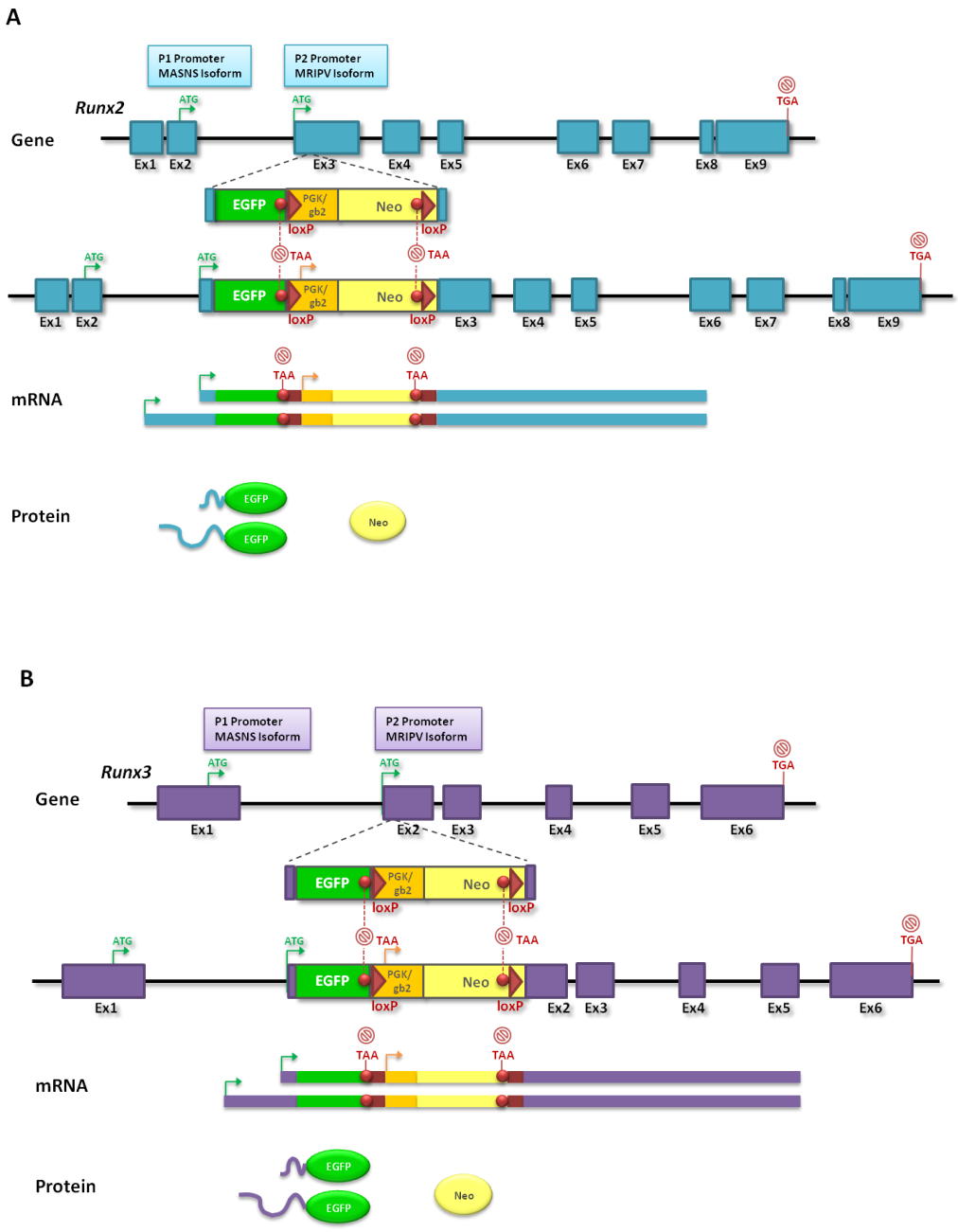


Figure 7. *Runx2* and *Runx3* Gene Knock-out Strategy

The *EGFP* followed by a *loxP*-flanking *neomycin* resistance marker *(Neo)* expressed under the control of a dual eukaryotic-prokaryotic (*PGK-gb2*) promoter were inserted in frame 6 amino acids after the P2 promoter in (A) the 3rd exon of the *Runx2* gene and (B) the 2nd exon of the *Runx3* gene, deleting a stretch of 90 bp DNA sequences concomitantly. The insertions capture the two major isoforms of Runx2 and Runx3 proteins. The flanking *loxP* sites enable the *Neo* to be excised upon exposure to Cre recombinase proteins.

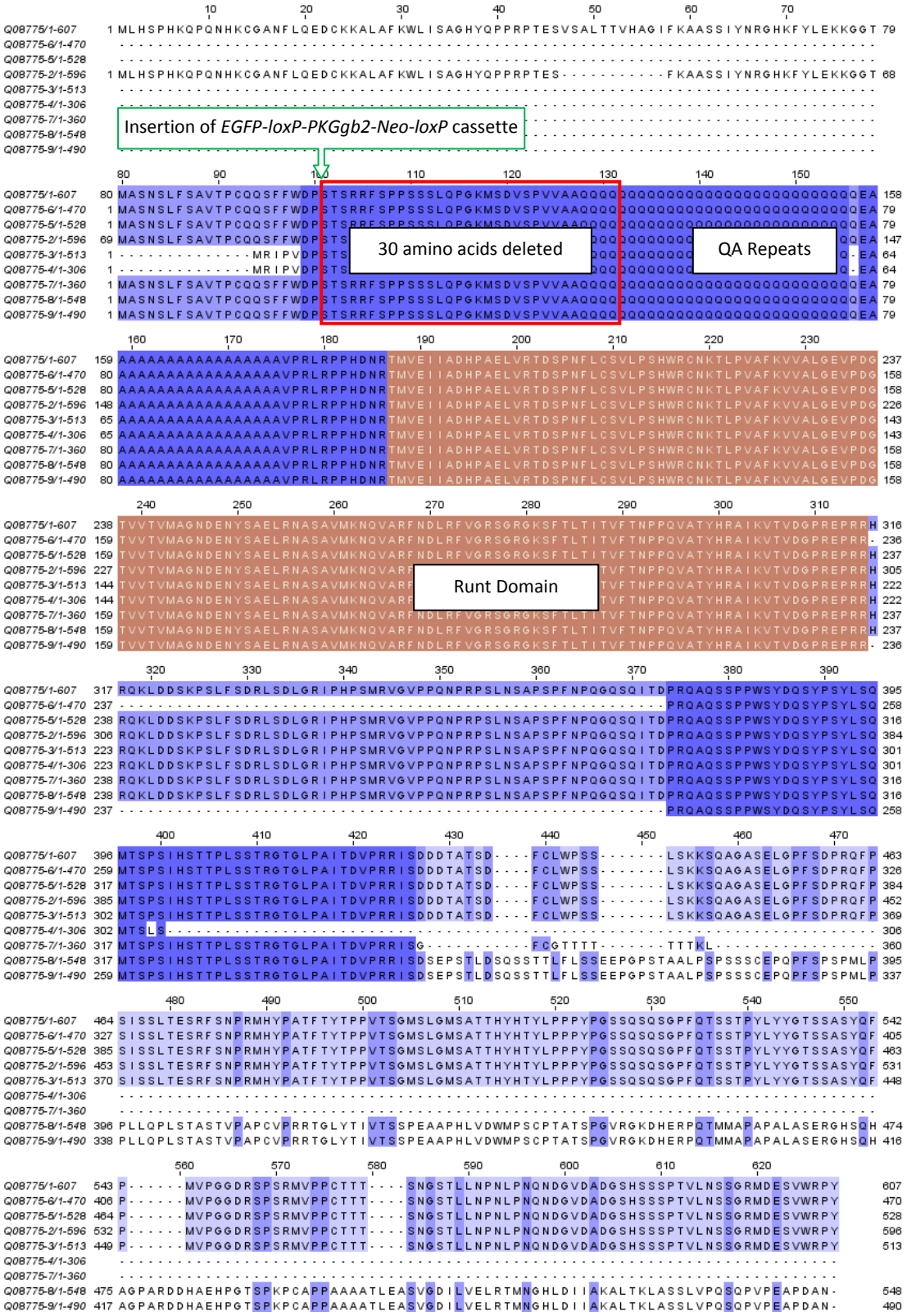


Figure 8. Runx2 Protein Isoforms and Illustration of Runx2 Modifications

The *EGFP-loxP-PGKgb2-Neo-loxP* cassette was inserted 6 a.a. after the ATG at the P2 promoter (MRIPV). Thirty amino acids after the insertion were deleted to ensure complete abrogation of the gene.

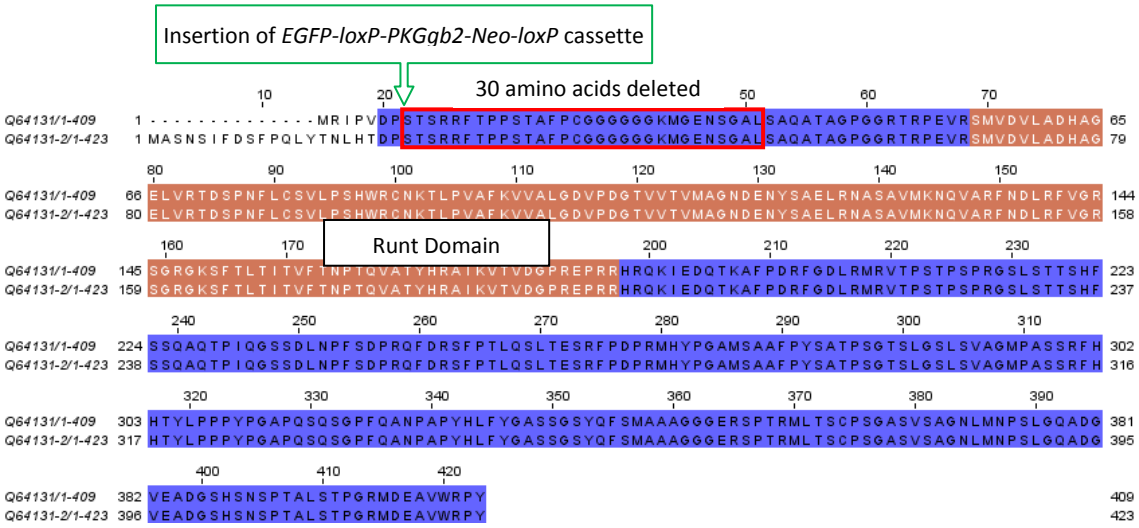


Figure 9. Runx3 Protein Isoforms and Illustration of Runx3 Modifications

The EGFP-loxP-PGKgb2-Neo-loxP cassette was inserted 6 a.a. after the ATG at the P2 promoter (MRIPV) of the *Runx3* gene. Thirty amino acids after the insertion were deleted to ensure complete abrogation of the gene.

Since *Runx3* haplo-sufficient mice are normal, viable and fertile, the *Runx3*$^{+/EGFP}$ mice could be taken as wild-type for transcriptomic comparison against the *Runx3*$^{EGFP/EGFP}$ knock-out mice. However, the *Runx2* haplo-sufficient mice present a skeletal phenotype that recapitulates the human cleidocranial dysplasia. Therefore, I had to establish a wild-type mouse line that expresses *EGFP* in the *Runx2*-expressing domains so that *Runx2*-expressing cells can be isolated by FACS for microarray comparing gene expression profiles of the wild-type, *Runx2* haplo-sufficient and *Runx2* homozygous knockout mice.

In order to co-express *Runx2* and *EGFP* as two functional discrete proteins, the 2A-peptide (also termed as cis-acting hydrolase elements “CHYSEL”) co-expression strategy was employed. This co-expression approach was chosen over the internal ribosome entry site (IRES) method to avoid the possible reduced expression of the downstream gene, which is a widely-known caveat linked to the use of IRES in multicistronic gene expression (de Felipe, 2002; Hellen and Sarnow, 2001; Licursi et

al., 2011). The 2A-peptide sequences function in a way that prevents the peptide bond formation between glycine and proline at the C-terminus of the 2A-peptide (de Felipe et al., 2003). Hence, when placed between two genes, the ribosome appears to skip at the glycine-proline junction before translating the downstream gene thus producing two discrete proteins that are expressed under the control of the same promoter. The compact 23-amino-acid 2A-oligopeptide sequences derived from the foot-and-mouth disease virus (FMDV) (F2A) were placed upstream of the *EGFP-Frt-PGKgb2-Neo-Frt* cassette. In addition, a furine protease recognition site (RAKR) followed by a Gly-Ser-Gly (GSG) spacer were added immediately upstream of the F2A sequences. The RAKR site served to trim the residual 2A peptide from the upstream Runx2 protein and the GSG spacer was included to enhance the ribosomal "skipping". The entire *RAKR-GSG-F2A-EGFP-Frt-PGKgb2-Neo-Frt* cassette was inserted in frame at the C-terminus of *Runx2* just before the stop codon using BAC containing the 3' region of *Runx2* genomic locus (RP23-7C18) (Fig. 10).

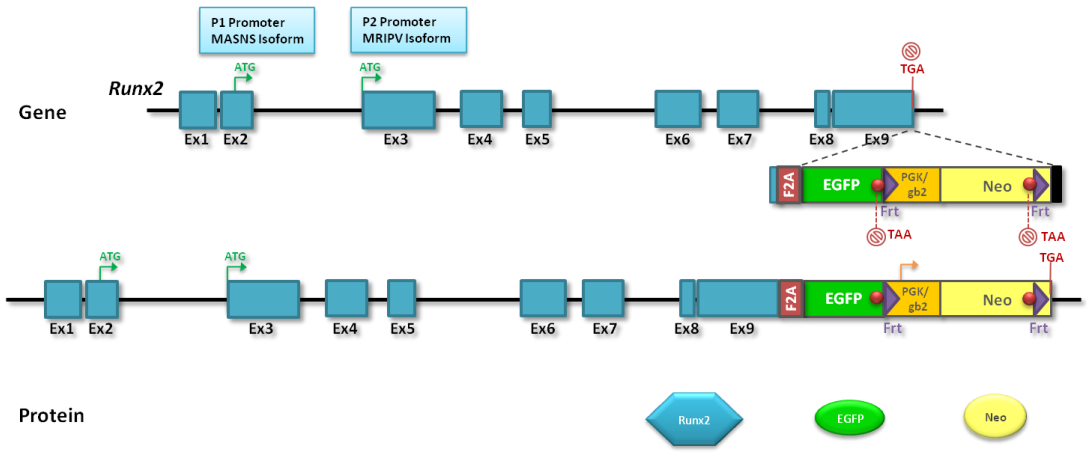


Figure 10. Targeting Strategy to Generate Wild-type Mice Expressing *EGFP* in *Runx2*-expressing Tissues

The *RAKR-GSG-F2A-EGFP-Frt-PGKgb2-Neo-Frt* cassette was inserted at the C-terminus of *Runx2* just before the stop codon. In this manner, both the *Runx2* and *EGFP* genes, concatenated by the F2A-peptide sequences can be expressed as discrete proteins under the control of the *Runx2* promoter.

The targeting constructs were introduced into mouse v6.4 ESCs (129/SvJ x C57BL/6J). Putative positive recombinants were selected for with G418 antibiotics over 8-10 days and these clones were further screened by Southern blotting with probes external to the homology arms to identify the authentic clones that have undergone homologous recombination in the correct locus (Figs. 11, 12 & 13).

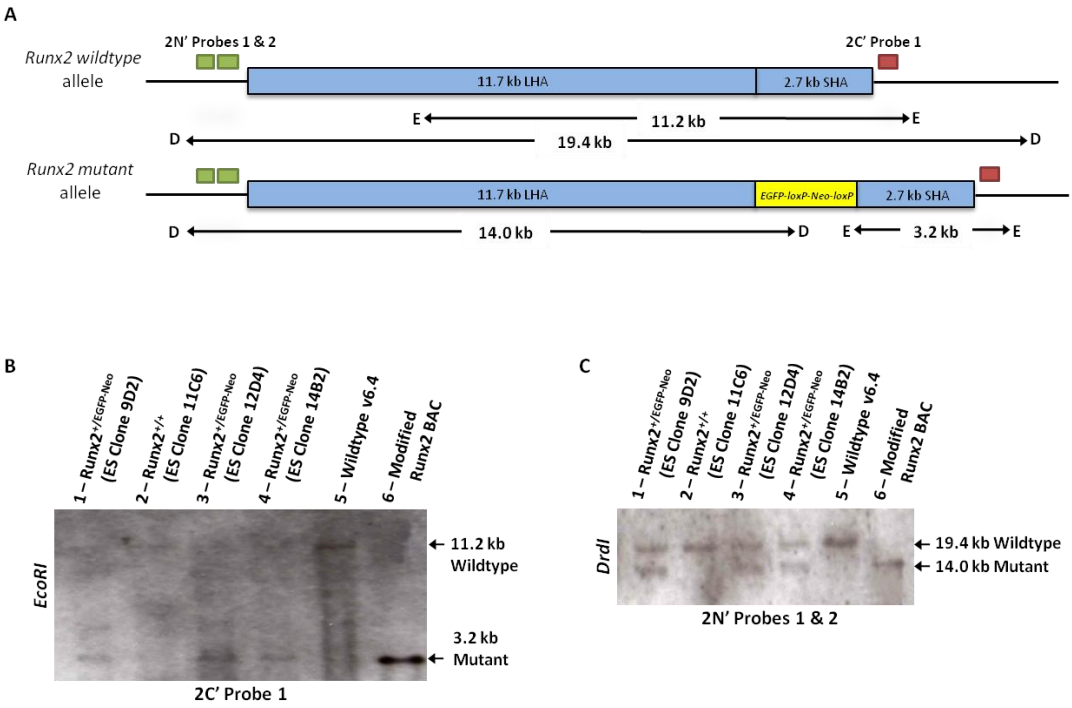


Figure 11. Southern Blot Screen for *Runx2*$^{+/EGFP-Neo}$ Positive Recombinant ES Cell Clones

Genomic DNA extracted from ES cell colonies resistant to G418 were digested with *EcoRI* restriction enzyme and probed with a C-terminal Probe 1 (2C'Probe1; 416 bp; 50 ng/ml) just outside the short homology arm. True homologous recombinant ES cell clones will present an 11.2 kb wild-type and a 3.2 kb mutant band (A; lanes 1, 3 & 4) while a false positive clone will only show the wild-type band (A; lane 2). These clones were confirmed by digesting with another restriction enzyme, *DrdI* and probed with N-terminal probes 1 and 2 (2N'Probe1 & 2; 400 bp & 450 bp; 6 ng/ml each). True clones will give a 19.4 kb wild-type and a 14 kb mutant band (B; lanes 1, 3 & 4) and the false positive clone will show only the wild-type band (B; lane 2). LHA, long homology arm; SHA, short homology arm; E, *EcoRI*; D, *DrdI*.

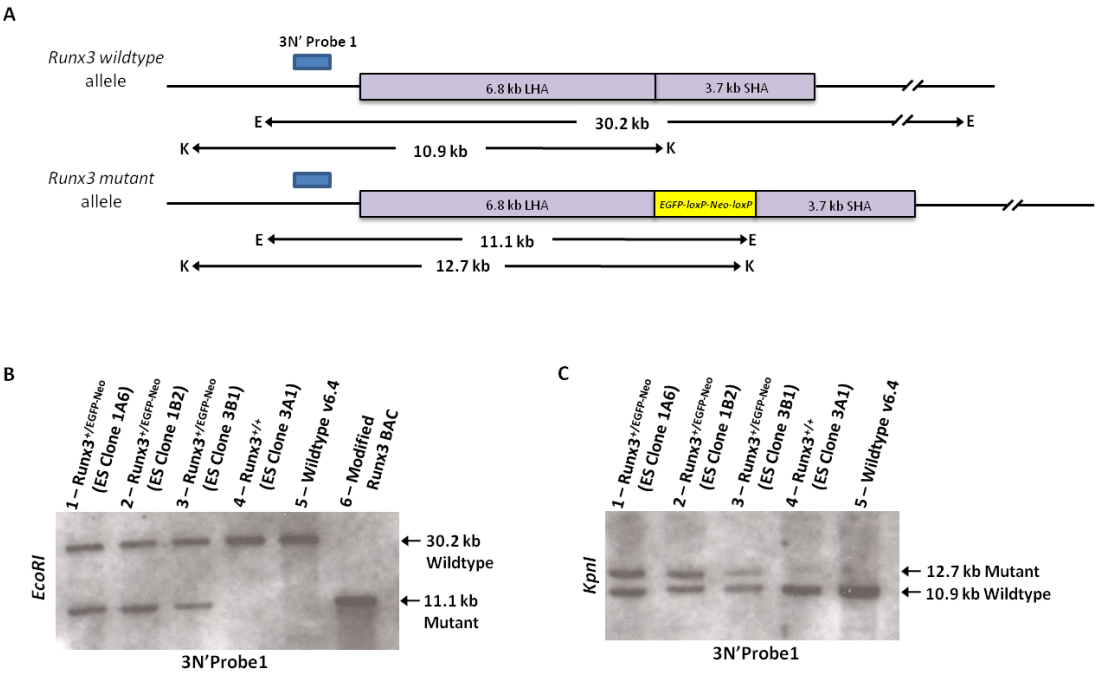


Figure 12. Southern Blot Screen for *Runx3*$^{+/EGFP-Neo}$ Positive Recombinant ES Cell Clones

Genomic DNA extracted from ES cell colonies resistant to G418 were digested with *EcoRI* (A) or *KpnI* (B) restriction enzyme and probed with an N-terminal Probe 1 (3N'Probe1; 801 bp; 25 ng/ml) just outside the long homology arm. True homologous recombinant ES cell clones will give a 30.2 kb wild-type and an 11.1 kb mutant band (A; lanes 1, 2 & 3) when cut with *EcoRI* or a 12.7 kb mutant and a 10.9 kb wild-type band (B; lanes 1, 2 & 3) when cut with *KpnI*. A false positive clone will only show the wild-type band (A; lane 4 and B; lane 4). LHA, long homology arm; SHA, short homology arm; E, *EcoRI*; K, *KpnI*.

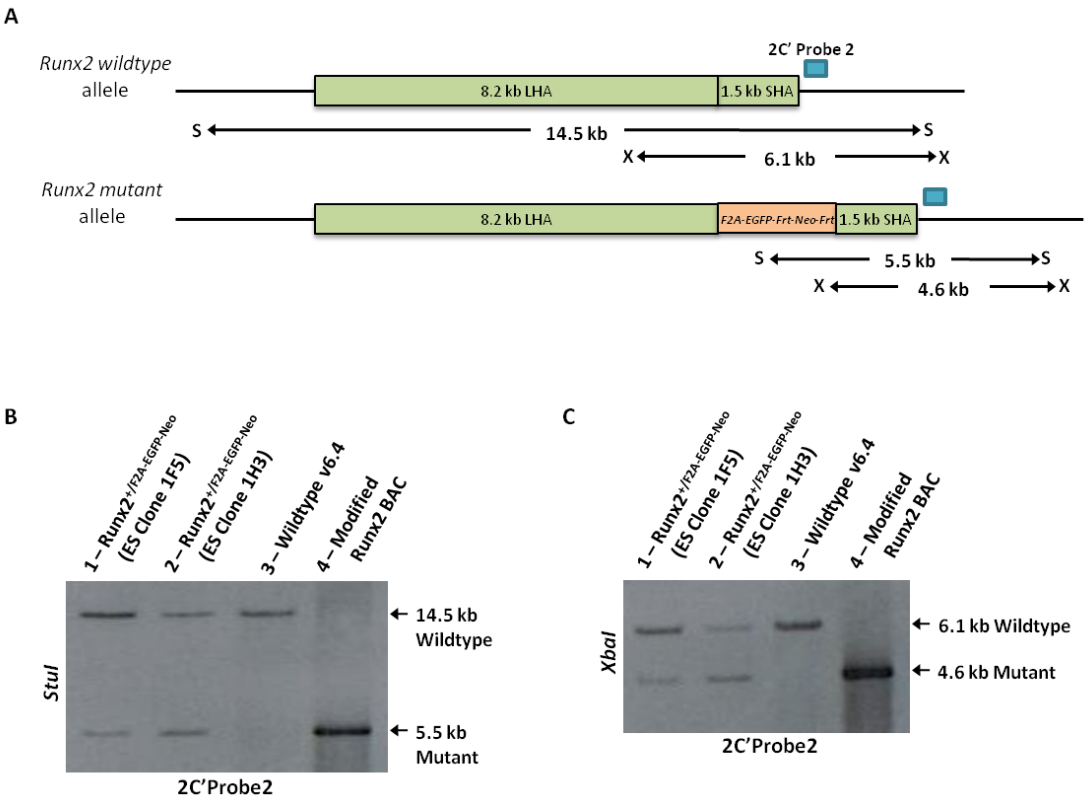


Figure 13. Southern Blot Screen for *Runx2*$^{+/F2A\text{-}EGFP\text{-}Neo}$ Positive Recombinant ES Cell Clones

Genomic DNA extracted from ES cell colonies resistant to G418 were digested with *StuI* (A) or *XbaI* (B) restriction enzyme and probed with a C-terminal Probe 2 (2C'Probe2; 590 bp; 30 ng/ml) just outside the short homology arm. True homologous recombinant ES cell clones will give a 14.5 kb wild-type and a 5.5 kb mutant band (A; lanes 1 & 2) when cut with *StuI* or a 6.1 kb wild-type and a 4.6 kb mutant band (B; lanes 1 & 2) when cut with *XbaI*. LHA, long homology arm; SHA, short homology arm; S, *StuI*; X, *XbaI*.

Table 1. Gene-Targeting Frequencies and ES Cell Lines used for Downstream Studies

Genotype	Targeting Frequency	Positive Clones	Clones that gave GLT Chimeras	ESC Clone used for Downstream Studies
Runx2$^{+/F2A\text{-}EGFP\text{-}Neo}$ *(Runx2WT-F2A-EGFP)*	8/192 (4.17%)	1B6 1C6 1F5 1H3 3B6 3E4 3F2 3F10	1B6 3B6 3E4	1B6
Runx2$^{+/EGFP\text{-}Neo}$ *(Runx2KO-EGFP)*	8/153 (5.23%)	8C2 9B5 9D2 10B6 11B6 12D4 13C6 14B2	9D2 12D4	9D2
Runx3$^{+/EGFP\text{-}Neo}$ *(Runx3KO-EGFP)*	26/93 (28.0%)	1A2 1A6 1B2 1C3 1D3 3A6 3B1 3C3 …	1A2 3B1	1A2

Gene-targeting frequencies at the *Runx2* and *Runx3* locus. Independent ESC clones derived from the same targeting constructs gave identical fluorescence pattern in the F1 mouse embryos of the chimeric mice generated. Only one mouse line was used to pursue downstream gene-expression studies. GLT, germ-line transmitting.

Positive ESC clones identified from the Southern blot screen were expanded for microinjections into 2- to 8-cell stage C57BL/6J mouse embryos (Kraus et al., 2010) to generate high percentage germ-line transmitting chimeras (Fig.14). The F1 offspring of the male chimeras when mated to C57BL/6J wild-type female mice were heterozygous for the mutant allele. These normal *Neo*$^{+}$ F1 mice were further mated to homozygous *Zp3*^{Cre} or *Rosa26R*^{Flpe} mice to flox out the *Neo* gene. For crosses with *Zp3*^{Cre} mice, only female F2 mice heterozygous for mutant *Runx2* or *Runx3* and *Zp3-Cre* alleles were mated to C57BL/6J wild-type male mice to obtain Neo^{-} and Cre^{-} mutant F3 *Runx2*$^{+/EGFP}$ or *Runx3*$^{+/EGFP}$ mice. The *Neo*$^{-}$ and *Cre*$^{-}$*/Flpe*$^{-}$ mutant mice were maintained by inter-crosses with C57BL/6J wild-type mice.

Figure 14. High-percentage *Runx2*$^{+/EGFP-Neo}$ Chimeras generated by ES Cell Microinjection

The mice were arranged in decreasing percentage of mutant ESC contribution from left to right. As the agouti allele from the 129/SvJ is dominant over the allele for black coat in the C57BL/6J, the hybrid v6.4 ESC lines microinjected into C57BL/6J mouse embryos would give rise to agouti-coated mice. These mice above (four on the left) are commonly referred to as high-percentage ESC-derived chimeras.

Tail tips from 3-week old mice or yolk sacs of embryos harvested for experiments were taken for genotyping by PCR. The following primers were used to determine the genotype:

Table 2. Primers for Genotyping

	Runx2WT-F2A-EGFP	***Runx2KO-EGFP***	***Runx3KO-EGFP***
Forward Primer	5' CCCAGCCACCTTT ACCTACA 3'	5' ACTCTGTCCGGTCT CCAGTC 3'	5' GCCACTTGATTCTCT AGGAT 3'
Reverse Primer	5' CTGCCTCTTGTCC CTTTCTG 3'	5' ACAGGAAGTTGGGA CTGTCG 3'	5' GGTGACATCCCCTT TCATGT 3'
Wild-type PCR band	809 bp	519bp	633 bp
Mutant PCR band	1,688 bp	1,298 bp	1,322 bp

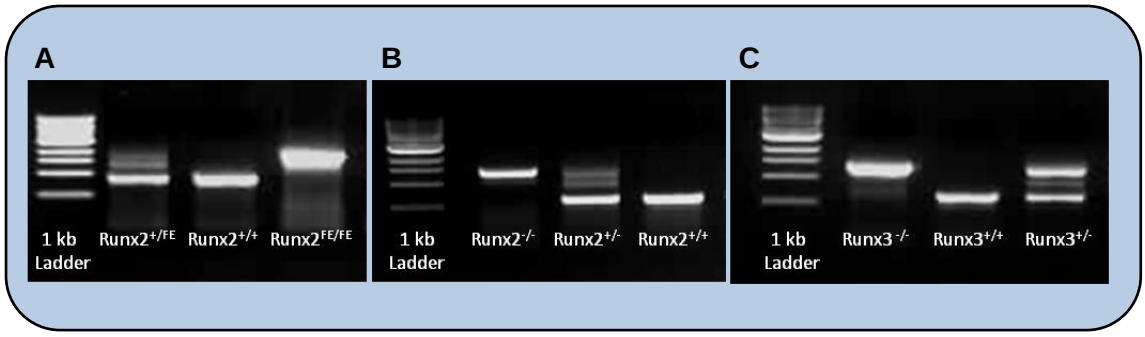


Figure 15. Mouse Genotype by PCR

(A) *Runx2WT-EGFP*; wild-type allele: 809 bp, mutant allele: 1.7 kb (B) *Runx2KO-EGFP*; wild-type allele: 519 bp, mutant allele: 1.3 kb (C) *Runx3KO-EGFP*; wild-type allele: 633 bp, mutant allele: 1.3 kb. *Runx2*$^{F2A\text{-}EGFP/F2A\text{-}EGFP}$ is referred as *Runx2*$^{FE/FE}$.

The homozygous *Runx2WT-F2A-EGFP* mice generated were normal, viable and fertile even when mated to a null background (i.e. *Runx2KO-EGFP*) and the genotypes of inter-heterozygous *Runx2WT-F2A-EGFP* mouse crosses were distributed according to the expected Mendelian ratio. Heterozygous *Runx2KO-EGFP* and *Runx3KO-EGFP* mice were viable and fertile. While homozygous *Runx2KO-EGFP* and *Runx3KO-EGFP* embryos up to 18.5 days post coitus (dpc) were obtained from crosses between heterozygous mice, no homozygous pups at

weaning age were identified by PCR genotyping. This was congruent with published literature reporting that *Runx2*^{null} mice died at birth due to respiratory failure and *Runx3*^{null} mice did not survive past a day after birth due to starvation (Komori et al., 1997; Li et al., 2002).

Mouse embryos of the various genotypes were harvested at different embryonic stages to assess the *EGFP* expression (Figs. 16-19). The fluorescence observed in all three mouse lines generated matched reported expression domains of *Runx2* or *Runx3* analyzed by RNA *in-situ* hybridizations respectively (Stricker et al., 2002). The expected outcome of the mouse crosses and the relevant fluorescence profiles indicated that the gene modifications carried out were correct and the desired wild-type and knockout mouse lines with fluorescence in the appropriate tissues were successfully created. Notably, the fluorescence of the *Runx2-F2A-EGFP* mouse embryos was much brighter than that of the *Runx2KO-EGFP* mouse embryos despite the same copy number of *EGFP*. This discrepancy in the fluorescence intensity between the two mouse lines could be attributed to the 21 amino acids fused at the N-terminus of EGFP that is derived from the P1 promoter (MASNS) in the *Runx2KO-EGFP* mice (refer to Fig. 8), thereby possibly preventing proper folding of EGFP and compromising its fluorescence. However, to determine the cause of the difference in fluorescence, it requires further experimentation that compares the transcript levels between the two mouse lines. Nevertheless, the weak fluorescence could still be detected by the FACS sorter.

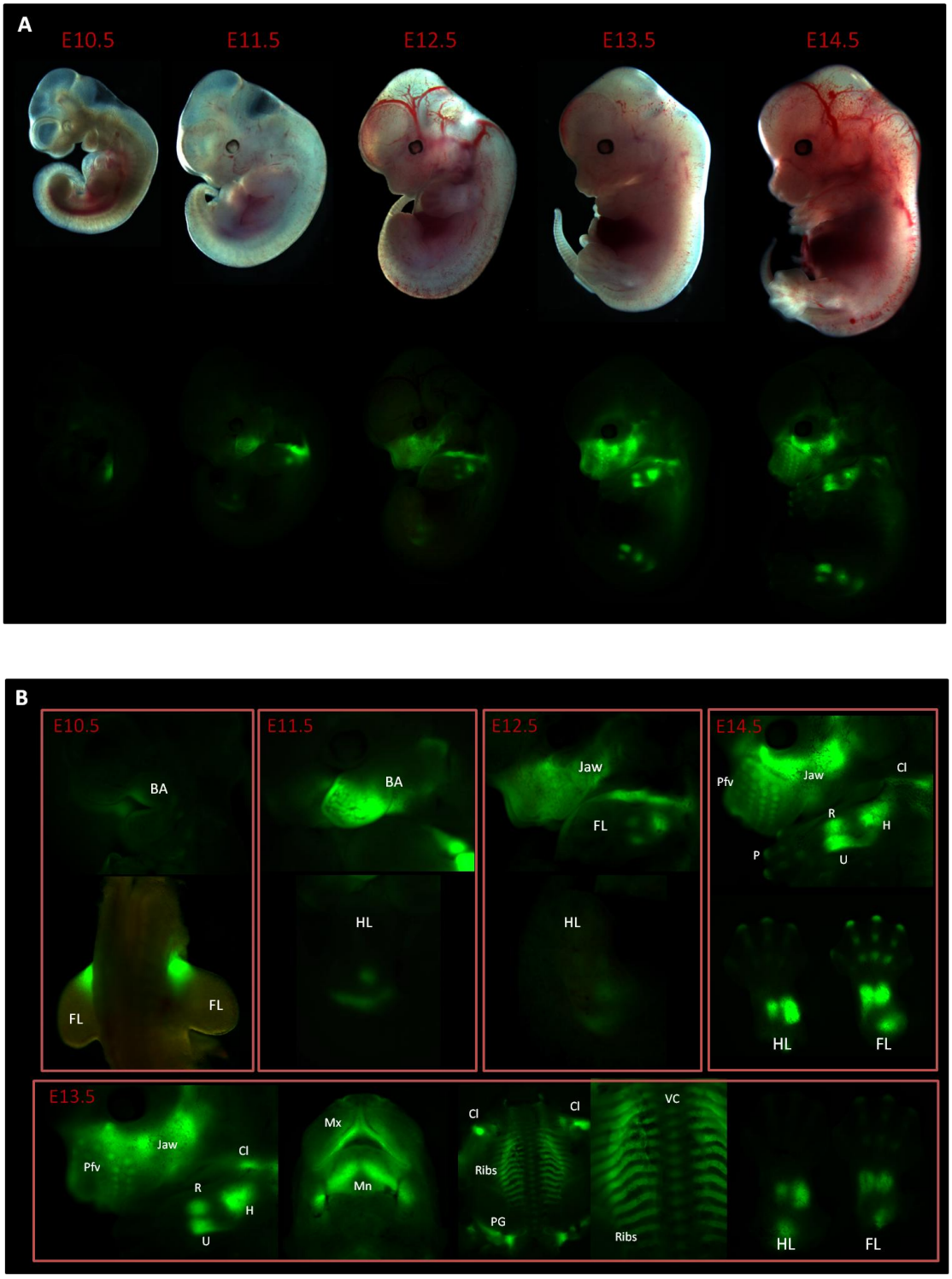


Figure 16. Green Fluorescence of *Runx2*$^{F2A\text{-}EGFP/F2A\text{-}EGFP}$ (*Runx2*$^{FE/FE}$) Wild-type Embryos

(A) E10.5 – E14.5 *Runx2*$^{FE/FE}$ mouse embryos taken under white light (top row) and ultraviolet light viewed through a GFP filter (bottom row) (B) E10.5 – E14.5 *Runx2*$^{FE/FE}$ mouse embryos at higher magnification (x80 to x120 magnification). *EGFP* expression mirrored the endogenous expression of *Runx2* in the mouse (Stricker et al., 2002). BA, branchial arch; FL, forelimb; HL, hindlimb; Pfv, primordium of follicle of vibrissae; P, phalanges; R, radius; U, ulna; H, humerus; Cl. Clavicle; Mx, maxilla; Mn, mandible; PG, pelvic girdle; VC, vertebral column.

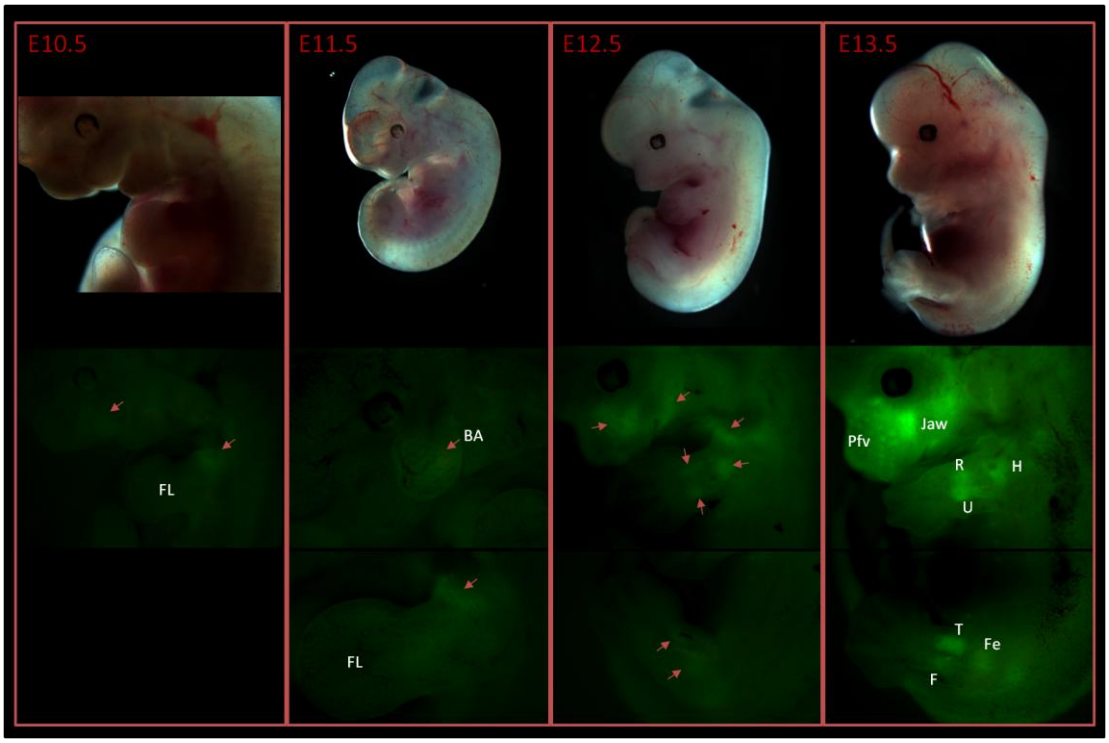


Figure 17. Green Fluorescence of E10.5 – E13.5 *Runx2*$^{EGFP/EGFP}$ Mutant Embryos

(A) E10.5 – E13.5 *Runx2*$^{EGFP/EGFP}$ mouse embryos taken under white light (top row) and ultraviolet light viewed through a GFP filter (bottom row). *EGFP* expression mirrored the endogenous expression of *Runx2* in the mouse (Stricker et al., 2002) but at a much lower level. BA, branchial arch; FL, forelimb; Pfv, primordium of follicle of vibrissae; R, radius; U, ulna; H, humerus.

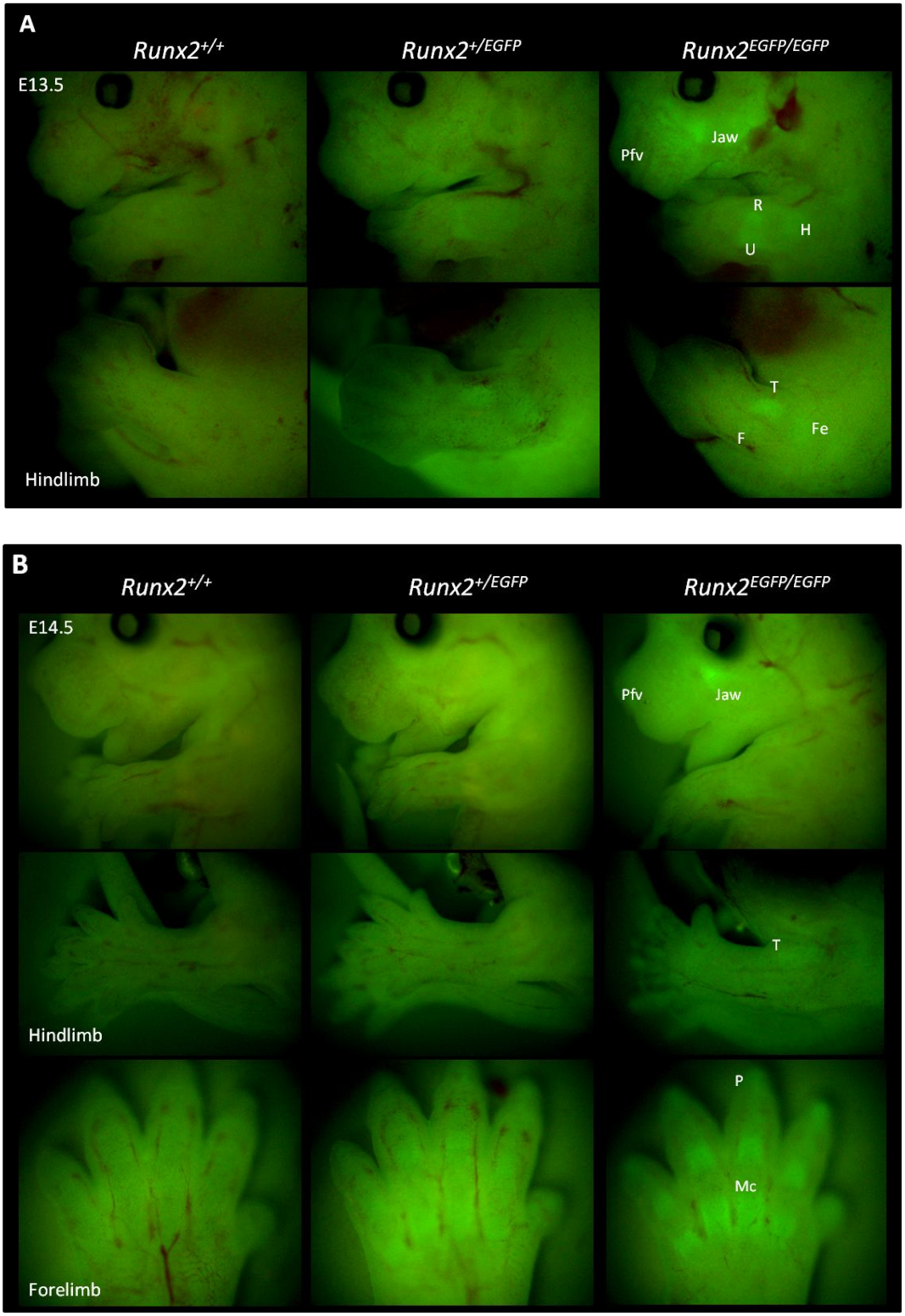


Figure 18. Green Fluorescence of *Runx2*$^{+/EGFP}$ and *Runx2*$^{EGFP/EGFP}$ Mutant Embryos

Fluorescence of *Runx2*$^{+/EGFP}$ and *Runx2*$^{EGFP/EGFP}$ mutant embryos at embryonic stages E13.5 (A) and E14.5 (B). Pfv, primordium of follicle of vibrissae; H, humerus; R, radius; U, ulna; T, tibia; F, fibula; Fe, femur; P, Phalanges; Mc, metacarpals.

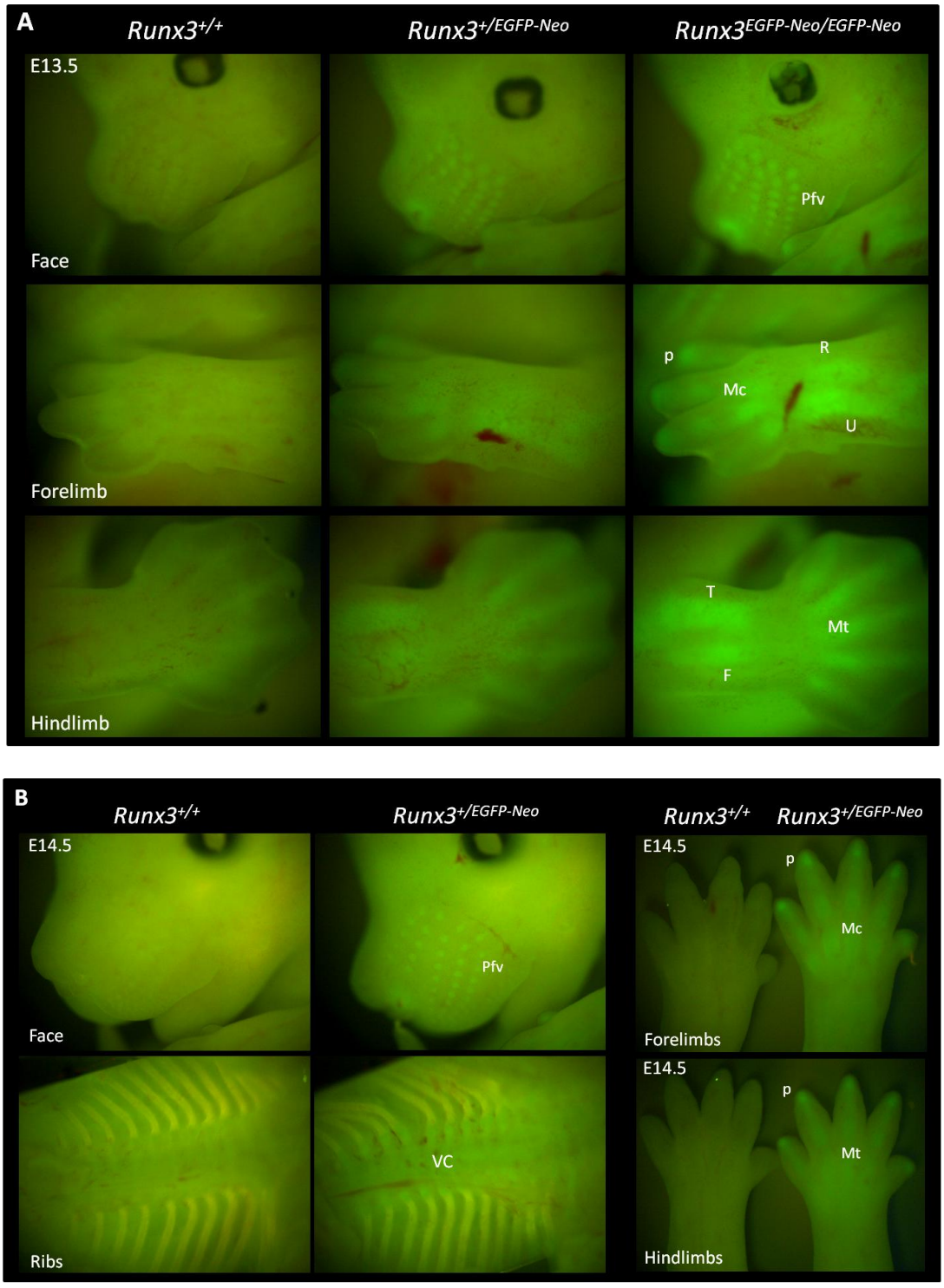


Figure 19. Green Fluorescence of *Runx3KO-EGFP-Neo* Mutant Embryos

Fluorescence of heterozygous and homozygous *Runx3KO-EGFP-Neo* mutant embryos at embryonic stages (A) E13.5 and (B) E14.5. Fluorescence detected recapitulated the endogenous expression of *Runx3* in the mouse (Stricker et al., 2002). Pfv, primordium of follicle of vibrissae; P, Phalanges; Mc, metacarpals; R, radius; U, ulna; T, tibia; F, fibula; Mt, metatarsals; VC, vertebral column.

3.1.2 Enrichment of Rare Population of *Runx2*- and/or *Runx3*-expressing Cells by FACS.

Ossification begins at E13.0, first osteoblasts emerge at E14.5 and bone mineralization commences at E15.5 in the mouse embryos. Initially, I wanted to unravel genes that were controlled by Runx2 and Runx3 at the onset of osteoblastogenesis and bone formation. In order to isolate cells that were embedded in the bone matrix for FACS, I tried to optimize the embryo dissociation protocol to break down the bone matrix of E15.5 mouse embryos into single cells by increasing the concentration of the collagenases and trypsin in the dissociation cocktail, adding varying amounts of EGTA, a calcium-chelating agent, into the dissociation buffer and incorporating 37°C incubation steps of varying durations. Unfortunately, the additional dissociation buffer components and 37°C incubations together with repeated manual pipetting were still inadequate to extricate the osteoblasts from the bone matrix of an E15.5 mouse embryo quick enough for FACS. Ultimately, harvesting osteoblasts from a late stage mouse embryo required manual pipetting and repeated incubations at 37°C for hours which were not feasible for my experiments.

I decided to work with E14.5 embryos for both *Runx2* and *Runx3* gene expression profiles since dissociation of E14.5 embryos with the buffer optimized in the laboratory was still achievable. As the Runx3 mutant mouse line was the first to be established, I started sorting $EGFP^+$ and $EGFP^-$ cells from heterozygous and homozygous E14.5 *Runx3* mutant mouse embryos. Isolation of $EGFP^+$ cells from E14.5 *Runx3* mutant mouse embryos was relatively easy and I could obtain sufficient cells for enough RNA to be put on the Illumina beadchip microarray. However, I could not obtain enough $EGFP^+$ cells from heterozygous E14.5 *Runx2* mutant mouse embryos for microarray. That could be because *Runx2* is highly expressed in the

osteoblasts at E14.5 and these cells are more tightly adhered in the bone matrix whereas *Runx3* is mainly expressed in the chondrocytes which are more easily dissociated than bone hence the *Runx3* mutant mouse embryos yielded more $EGFP^+$ cells after sorting. In view of that, I had to work with E13.5 *Runx2* mutant embryos instead.

Relevant skeletal tissues expressing EGFP (jaw, limbs, ribs and vertebral column) from E13.5 $Runx2^{+/FE}$, $Runx2^{FE/FE}$, $Runx2^{+/EGFP}$ and $Runx2^{EGFP/EGFP}$ mouse embryos as well as E14.5 $Runx3^{+/EGFP}$ and $Runx3^{EGFP/EGFP}$ mouse embryos were dissected out prior to dissociation and sorting with the fluorescence-activated cell sorter (FACSAria II). CD-1 wild-type embryos at the same developmental stage were dissected and dissociated in the same manner as the fluorescent embryos and used for gating the fluorescence. The genotype of each embryo was verified by PCR using the yolk sacs. The tissues were dissociated into single cells by manually pipetting up and down in the dissociation buffer and filtered through the 100 µm and 40 µm filter baskets to yield the first filtrate. Bits of hard bony tissue retained by the filter baskets were retrieved and subjected to repeated dissociation and filtration resulting in the subsequent filtrates. Figure 20 shows the FACS profiles and percentage of $EGFP^+$ cells corresponding to the different genotypes. As expected, there was a higher percentage of $EGFP^+$ cells found in the later filtrates compared to the first filtrate because *Runx2*- and *Runx3*-expressing cells were likely to be embeded in the bony tissues retained in the filter baskets. A much higher percentage of $EGFP^+$ cells were also found in $Runx2^{+/FE}$ and $Runx2^{FE/FE}$ mouse embryos (ranging from 6.3% - 14.7%) compared to $Runx2^{+/EGFP}$ and $Runx2^{EGFP/EGFP}$ mouse embryos (ranging from <0.1% - 2.4%) (Fig. 20). It was not surprising as the fluorescence in the $Runx2^{+/FE}$ and $Runx2^{FE/FE}$ mouse embryos were clearly higher than that in the $Runx2^{+/EGFP}$ and $Runx2^{EGFP/EGFP}$ mouse embryos (Figs. 16-18). Cells collected in 20% FBS/Leibovitz were pelleted and the collection medium was replaced with Trizol and stored in -80°C

until sufficient cells were collected for RNA extraction, amplification and array hybridization.

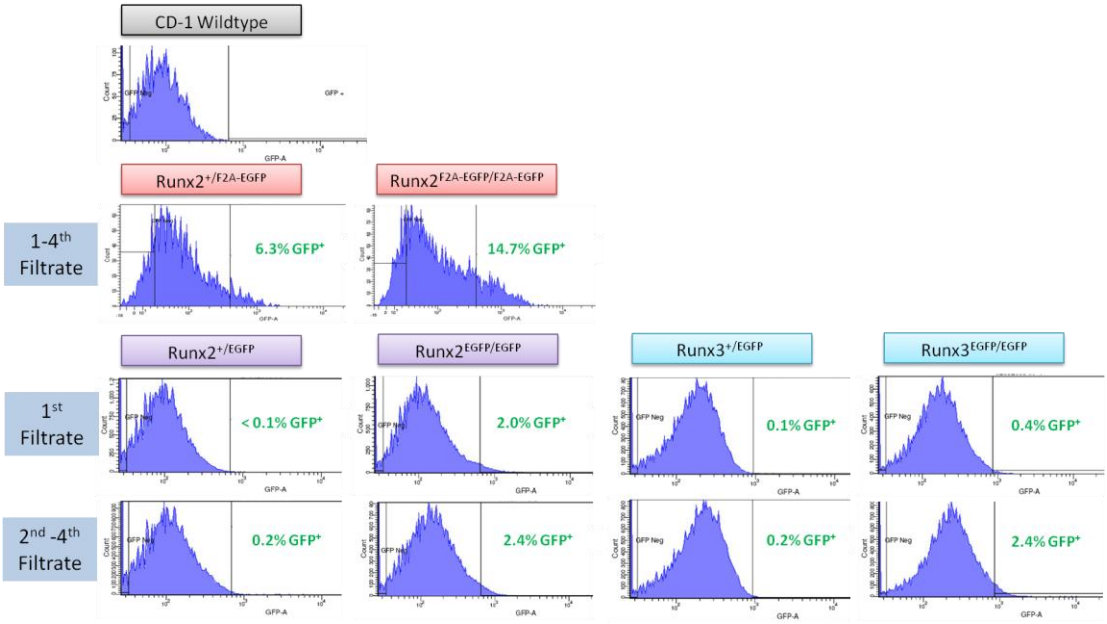


Figure 20. FACS Profiles of *Runx2* and *Runx3* Wild-type and Mutant Embryos

CD-1 wild-type mouse embryos were used to gate the fluorescence. Percentage of GFP^+ cells (in green) were calculated based on 100,000 cell counts.

3.1.3 Low RNA input amplification alternatives: Comparing two RNA amplification kits

One major challenge in exploiting microarray technology to compare transcriptomes is the large amount (1.5 μg) of labeled antisense RNA (aRNA) or cDNA required for hybridization on the high-density arrays. To overcome the obstacle of getting sufficient RNA material for hybridization, there is an assortment of RNA amplification kits requiring different amounts of starting RNA as input (ranging from 500 pg to 50 ng) commercially available to boost the amount of RNA obtained. These kits generally amplify RNA in a linear and 3'-initiated manner utilizing polyA primers. Considering that typical mammalian cells yield an average of 1 pg of total RNA per cell, I would need to collect 50,000 cells per biological replicate for RNA amplification

with the standard Illumina® TotalPrep RNA amplification kit that requires 50 ng of starting RNA as input. This would translate to 200,000 EGFP$^+$ cells to be collected for 4 biological replicates per genotype. With only 200 to 3,000 cells collected in an hour from each *Runx2*$^{+/EGFP}$ or *Runx3*$^{+/EGFP}$ mouse embryo, it would take me too long to acquire enough EGFP$^+$ cells for microarray. Therefore, I needed to find low RNA input amplification alternatives. As a pilot study, I compared two low input RNA amplification kits – TargetAmp™ 2-round Biotin-aRNA Amplification kit 3.0 and NuGEN® Ovation RNA Amplification System V2, using 1 ng of RNA extracted from EGFP$^+$ and EGFP$^-$ cells isolated from E13.5 *Runx2*$^{+/EGFP}$ mouse embryos as starting material for both kits (Table 3).

Table 3. Comparison of Two RNA Amplification Kits

	TargetAmp™ 2-round Biotin-aRNA Amplification kit 3.0	**NuGen® Ovation RNA Amplification System V2**
Starting Amount	50-500pg of RNA (Max. 1ng)	5-50ng of RNA. (Min. 1ng)
Amplification	Two-round linear amplification	Single-round linear amplification
Biotin Labeling	RNA amplified and labeled simultaneously to improve RNA yield.	Amplified cDNA labeled in 2 hours using a separate protocol.
Amplification Duration	2 days	0.5 days

Gene expression analysis was performed using the GeneSpring GX™ 11.0 software comparing EGFP$^+$ vs. EGFP$^-$ cells to find enrichment of relevant genes in the EGFP$^+$ fraction. Figure 21 clearly showed that there were more genes identified to be differentially expressed between the EGFP$^+$ and EGFP$^-$ fractions with the NuGEN® RNA amplification kit (9,830 genes) contrasted with the TargetAmp™ RNA amplification kit (4,073 genes) which could suggest that the NuGEN® kit is more

sensitive and is more able to amplify transcripts that are present at very low levels. More importantly, the *Runx2* transcript was found to be 39-fold enriched in the $EGFP^+$ compared to the $EGFP^-$ fraction when the NuGEN® RNA amplification kit was used (data not shown). However, *Runx2* transcript was not found to be enriched at all in the $EGFP^+$ fraction when the TargetAmp™ RNA amplification kit was used to amplify the same starting material.

The genes enriched in both the $EGFP^+$ fractions were further analyzed using an online functional annotation tool, DAVID Bioinformatics Resources 6.7 (NIAID/NIH) (Huang da et al., 2009a; Huang da et al., 2009b). GO terms were plotted against the enrichment scores for each GO annotation cluster derived from performing functional annotation clustering of the genes with DAVID (Figs. 21B & C). GO terms that were given enrichment scores higher than 1.3 (indicated by the red dotted line Fig. 21B & C) were considered significantly enriched. Genes enriched in the $EGFP^+$ fraction amplified by the NuGEN® kit were clustered as more precise skeletal annotations such as “bone and cartilage development” and “endochondral ossification” with enrichment scores of 2.15 and 1.51 respectively while those amplified by the TargetAmp™ were clustered as “bone development, osteoblast differentiation and ossification” with an enrichment score of only 1.51 (Figs. 21B & C). Although both the kits were able to amplify the starting RNA in a way that essentially preserves the original expression profile indicated by the relevant functional annotation clusters being above the significant threshold, the results presented in Figure 21 on the whole suggest that NuGEN® is more sensitive, produce less data noise and is less tedious to perform as the entire amplification and labeling procedure requires only 0.5 day compared to 2 days for TargetAmp™. For these reasons, I used the NuGEN® RNA amplification kits for all the subsequent RNA amplifications performed prior to the microarray studies. The results showing a 39-fold enrichment of *Runx2* transcripts as well as the significant enrichment of relevant skeletal genes in the $EGFP^+$ cells

compared to the EGFP$^-$ cells also demonstrated that the rare population of *Runx2*-expressing cells was truly enriched by FACS.

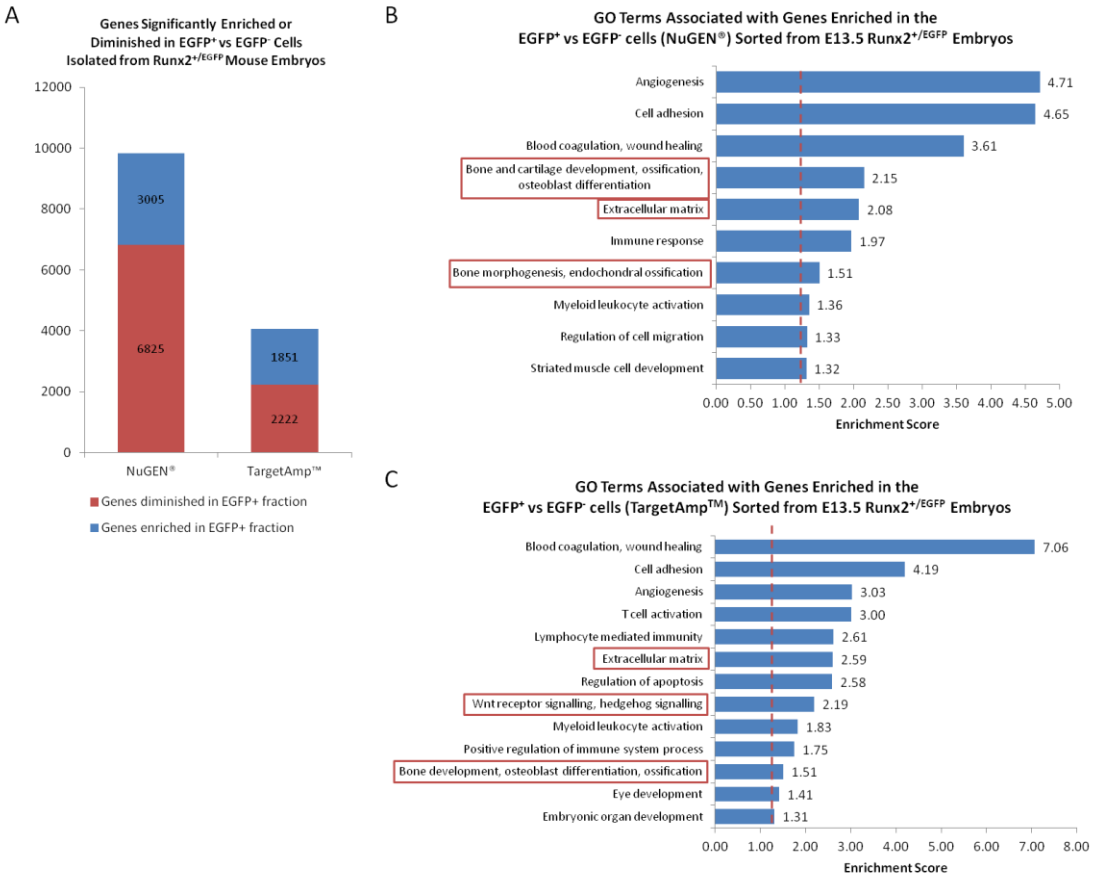


Figure 21. Comparing Results Produced by NuGEN® and TargetAmp™ RNA Amplification Kits

(A) A bar chart comparing the number of genes enriched or diminished in the EGFP$^+$ vs. EGFP$^-$ cells yielded by either the NuGEN® or the TargetAmp™ RNA amplification kits. (B and C) Each GO term associated with the genes enriched in the EGFP$^+$ cells was plotted against the enrichment score derived from performing functional annotation clustering of the genes on DAVID. The red dotted lines indicate the recommended cut-off point of 1.3. GO terms with enrichment scores above 1.3 indicate that genes with these functions are significantly enriched. The red boxes highlight relevant skeletal functions of Runx2.

After determining the RNA amplification kit to use and establishing that the FACS was able to achieve a good level of enrichment for the cell population of interest, total RNA was extracted, using the Trizol/Chloroform extraction method followed by a cleanup with the Qiagen® RNEasy® micro kit, from all the $EGFP^+$ cells isolated from *Runx2*$^{+/FE}$, *Runx2*$^{+/EGFP}$, *Runx2*$^{EGFP/EGFP}$, *Runx3*$^{+/EGFP}$ and *Runx3*$^{EGFP/EGFP}$ mouse embryos. This was followed by quantification with Ribogreen QuantIt™ fluorescent assay. RNA integrity was assessed using Agilent RNA 6000 Pico kit and analyzed with Agilent Bioanalyzer. Figures 22 & 23 contain the images of microcapillary electrophoretic separation of total RNA, extracted from $EGFP^+$ cells isolated from *Runx2*$^{+/FE}$, *Runx2*$^{+/EGFP}$, *Runx2*$^{EGFP/EGFP}$, *Runx3*$^{+/EGFP}$ and *Runx3*$^{EGFP/EGFP}$ mouse embryos, with the accompanying electropherograms on the right. The two prominent bands and peaks on the gel image and the electropherogram respectively represent the 28S and 18S ribosomal RNA (rRNA) and the Agilent Bioanalyzer uses the ratio of 28S:18S bands to assign a RNA integrity number (RIN) to each of the RNA samples. The RIN value of 10 indicates perfectly intact RNA while a value of 1 indicates completely degraded RNA (Schroeder et al., 2006). The RIN of the RNA samples were generally above 8 except for some which were below 5 as shown in Figures 22 & 23 and as summarized in Table 4. Currently, there is no consensus on what is an acceptable range for the various RNA applications. Even though the RNA extracted from the $EGFP^+$ cells of the *Runx2*$^{+/EGFP}$ samples had rather low RIN values and the RNA gel electrophoresis image did not match the standard ones, these samples were still used for microarray as it was extremely difficult to obtain sufficient $EGFP^+$ cells from this genotype due to the low fluorescence level of the embryos as evident in Figure 18. Besides, the NuGEN® Ovation RNA Amplification System V2 kit claimed to be able to amplify poor quality and partially degraded RNA.

Therefore, cDNA synthesis and amplification were performed with 1 ng of total RNA for all the samples using the NuGEN® Ovation RNA Amplification System V2 kit.

Three micrograms of the amplified cDNA were biotin-labeled and 1.5 μg of the labeled cDNA was applied onto the Illumina WG-6 Mouse Expression Beadchips for overnight hybridization. The beadchips were washed, blocked and hybridization signals were detected with Cy3-conjugated strepavidin (SA-Cy3) and scanned with the Illumina® BeadArray™ Reader the next day.

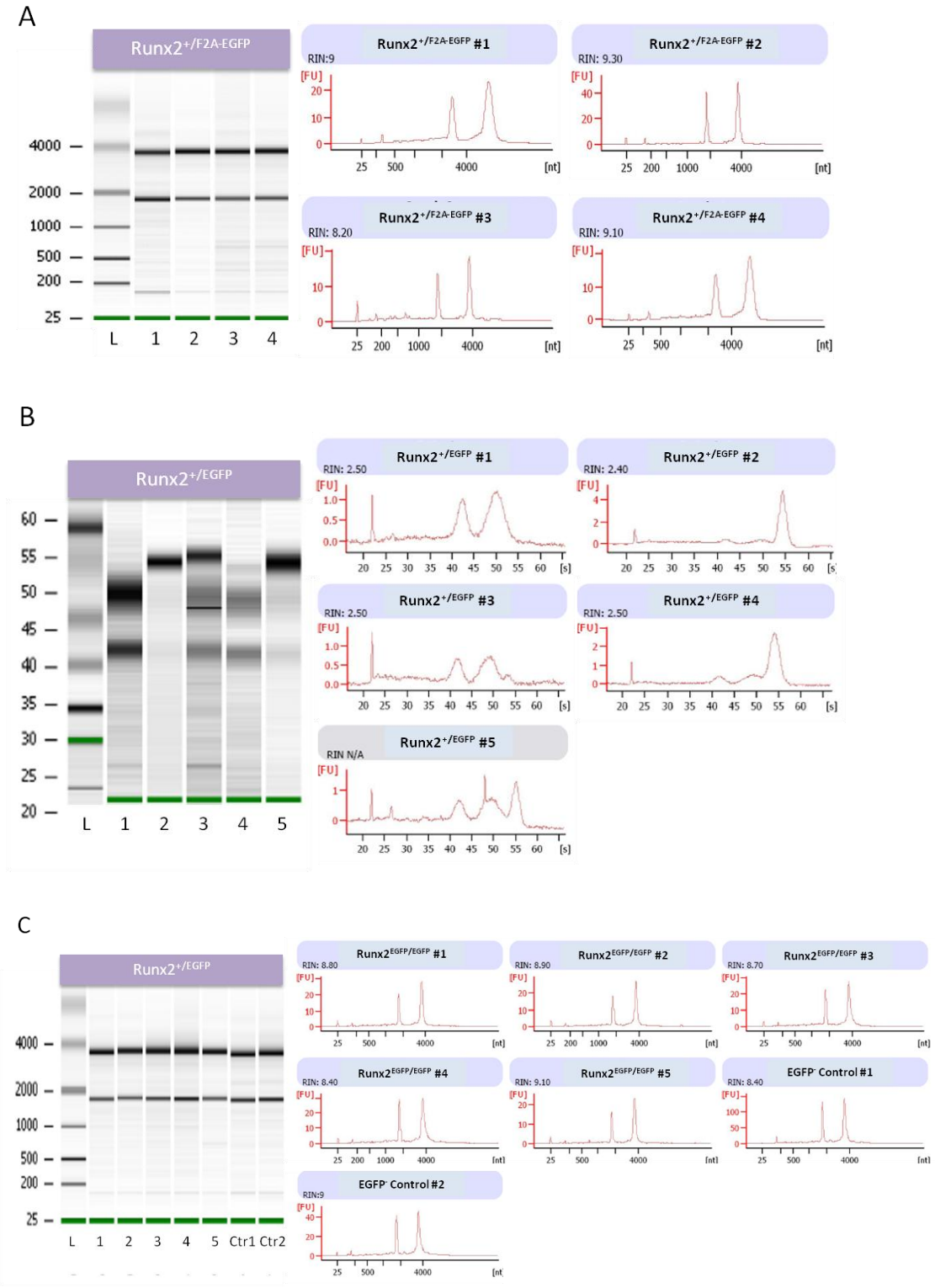


Figure 22. RNA Profiles of *Runx2-EGFP* Cells on Agilent Bioanalyzer Pico RNA Chip

Total RNA extracted from EGFP^{+} cells sorted out from E13.5 (A) *Runx2*$^{+/F2A\text{-}EGFP}$, (B) *Runx2*$^{+/EGFP}$ and (C) *Runx2*$^{EGFP/EGFP}$ mouse embryos were ran on a RNA Pico Chip using an Agilent Bioanalyzer. RNA from EGFP^{-} cells of *Runx2*$^{+/EGFP}$ were also checked for its RNA integrity. Left panel: Gel electrophoresis RNA profile. Right panel: Electropherogram with calculated RNA integrity number (RIN) to indicate the integrity of the RNA. The two prominent bands and peaks are 28S and 18S rRNA.

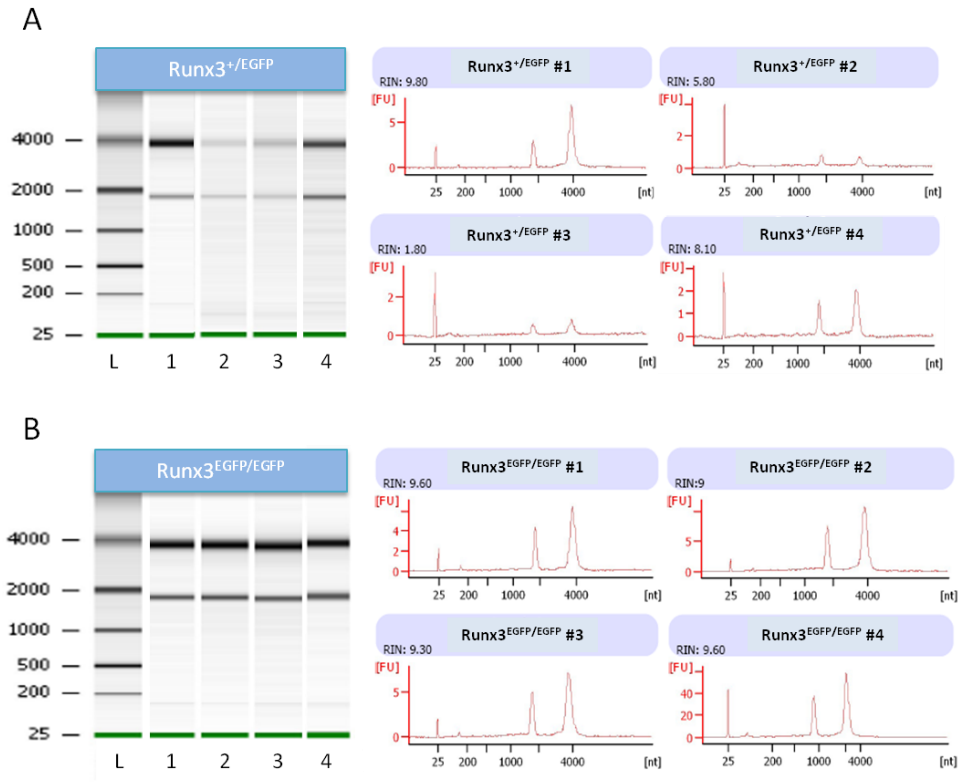


Figure 23. RNA Profiles of Runx3-EGFP Cells on Agilent Bioanalyzer Pico RNA Chip

Total RNA extracted from EGFP^{+} cells sorted out from (A) *Runx3*$^{+/EGFP}$ and (B) *Runx3*$^{EGFP/EGFP}$ mouse embryos were ran on a RNA Pico Chip using an Agilent Bioanalyzer. Left panel: Gel electrophoresis RNA profile. Right panel: Electropherogram with calculated RNA integrity number (RIN) to indicate the integrity of the RNA. The two prominent bands and peaks are 28S and 18S rRNA.

Table 4. Quantity and Quality of RNA extracted from EGFP^{+} Cells

Genotype	Biological Replicate *	No. of EGFP^{+} Cells	Total amt of RNA extracted (ng)	RIN
***Runx2*$^{+/F2A\text{-}EGFP}$**	# 1	59,050	30.30	9.0
	# 2	55,746	29.33	9.3
	# 3	49,412	16.82	8.2
	# 4	65,902	35.08	9.1
***Runx2*$^{+/EGFP}$**	# 1	5,398	2.67	2.5
	# 2	5,344	0.82	2.4
	# 3	5,629	2.69	2.5
	# 4	5,435	1.45	2.5
	# 5	5,395	2.30	N.A
***Runx2*$^{EGFP/EGFP}$**	# 1	28,298	14.45	8.8
	# 2	23,425	13.77	8.9
	# 3	26,189	17.59	8.7
	# 4	30,507	24.34	8.4
	# 5	20,786	10.73	9.1
***Runx3*$^{+/EGFP}$**	# 1	20,536	6.92	9.8
	# 2	1,394	3.43	5.8
	# 3	1,345	1.06	1.8
	# 4	4,500	3.37	8.1
***Runx3*$^{EGFP/EGFP}$**	# 1	35,207	30.60	9.6
	# 2	10,999	8.46	9.0
	# 3	13,075	17.30	9.3
	# 4	13,610	10.60	9.6

* Multiple embryos of the same genotype are pooled together to make each biological replicate

3.1.4 Runx2 and Runx3 Microarray Data Analysis

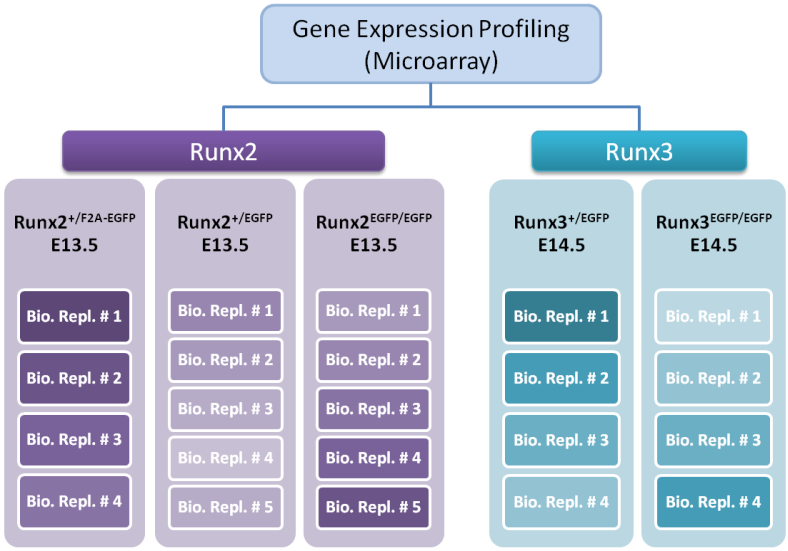


Figure 24. A Schematic Diagram of the Genotypes and the Number of Biological Replicates used for Microarray

Amplified and biotin-labelled RNA extracted from $EGFP^{+}$ cells of the various genotypes was loaded onto the Illumina Mouse WG-6 Beadarray chips. Although not clearly represented in this diagram, the biological replicates were randomized on the chips to overcome possible batch effects of the beadchips.

Raw image data files obtained from the Illumina® BeadArray™ Reader were downloaded into BeadStudio software to extract the probe lists coupled with the raw hybridization intensity values which were then exported as GeneSpring GX™ format files. Gene expression analyses were accomplished using the GeneSpring GX™ 11.0 software. Statistical tests such as the Analysis of Variance (ANOVA) and the unpaired Student's t-test for comparing $Runx2^{+/FE}(Runx2^{+/+})$ vs. $Runx2^{+/EGFP}(Runx2^{+/-})$ vs. $Runx2^{EGFP/EGFP}(Runx2^{-/-})$ and $Runx3^{+/EGFP}(Runx3^{+/-})$ vs. $Runx3^{EGFP/EGFP}(Runx3^{-/-})$ respectively were used to identify differentially expressing genes. All data were put through Benjamini-Hochberg FDR multiple testing corrections. Asymptotic p-values were computed and the significance threshold was set at p-value < 0.05 and fold change > 1.5. Figure 25 shows the number of genes differentially expressed in the various permutations of pair-wise comparisons. Notably, there were very few

significant Runx3 target genes identified when comparing *Runx3*$^{+/-}$ vs. *Runx3*$^{-/-}$. It has been observed in the laboratory that there is a strong correlation between overt phenotypes and large changes in gene expressions when a gene-of-interest is deleted. Hence, the small Runx3 target gene list could be due to the fact that there is a lack of an overt skeletal phenotype in *Runx3*$^{-/-}$. In all three *Runx2* genotype comparisons, there were more down-regulated genes in the genotype that had more *Runx2* allele knocked out. This could suggest that Runx2 acts as an activator more than a suppressor. In contrast, there were more up-regulated genes in *Runx3*$^{-/-}$ compared to *Runx3*$^{+/-}$ which likely implied that Runx3 serves as a suppressor more than an activator. This implication is supported by earlier reports that Runx3 has tumour suppressor roles in several cancers such as breast and gastric cancers (Li et al., 2002).

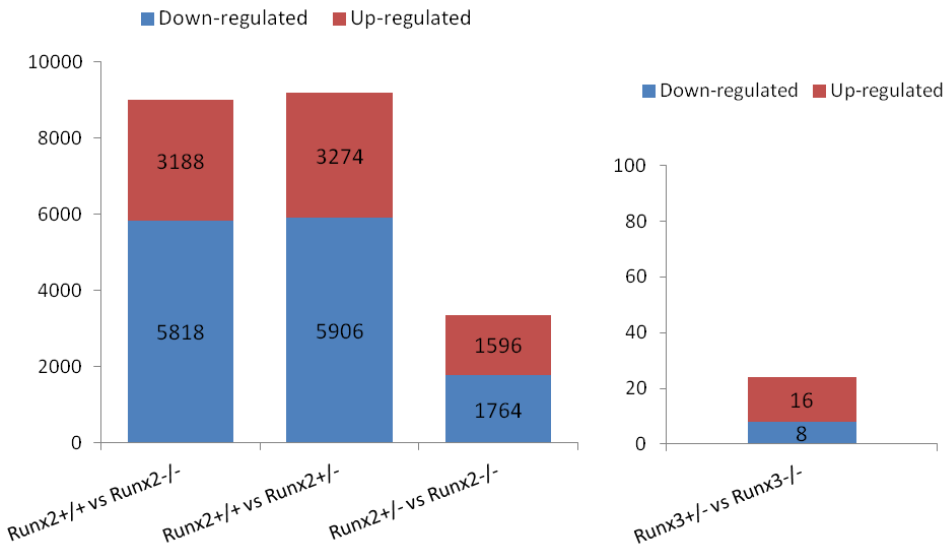


Figure 25. Number of Differentially Expressed Runx2 and Runx3 Target Genes

Bar charts summarizing the number of differentially expressed genes identified in the respective comparisons. Data were analyzed with GeneSpring GX™ 11.0 applying ANOVA test and unpaired Student's t-test for *Runx2* and *Runx3* comparisons respectively with a fold-change cut-off at >1.5 and p-value < 0.05. Down-regulated and up-regulated genes are genes with lower or higher expressions respectively in the latter genotype of the comparison.

The up- and down-regulated Runx2 putative target genes identified were analyzed with DAVID to investigate if they were relevant. Figures 26-28 reveal that genes giving rise to extracellular matrix proteins of the skeleton such as glycoproteins, collagens and fibronectins were highly enriched mainly in the down-regulated gene lists (3- to 22-fold enrichment in the *Runx2* knockout). Genes coding for proteins related to calcium-binding, glycosaminoglycan-binding, calmodulin-binding and collagen catabolic process were also significantly enriched in the down-regulated gene lists. Molecules involved in the Wnt signaling pathway were slightly enriched in the up-regulated gene list of $Runx2^{+/-}$ vs. $Runx2^{-/-}$ (Fig. 28). Genes particularly associated with skeletal and bone development, ossification, embryonic limb development, osteoblast differentiation, digit morphogenesis and bone mineralization were also enriched significantly ($1.3 < \text{enrichment score} < 5.44$) in both the up- and down-regulated gene lists. The observation that genes encoding skeletal extracellular matrix proteins were more highly enriched than genes specific to skeletal development could either indicate that Runx2 directly regulate these numerous structural genes leading to skeletal formation or the effect of knocking out *Runx2* by this developmental stage has mostly shifted to these secondary targets since *Runx2* expression had already begun several days before the embryos were harvested at E13.5 for microarray. The answer to this would require elucidation of Runx2 binding sites by performing ChIP-Seq of the embryonic skeletal tissue at this developmental stage.

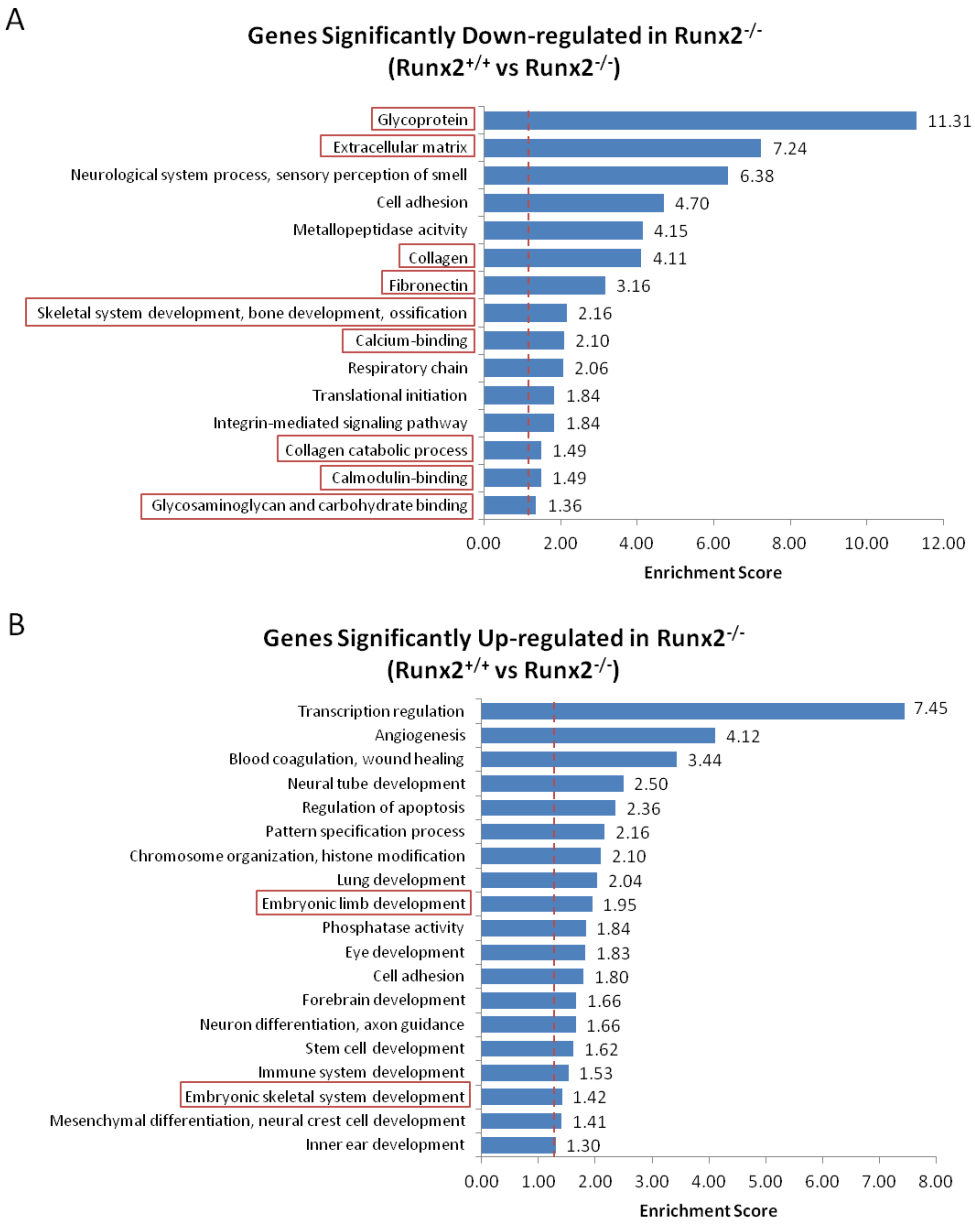


Figure 26. Functional Annotation Clusters of Differentially Expressed Runx2 Target Genes (*Runx2*$^{+/+}$ vs. *Runx2*$^{-/-}$)

Functional annotation clusters associated with the Runx2 target genes that were generated by DAVID were ranked according to their given enrichment scores. (A) Down-regulated and (B) Up-regulated genes in *Runx2*$^{-/-}$ compared to *Runx2*$^{+/+}$. The red dotted lines indicate the recommended cut-off point of 1.3. GO terms with enrichment scores above 1.3 indicate that genes with these functions are significantly enriched. The red boxes highlight relevant skeletal functions of Runx2.

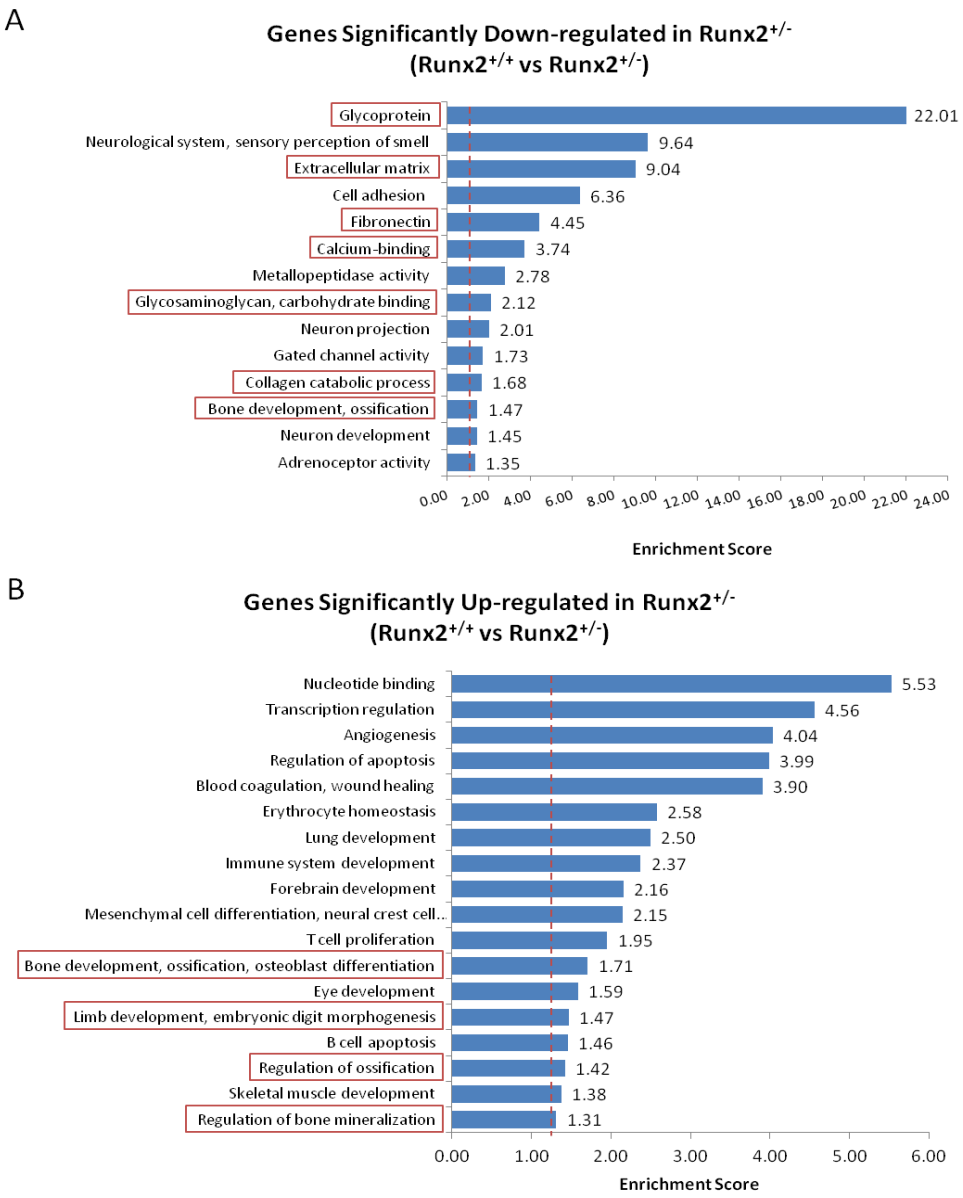


Figure 27. Functional Annotation Clusters of Differentially Expressed Runx2 Target Genes (*Runx2*$^{+/+}$ vs. *Runx2*$^{+/-}$)

Functional annotation clusters associated with the Runx2 target genes that were generated by DAVID were ranked according to their given enrichment scores. (A) Down-regulated and (B) Up-regulated genes in *Runx2*$^{+/-}$ compared to *Runx2*$^{+/+}$. The red dotted lines indicate the recommended cut-off point of 1.3. GO terms with enrichment scores above 1.3 indicate that genes with these functions are significantly enriched. The red boxes highlight relevant skeletal functions of Runx2.

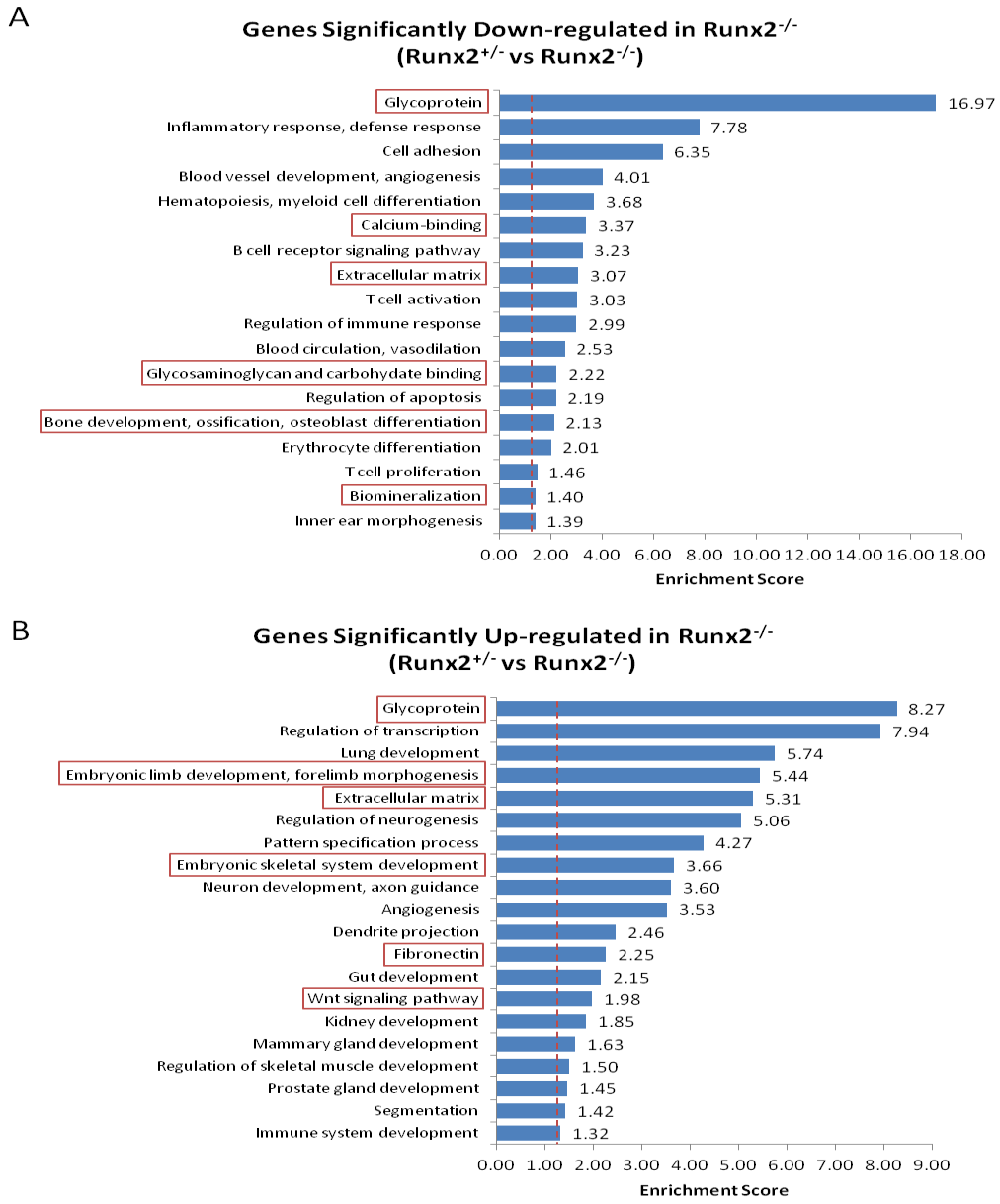


Figure 28. Functional Annotation Clusters of Differentially Expressed Runx2 Target Genes (*Runx2*$^{+/-}$ vs. *Runx2*$^{-/-}$)

Functional annotation clusters associated with the Runx2 target genes that were generated by DAVID were ranked according to their given enrichment scores. (A) Down-regulated and (B) Up-regulated genes in *Runx2*$^{-/-}$ compared to *Runx2*$^{+/-}$. The red dotted lines indicate the recommended cut-off point of 1.3. GO terms with enrichment scores above 1.3 indicate that genes with these functions are significantly enriched. The red boxes highlight relevant skeletal functions of Runx2.

To further examine how the *Runx2* gene lists relate to one another, they were compared using a Venn diagram tool, Venny. Figure 29 shows 82.7% (8,020 out of 9,692 genes) overlap between the *Runx2*$^{+/+}$ vs. *Runx2*$^{-/-}$ and the *Runx2*$^{+/+}$ vs. *Runx2*$^{+/-}$ gene lists while 24.5% (2,396 out of 9,765 genes) of the total number of differentially expressed genes were found in all three gene lists. When the genes in the core overlap were analyzed with DAVID, GO terms specific to embryonic skeletal development, ossification and osteoblast differentiation were enriched above the significant threshold (Fig. 29).

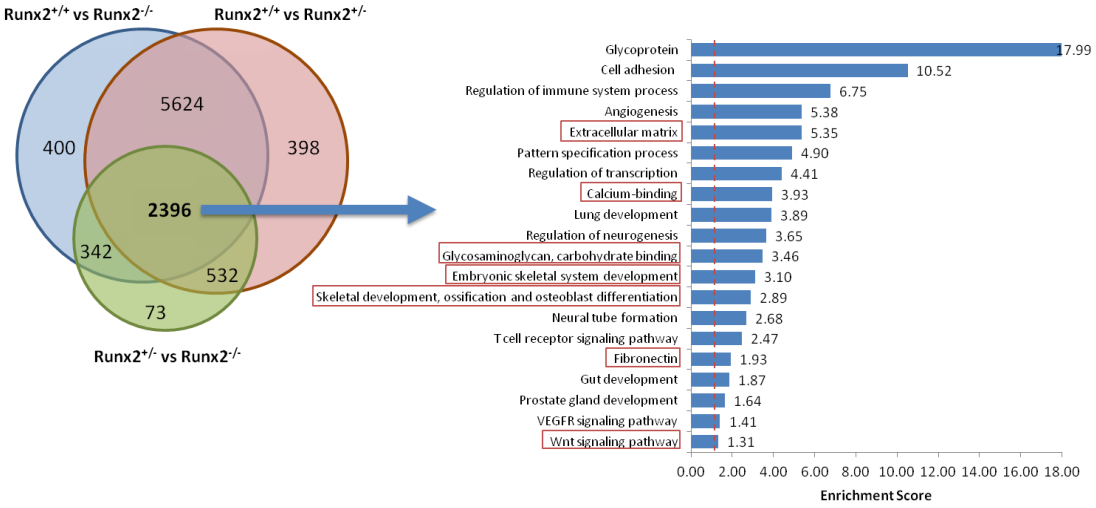
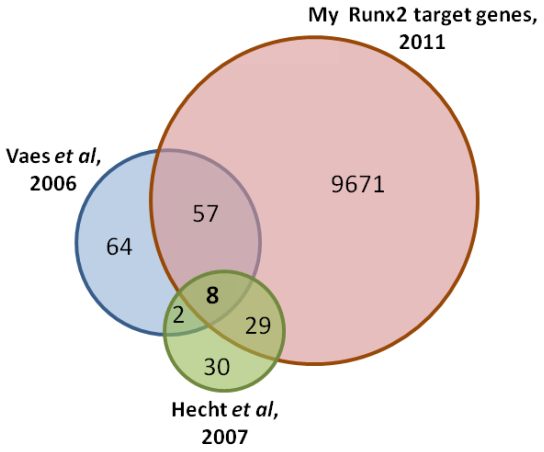


Figure 29. A Venn Diagram of Runx2 Target Genes and Enriched GO Terms Associated with the Core Overlap

Left: Lists of significant differentially expressed Runx2 target genes (p-value < 0.05; fold change >1.5) were compared using the interactive Venn diagram tool, VENNY. Venn diagram was re-drawn to proportions. Right: Functional annotation clusters associated with the core overlap of the Runx2 target genes were generated by DAVID and ranked according to the given enrichment scores. The red dotted lines indicate the recommended cut-off point of 1.3. GO terms with enrichment scores above 1.3 indicate that genes with these functions are significantly enriched. The red boxes highlight relevant skeletal functions of Runx2.

To ensure that the *Runx2* microarray data obtained were congruent with what had been reported in the literature, the Runx2 target genes were compared with previously published *in vivo* mouse *Runx2* microarray data by (Vaes et al., 2006) and (Hecht et al., 2007). Both studies used E14.5 mouse embryos and Affymetrix GeneChips. The only differences between the two studies are that Vaes *et al* (2006) compared the forelimbs, hindlimbs and calvariae of wild-type and *Runx2*$^{-/-}$ mice and analyzed both up- and down-regulated target genes in the *Runx2*$^{-/-}$ while Hecht *et al* (2007) compared only the humeri of wild-type and *Runx2*$^{-/-}$ mice and identified only the positively-regulated targets in the wild-type.

The Venn diagram in Figure 30 shows a good overlap between each of the previous independently reported Runx2 microarray studies and my data. This gave me confidence in my own data. The genes common in all three studies are mostly genes with known skeletal functions and/or are known direct targets of Runx2 (*Ibsp* and *Spp1*). Interestingly, there was a better overlap between my Runx2 target gene list and each of the two independent studies than the two studies had with each other even though both groups performed their microarrays with E14.5 mouse embryos and with the same Affymetrix GeneChips while my microarray was performed using E13.5 mouse embryos and Illumina WG-6 BeadChips. My Runx2 target gene list was also more extensive, consisting of 9,765 up- and down-regulated genes while Hecht *et al* (2007) identified only 69 positively-regulated Runx2 targets. Vaes *et al* (2006) originally found 1,277, 606 and 492 differentially modulated transcripts in the calvariae, forelimbs and hindlimbs respectively but assessed only the top 500 most significant transcripts based on the p-values for each tissue type and reported only 131 interesting target genes which were compared to my Runx2 data.



Common in all three studies	Common in Hecht *et al* (2007) and my study:		Common in Vaes *et al* (2006) and my study:		
Ibsp	***Smpd3***	*Apcdd1*	***Matn4***	***Tgfbi***	*Cxcl4*
Spp1	*Thy1*	***Phex***	***Col9a1***	*Tagln2*	*Gata2*
Satb2	*Fabp3*	*Ptprz1*	***Runx2***	***Sox9***	***Ghr***
Akp2	***S100a4***	***Hpgd***	*Ppp2r5d*	*2900010M23Rik*	*Lyzs*
Mmp9	***Npnt***	*Epha3*	*Ndufb10*	*Hbb-y*	***Ptn***
Cfh	***Tcf7***	***Mef2c***	*Fdps*	*Mif*	*Abca1*
Hck	***Ihh***	*Rasa3*	*Mfap5*	*Jub*	*Ebf3*
Lox	***Ccl9***	*Chst1*	***Dlk1***	***Col14a1***	*Gpm6b*
	Wnt5a	***Tmem119***	*Figf*	*Kitl*	***Ogn***
	Mmp13	*Gpx3*	*Sdpr*	***Dpt***	*Cxcl12*
	Atp6v0d2	*Capg*	*Cpxm2*	*Apoe*	*Rhoj*
	Cdo1	***Pthr1***	*Trim2*	*Fap*	***Igf1***
	Cnn2	*Gpc1*	*Abcf1*	*Mrc1*	*Ablim1*
	Ctsk	*Scd1*	*Hbb-bh1*	*Sepp1*	***Bgn***
		Anxa1	***Matn3***	***Matn2***	***Ppap2b***
			Matn1	*Cd36*	***Pdgfra***
			2310039H08Rik	***Dcn***	*Spnb2*
			Slc16a3	*Nova1*	*D0H4S114*
			Hba-x	*Igfbp7*	***Sparc***

Figure 30. Comparison between My Runx2 Target Genes and Two Published Runx2 Microarray Studies

Top: A Venn diagram comparing my list of Runx2 target genes and previously published Runx2 microarray analysis by Vaes *et al* (2006) and Hecht *et al* (2007). Bottom: Lists of genes found in the respective overlapping regions. Genes in bold have some form of known skeletal functions supported by literature and the Mouse Genome Informatics (MGI) database.

Not every cell in the entire limb or even humeri expresses *Runx2*; therefore, the effect of knocking out *Runx2* might be diluted out by transcripts in the other non *Runx2*-expressing cells within the same organ. As my microarray study utilized a pool of cells specifically enriched for *Runx2*-expressing cells obtained by FACS unlike the other two studies which used whole organs, I hypothesized that my data is more accurate and that the Runx2 target genes that could not be identified by the other two studies would be picked up by my microarray analysis. Hence, it would be interesting to examine the list of genes from my Runx2 microarray analysis that were not found by Vaes *et al* (2006) and Hecht *et al* (2007) to identify novel Runx2 targets that have not been previously discovered. In order to ensure that the Runx2 targets uncovered were genuine, validation of a few top Runx2 targets were performed by RNA section in situ hybridization which will be discussed in the next segment.

Akin to the functional annotation cluster analysis performed with the Runx2 target genes using DAVID, the Runx3 target genes were clustered to determine significantly enriched GO terms. However, the number of Runx3 target genes was too small and the enrichment score was below the significant level (data not shown). Currently, there are no Runx3 microarray data performed in the context of skeletal development reported for comparison. Thus, all 24 Runx3 target genes found in my microarray can be considered as novel Runx3 targets (Table 5). Genes with lower expression in the *Runx3*$^{-/-}$ are listed as down-regulated genes and vice versa. The genes in bold indicate that they were also found in the Runx2 target list and of the six Runx3 down-regulated genes in bold, three (*Ifi204*, *Clec7a* and *Lgmn*) were up-regulated and two (*Bcl2a1c* and *Use1*) were down-regulated in *Runx2*$^{-/-}$ and *Runx2*$^{+/-}$ compared to *Runx2*$^{+/+}$. *Adam15* was found in both up- and down-regulated lists of Runx2. Five out of sixteen Runx3 up-regulated genes were also found in the Runx2 target gene lists. *Lpar2*, *Ddr1* and *Zfp383* were regulated by Runx2 in the opposite direction as Runx3

while *Satb2* was regulated in the same direction. *Pgm5* was found in both the up- and down-regulated lists of Runx2.

Table 5. Runx3 Target Genes (*Runx3*$^{+/-}$ vs. *Runx3*$^{-/-}$).

Down-regulated	Fold Change	Up-regulated	Fold Change
Bcl2a1c	26.53	***Satb2***	126.92
Ifi204	20.43	*B230112C05Rik*	42.42
Adam15	14.99	*Myocd*	33.58
Clec7a	13.30	*Capn10*	27.08
1700003M02Rik	10.01	*E230017H14Rik*	20.96
Lgmn	7.30	*Ptpn20*	16.40
Ppfia3	7.00	***Lpar2***	15.99
Use1	6.07	***Ddr1***	15.38
		Gng5	14.55
		Cdadc1	13.52
		Pps	10.04
		Pgm5	8.81
		C230064E07Rik	4.80
		G3bp2	3.01
		LOC433261	2.04
		Zfp383	1.88

List of 24 differentially expressed Runx3 target genes (p-value < 0.05; fold change > 1.5) by comparing *Runx3*$^{+/-}$ vs. *Runx3*$^{-/-}$. Down-regulated and up-regulated genes were genes with lower or higher expression in *Runx3*$^{-/-}$ respectively compared to *Runx3*$^{+/-}$. The genes in bold were also found in the Runx2 target gene list.

Interferon activated gene 204 (Ifi204) protein, also known as p204, associates with Runx2 as a co-activator of osteoblast differentiation (Liu et al., 2005a) and pRB was found to link Ifi204 and Runx2 in a ternary complex at the promoters of *alkaline phosphatase* and *osteocalcin* to promote osteogenesis (Luan et al., 2007). The 5' cis-regulatory elements of *Ifi204* were also found to contain Runx2 and Sox5 binding sites. The binding of Runx2 at the promoter of *Ifi204* enhanced its expression and as a result induced chondrocyte hypertrophy in chondrocytic cell lines while Sox5 inhibited the expression of *Ifi204* (Zhang et al., 2008a). Thus, Ifi204 had been demonstrated in literature to cooperate with Runx2 primarily to regulate skeletogenesis but regulation between Runx3 and Ifi204 has not been suggested nor

shown. From the microarray data, it appears that Runx3 positively-regulates *Ifi204* and could be up-regulating it in the *Runx2*$^{-/-}$ to compensate the loss of Runx2. *Ifi204* might be the compensatory link between Runx2 and Runx3 in the skeletal network.

Satb2 was found to be a key node in regulating skeletal development. *Satb2*$^{-/-}$ mice exhibit craniofacial abnormalities resembling a cleft palate condition in humans with a translocation in the *SATB2* gene (Britanova et al., 2006; Dobreva et al., 2006). Satb2 was found to repress *Hoxa2* to release the inhibition on bone formation and directly interacts with and promotes the activity of Runx2 and ATF4 to drive osteoblast differentiation (Dobreva et al., 2006). It is not established if either Runx2 or Runx3 regulates *Satb2* as a negative feedback loop. From the microarray data, both Runx2 and/or Runx3 could be down-regulating *Satb2* to prevent an early onset of osteoblastogenesis that might lead to ectopic bone formation and limb defects if the osteoblast differentiation function of Runx2 were to be turned on too early before chondrocyte maturation could take place (Maeno et al., 2011).

Adam15 was suggested to have a homeostatic role in cartilage remodeling (Bohm et al., 2005). *Adam15* can be inferred to be positively regulated by Runx3 in cartilage development. *Clec7a, Lgmn, Bcl2a1c, Use1, Lpar2*, *Ddr1, Zfp383 and Pgm5* have no known skeletal functions as yet and could be potential novel targets of Runx3 in skeletogenesis and newly discovered factors contributing to skeletal development.

As mentioned earlier, the small Runx3 gene list is likely due to no overt skeletal phenotype in the *Runx3*$^{-/-}$ other than a slight delay in chondrocyte maturation. To reiterate, Runx3 is postulated to play a role in chondrocyte maturation during endochondral ossification that can be mostly compensated by Runx2 based on the study by (Yoshida et al., 2004) who reported that *Runx2*$^{-/-}$*Runx3*$^{-/-}$ mice had a more severe skeletal phenotype than *Runx2*$^{-/-}$ or *Runx3*$^{-/-}$ and chondrocyte maturation was

more impeded in *Runx2*$^{-/-}$ than in *Runx3*$^{-/-}$. Therefore, it is not surprising that there were not many differentially regulated genes uncovered from comparing *Runx3*$^{-/-}$ to *Runx3*$^{+/-}$ since the more dominant Runx member in skeletal development, Runx2, was likely to regulate most of Runx3 downstream target genes to compensate for the loss of Runx3. Henceforward, to elucidate the entire Runx3 targets involved in skeletogenesis, the transcriptional profiling of *Runx2*$^{-/-}$*Runx3*$^{-/-}$ had to be performed and compared to that of *Runx2*$^{-/-}$.

3.1.5 Validation of Runx2 Targets by SISH

Four top Runx2 targets (*Smpd3*, *Panx3*, *Ifitm5* and a novel *1200009I06Rik*) from the *Runx2*$^{+/+}$ vs. *Runx2*$^{-/-}$ list were validated by RNA in situ hybridizations on E13.5 mouse embryo sections. Figure 31 shows sagittal sections of the face, ribs, hindlimb and vertebral column. *Smpd3* expressions were seen in the mandible, maxilla and rib chondrocytes of the *Runx2*$^{+/+}$ and *Runx2*$^{+/-}$ mouse embryos. Weak signals were also detected in the femur and tibia of the *Runx2*$^{+/+}$ and *Runx2*$^{+/-}$ mouse embryos. The signals were completely abolished in the jaws, ribs and hindlimb of the *Runx2*$^{-/-}$ mouse embryo. Likewise, *Panx3* transcripts were detected in the clavicle, Meckel's cartilage, mandible, maxilla, rib chondrocytes and weakly in the hindlimb of the *Runx2*$^{+/+}$ and *Runx2*$^{+/-}$ mouse embryos but the expressions diminished drastically in the *Runx2*$^{-/-}$ embryos. Strong expressions of *Ifitm5* were found in the mandible and the basisphenoid bone of the *Runx2*$^{+/+}$ mouse embryo but not in the *Runx2*$^{+/-}$ and *Runx2*$^{-/-}$ mouse embryos. No expression of *Ifitm5* was found in the vertebral column and hindlimbs. Lastly, the expressions of a novel Riken gene *1200009I06Rik* were detected in the facial muscle, mandible, maxilla, intercostal muscles and rib perichondrium and chondrocytes of the *Runx2*$^{+/+}$ and *Runx2*$^{+/-}$ mouse embryos. In the hindlimb of the *Runx2*$^{+/+}$ and *Runx2*$^{+/-}$ mouse embryos, *1200009I06Rik* was detected in the perichondrium and maturing chondrocytes of the tibia, perichondrium of the metatarsals and mesenchymal condensations of the distal hindlimb digits. However,

the *1200009I06Rik* transcripts were visibly reduced in the facial muscle and the intercostal muscles of the *Runx2*$^{-/-}$ mouse embryos and expressions were obliterated in the mandible, maxilla, ribs, and hindlimbs of the *Runx2*$^{-/-}$ mouse embryos.

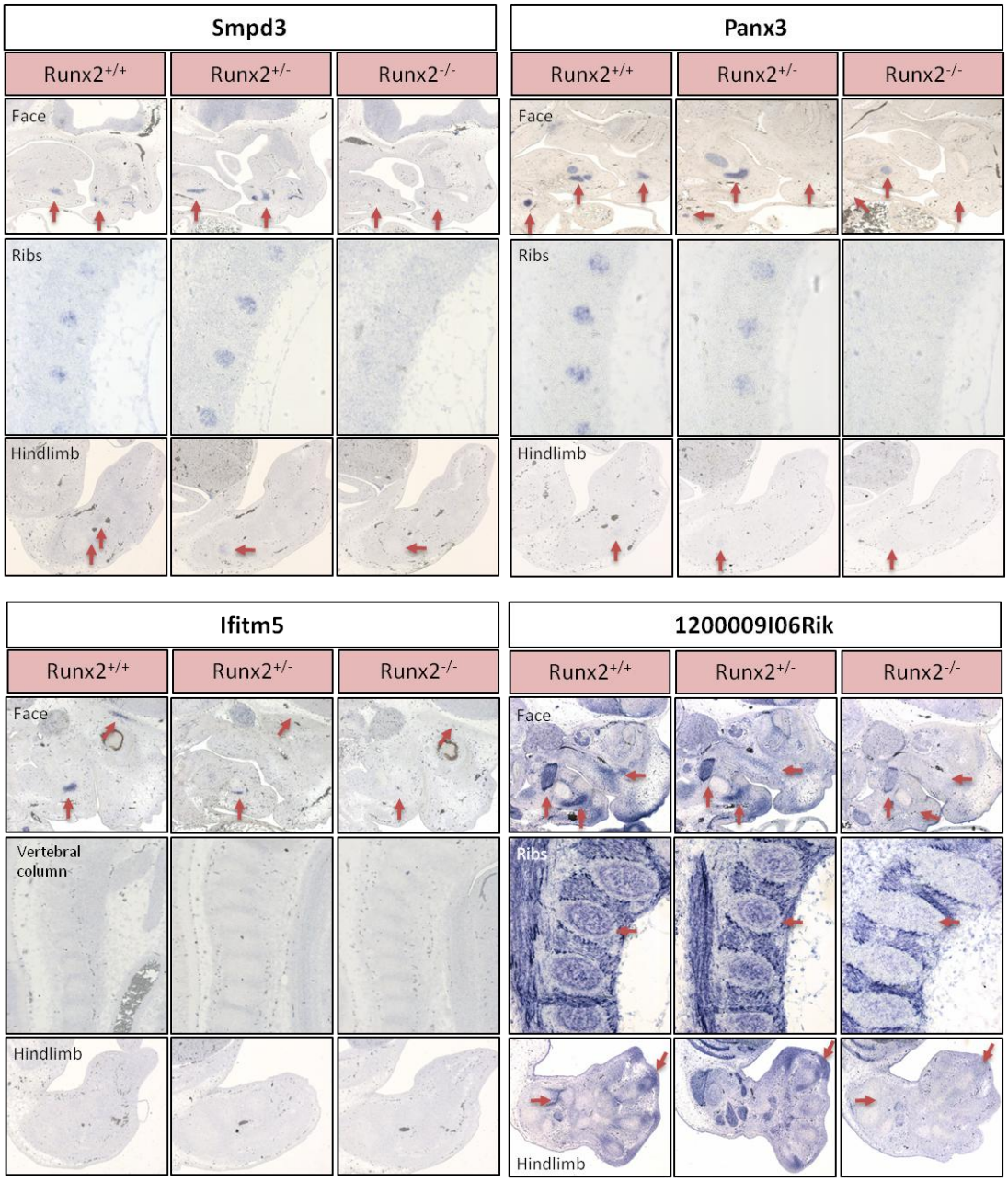


Figure 31. Validation of Runx2 Target Genes by RNA Section *In Situ* Hybridizations

In situ hybridizations of antisense *Smpd3*, *Panx3*, *Ifitm5* and a novel Riken gene *1200009I06Rik* RNA probes on sagittal E13.5 *Runx2*$^{+/+}$, *Runx2*$^{+/-}$ and *Runx2*$^{-/-}$ mouse sections. The red arrows highlight the RNA expression or the missing expression in the *Runx2*$^{-/-}$ mouse sections. Magnification: x200 (ribs) and x50 (face and hindlimb).

Sphingomyelin phosphodiesterase 3, neutral (*Smpd3*), also known as *Smase2*, is a member of the phosphodiesterase gene family of three: *Smpd1*, *Smpd2* and *Smpd3*. Smpd3 was originally identified as a brain-specific neutral sphingomyelinase which hydrolyzes sphingomyelin to ceramide and phosphorylcholine (Hofmann et al., 2000; Marchesini et al., 2003) until the year 2005 when Smpd3 was first discovered to cause osteogenesis and dentinogenesis imperfecta in mice upon its deletion (Aubin et al., 2005). Two years later, the *Smpd3*$^{-/-}$ mouse was again characterized as having a dwarf and chondrodysplasia phenotype that reflected common human achondrodysplasia (Stoffel et al., 2007). *Smpd3* was further demonstrated to respond positively to *Runx2* expression in an in vitro experiment and that Runx2 bound to the promoter of *Smpd3* in an electrophoretic mobility shift assay (EMSA) suggesting that Runx2 might regulate *Smpd3* directly during skeletal development (Chae et al., 2009). Hence, *Smpd3* found to be positively regulated by Runx2 in my Runx2 expression profiling and its expression in the skeletal elements which was obliterated in the *Runx2*$^{-/-}$ mouse embryos as illustrated by the RNA in situ hybridizations concurred with previous reports. Smpd3 is likely to have a role in skeletogenesis via Runx2 direct regulation although it still has to be shown if Runx2 regulates *Smpd3* directly *in vivo*.

Panx3, a member of the pannexin family which encodes for a class of gap junction proteins, was recently found to be highly expressed in the craniofacial flat bones and the long bones of the mouse axial and appendicular skeleton. It was detected specifically in the hypertrophic chondrocytes, perichondrium and osteoblasts (Bond et al., 2011; Iwamoto et al., 2010). Panx3 was demonstrated to regulate ATP/cAMP levels to promote chondrocyte differentiation in one study (Iwamoto et al., 2010) and shown to function as endoplasmic reticulum calcium channels, hemichannels and gap junctions to promote osteoblast differentiation in another study (Ishikawa et al., 2011). And very recently, it was established in luciferase and ChIP-PCR assays that

Runx2 was able to bind at the promoter of the *Panx3* between -275 and -283 bases to drive expression of *Panx3* thus putting forward the notion that *Panx3* is a direct target of Runx2 during bone development (Bond et al., 2011). The finding of *Panx3* transcripts greatly reduced in the *Runx2*$^{-/-}$ mouse embryos as revealed by both the microarray and RNA in situ hybridization in my experiments further supports the notion that *Panx3* might indeed be regulated by Runx2 and whether directly or not, it has yet to be determined *in vivo*.

Interferon-inducible transmembrane protein 5 (*Ifitm5*), also known as *Bril*, encodes for a small membrane protein that is highly expressed in osteoblasts (Moffatt et al., 2008). Its peak expression coincides with early bone mineralization during osteoblast maturation and knocking out both alleles of *Ifitm5* in the mouse resulted in smaller long bones. However, the phenotype became less severe in the adult mice possibly due to compensatory mechanisms by other factors or members of the interferon-inducible transmembrane protein family (Hanagata et al., 2010). There are no reports linking Runx2 and Ifitm5 in the regulation of bone formation so far. This could be a novel target of Runx2 especially in the craniofacial bone development as suggested by the RNA in situ hybridization results.

1200009I06Rik is a protein coding gene located on mouse chromosome 12. Currently, there is not much information on this gene. The expression of this gene is 15-fold decreased in the *Runx2*$^{-/-}$ mouse embryo when compared to the *Runx2*$^{+/-}$ mouse embryo in the Runx2 microarray and its significantly high transcript levels in both the muscles and skeletal elements suggest that this novel gene has a role in the development of both organs. It is potentially a downstream target of Runx2 and is likely to be involved in Runx2-modulated bone development as only its expressions in the skeletal elements were affected when *Runx2* was knocked out.

3.1.6 Transcriptional Profiling of *Runx2*$^{-/-}$*Runx3*$^{-/-}$ Mouse Embryos

Runx2$^{+/EGFP}$ (*Runx2*$^{+/-}$) and *Runx3*$^{+/EGFP}$ (*Runx3*$^{+/-}$) mice were crossed to produce *Runx2*$^{+/EGFP}$*Runx3*$^{+/EGFP}$ (*Runx2*$^{+/-}$*Runx3*$^{+/-}$) mice, which were grossly normal and fertile. These double heterozygotes were subsequently inter-mated to yield *Runx2*$^{EGFP/EGFP}$*Runx3*$^{EGFP/EGFP}$ (*Runx2*$^{-/-}$*Runx3*$^{-/-}$) mouse embryos at the Mendelian ratio of 1:16 (Fig. 32). Only the single and the double heterozygote mice survived and were normal in appearance. There were no live births of *Runx2*$^{-/-}$*Runx3*$^{+/-}$, *Runx2*$^{+/-}$*Runx3*$^{-/-}$ and *Runx2*$^{-/-}$*Runx3*$^{-/-}$ pups but E13.5 *Runx2*$^{-/-}$*Runx3*$^{+/-}$, *Runx2*$^{+/-}$*Runx3*$^{-/-}$ and *Runx2*$^{-/-}$ *Runx3*$^{-/-}$ mouse embryos could still be obtained.

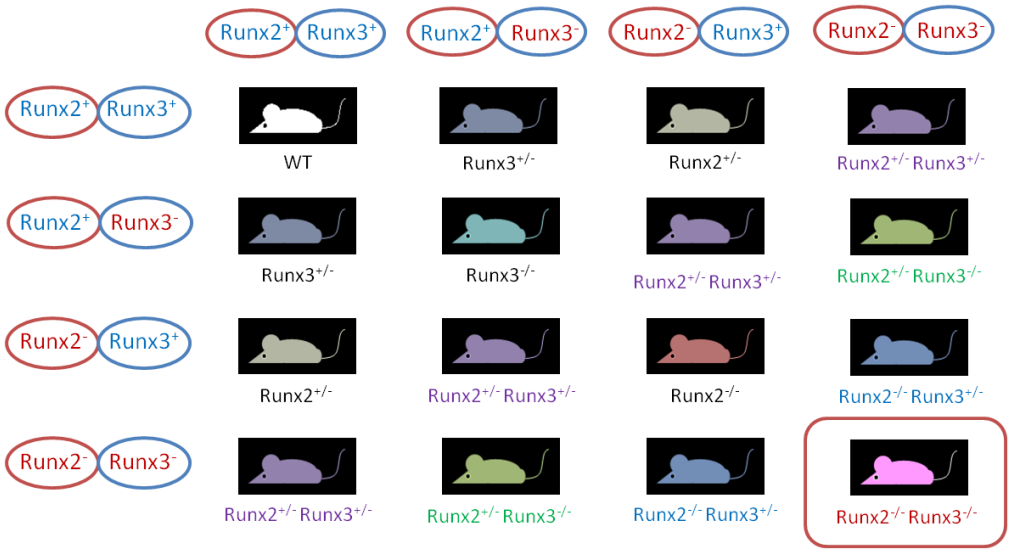
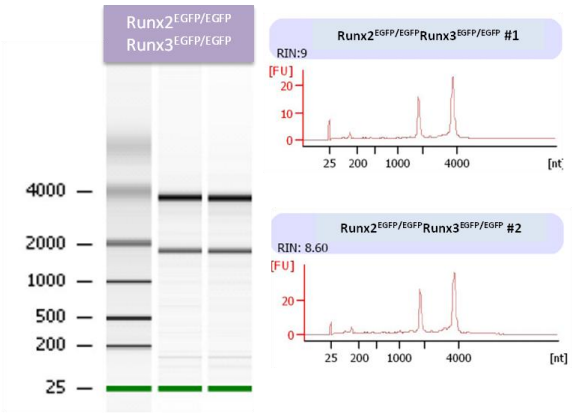


Figure 32. Inter-mating *Runx2*$^{+/-}$*Runx3*$^{+/-}$ Mice to Produce *Runx2*$^{-/-}$*Runx3*$^{-/-}$ Mouse Embryos

Runx2$^{-/-}$*Runx3*$^{-/-}$ mouse embryos were obtained by inter-mating *Runx2*$^{+/-}$*Runx3*$^{+/-}$ mice that were grossly normal and fertile. *Runx2*$^{-/-}$*Runx3*$^{-/-}$ mouse embryos were produced below the expected Mendelian inheritance ratio of 1:16.

EGFP^{+} cells from enriched tissues (jaw, limbs and ribs) of the brightest E13.5 mouse embryos produced by inter-mating *Runx2*$^{+/-}$*Runx3*$^{+/-}$ mice were isolated by FACS and yolk sacs were kept for PCR genotyping to determine the genotype of the embryos dissociated and sorted. After 6 months of sorting, only 2 out of 70 embryos sorted were *Runx2*$^{-/-}$*Runx3*$^{-/-}$. From PCR genotyping of all the embryos produced by the

double heterozygote crosses (data not shown), the *Runx2*$^{-/-}$*Runx3*$^{-/-}$ mouse embryos were observed to be produced at less than the expected Mendelian inheritance ratio of 1:16 indicating that many *Runx2*$^{-/-}$*Runx3*$^{-/-}$ mouse embryos die even before E13.5. Owing to time constraints, RNA was extracted (Fig. 33), amplified with the NuGEN® Ovation RNA Amplification System V2 kit and microarray was proceeded with the two *Runx2*$^{-/-}$*Runx3*$^{-/-}$ mouse embryos obtained. Technical replicates of the *Runx2*$^{-/-}$ *Runx3*$^{-/-}$ mouse embryos were used instead to give a total of 4 samples for the microarray.



Genotype	Embryo	No. of EGFP$^+$ Cells	Total amt of RNA extracted (ng)	RIN
Runx2$^{EGFP/EGFP}$ ***Runx3***$^{EGFP/EGFP}$ (***Runx2***$^{-/-}$***Runx3***$^{-/-}$)	# 1	9,110	12.0	9.0
	# 2	19,051	21.7	8.6

Figure 33. Profiles of RNA from EGFP$^+$ Cells of *Runx2*$^{-/-}$*Runx3*$^{-/-}$ mouse embryos

Total RNA extracted from EGFP$^+$ cells sorted out from E13.5 *Runx2*$^{EGFP/EGFP}$*Runx3*$^{EGFP/EGFP}$ (*Runx2*$^{-/-}$*Runx3*$^{-/-}$) mouse embryos were ran on a RNA Pico Chip using an Agilent Bioanalyzer. Top left panel: Gel electrophoresis RNA profile. Top right panel: Electropherogram with calculated RNA integrity number (RIN) to indicate the integrity of the RNA. The two prominent bands and peaks are 28S and 18S rRNA. Bottom: A table showing the number of EGFP$^+$ cells, quantity and quality of RNA obtained per embryo.

In order to circumvent the problem of not being able to pick out Runx3-regulated genes due to the likely compensation by Runx2 during chondrocyte maturation, it was hypothesized that by comparing the *Runx2*$^{-/-}$*Runx3*$^{-/-}$ microarray profile against

the *Runx2*$^{-/-}$ profile, Runx3 targets could be elucidated. Hence, the raw microarray profile for *Runx2*$^{-/-}$*Runx3*$^{-/-}$ was extracted using BeadStudio software and exported as GeneSpring GX™ format files for differential gene expression analysis by comparing *Runx2*$^{-/-}$*Runx3*$^{-/-}$ against the previous *Runx2*$^{-/-}$ microarray data using GeneSpring GX™ 11.0 software. In addition, *Runx2*$^{-/-}$*Runx3*$^{-/-}$ was compared against the previous *Runx3*$^{-/-}$ microarray data to investigate if there were new Runx2 target genes that were not identified in the *Runx2*$^{+/+}$ vs. *Runx2*$^{+/-}$ vs. *Runx2*$^{-/-}$ comparisons owing to compensation by Runx3. Unpaired Student's t-test and a volcano plot with significance threshold set at p-value < 0.05 and fold change > 2 was applied. All data were put through Benjamini-Hochberg FDR multiple testing corrections.

Figure 34 shows the number of differentially expressed genes identified from comparing *Runx2*$^{-/-}$*Runx3*$^{-/-}$ vs. *Runx2*$^{-/-}$ and *Runx2*$^{-/-}$*Runx3*$^{-/-}$ vs. *Runx3*$^{-/-}$. Genes with lower expressions in *Runx2*$^{-/-}$*Runx3*$^{-/-}$ were listed as down-regulated genes and vice versa in both the comparisons. There were more up-regulated genes in *Runx2*$^{-/-}$ *Runx3*$^{-/-}$ vs. *Runx2*$^{-/-}$ and more down-regulated genes in *Runx2*$^{-/-}$*Runx3*$^{-/-}$ vs. *Runx3*$^{-/-}$. In total, there were also more significant differentially modulated genes by Runx2 (6512) than Runx3 (7141). These observations support what has been established in literature that Runx2 plays a more prominent role in skeletal development and reiterate the connotation made earlier that Runx3 probably acts more as a repressor while Runx2 function more as an activator.

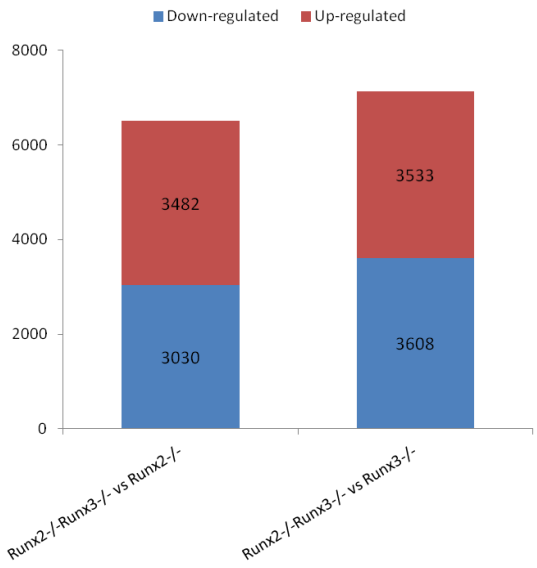


Figure 34. Number of Exclusively Runx2 or Runx3 Target Genes

A bar chart summarizing the number of differentially expressed genes identified in the *Runx2*$^{-/-}$*Runx3*$^{-/-}$ vs. *Runx2*$^{-/-}$ and *Runx2*$^{-/-}$*Runx3*$^{-/-}$ vs. *Runx3*$^{-/-}$ comparisons. Data was analyzed with GeneSpring GX™ 11.0 using unpaired Student's t-test with a fold-change cut-off at >1.5 and p-value < 0.05. Down-regulated and up-regulated genes are genes with lower or higher expressions in *Runx2*$^{-/-}$*Runx3*$^{-/-}$ for both comparisons.

The Runx3 targets identified were analyzed with DAVID to examine if they were relevant to skeletal development or known Runx3 functions (Fig. 35). Intriguingly, the top 2 enriched GO terms associated with genes positively-regulated by Runx3 were "transcription" and "transcription repressor activity". Evidently, Runx3 represses genes and activates repressor genes to establish itself as an overall transcription repressor. Genes associated with embryonic limb development, skeletal system development and axis patterning were significantly enriched between the enrichment scores of 1.93 to 2.48. Genes associated specifically with osteoblast differentiation and ossification were not enriched at all; suggesting that Runx3 is only involved in chondrocyte maturation and has no role in bone ossification. Genes associated with other known Runx3 functions such as neural development (Inoue et al., 2002; Levanon et al., 2002), lung development (Lee et al., 2010), immune system development (Woolf et al., 2003) and various cancers (Xiao and Liu, 2004; Yano et al., 2006) were also enriched significantly (Fig. 35A) in the down-regulated gene list.

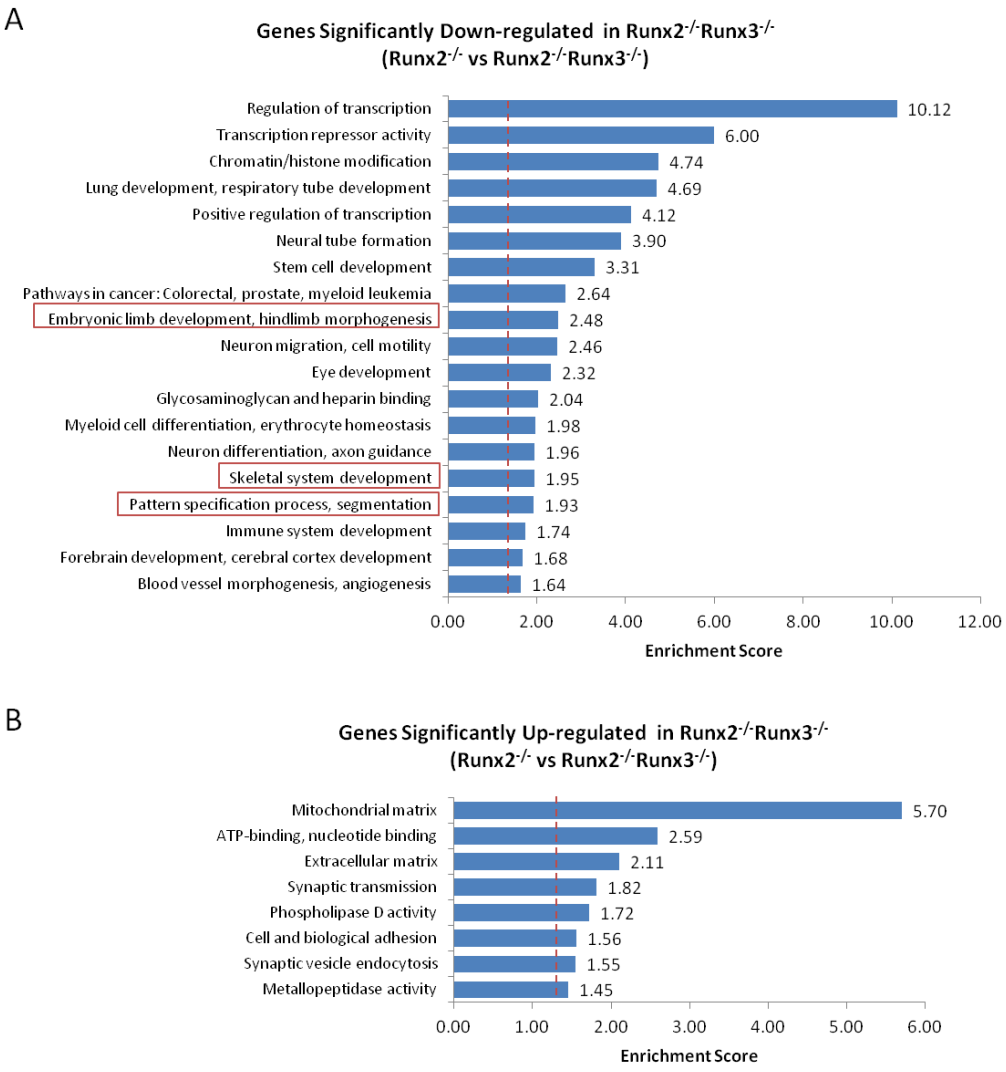


Figure 35. Functional Annotation Clusters of Differentially Expressed Runx3 Target Genes (*Runx2*$^{-/-}$*Runx3*$^{-/-}$ vs. *Runx2*$^{-/-}$)

Functional annotation clusters associated with the Runx3 target genes that were generated by DAVID were ranked according to their given enrichment scores. (A) Down-regulated and (B) Up-regulated genes in *Runx2*$^{-/-}$*Runx3*$^{-/-}$ compared to *Runx2*$^{-/-}$. The red dotted lines indicate the recommended cut-off point of 1.3. GO terms with enrichment scores above 1.3 indicate that genes with these functions are significantly enriched. The red boxes highlight skeletal functions.

Likewise, the Runx2 targets identified from *Runx2*$^{-/-}$*Runx3*$^{-/-}$ vs. *Runx3*$^{-/-}$ comparison were analyzed with DAVID to examine if they were relevant to skeletal development or known Runx2 functions (Fig. 36). The GO term specifically pertaining to bone development and ossification that was associated with the down-regulated gene list was ranked third with a significant enrichment score of 3.60 (Fig. 36A). The functional

annotation analysis concurs with the role of Runx2 in osteoblast differentiation and ossification.

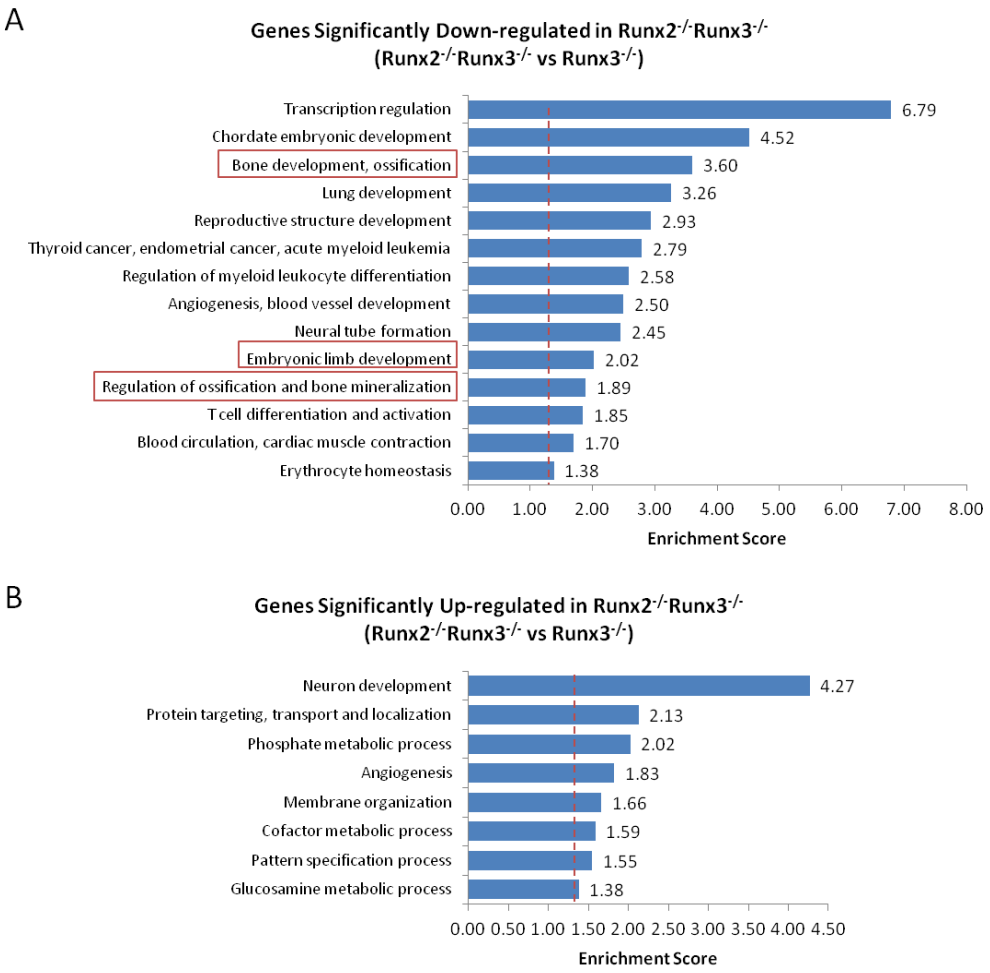


Figure 36. Functional Annotation Clusters of Differentially Expressed Runx2 Target Genes (*Runx2*$^{-/-}$*Runx3*$^{-/-}$ vs. *Runx3*$^{-/-}$)

Functional annotation clusters associated with the Runx2 target genes that were generated by DAVID were ranked according to their given enrichment scores. (A) Down-regulated and (B) Up-regulated genes in *Runx2*$^{-/-}$*Runx3*$^{-/-}$ compared to *Runx3*$^{-/-}$. The red dotted lines indicate the recommended cut-off point of 1.3. GO terms with enrichment scores above 1.3 indicate that genes with these functions are significantly enriched. The red boxes highlight relevant skeletal functions of Runx2.

Unfortunately, not all the genes originally found to be differentially expressed in the *Runx3*$^{+/-}$ vs. *Runx3*$^{-/-}$ comparison were differentially expressed in the *Runx2*$^{-/-}$*Runx3*$^{-/-}$ vs. *Runx2*$^{-/-}$ comparison. Of the 24 genes, only 11 were differentially expressed in the *Runx2*$^{-/-}$*Runx3*$^{-/-}$ compared to the *Runx2*$^{-/-}$. Out of the 11 gene, only 3 (*Lgmn*, *Ddr1* and *Pps*) were regulated in the same direction and the remaining 9 genes (*Adam15*, *Satb2*, *B230112C05Rik*, *Gng5*, *Pgm5*, *G3bp2*, *LOC433261*, *Zfp383*) were regulated in the opposite direction. *Ifi204*, thought to be the linking gene between the regulatory mechanisms of Runx2 and Runx3 was not found to be differentially expressed in the *Runx2*$^{-/-}$*Runx3*$^{-/-}$ vs. *Runx2*$^{-/-}$ comparison. This discrepancy could be due to the fact that the earlier Runx3 microarray was performed with E14.5 *Runx3*$^{+/-}$ and *Runx3*$^{-/-}$ mouse embryos while the *Runx2*$^{-/-}$*Runx3*$^{-/-}$ and the *Runx2*$^{-/-}$ mouse embryos were harvested at E13.5 for microarray. To maintain consistency, only the results obtained from E13.5 embryos were used for building the gene regulatory network.

3.2 Genome-wide Mapping of Runx2 and Runx3 Binding Sites

3.2.1 Introduction to Chromatin Immunoprecipitation-Sequencing (ChIP-Seq)

At the core of a transcriptional regulation network is the complex interaction amongst multiple transcription factors (TFs), chromatin modifiers and the polymerase complex to initiate or repress transcription at the transcriptional start site of target genes (MacQuarrie et al., 2011). This interaction is facilitated by the binding of TFs and/or chromatin modifiers at nearby or distant *cis*-regulatory DNA sites to induce proximity amongst the core complexes for interaction (Cosma, 2002; Fry and Peterson, 2001). Hence, determining where the TFs bind on the DNA to regulate gene expression complemented by genome-wide gene expression of these TFs enable the depiction of both direct and indirect links between the TFs and their target genes in a more comprehensive gene regulatory network.

Chromatin immunoprecipitation followed by hybridization onto DNA arrays (ChIP-chip) or high-throughput sequencing (ChIP-Seq) are two powerful methods used to detail genome-wide distribution of TF binding sites. With the lowering cost of high-throughput sequencing, ChIP-Seq, however, has become a more popular alternative to ChIP-chip to map the genome-wide physical DNA-TF interactions. The latter which requires hybridization on an array has inherent hybridization biases and its whole-genome mapping power and resolution is restricted by a fixed number of probes and their predetermined lengths on the array. ChIP-seq, on the other hand, requires less input material and can generate more precise binding site information as data is presented in short sequence reads (~30 base pairs) (Robertson et al., 2007). As a sequencing-based technology, it is also able to capture mutation in binding site sequences (Mardis, 2007) and identify novel binding sites that are not limited to an array. Nevertheless, the more superior ChIP-Seq with all its advantages has a caveat

in the form of sequencing bias that is yet to be well studied and understood (Mardis, 2007).

The foremost part of ChIP-seq entails the enrichment of the DNA sequences bound by the TF of interest in living cells. Whole mouse embryos or enriched mouse tissue parts are first homogenized into single cells with a Douncer. Proteins and their target DNA are crosslinked with formaldehyde before the chromatin and the bound proteins are lysed from the cells. The chromatin is then sheared to fragments between 100 to 500 base pairs by sonication. The DNA regions bound by proteins remain intact. An antibody specific to the TF of interest is used to exclusively immunoprecipitate the TF-bound DNA fragments. The TF-DNA complexes are next purified and reverse-crosslinked. The retrieved DNA fragments are further purified, ligated with adaptors, amplified and size selected (200-400bp) to produce a library of target DNA binding sites for the second half of ChIP-Seq and that is parallel sequencing with a high-throughput Solexa sequencer. Short reads of 30 to 50 bases from either ends of the DNA fragments are subsequently mapped back to the genome. The number of reads associated with the genomic locus will be proportional to the occupancy of the TF at that locus.

3.2.2 Generation of HA_3-tagged Mice for ChIP-Seq

As previously discussed, Runx proteins are highly similar in structure and commercially available anti-Runx2 and anti-Runx3 antibodies may not be specific enough to discriminate Runx2 or Runx3 from the other Runx members. Therefore, to avoid issues of antibody cross-reactivity and antibody promiscuity, C-terminally HA-tagged Runx2 and Runx3 mice were generated via locus-specific homologous recombination for ChIP experiments with HA antibodies. This would enable the specific tagged and DNA-bound protein-of-interest to be pulled down in ChIP

experiments prior to sequencing. The tag was intentionally placed at the C-terminal end so that all the Runx2 and Runx3 protein isoforms which differ at the N-terminus would be tagged. Three HA epitopes were added instead of one to increase the level of recognition by the antibody.

A long forward primer containing a 50bp homology arm followed by a triple HA tag (HA_3), a stop codon (TGA) and a 20bp sequence complementary to the 5' region of the *PGK-gb2-Neo* fragment together with a 70bp reverse primer (a 50bp homology arm and a 20bp complementing the 3' region of *Neo*) were used to synthesize the *HA_3-TGA-PGK-gb2-Neo* insertion cassette. This cassette with the flanking 50bp homology arms was inserted just before the stop codon of *Runx2* and *Runx3* via BAC modifications done with the Gene Bridges Quick and Easy BAC Modification kit (Fig. 37). BACs containing the 3' region of the genomic locus of *Runx2* (RP23-7C18) and the entire genomic locus of *Runx3* (RP24-118B14) were used for the modifications. Upon modification, the region-of-interest was amplified by PCR and verified by double-stranded sequencing to ensure that there were no errors in the modification. Successfully-modified and mutation-free BAC clones were subcloned into a minimal vector using the Gene Bridges Quick and Easy BAC Subcloning kit. The subcloned plasmids comprised the modified region flanked by short and long homology arms whereby the combined homology arms added up to at least 10 kb in length. Subsequently, these subclones were linearized with *XmnI (*for modified *Runx2)* and *AcI (*for modified *Runx3)* restriction enzymes which cut only within the minimal vector before being electroporated into the mouse ESCs for gene-targeting.

The targeting constructs were introduced into mouse v6.4 ESCs (129/SvJ x C57BL/6J) and putative positive recombinants were selected as mentioned and screened first by PCR with a forward primer sitting within the insertion cassette and a reverse primer outside the short homology arm. This was followed by verification of

the PCR-positive clones using Southern blotting with probes external to the homology arms (Figs. 38 & 39).

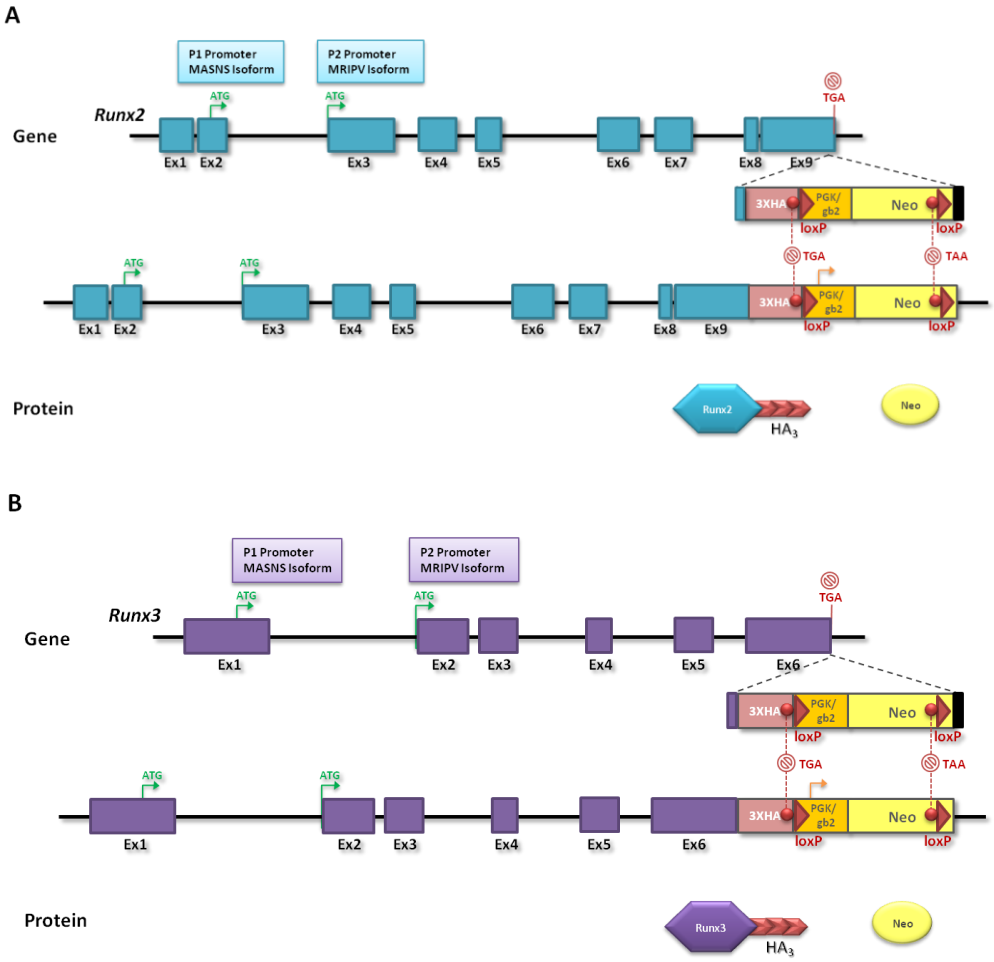


Figure 37. Targeting Strategy to Generate HA_3-tagged Mice

The *HA_3-TGA-loxP-Neo-loxP* cassette was inserted at the C-terminus of (A) *Runx2* and (B) *Runx3* just before the stop codon. The *Neo* gene is excised by crossing with the *Zp3-Cre* mice. HA. Hemaglutinin A.

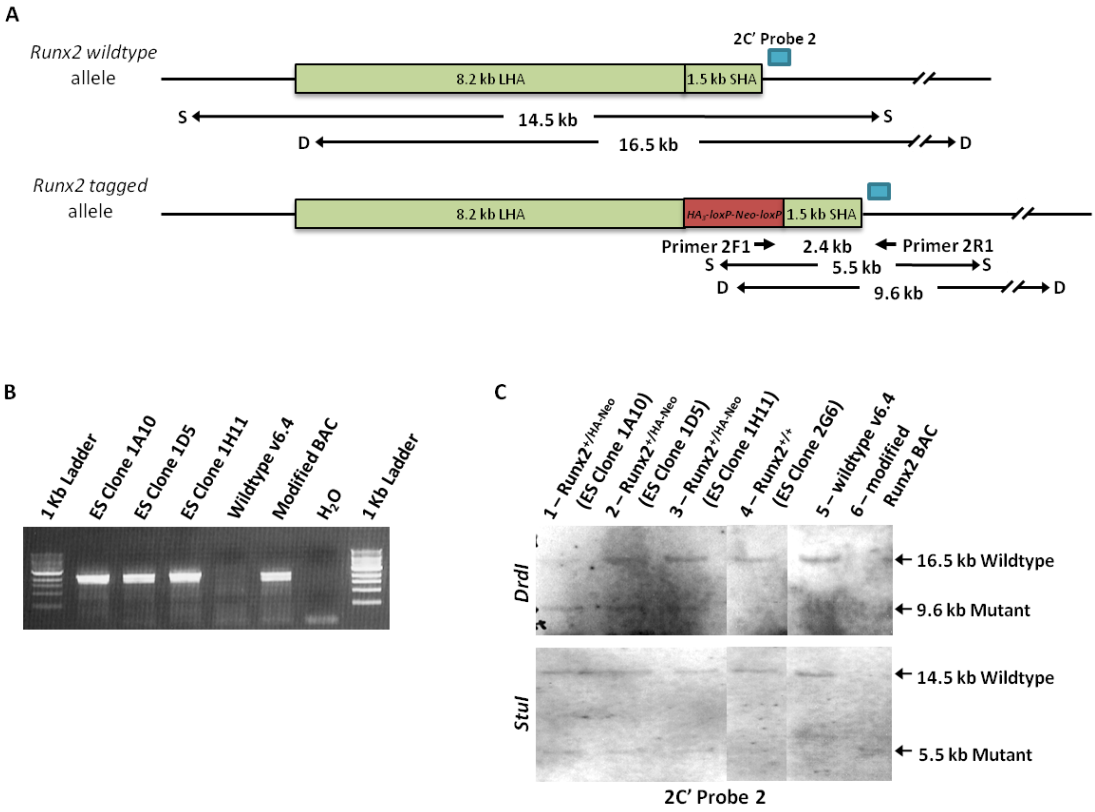


Figure 38. PCR and SB Screen for *Runx2*$^{+/HA\text{-}Neo}$ Positive Recombinant ES Cell Clones

(A) Correctly targeted ES clones would have a wild-type and a tagged allele as shown. (B) PCR screen for positive recombinant clones was carried out with a forward primer complementary to the 3' end of *Neo* and a reverse primer outside the short homology arm yielding a 2.4kb PCR product (primers indicated with black arrows in A). (C) Genomic DNA extracted from putative positive ES clones identified by PCR were digested with *StuI* (top panel) or *DrdI* (bottom panel) restriction enzyme and probed with a C-terminal Probe 2 (2C'Probe2; 590 bp; 25 ng/ml) just outside the short homology arm. True homologous recombinant ES cell clones will give a 16.5 kb wild-type and a 9.6 kb mutant band (top panel; lanes 1, 2 & 3) when cut with *StuI* or a 14.5 kb wild-type and a 5.5 kb mutant band (bottom panel; lanes 1, 2 & 3) when cut with *DrdI*. A false positive clone will only show the wild-type band (lane 4). LHA, long homology arm; SHA, short homology arm; S, *StuI*, D, *DrdI*.

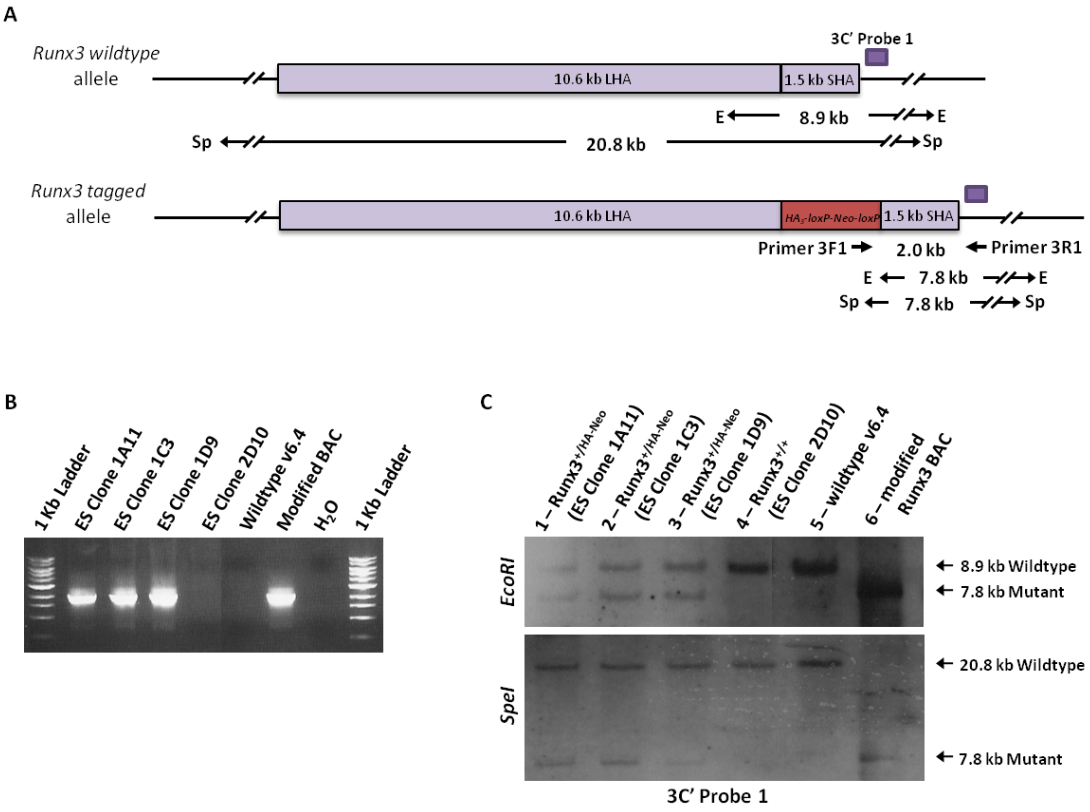


Figure 39. PCR and SB Screen for *Runx3*$^{+/HA\text{-}Neo}$ Positive Recombinant ES Cell Clones

(A) Correctly targeted ES clones would have a wild-type and a tagged allele as shown. (B) PCR screen for positive recombinant clones was carried out with a forward primer complementary to the 3' end of *Neo* and a reverse primer outside the short homology arm yielding a 2.4kb PCR product (primers indicated with black arrows in A). (C) Genomic DNA extracted from putative positive ES clones identified by PCR were digested with *StuI* (top panel) or *DrdI* (bottom panel) restriction enzyme and probed with a C-terminal Probe 2 (2C'Probe2; 590 bp; 25 ng/ml) just outside the short homology arm. True homologous recombinant ES cell clones will give a 16.5 kb wild-type and a 9.6 kb mutant band (top panel; lanes 1, 2 & 3) when cut with *StuI* or a 14.5 kb wild-type and a 5.5 kb mutant band (bottom panel; lanes 1, 2 & 3) when cut with *DrdI*. A false positive clone will only show the wild-type band (lane 4). LHA, long homology arm; SHA, short homology arm; S, *StuI*, D, *DrdI*.

Table 6. Gene-Targeting Frequencies of Runx2-HA$_3$ and Runx3-HA$_3$ ES Cells

Genotype	Targeting Frequency	Positive Clones	ESC Clone that gave GLT chimeras
Runx2$^{+/HA_3\text{-}Neo}$	7/170 (4.12%)	1A10 1D5 1H11 2E12 2F11 2G6 2G8	1A10
Runx3$^{+/HA_3\text{-}Neo}$	4/148 (2.70%)	1A11 1C3 1D9 2D10	1A11

Gene-targeting frequencies at the 3' region of *Runx2* and *Runx3* locus. Only one mouse line was used to pursue downstream ChIP studies. GLT, germ-line transmitting.

Positive ESC clones identified from the Southern blot screen were expanded for microinjections into 2- to 8-cell stage C57BL/6J mouse embryos (Kraus et al., 2010) to generate high percentage germ-line transmitting chimeras. The F1 offspring of the male chimeras when mated to C57BL/6J wild-type female mice were heterozygous for the tagged allele. These *Runx2$^{+/HA\text{-}Neo}$* and *Runx3$^{+/HA\text{-}Neo}$* mice were normal but when they were mated to homozygosity, *Runx2$^{HA\text{-}Neo/HA\text{-}Neo}$* mice were smaller and had severe malocclusion while *Runx3$^{HA\text{-}Neo/HA\text{-}Neo}$* pups were absent. The malocclusion in *Runx2$^{HA\text{-}Neo/HA\text{-}Neo}$* mice probably prevented proper feeding and in turn resulted in the poor growth (Fig. 40). The *Runx2$^{HA\text{-}Neo/HA\text{-}Neo}$* mice were crossed to *Runx2$^{+/HA\text{-}Neo}$* mice but no pups dropped even after several months. The *Runx2$^{HA\text{-}Neo/HA\text{-}Neo}$* mice were either infertile or were too weak to mate.

The *Neo* gene was subsequently excised by mating the *Runx2$^{+/HA\text{-}Neo}$* and the *Runx3$^{+/HA\text{-}Neo}$* mice to homozygous *Zp3^{Cre}* mice. *Runx2$^{+/HA}$* and *Runx3$^{+/HA}$* mice were

inter-mated to homozygosity. Consequently, there were no *Runx2*$^{HA/HA}$ nor *Runx3*$^{HA/HA}$ mice produced among the littermates. There were two occasions where a very small three-week old mouse with a foreshortened snout was genotyped to be *Runx2*$^{HA/HA}$ but the mouse did not survive past a month. Unfortunately, the *Neo* gene was not the cause of the malocclusion in the *Runx2*$^{HA\text{-}Neo/HA\text{-}Neo}$ mice and the lethality in the *Runx3*$^{HA\text{-}Neo/HA\text{-}Neo}$ embryos. On the contrary, removing the Neo appeared to aggravate the condition of the tagged mice for unexplained reasons.

Next, I tried to harvest E13.5 *Runx2*$^{HA/HA}$ and *Runx3*$^{HA/HA}$ mouse embryos from intra-heterozygous crosses of *Runx2*$^{+/HA}$ and *Runx3*$^{+/HA}$ mice and apparently normal E13.5 *Runx2*$^{HA/HA}$ mouse embryos were obtained but *Runx3*$^{HA/HA}$ embryos were absent even as early as E10.5. The phenotype observed in these homozygous tagged embryos did not match that of the *Runx2* and *Runx3* homozygous mouse knockouts although the phenomenon seen in the *Runx2*$^{HA\text{-}Neo/HA\text{-}Neo}$ mice could be associated with impaired skeletal development of the jaw and the occasional *Runx2*$^{HA/HA}$ mice resembled one with severe abnormalities in the facio-cranial and axial skeleton.

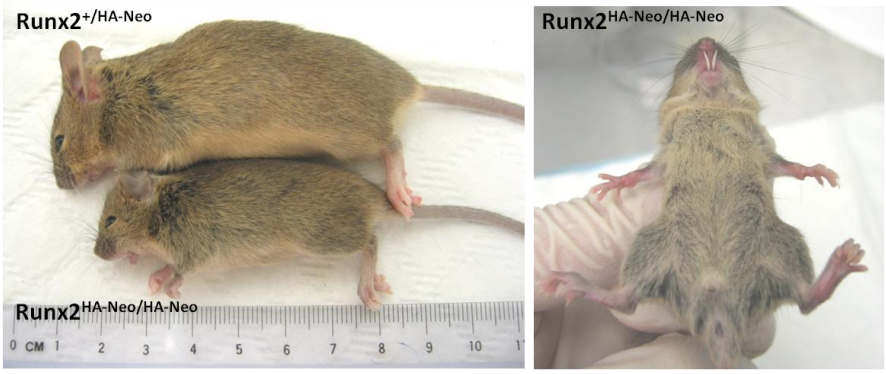


Figure 40. Phenotype of Runx2$^{+/HA\text{-}Neo}$ and Runx2$^{HA\text{-}Neo/HA\text{-}Neo}$ Mice

Runx2$^{+/HA\text{-}Neo}$ mice were grossly normal and fertile. *Runx2*$^{HA\text{-}Neo/HA\text{-}Neo}$ mice were smaller and had severe malocclusion resulting in the inability to feed well and consequently poor growth.

The genotype of mice or embryos was determined by PCR using tail tips from 3-week old mice or yolk sacs of embryos harvested for experiments. The following primers were used to determine the genotype:

Table 7. Primers for Genotyping *Runx2-HA$_3$* and *Runx3-HA$_3$* mice or embryos

	Runx2-HA$_3$	***Runx3-HA$_3$***
Forward Primer	5' CCCAGCCACCTTTACCTACA 3'	5' GGACCCCAGGATGCACTAC 3'
Reverse Primer	5' CTGCCTCTTGTCCCTTTCTG 3'	5' GGGAGGGAGAGAAAGTCCAG 3'
Wild-type PCR band	809 bp	782 bp
Mutant PCR band	944 bp	917 bp

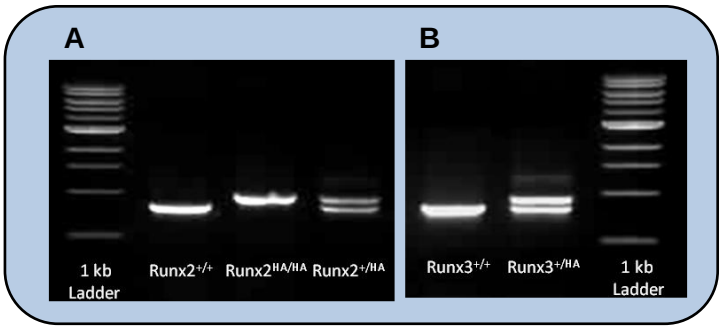


Figure 41. Genotyping *Runx2-HA$_3$* and *Runx3-HA$_3$* Mice by PCR

(A) *Runx2-HA$_3$*; wild-type allele: 809 bp, mutant allele: 944 bp (B) *Runx3-HA$_3$*; wild-type allele: 782 bp, mutant allele: 917 bp.

In order to find out if the mice were correctly tagged and the HA epitopes were exposed for detection by anti-HA antibodies, immunohistochemistry (IHC) was performed with IHC-grade anti-HA antibody (AP09230PU-N; Acris Antibodies) on sagitally sectioned E13.5 *Runx2$^{+/HA}$* and *Runx3$^{+/HA}$* mouse embryos embedded in paraffin.

Figure 42 details the expression of Runx2-HA_3 protein detected by both the anti-HA and the anti-Runx2 antibodies in the Meckel's cartilage of the lower jaw, clavicle and the perichondrium of the ribs, humerus and hindlimb digits. Runx2-HA_3 protein expression was specifically detected by the anti-HA antibody (Fig. 42 top right panel) which was not found in the wild-type embryo (Fig. 42 top left panel) and the expression recapitulates the endogenous expression of Runx2. (Fig. 42 bottom left panel). Anti-Runx2 antibody was also able to detect the Runx2 and the Runx2-HA_3 proteins in the heterozygous tagged embryo in the corresponding expression domains (Fig. 42 bottom right panel).

In Figure 43, it is evident that the HA epitope of Runx3-HA_3 protein could not be detected by the anti-HA antibody. This could be because (1) the HA epitope at the C-terminus was folded within the Runx3-HA_3 protein and could not be recognized, (2) the HA_3 tag was cleaved off by a protease due to the presence of an unidentified protease site at the C-terminus or (3) the C-terminally tagged Runx3-HA_3 was prone to degradation for some unknown reason (Munro and Pelham, 1987). Although signals were present in the *Runx3-HA_3* embryo sections when probed with the anti-Runx3 antibody, it could not be verified if the antibody detected both the Runx3 and the Runx3-HA_3 proteins or just the Runx3 protein in the heterozygous Runx3-HA_3 embryo. Unfortunately, E13.5 homozygous *Runx3-HA_3* mouse embryos could not be obtained for IHC with anti-Runx3 antibody to investigate if the Runx3-HA_3 proteins were present or absent.

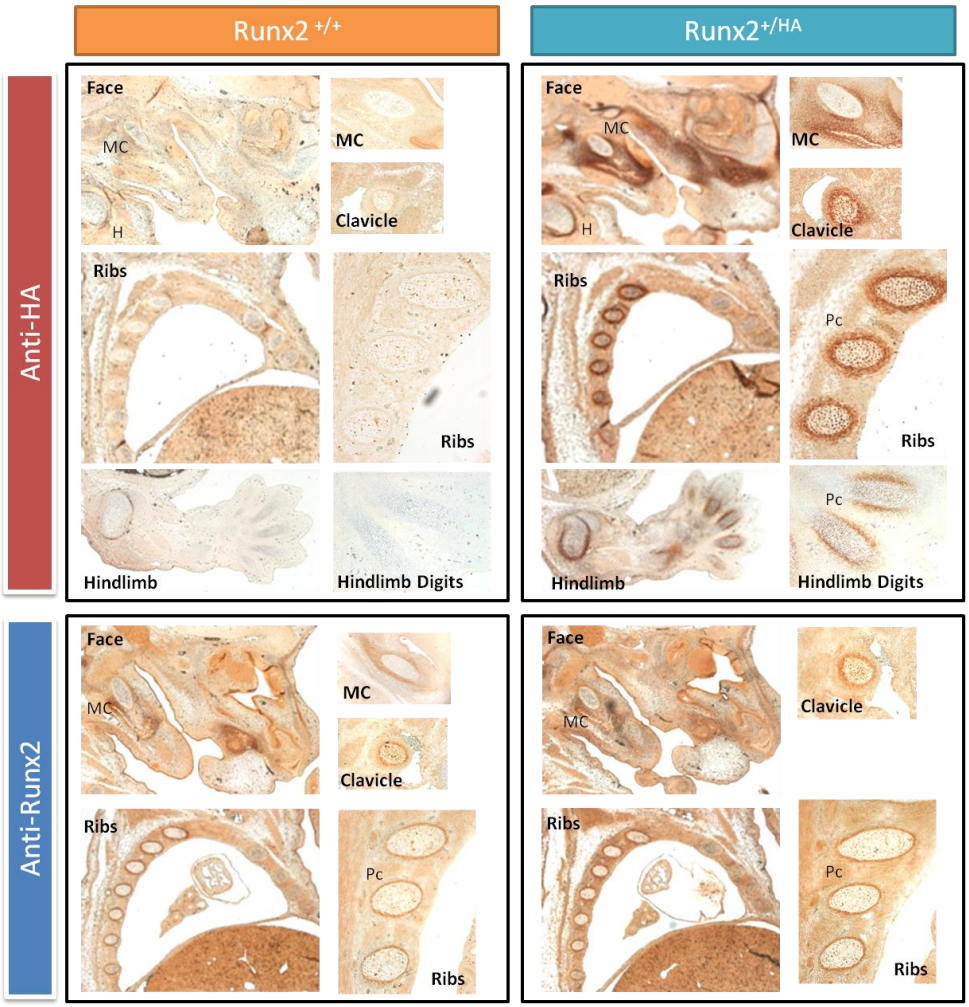


Figure 42. Runx2-HA$_3$ Protein Expression Recapitulates Endogenous *Runx2* Expression

Immunohistochemistry was performed using IHC-grade anti-HA and anti-Runx2 antibodies on sagitally sectioned E13.5 wild-type and *Runx2*$^{+/HA}$ mouse embryos embedded in paraffin. Images on the left and right of each quarter panel were taken at x50 and x200 magnifications respectively. MC, Meckel's cartilage; H, humerus; Pc, perichondrium.

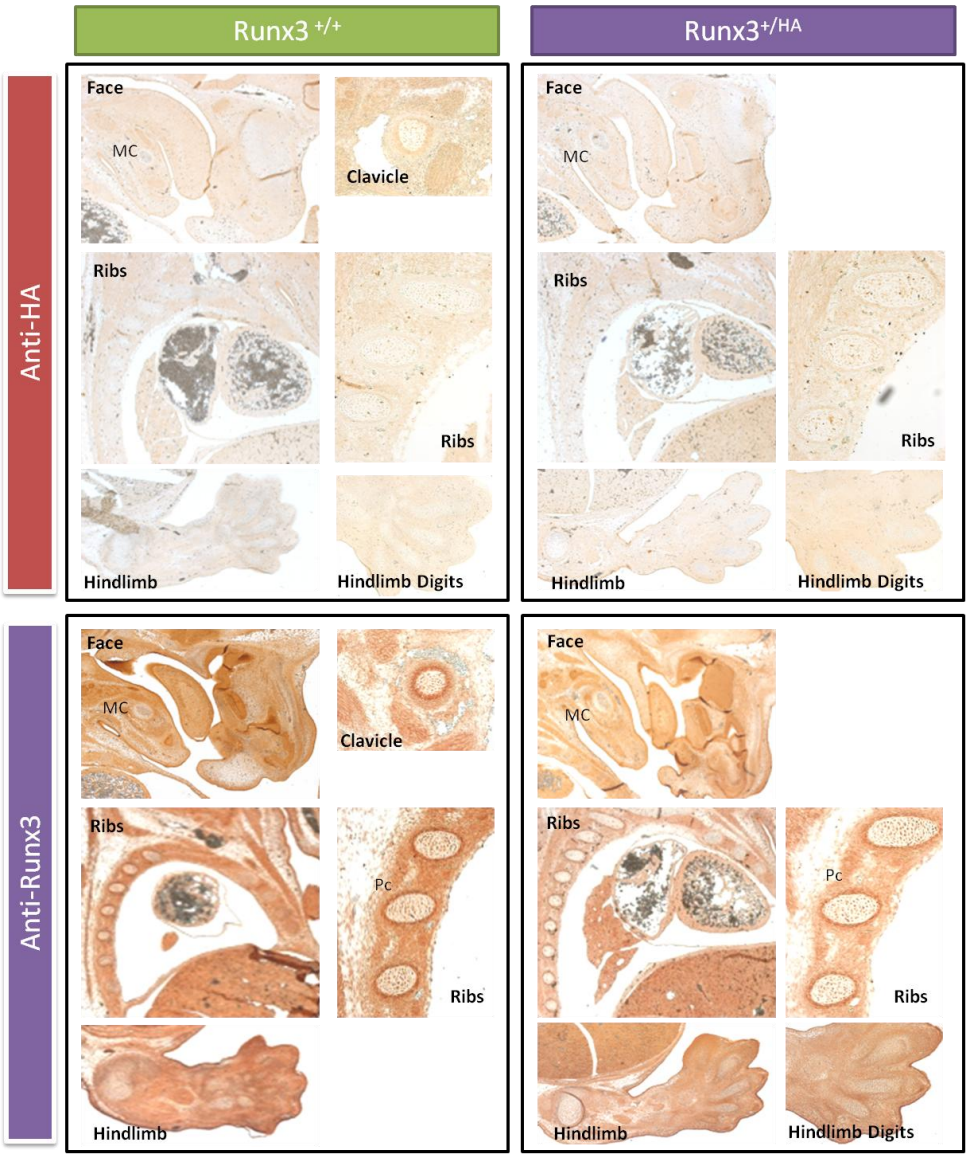


Figure 43. Runx3-HA$_3$ Protein Expression

Immunohistochemistry was performed using IHC-grade anti-HA and anti-Runx2 antibodies on sagitally sectioned E13.5 wild-type and *Runx2*$^{+/HA}$ mouse embryos embedded in paraffin. Images on the left and right of each quarter panel were taken at x50 and x200 magnifications respectively. MC, Meckel's cartilage; H, humerus; Pc, perichondrium.

In summary, Runx2 was successfully tagged with the HA epitope and the expression of the Runx2-HA_3 proteins reflected endogenous *Runx2* expressions. The HA epitope could also be recognized by the anti-HA antibody. The only problem was that the *Runx2*$^{HA/HA}$ mice frequently did not survive and a colony of *Runx2*$^{HA/HA}$ mice could not be maintained. Hence, E13.5 *Runx2*$^{HA/HA}$ embryos could only be obtained via intra-heterozygous crosses for ChIP experiments. The rationale for proceeding to use the *Runx2*$^{HA/HA}$ mouse embryos for ChIP despite the phenotype observed was that the E13.5 *Runx2*$^{HA/HA}$ embryos appeared normal and the tagged protein could be recognized by anti-HA antibodies. It was also postulated that the Runx2-HA_3 was likely to retain its DNA-binding ability. This could be further determined by future EMSA experiments with purified Runx2-HA_3 protein.

On the other hand, the tagging of Runx3 in the mouse was not successful. The HA epitope of the Runx3-HA_3 protein in the *Runx3*$^{+/HA}$ mouse embryos could not be recognized for one of the several possible reasons. Firstly, the tag could have been folded within the protein and hidden away from the antibody. Secondly, the tagged protein could have undergone post-translational modifications resulting in the HA epitope being cleaved off and thirdly, the C-terminal tag could expose the 3' end domains to phosphorylation and subsequent protein degradation. Furthermore, it could not be determined if Runx3-HA_3 protein was either present but undetectable or degraded as *Runx3*$^{HA/HA}$ mouse embryos could not be obtained for further experimentation. The lethality of the *Runx3*$^{HA/HA}$ mouse embryos exceptionally early in developmental stage could not be explained. Although, from the IHC experiments, it was anticipated that the Runx3-HA_3 proteins could not be immunoprecipitated, HA-tagged Runx3 mouse embryos were still collected for ChIP-Western blot experiments described in the following segment to assess if Runx3-HA_3 proteins could be pulled down by another anti-HA antibody. Since *Runx3*$^{HA/HA}$ mouse embryos could not be

acquired, a pool of wild-type and *Runx3*$^{+/HA}$ embryos were collected from inter-*Runx3*$^{+/HA}$ crosses for ChIP experiments.

3.2.3 ChIP-Western blot (ChIP-WB) to Assess Ability to Immunoprecipitate Runx2-HA$_3$ and Runx3-HA$_3$ Proteins from Tagged mice.

ChIP with an anti-HA antibody (Abcam 9110) followed by Western blot (WB) detection of the immunoprecipitated proteins with another WB-grade HRP-conjugated anti-HA antibody (Bethyl) was first performed using the tagged mouse tissue to assess if the anti-HA antibody (Abcam 9110) was sensitive and specific enough for ChIP experiments. It was also done to test if the tagged proteins from the mouse tissue could be immunoprecipitated and enriched. A HRP-conjugated primary anti-HA antibody (Bethyl) was chosen for WB detection to avoid overwhelming background signals from the binding of secondary antibodies to the heavy and light chains of the primary antibody used for immunoprecipitation.

E13.5 mouse embryos were harvested from intra-heterozygous mouse crosses of *Runx2*$^{+/HA}$ or *Runx3*$^{+/HA}$ mice. Different tissues such as the jaw, the ribs and the limbs were dissected from the embryos on ice to enrich for *Runx2*- or *Runx3*-expressing cells. Similar tissues were also harvested from E13.5 CD-1 wild-type mouse embryos as controls for the ChIP experiments with anti-HA antibody. The same tissues from all the embryos with the same tagged protein were pooled together. The different jaw, rib and limb tissues were kept as separate samples during homogenization, cross-linking, lysis, chromatin shearing, immunoprecipitation and WB or sequencing. The yolk sacs of each embryo were kept for genotyping. Among the 74 E13.5 embryos harvested from the *Runx2*$^{+/HA}$ mouse crosses, there were 15 wild-type, 43 *Runx2*$^{+/HA}$ and 16 *Runx2*$^{HA/HA}$ mouse embryos which corresponded to the Mendelian inheritance ratio. This would imply that 50.7% of the Runx2 proteins were tagged in

each pool of Runx2-HA_3 tissues collected. Correspondingly, out of the 41 E13.5 embryos collected from the *Runx3*$^{+/HA}$ mouse crosses, there were 10 wild-type, 31 *Runx3*$^{+/HA}$ and no *Runx3*$^{HA/HA}$ mouse embryos. This would translate to 37.8% of the Runx3 proteins being tagged.

Figure 44 made evident that the Runx2-HA_3 protein from the tagged mouse tissues was successfully immunoprecipitated. This also demonstrated that the anti-HA antibody was sufficiently specific and sensitive for ChIP and a good enrichment of the DNA-bound TF of interest was attained. As expected, the Runx3-HA_3 protein could not be pulled down. One can interpret that either the HA tag was hidden, the Runx3-HA_3 protein was absent or that the proportion of tagged proteins was too low to be successfully pulled down by the anti-HA antibody since only 37% of the Runx3 proteins used for ChIP was tagged. Notably, the ChIP-WB experiment results corroborated with the IHC results even though a different anti-HA antibody was used. As a result, only the *Runx2-HA*$_3$ tissues were used in the subsequent ChIP-Seq experiments.

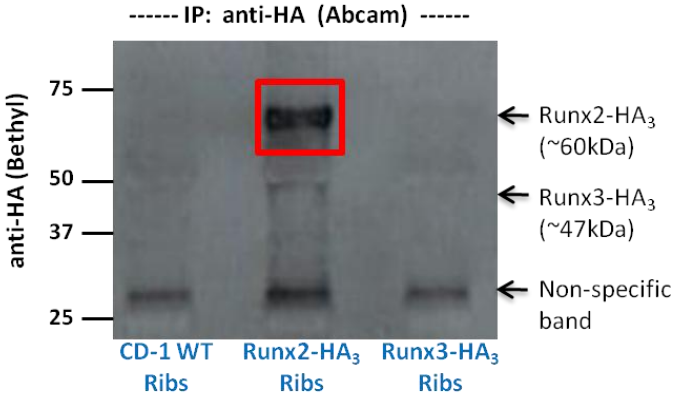


Figure 44. ChIP-WB of Runx2-HA_3 and Runx3-HA_3 proteins

Crossed-linked single cells from CD-1 wild-type, *Runx2-HA*$_3$ and *Runx3-HA*$_3$ mouse rib tissue were lysed, sonicated and DNA-bound Runx2-HA_3 and Runx3-HA_3 proteins were immunoprecipitated with anti-HA antibody (Abcam 9110; 10ug) overnight and loaded onto a Western gel for detection with HRP-conjugated anti-HA antibody (Bethyl; 1:20,000). The red box highlighted the Runx2-HA_3 protein band that was successfully immunoprecipitated. CD-1 wild-type mouse rib tissue serves as a negative control. IP, immunoprecipitation.

3.2.4 *In Vivo* ChIP-Seq of Runx2-specific Binding Sites

After verifying by ChIP-WB that Runx2-HA_3 proteins could be pulled down with the anti-HA antibody, jaw, rib and limb tissues from *Runx2-HA_3* and CD-1 wild-type mouse embryos went on to be processed separately for ChIP-Seq. DNA libraries of Runx2-HA_3-bound DNA fragments immunoprecipitated by anti-HA antibodies were synthesized using the NEBNext® ChIP-Seq Sample Prep kit. CD-1 wild-type tissue samples immunoprecipitated with anti-HA antibody were used as a background control for the corresponding *Runx2-HA_3* tissue samples. The quantity and quality of the size-selected DNA ChIP libraries were analyzed using the Agilent DNA chip and the Agilent Bioanalyzer before sequencing. Figure 45 illustrates that the DNA libraries were indeed between a good range of 200-400 bp and the amount of DNA was also sufficient for sequencing with the Illumina Solexa sequencing technology.

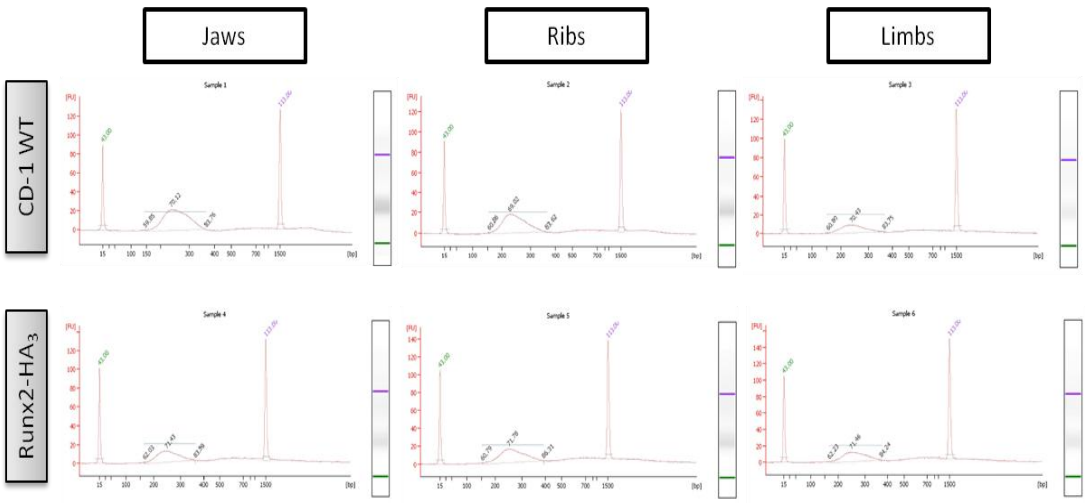


Figure 45. Runx2-HA_3 *in Vivo* ChIP Libraries Assessed with Agilent DNA BioAnalyzer

DNA libraries were prepared for each of the immunoprecipitated Runx2-HA_3-bound DNA fragments as well as the corresponding CD-1 WT tissue samples. Fragment size, quantity and quality of the synthesized DNA libraries were analyzed using Agilent DNA chip and Agilent Bioanalyzer.

Short sequence reads were mapped back to the mouse genome (NCBI build 37/mm9) to determine the genome-wide Runx2-DNA interaction sites. Model-based Analysis of ChIP-Seq (MACS) (Zhang et al., 2008b) was applied to the Runx2-HA_3 ChIP-Seq datasets to predict Runx2 binding peaks with high resolution. The peaks were analyzed using the Genomic Regions Enrichment of Annotations Tool (GREAT) (McLean et al., 2010). GREAT first assigns a regulatory domain to each gene. The genomic regions which coincide with the Runx2 binding peaks are next associated with all the genes whose regulatory regions overlap with the Runx2 binding peaks. The association rule used to identify genes which the Runx2-bound genomic regions possibly regulate was “Approach 2: Two Nearest Genes”.

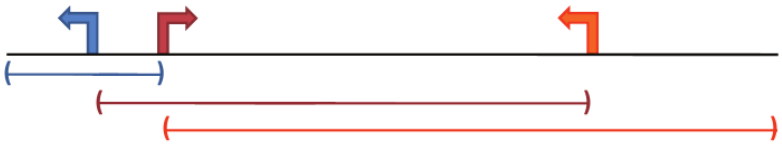


Figure 46. GREAT Association Rule Setting Approach 2: Two Nearest Genes

Figure from GREAT website. Arrows represent transcription start site. Each gene is assigned a regulatory domain that extends in both directions to the nearest gene's TSS but no more than a maximum extension in one direction.

Using the association rule mentioned above (Fig. 46), the two nearest genes within 1kb, 5kb, 10kb, 100kb and 1000kb from either side of the Runx2 binding peaks were identified by GREAT for each of the different tissues. The number of genes identified was plotted against the distance of the Runx2 binding site from the TSS (Fig. 47). In other words, Figure 47 shows the distribution of Runx2 binding sites from the TSS of any gene. The most number of genes found associated with the Runx2 binding sites were between 10-1000kb away from the TSS of any gene. This implies that Runx2 at this stage of development likely regulates most of its target genes directly at a distal enhancer rather than at the promoter of target genes. There have been evidences where target genes are controlled at enhancer sites a megabase away and a looping

mechanism brings these distal enhancers in proximity to the promoter of the target gene (Hill and van Heyningen, 2008; Ptashne, 1986; Yoshida et al., 1999). One example is the long range enhancer of the Shh gene (Lettice et al., 2003) regulating expression in the limb. Hence, it could be postulated that there might be the same looping mechanism bringing Runx2 bound at the distal enhancer to the transcription machinery at the promoter.

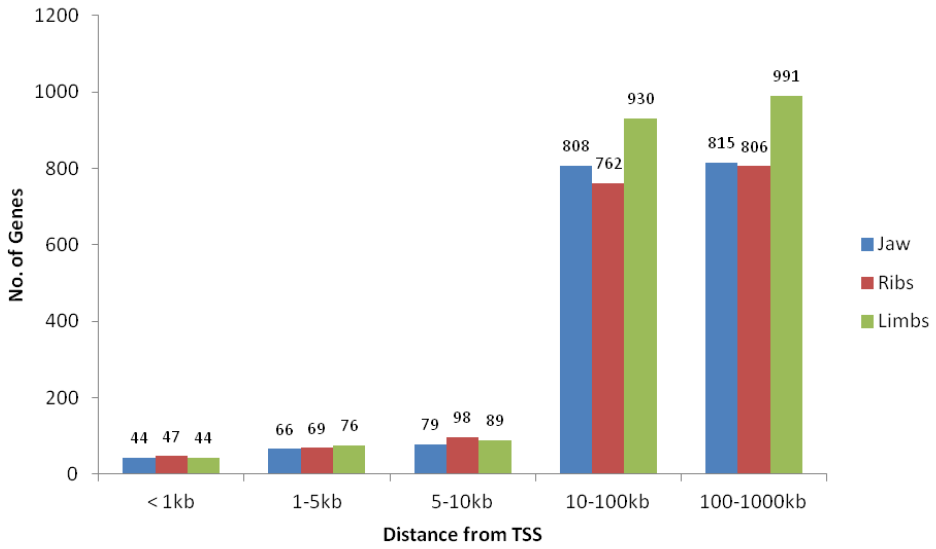


Figure 47. Distribution of Runx2 Binding Sites from Jaw, Rib and Limb Tissues

The number of genes identified within 1 kb, 1 - 5 kb, 5 -10 kb, 10 -100 kb and 100 -1000 kb from the significant Runx2 binding peaks was plotted against the distance between the Runx2 binding sites and the transcription start site (TSS) of the genes.

The genes were next analyzed by DAVID to examine what GO terms associated with the genes were significantly enriched. Figure 48A shows significant enrichment scores of 4.92 and 2.05 for GO terms such as limb morphogenesis and cartilage development respectively for genes found within 1 kb from the Runx2 binding site. In fact, the highest ranking GO term that came up in the DAVID analysis was limb morphogenesis. The result confers confidence that the ChIP-seq data was reliable and relevant. The genes clustered under “limb morphogenesis” were *Lnp*, *Evx2*, *Shh*,

Hoxa11, *Hoxd10*, *Hoxd11*, *Hoxd12* and *Hoxd13* and those clustered under “cartilage development” were *Hoxa11*, *Runx2* and *Hoxd11*. Interestingly, when the binding sites of these genes were examined more closely, most of them identified as within 1 kb from Runx2 binding sites were in fact more than 1 kb away. To supplement the association rules mentioned above, GREAT utilizes literature-curated regulatory domains where it had been shown that a gene is directly regulated by a regulatory element that falls outside the basal regulatory domain i.e. the promoter region. Hence, GREAT overrides the restricted rules to include curated regulatory domains. Genes currently with curated regulatory domains are *Shh*, *Lnp*, *Evx2*, *Hoxd10*, *Hox11*, *Hoxd12* and *Hoxd13* which explains why these genes were still associated with the Runx2 binding sites even when the 1 kb rule was applied during the analysis.

GO terms associated with genes found within 5 kb from the Runx2 binding site were still relevant such as limb development and cartilage development albeit the enrichment scores were lower (Fig. 48B) and *Ptch1* was the only additional skeletal gene identified but was not found in any of the Runx2 microarray lists. Table 8 shows the relevant skeletal genes identified as within 1 kb or 5 kb from Runx2 binding sites, the location of the binding site from the TSS, the tissue which the genes are identified in and if they were also found in the Runx2 microarray data.

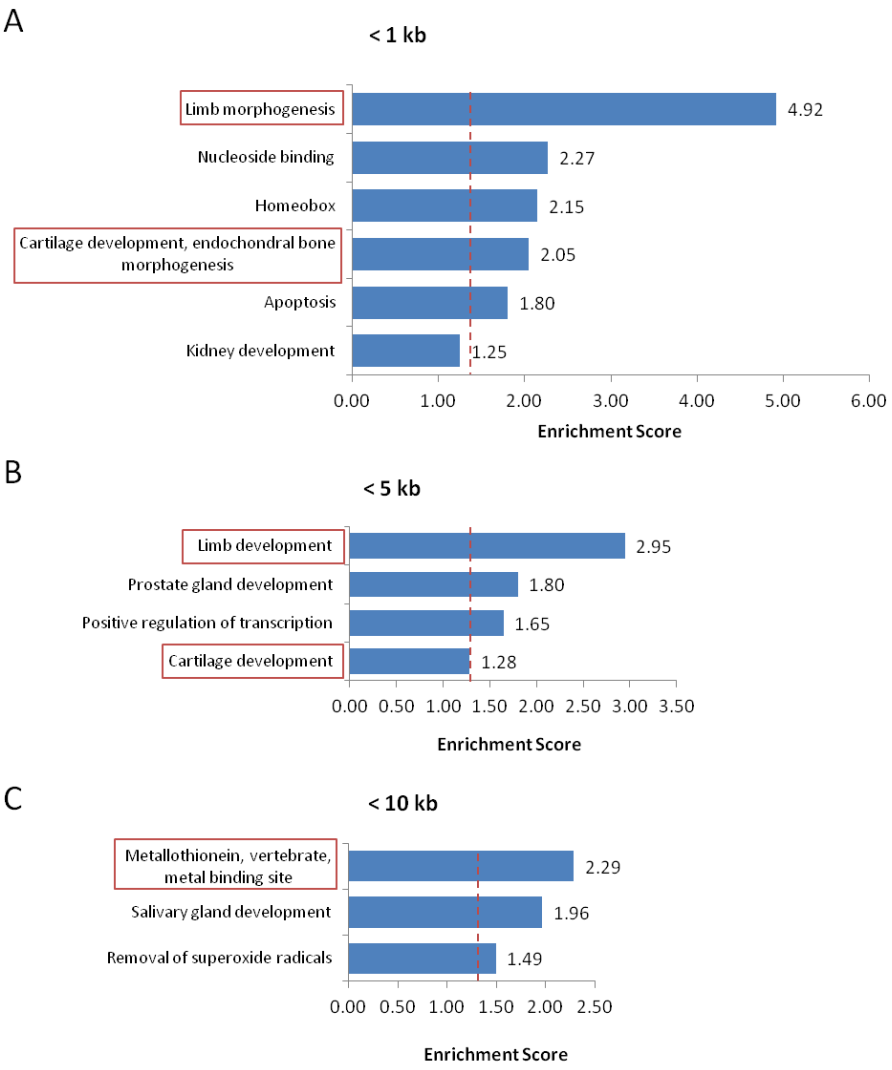


Figure 48. GO Terms Associated with Genes Nearest to Runx2 Binding Sites

Enriched GO terms associated with nearest genes found within (A) 1 kb, (B) 5 kb and (C) 10 kb from the significant Runx2 binding peaks in all the tissues combined. GO terms with enrichment scores above 1.3 are significantly enriched. The red dotted lines indicate the recommended cut-off point of 1.3. The red boxes highlight skeletal functions.

Table 8. Genes Associated within 1 kb and 5 kb of Runx2 Binding Sites and Clustered under Relevant Skeletal GO Terms.

Gene Symbol	Gene Name	Binding Sites (Distance from TSS)	Tissue	Microarray (Regulation by Runx2)
Genes identified as within 1 kb from binding sites:				
Lnp	*Limb and neural patterns*	(+377, 321) (+489,709) (+2,076)	Jaw Ribs Limbs	Yes (Negative)
Evx2	*Even skipped homeotic gene 2 homolog*	(+82,547)	Limbs	No
Shh	*Sonic hedgehog*	(-417,351) (-366,989), (+142,089)	Jaw Limbs	Yes (Positive and negative)
Hoxa11	*Homeobox A11*	(-744)	Ribs	Yes (Positive)
Hoxd10	*Homeobox D10*	(-115,076)	Limbs	No
Hoxd11	*Homeobox D11*	(-105,521)	Limbs	No
Hoxd12	*Homeobox D12*	(-98,158)	Limbs	No
Hoxd13	*Homeobox D13*	(-91,438)	Limbs	Yes (Positive)
Runx2	*Runt related transcription factor 2*	(+616)	Ribs	Yes (Negative)
Gene identified as within 5 kb from binding sites:				
Ptch1	*Patched homolog 1*	(-2674), (+2298)	Ribs	No

When genes found within 10 kb of Runx2 binding sites were assessed, no relevant skeletal functions were found. Instead, the GO term “Metallothionein, vertebrate, metal binding site” was given a significant enrichment score of 2.29.

Metallothioneins are cysteine-rich metal-binding proteins. Mammals express at least four types of metallothioneins – metallothionein-1 to metallothionein-4 (Mt1 to Mt4). They bind and regulate the cellular metabolism of physiological metals such as zinc and copper. In addition, they have detoxifying functions by binding to toxic metals such as cadmium, silver, copper and mercury (Nordberg, 1998). Furthermore, Mt3 is a brain-specific metallothionein and has been detected in zinc-enriched neurons and

astrocytes (Masters et al., 1994; Palmiter et al., 1992). It was suggested to be involved in brain repair (Hozumi et al., 1998). However, the full functions and functional mechanisms of metallothioneins are not well-understood. The *metallothionein* genes found within 10 kb of Runx2 binding sites were *Mt1*, *Mt3* and *Mt4* while only *Mt1*, *Mt2* and *Mt4* genes were found in the Runx2 microarray gene lists. The direction of regulation by Runx2 of these was conflicting; hence, it is not possible to determine if Runx2 positively or negatively regulated these *metallothionein* genes. Perhaps Runx2 has a role in the detoxification function of bone by regulating the metallothionein genes or that the metallothionein genes have yet to be discovered functions in bone development.

Next, the genes whose regulatory domains, set at 100 kb and beyond, overlap with Runx2 binding sites were assessed with DAVID and many GO terms related with skeletal functions were found significant such as “Extracellular matrix”, “Embryonic limb development”, “Wnt signalling”, “Osteoclast differentiation” and “Bone development, ossification and osteoblast differentiation” (Fig. 49). As mentioned earlier, it is likely that Runx2 directly regulates most of its skeletal target genes at a distal enhancer cooperating with other transcription factors bound at the promoter of the target genes at this stage of development.

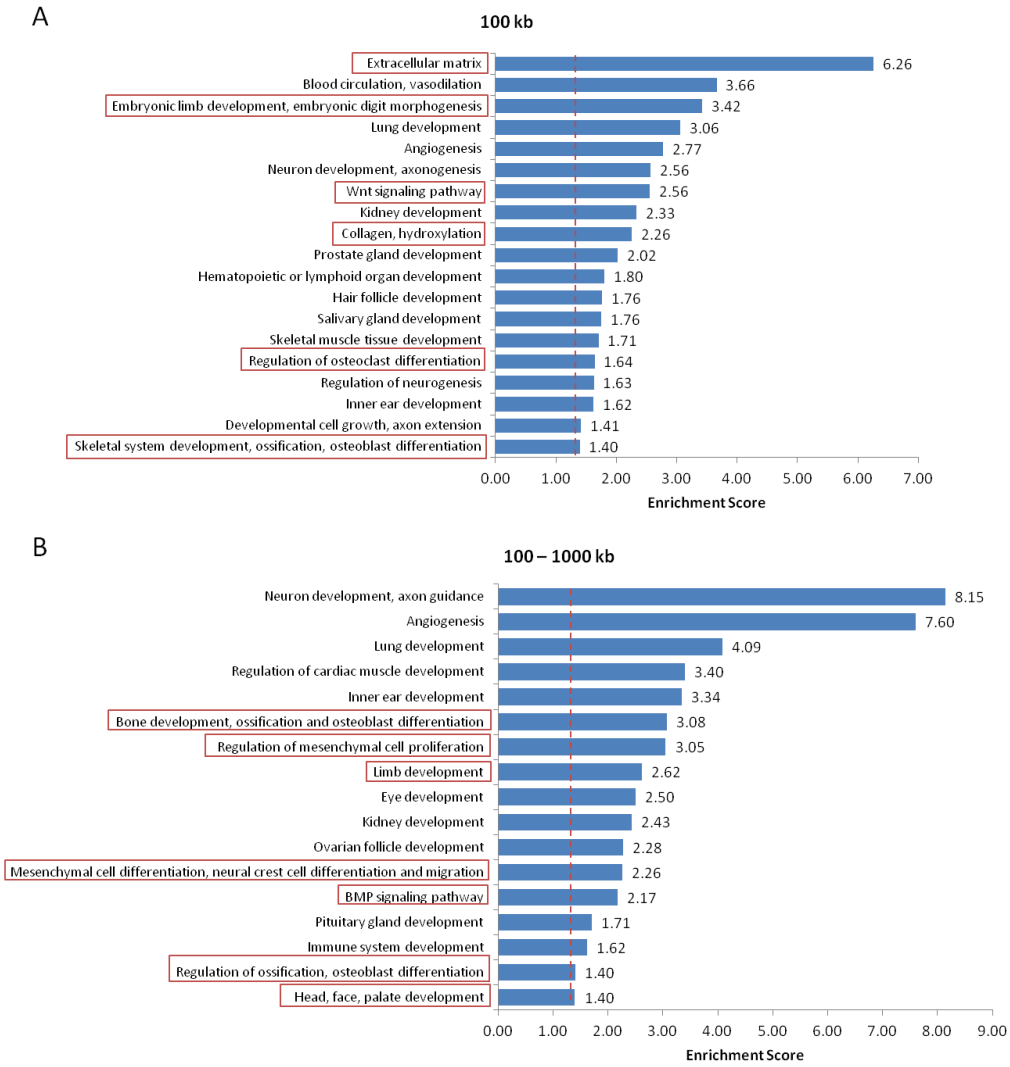


Figure 49. GO Terms Associated with Genes Identified with Distal Runx2 Binding Sites

Enriched GO terms associated with nearest genes found within (A) 100 kb and (B) 100-1000 kb from the significant Runx2 binding peaks in all the tissues combined. GO terms with enrichment scores above 1.3 are significantly enriched. The red dotted lines indicate the recommended cut-off point of 1.3. The red boxes highlight skeletal functions and pathways.

Lastly, I tried to find out how many of the potential Runx2 targets identified from the microarray were direct targets. Figure 50 shows two Venn diagrams illustrating the number of genes found in both the Runx2 microarray gene list and the Runx2 ChIP-Seq list of genes with Runx2 binding sites within 1 kb or 1 Mb from the TSS. A list of Runx2 targets with Runx2 binding at the promoter region is found in Table 9. In addition, Figure 51 illustrates the peaks found at the promoter of some of these

genes. Approximately two-thirds of the genes with their promoters or distal enhancers bound by Runx2 transcription factor were not differentially regulated. The regulation of these genes could require other co-factors to bind to Runx2 or other transcription factors to sit along the DNA near the Runx2 binding site in order to synergistically regulate the transcription. Hence, the Runx2 transcription factor may be situated on these binding sites poised to be activated by other factors at a different developmental stage.

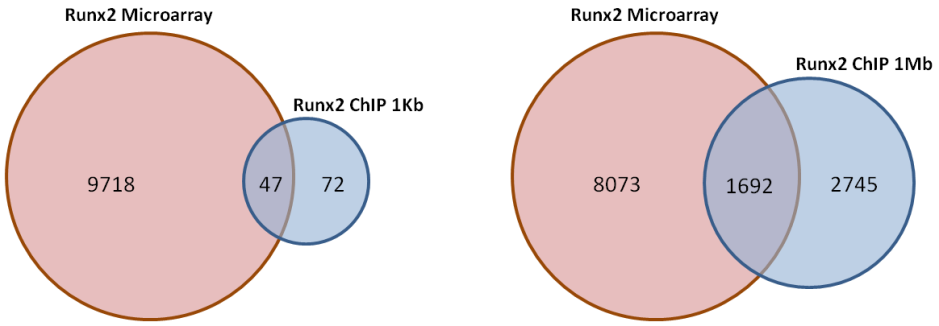


Figure 50. Putative Runx2 direct targets

Overlap of potential Runx2 targets from microarray data with genes identified within 1 kb from Runx2 binding site (left Venn diagram) and with genes identified within 1 Mb of Runx2 binding site (right Venn diagram).

Table 9. List of Runx2 direct targets with Runx2 binding at the promoter region

No.	Runx2 Direct Targets	Microarray $Runx2^{+/+}$ vs. $Runx2^{-/-}$		Runx2 Binding Site(s)	Tissue
1	*Ptp4a3*	8.6-fold	↓	(+146)	Jaw
				(+149223)	Ribs
2	*V1rc6*	8.6-fold	↓	(-600)	Jaw
3	*4930572J05Rik*	7.4-fold	↓	(+319)	Jaw
4	*Lrrc61*	6.7-fold	↓	(-673)	Ribs
5	*Zfp316*	6.2-fold	↓	(+293)	Jaw
6	*Olfr800*	4.9-fold	↓	(-674)	Ribs
7	*Hoxd13*	4.4-fold	↓	(-91438)	Limbs
8	*Ptgis*	4.0-fold	↓	(-192)	Jaw
9	*Tsc22d2*	3.9-fold	↓	(-963)	Limbs
10	*Eif4a1*	3.2-fold	↓	(+900)	Ribs
11	*Hoxa11*	3.1-fold	↓	(-744)	Ribs
12	*Eif3g*	2.9-fold	↓	(-95)	Limbs
13	*Scara3*	2.6-fold	↓	(-797)	Jaw

14	*Abcb8*	2.6-fold	↓	(+740)	Limbs
15	*Mapk7*	2.6-fold	↓	(+741)	Limbs
				(-58922)	Ribs
16	*Midn*	2.3-fold	↓	(+738)	Ribs
17	*Ppp1r13b*	2.3-fold	↓	(+467)	Jaw
18	*Nuak1*	2.2-fold	↓	(-387)	Ribs
				(+647780)	Limbs
19	*Adamts14*	1.7-fold	↓	(-534)	Ribs
20	*Scrib*	1.6-fold	↓	(-54)	Jaw
21	*Parva*	1.5-fold	↓	(+200)	Ribs
				(+243614)	Jaw
22	*T*	4.9-fold	↓	(-45196), (+294), (+29544)	Limbs
				(+764)	Jaw
				(+7303), (+153857)	Ribs
23	*Odz2*	4.2-fold	↓	(+552)	Limbs
24	*Kti12*	2.1-fold	↓	(+101)	Ribs
25	*2010001M09Rik*	1.9-fold	↓	(+402)	Limbs
26	*Smarcd1*	1.8-fold	↓	(-13)	Limbs
27	*Pip5k1a*	1.7-fold	↓	(+837)	Jaw
28	*Rwdd1*	1.6-fold	↓	(-95)	Limbs
29	*Rrm1*	1.6-fold	↓	(+561)	Ribs
30	*Exosc8*	1.6-fold	↓	(+890)	Jaw
31	*Mpzl1*	3.0-fold	↓	(+54), (+225)	Jaw
		9.7-fold	↑	(+140)	Limbs
32	*Lnp*	34.8-fold	↑	(+2076)	Limbs
				(+377321)	Jaw
				(+489709)	Ribs
33	*Ppp1r13l*	8.4-fold	↑	(-101), (+4)	Jaw
				(-14)	Limbs
34	*Runx2*	8.4-fold	↑	(+616)	Ribs
35	*Shh*	7.2-fold	↑	(-366989), (+142089)	Limbs
				(-417351)	Jaw
36	*Acox2*	7.5-fold	↑	(-977)	Limbs
37	*Amotl2*	5.3-fold	↑	(-133)	Ribs
				(-129)	Jaw
38	*Rcn2*	4.3-fold	↑	(+210)	Ribs
39	*Gabpb1*	2.9-fold	↑	(-306), (-69)	Ribs
40	*Glcci1*	2.8-fold	↑	(-19813), (-142)	Jaw
				(-67312)	Ribs
41	*H2afz*	2.2-fold	↑	(+643)	Jaw
42	*Abcg2*	2.1-fold	↑	(-355)	Ribs
				(-360)	Jaw
43	*Hdgfrp3*	2.0-fold	↑	(-7)	Jaw
44	*Parp12*	1.8-fold	↑	(+857)	Limbs

Figure 51. Runx2 ChIP-Seq Peaks

Runx2 ChIP-Seq peaks at promoters of different genes underlined in red. Peaks are highlighted with the red boxes. ChIP was performed separately with three different tissue types (jaw, ribs and limbs) harvested from the *Runx2-HA*$_3$ mouse embryos. ChIP-Seq library tracks were uploaded to the USCS browser and images were captured on the browser.

It is noted that only 0.5% of all the Runx2 differentially regulated genes have Runx2 binding sites at its promoter and 17.3% have Runx2 binding within 1 Mb of its TSS. This could mean that Runx2 indirectly regulates most of its downstream genes or that the ChIP experiment was not able to pick up all the Runx2 binding sites. On the whole, the Runx2 ChIP-Seq peaks were not very enriched and there was a lot of background noise evident by the small random peaks seen in Figure 51 which probably masked out many true peaks. Mouse tissues, unlike cells from tissue cultures, consist of a heterogeneous population of cells and the endogenous expression of transcription factor Runx2 is relatively low compared to histones which are abundant in tissues. Thus, the noisy ChIP-Seq results may be due to the

insufficient enrichment of the Runx2 protein for immunoprecipitation. Fluorescence-activated cell sorting of the *Runx2*$^{F2A\text{-}EGFP/HA}$ mouse embryos prior to immunoprecipitation was considered, however, the time taken to sort the cells before cross-linking would have changed the binding dynamics. Another group in the institute attempted to cross-link mouse tissues before sorting for ChIP but that did not yield good ChIP-Seq results either. Though a sensitive and specific ChIP-grade antibody is tantamount to a successful ChIP experiment, it is also believed that the protein of interest needs to be sufficiently enriched for prior to immunoprecipitation to reduce the noise level and to enhance the binding peak so as to be detected as significant. So far, there are no reports of a successful ChIP-seq to map binding sites of a transcription factor using mouse tissue. Though my Runx2 ChIP-Seq result is far from ideal, some of the relevant binding sites identified are nonetheless informative and useful.

CHAPTER 4 – BUILDING THE GENE REGULATORY NETWORK

The task of integrating the massive amount of data generated from gene expression profiling and ChIP-Seq requires a huge amount of time and bioinformatics expertise. Furthermore, a lot more work is required to validate the targets elucidated. As a preliminary attempt to map out the gene regulatory network governed by Runx2 and Runx3 during skeletogenesis, the top genes from the microarray profiles were screened through for relevant skeletal functions or phenotype. Runx2 targets with Runx2 binding at the promoter were also given more focus. Figure 52 illustrates a plausible gene regulatory network connecting Runx2, Runx3 and novel skeletal targets. Novel regulatory relationships between the Runx proteins and known skeletal factors are also depicted below.

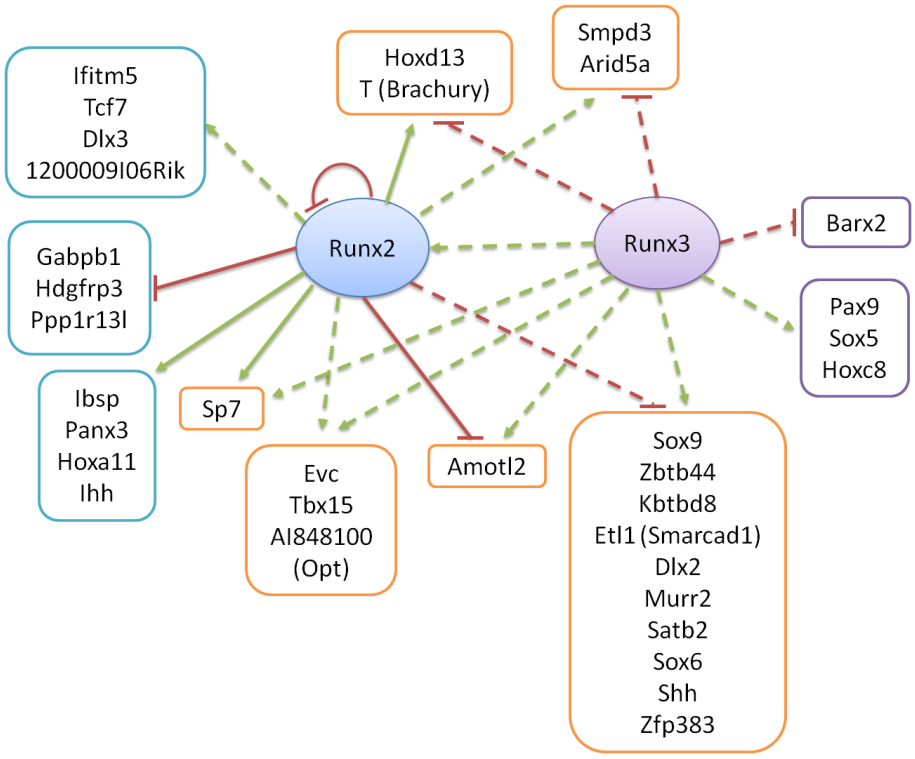


Figure 52. A GRN Centred on Runx2 and Runx3

Green arrows depict positive influence while blunted red lines depict inhibitory regulation. Solid and dotted lines represent direct and indirect regulation respectively. All regulations by Runx3 are depicted with dotted lines since no ChIP-Seq was done to unravel the Runx3 direct targets.

Genes with probable connection to skeletal development found to be negatively-regulated by Runx2 and positively-regulated by Runx3 in my microarray studies were *Amotl2*, *Sox9*, *Zbtb44*, *Kbtbd8*, *Etl1* (*Smarcad1*), *Dlx2*, *Murr2*, *Satb2*, *Sox6*, *Shh* and *Zfp383* while genes regulated vice versa by Runx2 and Runx3 were *Hoxd13*, *T* (*Brachyury)*, *Smpd3* and *Arid5a*. *Angiomotin-like 2*, *Amotl2*, encodes for a protein receptor for angiostatin, an angiogenesis inhibitor (Aase et al., 2007). Runx2 and Runx3 might be regulating this gene to control vascular invasion of the endochondral bone. The gene locus of *Kelch repeat and BTB (POZ) domain containing 8*, *Kbtbd8* and *Zinc finger and BTB domain containing 38, Zbtb38,* were recently discovered to be strongly associated with short stature in humans (Hong et al., 2011; Kim et al., 2010; Kim et al., 2011). Furthermore, strong expressions of *Kbtbd8* were detected in the forelimb bud and facial regions of an E11.5 mouse embryo (MGI database; Table 10). Although *Zbtb38* was not differentially expressed in the *Runx2* and *Runx3* expression profile, a close member *Zbtb44* was significantly down-regulated (94.6-fold) in the $Runx2^{-/-}Runx3^{-/-}$ transcriptome compared to that of the $Runx2^{-/-}$ and up-regulated (8.2-fold) in the $Runx2^{-/-}$ compared to $Runx2^{+/+}$. Deletion of *Etl1* (*Smarcad1*), *Dlx2* and *Satb2* in the mouse presented a relevant skeletal phenotype (Dobreva et al., 2006; Harris et al., 2003; Schoor et al., 1999). The expression of Etl1 was also detected in the forelimb and head mesenchyme of an E12.5 mouse embryo (MGI database) and *Murr2* is a positive regulator of Wnt signalling pathway (Lui et al., 2011) (Table 10). Therefore, these targets were considered interesting and germane to bone development.

Apart from the known committed osteoblast marker Sp7, deletion of *Evc*, *Tbx15*, *AI848100 (Opt)* in the mouse resulted in impaired bone formation (Ruiz-Perez et al., 2007; Singh et al., 2005; Sohaskey et al., 2010) (Table 10). *Evc* is also a positive modulator of Ihh signalling (Ruiz-Perez et al., 2007). These genes were found to be down-regulated in $Runx2^{-/-}$ and $Runx2^{-/-}Runx3^{-/-}$ compared to $Runx2^{+/+}$ and $Runx2^{-/-}$

respectively suggesting that they are possibly activated by Runx2 and Runx3 in a synergistic or compensatory manner to regulate chondrocyte maturation and ossification. Besides working cooperatively, *Runx3* positively regulates *Runx2* which concurs with the data published (Otto et al., 2003). *Runx2* transcripts were 8.4-fold up-regulated in the *Runx2*$^{-/-}$ compared to *Runx2*$^{+/+}$ but when both alleles of *Runx2* and *Runx3* were disrupted, *Runx2* transcripts were down-regulated by 13.9-fold in the *Runx2*$^{-/-}$*Runx3*$^{-/-}$ compared to that in the *Runx2*$^{-/-}$. It is postulated that when *Runx2* was disrupted, the change in levels of the Runx2 downstream targets triggered Runx3 to upregulate *Runx2* transcripts in a negative feedback manner and when both *Runx2* and *Runx3* were disrupted, there was no functional Runx3 protein to maintain the *Runx2* transcript levels resulting in a greatly reduced *Runx2* transcript level. However, it could not be determined if Runx3 directly regulates *Runx2*.

Genes that were exclusively regulated by Runx3 and not Runx2 were *Pax9*, *Sox5*, *Hoxc8* and *Barx2* of which *Barx2* was repressed by Runx3. Pax9, Hoxc8, Sox5 and Barx2 all have roles in skeletal development based on their knockout phenotype and literature. Pax9 directly activates *Bapx1* transcription to induce chondrogenic differentiation in the vertebral column (Rodrigo et al., 2003), Sox5 is essential for cartilage formation and maintenance of hypertrophic chondrocytes (Smits et al., 2004; Smits et al., 2001) and Hoxc8 negatively regulates osteoblast differentiation (Zheng et al., 2009). Runx3 might be positively regulating *Pax9* and *Sox5* to promote chondrogenesis and chondrocyte maturation while up-regulating *Hoxc8* to inhibit osteoblast differentiation and induce chondrogenic differentiation of osteochondroprogenitors. Barx2 cooperates with the Sox proteins to regulate expression of *Col2a1* during limb chondrogenesis particularly in the joint and articular cartilage of the developing digits (Meech et al., 2005). Runx3 might be inhibiting *Barx2* to advance chondrocyte hypertrophy in the mesenchymal condensations of the limb digits.

Runx2 exclusively activated *Ibsp*, *Panx3*, *Hoxa11*, *Ihh*, *Tcf7*, *Dlx3, Ifitm5* and *1200009I06Rik* to promote bone formation. *Ibsp, Panx3* and *Ihh* were found to be direct targets of Runx2 in *in vitro* studies conducted by (Bond et al., 2011; Roca et al., 2005; Yoshida et al., 2004). It was never demonstrated *in vivo* that Runx2 directly regulates these genes through binding at its promoter. I did not find any Runx2 bound to the promoter or any distal enhancer sites of *Ibsp, Panx3* or *Ihh* in my *in vivo* ChIP-Seq study. Maybe Runx2 does not directly regulate *Ibsp, Panx3* and *Ihh* at this developmental stage or the *in vivo* ChIP-Seq study performed was not sensitive enough to pick up the binding sites. Nevertheless, I cannot rule out the possibility that Runx2 might be directly regulating these genes during skeletal development. *Hoxa11* was identified as a direct target of Runx2 in the ribs (Table 9) and is a potential direct target of Runx2. My finding of *Tcf7* being positively regulated by Runx2 concurred with the recent literature reporting that Runx2 regulates chondrocyte maturation through *Tcf7* (Mikasa et al., 2011). Although Dlx3 was reported to be upstream of Runx2 during osteoblastogenesis (Hassan et al., 2006), the elimination of functional Runx2 in the mice resulted in down-regulation of *Dlx3* transcripts in my microarray study indicating that Runx2 is likely to regulate *Dlx3* as a form of positive feedback loop. *Ifitm5* and *1200009I06Rik* are hypothesized to be two novel targets of Runx2 found in my microarray study. There are no reports of them being linked to Runx2 as yet but the RNA in-situ hybridization validation results (Fig. 31) strongly support the hypothesis.

Genes found to be exclusively repressed directly by Runx2 were *Gabpb1*, *Hdgfrp3* and *Ppp1r13l*. *Gabpb1*, *Hdgfrp3* and *Ppp1r13l* have no known skeletal roles but their promoters were bound by Runx2 in the various tissues and were up-regulated in the *Runx2*$^{-/-}$ compared to *Runx2*$^{+/+}$ (Table 9). In order to determine their involvement in skeletal development, their spatio-temporal expressions need to be further assessed.

Table 10. List of targets with relevant skeletal functions

No.	Gene Symbol	Gene Name	Remarks	Reference
1	*Sox9*	SRY-box containing gene 9	Sox9 is the master regulator for chondrogenesis	(Wright et al., 1995)
2	*BC038156 (Zbtb44)*	Zinc finger and BTB domain containing 44	No known skeletal function but a close family member Zbtb38 was strongly associated with human height	(Hong et al., 2011) (Kim et al., 2011) (Kim et al., 2010)
3	*Kbtbd8*	Kelch repeat and BTB (POZ) domain containing 8	Associated with short stature in humans. Strong expression in FL bud and facial regions of E11.5 mouse embryo (MGI database)	(Kim et al., 2010)
4	*Etl1 (Smarcad1)*	SWI/SNF-related, matrix-associated actin-dependent regulator of chromatin, subfamily a, containing DEAD/H box 1`	Homozygotes for a targeted null mutation exhibit retarded growth, impaired fertility, skeletal dysplasias, and peri- and postnatal lethality. Strong expression in the forelimb and head mesenchyme of E12.5 mouse embryos (MGI database)	(Schoor et al., 1999)
5	*Dlx2*	Distal-less homeobox 2	Homozygous null mutants show morphogenetic abnormalities in first and second branchial arch-derived proximal skeletal and soft tissue structures.	(Harris et al., 2003)
6	*Evc*	Ellis van Creveld gene homolog (human)	Mice homozygous for a null allele exhibit some lethality shortly after birth and exhibit aphagia, infertile, teeth abnormalities, short limbs and long bones, delays in ossification, and short ribs. Evc is a positive mediator of Ihh-regulated bone growth.	(Ruiz-Perez et al., 2007)
7	*Murr2 (Usp34)*	Ubiquitin specific peptidase 34	Positive regulator of canonical Wnt receptor signalling pathway	(Lui et al., 2011)
8	*AI848100 (Opt)*	Osteopotentia	Mice homozygous for a mutation in this gene display background strain dependent neonatal and postnatal lethality with impaired osteoblast differentiation resulting in impaired bone formation, brittle bones, and impaired fracture repair.	(Sohaskey et al., 2010)
9	*Satb2*	Special AT-rich sequence binding protein 2	SATB2 is a multifunctional determinant of craniofacial patterning and osteoblast differentiation	(Dobreva et al., 2006)
10	*Sox6*	SRY-box containing gene 6	The transcription factors L-Sox5 and Sox6 are essential for cartilage formation	(Smits et al., 2001) (Smits et al., 2004)
11	*Tbx15*	T-box 15	Homozygous mutants have low set ears that project laterally, skeletal abnormalities and distinctive dorsoventral coat color patterning. Expression found in prehypertrophic chondrocytes.	(Singh et al., 2005)
12	*Amotl2*	Angiomotin-like 2	Runx2 and Runx3 might be regulating this gene to control vascular invasion of the endochondral bone.	(Aase et al., 2007)

Table 10. Continued…

No.	Gene Symbol	Gene Name	Remarks	Reference
13	*Barx2*	BarH-like homeobox 2	Barx2 regulates chondrogenesis during limb development	(Meech et al., 2005)
14	*Hoxc8*	Homeobox C8	Mice homozygous for a hypomorphic allele exhibit abnormal growth and axial skeleton morphology. Mice homozygous for a knock-out allele exhibit postnatal lethality, axial skeletal defects, abnormal growth, and abnormal gait.	(Juan et al., 2006) (Zheng et al., 2009)
15	*Pax9*	Paired box gene 9	Mice homozygous for knock-out allele exhibit neonatal lethality, abnormal cranium morphology, arrested tooth development, cleft secondary palate, athymia, and polydactyly.	(Peters et al., 1999) (Rodrigo et al., 2003)
16	*Sox5*	SRY-box containing gene 5	Homozygous null mice fail to breathe and die at birth exhibiting a narrow thoracic cage, irregularly mineralized sternum, cleft secondary palate, and delayed bone mineralization.	(Smits et al., 2001) (Smits et al., 2004)
17	*T*	Brachyury	Homozygotes exhibit defects in notochord differentiation and mesoderm formation, lack a trunk and tail, and die around 10 dpc. Heterozygotes have a shortened tail and abnormal sacral vertebrae.	(Hoffmann et al., 2002) (Ghebranious et al., 2008) (Wu et al., 2010)
18	*Smpd3*	Sphingomyelin phosphodiesterase 3, neutral	Homozygous null mice exhibit dwarfism, delayed sexual and gonad maturation, delayed ossification of the long bones, and reduced serum levels of thyroxine, triiodothyronine, cortisol, and insulin-like growth factor.	(Aubin et al., 2005) (Stoffel et al., 2007) (Chae et al., 2009)
19	*Arid5a*	AT rich interactive domain 5A (MRF1-like)	Arid5a cooperates with Sox9 to stimulate chondrocyte differentiation	(Amano et al., 2011)
20	*Hoxd13*	Homeobox D13	Homozygotes for targeted and spontaneous mutations exhibit abnormalities of the axial and appendicular skeleton especially the limbs.	(Debeer et al., 2002) (Salsi et al., 2008)
21	*Panx3*	Pannexin 3	Panx3 promotes chondrocyte and osteoblast differentiation. Found to be a direct target of Runx2 in osteoblasts and mature chondrocytes.	(Iwamoto et al., 2010) (Bond et al., 2011) (Ishikawa et al., 2011)
22	*Hoxa11*	Homeobox A11	Homozygotes for targeted null mutations exhibit homeotic transformations affecting thoracic and sacral vertebrae, and forelimb defects.	(Boulet and Capecchi, 2004)
23	*Ifitm5 (Bril)*	Interferon induced transmembrane protein 5	A novel osteoblast-specific protein. Long bones of homozygous mutant mice exhibit reduction in length during prenatal development.	(Moffatt et al., 2008) (Hanagata et al., 2010)
24	*Tcf7*	Transcription factor 7, T-cell specific	Regulation of Tcf7by Runx2 during chondrocyte maturation and proliferation	(Mikasa et al., 2011)

Table 10. Continued 2…

No.	Gene Symbol	Gene Name	Remarks	Reference
25	*Dlx3*	Distal-less homeobox 3	Dlx3 together with Dlx5 directly activate Runx2 activity in osteoblasts	(Hassan et al., 2006)

CHAPTER 5 – CONCLUSION

By coupling mouse transgenic techniques with high throughput genomic studies such as gene expression profiling and ChIP-Seq for the study of two skeletal genes, *Runx2* and *Runx3*, I was able to more accurately identify new factors that are potentially involved in skeletogenesis directly or indirectly governed by Runx2 and/or Runx3 transcription factors. The generation of three fluorescing mouse lines – wild-type *Runx2* tagged with *EGFP* and mutant *Runx2* and *Runx3* whereby the genes were disrupted by the insertion of *EGFP* – allowed for the purification of *Runx2*- or *Runx3*-expressing cells prior to microarray in order to specifically tease out the authentic targets of Runx2 and Runx3. The generation of the C-terminally *HA-tagged Runx2* mouse line also enabled ChIP-seq to be performed with anti-HA antibody that was known to work well for ChIP experiments and allowed for the specific immunoprecipitation of Runx2 proteins in order to map out Runx2 binding sites without concerns of antibody cross-reactivity which may confound the ChIP-seq data. Unfortunately, the tagging of *Runx3* at the C-terminus was not successful for reasons that were explored in chapter 3.2.2.

As validated and discussed in chapter 3.1.5, there were already evidences that my microarray data was more accurate. Some genes that were not found in the previous two Runx2 microarray studies (Hecht et al., 2007; Vaes et al., 2006) such as *Panx3* and *Ifitm5* were recently found to have skeletal functions and *Panx3* was even demonstrated to be a direct target of Runx2 in an *in vitro* setting (Bond et al., 2011). Unfortunately, my ChIP-seq experiment performed with anti-HA antibody on DNA-bound Runx2-HA proteins from the HA-tagged mice did not uncover any Runx2 binding sites within 1 megabase of *Panx3* in any of the three tissues (jaw, ribs and limbs). Runx2 might not be regulating *Panx3* directly at this stage or the binding site

was not picked up for the reasons discussed in chapter 3.2.4. Ifitm5 has been implicated as a novel osteoblast protein and could be a novel target of Runx2 during osteoblastogenesis. In addition, a novel gene *1200009I06Rik* was discovered to be a potential target of Runx2 in the microarray data. Its expression was specifically obliterated in the skeletal elements of the *Runx2*$^{-/-}$ embryos. Other genes specifically regulated by Runx2 and not Runx3 were also identified such as *Ibsp, Tcf7, Hoxa11, Dlx3, Ihh, Gabpb1, Hdgfrp3* and *Ppp1r13l*.

Runx3 is the least studied of the three *Runx* members. As mentioned, there is currently no gene expression profiling of *Runx3* performed in the context of skeletal development for comparison. This is the first study undertaken to uncover Runx3 skeletal targets by comparing the transcriptomes of an enriched pool of *Runx3*-expressing cells from the *Runx3*$^{+/-}$ and *Runx3*$^{-/-}$ mouse embryos. However, only 24 factors emerged to be regulated by Runx3. From the skeletal phenotype of the *Runx2*$^{-/-}$ and *Runx3*$^{-/-}$ mice, it is recognized that Runx2 is the dominant skeletal player of the two as it is able to compensate almost fully for the lack of Runx3 resulting in a Runx3$^{-/-}$ mouse that has no overt skeletal phenotype. In order to circumvent the compensation of Runx3 by Runx2 and to successfully elucidate all the Runx3 skeletal targets, the transcriptome of enriched populations of *Runx2*- and *Runx3*-expressing cells from *Runx2*$^{-/-}$*Runx3*$^{-/-}$ mouse embryos were profiled and compared against that from *Runx2*$^{-/-}$ mouse embryos. As a result, novel skeletal factors such as *Zbtb44*, *Kbtbd8*, *Etl1 (Smarcad1)*, *Murr2 (Usp34)*, *Zfp383*, *Evc*, *Tbx15*, *AI848100 (Opt)*, *Arid5a, Smpd3* and *Brachyury* were discovered to be regulated by both Runx2 and Runx3. Furthermore, well-established skeletal genes such as *Sox9*, *Satb2*, *Sox6, Dlx2*, *Hoxd13*, *Sp7* and *Shh* were also identified as putative novel targets of both Runx2 and Runx3. Genes such as *Pax9*, *Sox5*, *Hoxc8* and *Barx2* were also found to be exclusively regulated by Runx3 as they were not found differentially expressed in my Runx2 microarray data. These could be the genes that Runx2 is

unable to compensate at this particular stage to produce a mild delay in chondrocyte hypertrophy of the E15.5 *Runx3*$^{-/-}$ mouse embryos (Yoshida et al., 2004). Genes that were thought to be expressed early in development such as *Shh*, *Sox9* and *Pax9* might still be involved in later events of skeletogenesis and be indirectly or directly regulated by Runx3 as a positive feedback mechanism to fine tune skeletal events. Regrettably, the tagging of Runx3 protein in the mouse was unsuccessful and ChIP-seq could not be performed with anti-HA antibody to identify Runx3-specific direct targets. Alternative tagging strategy might need to be explored with Runx3. Other tags such as the V5 or the S-peptide tag might work better with the Runx3 protein. The tagging position other than the C-terminus can also be experimented to circumvent the problem encountered with the C-terminal HA-tagging of Runx3.

While extensive validation of individual targets and their significance to bone formation needs to follow, the strength of this study lies in the vast number of potential Runx2 and Runx3 downstream targets that have not been previously identified and in the fact that the study was performed entirely *in vivo* thus accounting for the spatio-temporal factors that come into play during actual developmental events. Furthermore, sufficient cellular resolution of the specific investigated gene activity was achieved prior to transcriptome profiling. The known targets of Runx2 that emerged in my study also conveyed confidence to the method and technique used. With more targets identified to be involved in skeletogenesis, greater insights into bone development are gleaned and more of these novel skeletal factors potentially serve as targets for preventive and regenerative therapy as well as bone tissue engineering. Given the prevailing uncharted transcriptional control of chondrocyte hypertrophy, the targets of Runx3 are likely to serve as valuable nodes in the pursuit of generating a comprehensive network that represents the complex operations driving skeletal development.

Moving forward, the results serve as a platform to venture into more specific studies such as investigating osteoblastogenesis in intramembranous bones and chondrocyte maturation in endochondral bones or moving to earlier time-points to unravel the first targets of skeletal development. The main challenge facing these more specific or earlier time-point studies is the availability of material for microarray. However, with the advent of RNA-Seq which requires as little material as RNA from a single cell (Tang et al., 2010), the obstacles are overcome. RNA-Seq is not limited to probes placed on the microarray and differentially regulated isoforms can also be identified. In the same way as ChIP-chip has given way to ChIP-Seq, RNA-Seq is the future avenue to expression profiling.

Lastly, the approach in this study to uncover downstream targets of known key factors to map out the network controlling skeletogenesis presents a promising method that can be applied to unravel downstream targets of known factors in any investigated developmental system.

REFERENCES

Aase, K., Ernkvist, M., Ebarasi, L., Jakobsson, L., Majumdar, A., Yi, C., Birot, O., Ming, Y., Kvanta, A., Edholm, D.*, et al.* (2007). Angiomotin regulates endothelial cell migration during embryonic angiogenesis. Genes Dev *21*, 2055-2068.

Akiyama, H., Chaboissier, M.C., Martin, J.F., Schedl, A., and de Crombrugghe, B. (2002). The transcription factor Sox9 has essential roles in successive steps of the chondrocyte differentiation pathway and is required for expression of Sox5 and Sox6. Genes Dev *16*, 2813-2828.

Amano, K., Hata, K., Muramatsu, S., Wakabayashi, M., Takigawa, Y., Ono, K., Nakanishi, M., Takashima, R., Kogo, M., Matsuda, A.*, et al.* (2011). Arid5a cooperates with Sox9 to stimulate chondrocyte-specific transcription. Mol Biol Cell *22*, 1300-1311.

Arnold, M.A., Kim, Y., Czubryt, M.P., Phan, D., McAnally, J., Qi, X., Shelton, J.M., Richardson, J.A., Bassel-Duby, R., and Olson, E.N. (2007). MEF2C transcription factor controls chondrocyte hypertrophy and bone development. Dev Cell *12*, 377-389.

Aubin, I., Adams, C.P., Opsahl, S., Septier, D., Bishop, C.E., Auge, N., Salvayre, R., Negre-Salvayre, A., Goldberg, M., Guenet, J.L.*, et al.* (2005). A deletion in the gene encoding sphingomyelin phosphodiesterase 3 (Smpd3) results in osteogenesis and dentinogenesis imperfecta in the mouse. Nat Genet *37*, 803-805.

Avraham, K.B., Levanon, D., Negreanu, V., Bernstein, Y., Groner, Y., Copeland, N.G., and Jenkins, N.A. (1995). Mapping of the mouse homolog of the human runt domain gene, AML2, to the distal region of mouse chromosome 4. Genomics *25*, 603-605.

Bae, S.C., Takahashi, E., Zhang, Y.W., Ogawa, E., Shigesada, K., Namba, Y., Satake, M., and Ito, Y. (1995). Cloning, mapping and expression of PEBP2 alpha C, a third gene encoding the mammalian Runt domain. Gene *159*, 245-248.

Banerjee, C., Javed, A., Choi, J.Y., Green, J., Rosen, V., van Wijnen, A.J., Stein, J.L., Lian, J.B., and Stein, G.S. (2001). 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.

Banerjee, C., McCabe, L.R., Choi, J.Y., Hiebert, S.W., Stein, J.L., Stein, G.S., and Lian, J.B. (1997). Runt homology domain proteins in osteoblast differentiation: AML3/CBFA1 is a major component of a bone-specific complex. J Cell Biochem *66*, 1-8.

Bangsow, C., Rubins, N., Glusman, G., Bernstein, Y., Negreanu, V., Goldenberg, D., Lotem, J., Ben-Asher, E., Lancet, D., Levanon, D.*, et al.* (2001). The RUNX3 gene--sequence, structure and regulated expression. Gene *279*, 221-232.

Bar-Shavit, Z. (2007). The osteoclast: a multinucleated, hematopoietic-origin, bone-resorbing osteoimmune cell. J Cell Biochem *102*, 1130-1139.

Bell, D.M., Leung, K.K., Wheatley, S.C., Ng, L.J., Zhou, S., Ling, K.W., Sham, M.H., Koopman, P., Tam, P.P., and Cheah, K.S. (1997). SOX9 directly regulates the type-II collagen gene. Nat Genet *16*, 174-178.

Bialek, P., Kern, B., Yang, X., Schrock, M., Sosic, D., Hong, N., Wu, H., Yu, K., Ornitz, D.M., Olson, E.N.*, et al.* (2004). A twist code determines the onset of osteoblast differentiation. Dev Cell *6*, 423-435.

Blyth, K., Cameron, E.R., and Neil, J.C. (2005). The RUNX genes: gain or loss of function in cancer. Nat Rev Cancer *5*, 376-387.

Bohm, B.B., Aigner, T., Roy, B., Brodie, T.A., Blobel, C.P., and Burkhardt, H. (2005). Homeostatic effects of the metalloproteinase disintegrin ADAM15 in degenerative cartilage remodeling. Arthritis Rheum *52*, 1100-1109.

Bond, S.R., Lau, A., Penuela, S., Sampaio, A.V., Underhill, T.M., Laird, D.W., and Naus, C.C. (2011). Pannexin 3 is a novel target for Runx2, expressed by osteoblasts and mature growth plate chondrocytes. J Bone Miner Res *26*, 2911-2922.

Boulet, A.M., and Capecchi, M.R. (2004). Multiple roles of Hoxa11 and Hoxd11 in the formation of the mammalian forelimb zeugopod. Development *131*, 299-309.

Brady, G., and Farrell, P.J. (2009). RUNX3-mediated repression of RUNX1 in B cells. J Cell Physiol *221*, 283-287.

Britanova, O., Depew, M.J., Schwark, M., Thomas, B.L., Miletich, I., Sharpe, P., and Tarabykin, V. (2006). Satb2 haploinsufficiency phenocopies 2q32-q33 deletions, whereas loss suggests a fundamental role in the coordination of jaw development. Am J Hum Genet *79*, 668-678.

Britten, R.J., and Davidson, E.H. (1969). Gene regulation for higher cells: a theory. Science *165*, 349-357.

Calabi, F., Rhodes, M., Williamson, P., and Boyd, Y. (1995). Identification and chromosomal mapping of a third mouse runt-like locus. Genomics *26*, 607-610.

Chae, Y.M., Heo, S.H., Kim, J.Y., Lee, J.M., Ryoo, H.M., and Cho, J.Y. (2009). Upregulation of smpd3 via BMP2 stimulation and Runx2. BMB Rep *42*, 86-90.

Chi, X.Z., Yang, J.O., Lee, K.Y., Ito, K., Sakakura, C., Li, Q.L., Kim, H.R., Cha, E.J., Lee, Y.H., Kaneda, A.*, et al.* (2005). RUNX3 suppresses gastric epithelial cell growth by inducing p21(WAF1/Cip1) expression in cooperation with transforming growth factor {beta}-activated SMAD. Mol Cell Biol *25*, 8097-8107.

Chin, H.J., Fisher, M.C., Li, Y., Ferrari, D., Wang, C.K., Lichtler, A.C., Dealy, C.N., and Kosher, R.A. (2007). Studies on the role of Dlx5 in regulation of chondrocyte differentiation during endochondral ossification in the developing mouse limb. Dev Growth Differ *49*, 515-521.

Choi, K.Y., Lee, S.W., Park, M.H., Bae, Y.C., Shin, H.I., Nam, S., Kim, Y.J., Kim, H.J., and Ryoo, H.M. (2002). Spatio-temporal expression patterns of Runx2 isoforms in early skeletogenesis. Exp Mol Med *34*, 426-433.

Cobb, J., Dierich, A., Huss-Garcia, Y., and Duboule, D. (2006). A mouse model for human short-stature syndromes identifies Shox2 as an upstream regulator of Runx2 during long-bone development. Proc Natl Acad Sci U S A *103*, 4511-4515.

Cohen, M.M., Jr. (2001). RUNX genes, neoplasia, and cleidocranial dysplasia. Am J Med Genet *104*, 185-188.

Cosma, M.P. (2002). Ordered recruitment: gene-specific mechanism of transcription activation. Mol Cell *10*, 227-236.

Crute, B.E., Lewis, A.F., Wu, Z., Bushweller, J.H., and Speck, N.A. (1996). Biochemical and biophysical properties of the core-binding factor alpha2 (AML1) DNA-binding domain. J Biol Chem *271*, 26251-26260.

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

Davidson, E.H., Rast, J.P., Oliveri, P., Ransick, A., Calestani, C., Yuh, C.H., Minokawa, T., Amore, G., Hinman, V., Arenas-Mena, C.*, et al.* (2002). A provisional regulatory gene network for specification of endomesoderm in the sea urchin embryo. Dev Biol *246*, 162-190.

de Bruijn, M.F., and Speck, N.A. (2004). Core-binding factors in hematopoiesis and immune function. Oncogene *23*, 4238-4248.

de Felipe, P. (2002). Polycistronic viral vectors. Curr Gene Ther *2*, 355-378.

de Felipe, P., Hughes, L.E., Ryan, M.D., and Brown, J.D. (2003). Co-translational, intraribosomal cleavage of polypeptides by the foot-and-mouth disease virus 2A peptide. J Biol Chem *278*, 11441-11448.

Debeer, P., Bacchelli, C., Scambler, P.J., De Smet, L., Fryns, J.P., and Goodman, F.R. (2002). Severe digital abnormalities in a patient heterozygous for both a novel missense mutation in HOXD13 and a polyalanine tract expansion in HOXA13. J Med Genet *39*, 852-856.

Dobreva, G., Chahrour, M., Dautzenberg, M., Chirivella, L., Kanzler, B., Farinas, I., Karsenty, G., and Grosschedl, R. (2006). SATB2 is a multifunctional determinant of craniofacial patterning and osteoblast differentiation. Cell *125*, 971-986.

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

Drissi, H., Pouliot, A., Koolloos, C., Stein, J.L., Lian, J.B., Stein, G.S., and van Wijnen, A.J. (2002). 1,25-(OH)2-vitamin D3 suppresses the bone-related Runx2/Cbfa1 gene promoter. Exp Cell Res *274*, 323-333.

Drissi, M.H., Li, X., Sheu, T.J., Zuscik, M.J., Schwarz, E.M., Puzas, J.E., Rosier, R.N., and O'Keefe, R.J. (2003). Runx2/Cbfa1 stimulation by retinoic acid is potentiated by BMP2 signaling through interaction with Smad1 on the collagen X promoter in chondrocytes. J Cell Biochem *90*, 1287-1298.

Ducy, P., Starbuck, M., Priemel, M., Shen, J., Pinero, G., Geoffroy, V., Amling, M., and Karsenty, G. (1999). A Cbfa1-dependent genetic pathway controls bone formation beyond embryonic development. Genes Dev *13*, 1025-1036.

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

Eferl, R., Hoebertz, A., Schilling, A.F., Rath, M., Karreth, F., Kenner, L., Amling, M., and Wagner, E.F. (2004). The Fos-related antigen Fra-1 is an activator of bone matrix formation. EMBO J *23*, 2789-2799.

Eggan, K., Akutsu, H., Loring, J., Jackson-Grusby, L., Klemm, M., Rideout, W.M., 3rd, Yanagimachi, R., and Jaenisch, R. (2001). Hybrid vigor, fetal overgrowth, and viability of mice derived by nuclear cloning and tetraploid embryo complementation. Proc Natl Acad Sci U S A *98*, 6209-6214.

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

Fry, C.J., and Peterson, C.L. (2001). Chromatin remodeling enzymes: who's on first? Curr Biol *11*, R185-197.

Furlong, E.E. (2004). Integrating transcriptional and signalling networks during muscle development. Curr Opin Genet Dev *14*, 343-350.

Garofalo, S., Metsaranta, M., Ellard, J., Smith, C., Horton, W., Vuorio, E., and de Crombrugghe, B. (1993). Assembly of cartilage collagen fibrils is disrupted by overexpression of normal type II collagen in transgenic mice. Proc Natl Acad Sci U S A *90*, 3825-3829.

Gaur, T., Lengner, C.J., Hovhannisyan, H., Bhat, R.A., Bodine, P.V., Komm, B.S., Javed, A., van Wijnen, A.J., Stein, J.L., Stein, G.S.*, et al.* (2005). Canonical WNT signaling promotes osteogenesis by directly stimulating Runx2 gene expression. J Biol Chem *280*, 33132-33140.

Geoffroy, V., Kneissel, M., Fournier, B., Boyde, A., and Matthias, P. (2002). High bone resorption in adult aging transgenic mice overexpressing cbfa1/runx2 in cells of the osteoblastic lineage. Mol Cell Biol *22*, 6222-6233.

Gerber, H.P., Vu, T.H., Ryan, A.M., Kowalski, J., Werb, Z., and Ferrara, N. (1999). VEGF couples hypertrophic cartilage remodeling, ossification and angiogenesis during endochondral bone formation. Nat Med *5*, 623-628.

Gergen, J.P., and Butler, B.A. (1988). Isolation of the Drosophila segmentation gene runt and analysis of its expression during embryogenesis. Genes Dev *2*, 1179-1193.

Ghebranious, N., Blank, R.D., Raggio, C.L., Staubli, J., McPherson, E., Ivacic, L., Rasmussen, K., Jacobsen, F.S., Faciszewski, T., Burmester, J.K.*, et al.* (2008). A missense T (Brachyury) mutation contributes to vertebral malformations. J Bone Miner Res *23*, 1576-1583.

Gilbert, L., He, X., Farmer, P., Rubin, J., Drissi, H., van Wijnen, A.J., Lian, J.B., Stein, G.S., and Nanes, M.S. (2002). Expression of the osteoblast differentiation factor RUNX2 (Cbfa1/AML3/Pebp2alpha A) is inhibited by tumor necrosis factor-alpha. J Biol Chem *277*, 2695-2701.

Goel, A., Arnold, C.N., Tassone, P., Chang, D.K., Niedzwiecki, D., Dowell, J.M., Wasserman, L., Compton, C., Mayer, R.J., Bertagnolli, M.M.*, et al.* (2004). Epigenetic inactivation of RUNX3 in microsatellite unstable sporadic colon cancers. Int J Cancer *112*, 754-759.

Gollner, H., Dani, C., Phillips, B., Philipsen, S., and Suske, G. (2001). Impaired ossification in mice lacking the transcription factor Sp3. Mech Dev *106*, 77-83.

Gu, T.L., Goetz, T.L., Graves, B.J., and Speck, N.A. (2000). Auto-inhibition and partner proteins, core-binding factor beta (CBFbeta) and Ets-1, modulate DNA binding by CBFalpha2 (AML1). Mol Cell Biol *20*, 91-103.

Gutierrez, S., Javed, A., Tennant, D.K., van Rees, M., Montecino, M., Stein, G.S., Stein, J.L., and Lian, J.B. (2002). CCAAT/enhancer-binding proteins (C/EBP) beta and delta activate osteocalcin gene transcription and synergize with Runx2 at the C/EBP element to regulate bone-specific expression. J Biol Chem *277*, 1316-1323.

Hanagata, N., Li, X., Morita, H., Takemura, T., Li, J., and Minowa, T. (2010). Characterization of the osteoblast-specific transmembrane protein IFITM5 and analysis of IFITM5-deficient mice. J Bone Miner Metab *29*, 279-290.

Harada, H., Tagashira, S., Fujiwara, M., Ogawa, S., Katsumata, T., Yamaguchi, A., Komori, T., and Nakatsuka, M. (1999). Cbfa1 isoforms exert functional differences in osteoblast differentiation. J Biol Chem *274*, 6972-6978.

Harris, S.E., Guo, D., Harris, M.A., Krishnaswamy, A., and Lichtler, A. (2003). Transcriptional regulation of BMP-2 activated genes in osteoblasts using gene expression microarray analysis: role of Dlx2 and Dlx5 transcription factors. Front Biosci *8*, s1249-1265.

Hartmann, C. (2009). Transcriptional networks controlling skeletal development. Curr Opin Genet Dev *19*, 437-443.

Hassan, M.Q., Javed, A., Morasso, M.I., Karlin, J., Montecino, M., van Wijnen, A.J., Stein, G.S., Stein, J.L., and Lian, J.B. (2004). Dlx3 transcriptional regulation of osteoblast differentiation: temporal recruitment of Msx2, Dlx3, and Dlx5 homeodomain proteins to chromatin of the osteocalcin gene. Mol Cell Biol *24*, 9248-9261.

Hassan, M.Q., Tare, R.S., Lee, S.H., Mandeville, M., Morasso, M.I., Javed, A., van Wijnen, A.J., Stein, J.L., Stein, G.S., and Lian, J.B. (2006). BMP2 commitment to the osteogenic lineage involves activation of Runx2 by DLX3 and a homeodomain transcriptional network. J Biol Chem *281*, 40515-40526.

Hecht, J., Seitz, V., Urban, M., Wagner, F., Robinson, P.N., Stiege, A., Dieterich, C., Kornak, U., Wilkening, U., Brieske, N.*, et al.* (2007). Detection of novel skeletogenesis target genes by comprehensive analysis of a Runx2(-/-) mouse model. Gene Expr Patterns *7*, 102-112.

Hellen, C.U., and Sarnow, P. (2001). Internal ribosome entry sites in eukaryotic mRNA molecules. Genes Dev *15*, 1593-1612.

Herbrand, H., Pabst, O., Hill, R., and Arnold, H.H. (2002). Transcription factors Nkx3.1 and Nkx3.2 (Bapx1) play an overlapping role in sclerotomal development of the mouse. Mech Dev *117*, 217-224.

Hess, J., Porte, D., Munz, C., and Angel, P. (2001). AP-1 and Cbfa/runt physically interact and regulate parathyroid hormone-dependent MMP13 expression in osteoblasts through a new osteoblast-specific element 2/AP-1 composite element. J Biol Chem *276*, 20029-20038.

Higashikawa, A., Saito, T., Ikeda, T., Kamekura, S., Kawamura, N., Kan, A., Oshima, Y., Ohba, S., Ogata, N., Takeshita, K.*, et al.* (2009). Identification of the core element responsive to runt-related transcription factor 2 in the promoter of human type X collagen gene. Arthritis Rheum *60*, 166-178.

Hill, R.E., and van Heyningen, V. (2008). Long-range control of gene expression. Preface. Adv Genet *61*, xiii-xv.

Hill, T.P., Spater, D., Taketo, M.M., Birchmeier, W., and Hartmann, C. (2005). Canonical Wnt/beta-catenin signaling prevents osteoblasts from differentiating into chondrocytes. Dev Cell *8*, 727-738.

Hinoi, E., Bialek, P., Chen, Y.T., Rached, M.T., Groner, Y., Behringer, R.R., Ornitz, D.M., and Karsenty, G. (2006). Runx2 inhibits chondrocyte proliferation and hypertrophy through its expression in the perichondrium. Genes Dev *20*, 2937-2942.

Hoffmann, A., Czichos, S., Kaps, C., Bachner, D., Mayer, H., Kurkalli, B.G., Zilberman, Y., Turgeman, G., Pelled, G., Gross, G.*, et al.* (2002). The T-box transcription factor Brachyury mediates cartilage development in mesenchymal stem cell line C3H10T1/2. J Cell Sci *115*, 769-781.

Hofmann, K., Tomiuk, S., Wolff, G., and Stoffel, W. (2000). Cloning and characterization of the mammalian brain-specific, Mg2+-dependent neutral sphingomyelinase. Proc Natl Acad Sci U S A *97*, 5895-5900.

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

Hong, K.W., Shin, Y.B., Jin, H.S., Lim, J.E., Choi, J.Y., Chang, K.T., Kim, H.S., and Oh, B. (2011). Alternative splicing of human height-related zinc finger and BTB domain-containing 38 gene through Alu exonization. Biochem Genet *49*, 283-291.

Hozumi, I., Inuzuka, T., and Tsuji, S. (1998). Brain injury and growth inhibitory factor (GIF)--a minireview. Neurochem Res *23*, 319-328.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009a). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc *4*, 44-57.

Huang da, W., Sherman, B.T., Zheng, X., Yang, J., Imamichi, T., Stephens, R., and Lempicki, R.A. (2009b). Extracting biological meaning from large gene lists with DAVID. Curr Protoc Bioinformatics *Chapter 13*, Unit 13 11.

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

Inoue, K., Ozaki, S., Shiga, T., Ito, K., Masuda, T., Okado, N., Iseda, T., Kawaguchi, S., Ogawa, M., Bae, S.C.*, et al.* (2002). Runx3 controls the axonal projection of proprioceptive dorsal root ganglion neurons. Nat Neurosci *5*, 946-954.

Inoue, T., Wang, M., Ririe, T.O., Fernandes, J.S., and Sternberg, P.W. (2005). Transcriptional network underlying Caenorhabditis elegans vulval development. Proc Natl Acad Sci U S A *102*, 4972-4977.

Ishikawa, M., Iwamoto, T., Nakamura, T., Doyle, A., Fukumoto, S., and Yamada, Y. (2011). Pannexin 3 functions as an ER Ca(2+) channel, hemichannel, and gap junction to promote osteoblast differentiation. J Cell Biol *193*, 1257-1274.

Ito, K., Lim, A.C., Salto-Tellez, M., Motoda, L., Osato, M., Chuang, L.S., Lee, C.W., Voon, D.C., Koo, J.K., Wang, H.*, et al.* (2008). RUNX3 attenuates beta-catenin/T cell factors in intestinal tumorigenesis. Cancer Cell *14*, 226-237.

Iwamoto, M., Kitagaki, J., Tamamura, Y., Gentili, C., Koyama, E., Enomoto, H., Komori, T., Pacifici, M., and Enomoto-Iwamoto, M. (2003). Runx2 expression and action in chondrocytes are regulated by retinoid signaling and parathyroid hormone-related peptide (PTHrP). Osteoarthritis Cartilage *11*, 6-15.

Iwamoto, T., Nakamura, T., Doyle, A., Ishikawa, M., de Vega, S., Fukumoto, S., and Yamada, Y. (2010). Pannexin 3 regulates intracellular ATP/cAMP levels and promotes chondrocyte differentiation. J Biol Chem *285*, 18948-18958.

Javed, A., Barnes, G.L., Jasanya, B.O., Stein, J.L., Gerstenfeld, L., Lian, J.B., and Stein, G.S. (2001). runt homology domain transcription factors (Runx, Cbfa, and AML) mediate repression of the bone sialoprotein promoter: evidence for promoter context-dependent activity of Cbfa proteins. Mol Cell Biol *21*, 2891-2905.

Javed, A., Guo, B., Hiebert, S., Choi, J.Y., Green, J., Zhao, S.C., Osborne, M.A., Stifani, S., Stein, J.L., Lian, J.B.*, et al.* (2000). Groucho/TLE/R-esp proteins associate with the nuclear matrix and repress RUNX (CBF(alpha)/AML/PEBP2(alpha)) dependent activation of tissue-specific gene transcription. J Cell Sci *113 (Pt 12)*, 2221-2231.

Jensen, E.D., Gopalakrishnan, R., and Westendorf, J.J. (2010). Regulation of gene expression in osteoblasts. Biofactors *36*, 25-32.

Jensen, E.D., Schroeder, T.M., Bailey, J., Gopalakrishnan, R., and Westendorf, J.J. (2008). Histone deacetylase 7 associates with Runx2 and represses its activity during osteoblast maturation in a deacetylation-independent manner. J Bone Miner Res *23*, 361-372.

Jimenez, M.J., Balbin, M., Alvarez, J., Komori, T., Bianco, P., Holmbeck, K., Birkedal-Hansen, H., Lopez, J.M., and Lopez-Otin, C. (2001). A regulatory cascade involving retinoic acid, Cbfa1, and matrix metalloproteinases is coupled to the development of a process of perichondrial invasion and osteogenic differentiation during bone formation. J Cell Biol *155*, 1333-1344.

Jimenez, M.J., Balbin, M., Lopez, J.M., Alvarez, J., Komori, T., and Lopez-Otin, C. (1999). Collagenase 3 is a target of Cbfa1, a transcription factor of the runt gene family involved in bone formation. Mol Cell Biol *19*, 4431-4442.

Jones, D.C., Wein, M.N., Oukka, M., Hofstaetter, J.G., Glimcher, M.J., and Glimcher, L.H. (2006). Regulation of adult bone mass by the zinc finger adapter protein Schnurri-3. Science *312*, 1223-1227.

Juan, A.H., Lei, H., Bhargava, P., Lebrun, M., and Ruddle, F.H. (2006). Multiple roles of hoxc8 in skeletal development. Ann N Y Acad Sci *1068*, 87-94.

Kagoshima, H., Shigesada, K., Satake, M., Ito, Y., Miyoshi, H., Ohki, M., Pepling, M., and Gergen, P. (1993). The Runt domain identifies a new family of heteromeric transcriptional regulators. Trends Genet *9*, 338-341.

Kanatani, N., Fujita, T., Fukuyama, R., Liu, W., Yoshida, C.A., Moriishi, T., Yamana, K., Miyazaki, T., Toyosawa, S., and Komori, T. (2006). Cbf beta regulates Runx2 function isoform-dependently in postnatal bone development. Dev Biol *296*, 48-61.

Kang, G.H., Lee, S., Lee, H.J., and Hwang, K.S. (2004). Aberrant CpG island hypermethylation of multiple genes in prostate cancer and prostatic intraepithelial neoplasia. J Pathol *202*, 233-240.

Kanno, T., Kanno, Y., Chen, L.F., Ogawa, E., Kim, W.Y., and Ito, Y. (1998). Intrinsic transcriptional activation-inhibition domains of the polyomavirus enhancer binding protein 2/core binding factor alpha subunit revealed in the presence of the beta subunit. Mol Cell Biol *18*, 2444-2454.

Karreth, F., Hoebertz, A., Scheuch, H., Eferl, R., and Wagner, E.F. (2004). The AP1 transcription factor Fra2 is required for efficient cartilage development. Development *131*, 5717-5725.

Karsenty, G. (2001). Minireview: transcriptional control of osteoblast differentiation. Endocrinology *142*, 2731-2733.

Karsenty, G. (2008). Transcriptional control of skeletogenesis. Annu Rev Genomics Hum Genet *9*, 183-196.

Karsenty, G., Kronenberg, H.M., and Settembre, C. (2009). Genetic control of bone formation. Annu Rev Cell Dev Biol *25*, 629-648.

Karsenty, G., and Wagner, E.F. (2002). Reaching a genetic and molecular understanding of skeletal development. Dev Cell *2*, 389-406.

Kauffman, S. (1969). Homeostasis and differentiation in random genetic control networks. Nature *224*, 177-178.

Kaufman, J.M., Taelman, P., Vermeulen, A., and Vandeweghe, M. (1992). Bone mineral status in growth hormone-deficient males with isolated and multiple pituitary deficiencies of childhood onset. J Clin Endocrinol Metab *74*, 118-123.

Kern, B., Shen, J., Starbuck, M., and Karsenty, G. (2001). Cbfa1 contributes to the osteoblast-specific expression of type I collagen genes. J Biol Chem *276*, 7101-7107.

Kim, I.S., Otto, F., Zabel, B., and Mundlos, S. (1999). Regulation of chondrocyte differentiation by Cbfa1. Mech Dev *80*, 159-170.

Kim, J.J., Lee, H.I., Park, T., Kim, K., Lee, J.E., Cho, N.H., Shin, C., Cho, Y.S., Lee, J.Y., Han, B.G.*, et al.* (2010). Identification of 15 loci influencing height in a Korean population. J Hum Genet *55*, 27-31.

Kim, J.J., Park, Y.M., Baik, K.H., Choi, H.Y., Yang, G.S., Koh, I., Hwang, J.A., Lee, J., Lee, Y.S., Rhee, H.*, et al.* (2011). Exome sequencing and subsequent association studies identify five amino acid-altering variants influencing human height. Hum Genet.

Kim, S., Koga, T., Isobe, M., Kern, B.E., Yokochi, T., Chin, Y.E., Karsenty, G., Taniguchi, T., and Takayanagi, H. (2003). Stat1 functions as a cytoplasmic attenuator of Runx2 in the transcriptional program of osteoblast differentiation. Genes Dev *17*, 1979-1991.

Kim, W.J., Kim, E.J., Jeong, P., Quan, C., Kim, J., Li, Q.L., Yang, J.O., Ito, Y., and Bae, S.C. (2005). RUNX3 inactivation by point mutations and aberrant DNA methylation in bladder tumors. Cancer Res *65*, 9347-9354.

Kimura, A., Inose, H., Yano, F., Fujita, K., Ikeda, T., Sato, S., Iwasaki, M., Jinno, T., Ae, K., Fukumoto, S.*, et al.* (2010). Runx1 and Runx2 cooperate during sternal morphogenesis. Development *137*, 1159-1167.

Koga, T., Matsui, Y., Asagiri, M., Kodama, T., de Crombrugghe, B., Nakashima, K., and Takayanagi, H. (2005). NFAT and Osterix cooperatively regulate bone formation. Nat Med *11*, 880-885.

Koide, T., Hayata, T., and Cho, K.W. (2005). Xenopus as a model system to study transcriptional regulatory networks. Proc Natl Acad Sci U S A *102*, 4943-4948.

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

Kramer, I., Sigrist, M., de Nooij, J.C., Taniuchi, I., Jessell, T.M., and Arber, S. (2006). A role for Runx transcription factor signaling in dorsal root ganglion sensory neuron diversification. Neuron *49*, 379-393.

Kraus, P., Leong, G., Tan, V., Xing, X., Goh, J.W., Yap, S.P., and Lufkin, T. (2010). A more cost effective and rapid high percentage germ-line transmitting chimeric mouse generation procedure via microinjection of 2-cell, 4-cell, and 8-cell embryos with ES and iPS cells. Genesis *48*, 394-399.

Kronenberg, H.M. (2003). Developmental regulation of the growth plate. Nature *423*, 332-336.

Lau, Q.C., Raja, E., Salto-Tellez, M., Liu, Q., Ito, K., Inoue, M., Putti, T.C., Loh, M., Ko, T.K., Huang, C.*, et al.* (2006). RUNX3 is frequently inactivated by dual mechanisms of protein mislocalization and promoter hypermethylation in breast cancer. Cancer Res *66*, 6512-6520.

Le, X.F., Groner, Y., Kornblau, S.M., Gu, Y., Hittelman, W.N., Levanon, D., Mehta, K., Arlinghaus, R.B., and Chang, K.S. (1999). Regulation of AML2/CBFA3 in hematopoietic cells through the retinoic acid receptor alpha-dependent signaling pathway. J Biol Chem *274*, 21651-21658.

Lee, B., Thirunavukkarasu, K., Zhou, L., Pastore, L., Baldini, A., Hecht, J., Geoffroy, V., Ducy, P., and Karsenty, G. (1997). Missense mutations abolishing DNA binding of the osteoblast-specific transcription factor OSF2/CBFA1 in cleidocranial dysplasia. Nat Genet *16*, 307-310.

Lee, K.S., Kim, H.J., Li, Q.L., Chi, X.Z., Ueta, C., Komori, T., Wozney, J.M., Kim, E.G., Choi, J.Y., Ryoo, H.M.*, et al.* (2000). Runx2 is a common target of transforming growth factor beta1 and bone morphogenetic protein 2, and cooperation between Runx2 and Smad5 induces osteoblast-specific gene expression in the pluripotent mesenchymal precursor cell line C2C12. Mol Cell Biol *20*, 8783-8792.

Lee, K.S., Lee, Y.S., Lee, J.M., Ito, K., Cinghu, S., Kim, J.H., Jang, J.W., Li, Y.H., Goh, Y.M., Chi, X.Z.*, et al.* (2010). Runx3 is required for the differentiation of lung epithelial cells and suppression of lung cancer. Oncogene *29*, 3349-3361.

Lee, M.H., Kim, Y.J., Yoon, W.J., Kim, J.I., Kim, B.G., Hwang, Y.S., Wozney, J.M., Chi, X.Z., Bae, S.C., Choi, K.Y.*, et al.* (2005). Dlx5 specifically regulates Runx2 type II expression by binding to homeodomain-response elements in the Runx2 distal promoter. J Biol Chem *280*, 35579-35587.

Lee, N.K., and Karsenty, G. (2008). Reciprocal regulation of bone and energy metabolism. Trends Endocrinol Metab *19*, 161-166.

Lefebvre, V., and Bhattaram, P. (2010). Vertebrate skeletogenesis. Curr Top Dev Biol *90*, 291-317.

Lefebvre, V., Li, P., and de Crombrugghe, B. (1998). A new long form of Sox5 (L-Sox5), Sox6 and Sox9 are coexpressed in chondrogenesis and cooperatively activate the type II collagen gene. EMBO J *17*, 5718-5733.

Leiden, J.M., and Thompson, C.B. (1994). Transcriptional regulation of T-cell genes during T-cell development. Curr Opin Immunol *6*, 231-237.

Lengner, C.J., Hassan, M.Q., Serra, R.W., Lepper, C., van Wijnen, A.J., Stein, J.L., Lian, J.B., and Stein, G.S. (2005). Nkx3.2-mediated repression of Runx2 promotes chondrogenic differentiation. J Biol Chem *280*, 15872-15879.

Lettice, L.A., Heaney, S.J., Purdie, L.A., Li, L., de Beer, P., Oostra, B.A., Goode, D., Elgar, G., Hill, R.E., and de Graaff, E. (2003). A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. Hum Mol Genet *12*, 1725-1735.

Levanon, D., Bernstein, Y., Negreanu, V., Ghozi, M.C., Bar-Am, I., Aloya, R., Goldenberg, D., Lotem, J., and Groner, Y. (1996). A large variety of alternatively spliced and differentially expressed mRNAs are encoded by the human acute myeloid leukemia gene AML1. DNA Cell Biol *15*, 175-185.

Levanon, D., Bettoun, D., Harris-Cerruti, C., Woolf, E., Negreanu, V., Eilam, R., Bernstein, Y., Goldenberg, D., Xiao, C., Fliegauf, M.*, et al.* (2002). The Runx3 transcription factor regulates development and survival of TrkC dorsal root ganglia neurons. EMBO J *21*, 3454-3463.

Levanon, D., Brenner, O., Negreanu, V., Bettoun, D., Woolf, E., Eilam, R., Lotem, J., Gat, U., Otto, F., Speck, N.*, et al.* (2001). Spatial and temporal expression pattern of Runx3 (Aml2) and Runx1 (Aml1) indicates non-redundant functions during mouse embryogenesis. Mech Dev *109*, 413-417.

Levanon, D., and Groner, Y. (2004). Structure and regulated expression of mammalian RUNX genes. Oncogene *23*, 4211-4219.

Levanon, D., Negreanu, V., Bernstein, Y., Bar-Am, I., Avivi, L., and Groner, Y. (1994). AML1, AML2, and AML3, the human members of the runt domain gene-family: cDNA structure, expression, and chromosomal localization. Genomics *23*, 425-432.

Levine, M., and Davidson, E.H. (2005). Gene regulatory networks for development. Proc Natl Acad Sci U S A *102*, 4936-4942.

Li, J., Kleeff, J., Guweidhi, A., Esposito, I., Berberat, P.O., Giese, T., Buchler, M.W., and Friess, H. (2004a). RUNX3 expression in primary and metastatic pancreatic cancer. J Clin Pathol *57*, 294-299.

Li, Q.L., Ito, K., Sakakura, C., Fukamachi, H., Inoue, K., Chi, X.Z., Lee, K.Y., Nomura, S., Lee, C.W., Han, S.B.*, et al.* (2002). Causal relationship between the loss of RUNX3 expression and gastric cancer. Cell *109*, 113-124.

Li, T.F., Dong, Y., Ionescu, A.M., Rosier, R.N., Zuscik, M.J., Schwarz, E.M., O'Keefe, R.J., and Drissi, H. (2004b). Parathyroid hormone-related peptide (PTHrP) inhibits Runx2 expression through the PKA signaling pathway. Exp Cell Res *299*, 128-136.

Licursi, M., Christian, S.L., Pongnopparat, T., and Hirasawa, K. (2011). In vitro and in vivo comparison of viral and cellular internal ribosome entry sites for bicistronic vector expression. Gene Ther *18*, 631-636.

Liu, C.J., Chang, E., Yu, J., Carlson, C.S., Prazak, L., Yu, X.P., Ding, B., Lengyel, P., and Di Cesare, P.E. (2005a). The interferon-inducible p204 protein acts as a transcriptional coactivator of Cbfa1 and enhances osteoblast differentiation. J Biol Chem *280*, 2788-2796.

Liu, H., Carlsson, L., and Grundstrom, T. (2006). Identification of an N-terminal transactivation domain of Runx1 that separates molecular function from global differentiation function. J Biol Chem *281*, 25659-25669.

Liu, J., Yang, H., Liu, W., Cao, X., and Feng, X. (2005b). Sp1 and Sp3 regulate the basal transcription of receptor activator of nuclear factor kappa B ligand gene in osteoblasts and bone marrow stromal cells. J Cell Biochem *96*, 716-727.

Liu, W., Toyosawa, S., Furuichi, T., Kanatani, N., Yoshida, C., Liu, Y., Himeno, M., Narai, S., Yamaguchi, A., and Komori, T. (2001). Overexpression of Cbfa1 in osteoblasts inhibits osteoblast maturation and causes osteopenia with multiple fractures. J Cell Biol *155*, 157-166.

Loose, M., and Patient, R. (2004). A genetic regulatory network for Xenopus mesendoderm formation. Dev Biol *271*, 467-478.

Luan, Y., Yu, X.P., Xu, K., Ding, B., Yu, J., Huang, Y., Yang, N., Lengyel, P., Di Cesare, P.E., and Liu, C.J. (2007). The retinoblastoma protein is an essential mediator of osteogenesis that links the p204 protein to the Cbfa1 transcription factor thereby increasing its activity. J Biol Chem *282*, 16860-16870.

Lui, T.T., Lacroix, C., Ahmed, S.M., Goldenberg, S.J., Leach, C.A., Daulat, A.M., and Angers, S. (2011). The ubiquitin-specific protease USP34 regulates axin stability and Wnt/beta-catenin signaling. Mol Cell Biol *31*, 2053-2065.

MacLean, H.E., Kim, J.I., Glimcher, M.J., Wang, J., Kronenberg, H.M., and Glimcher, L.H. (2003). Absence of transcription factor c-maf causes abnormal terminal differentiation of hypertrophic chondrocytes during endochondral bone development. Dev Biol *262*, 51-63.

MacQuarrie, K.L., Fong, A.P., Morse, R.H., and Tapscott, S.J. (2011). Genome-wide transcription factor binding: beyond direct target regulation. Trends Genet *27*, 141-148.

Maeno, T., Moriishi, T., Yoshida, C.A., Komori, H., Kanatani, N., Izumi, S., Takaoka, K., and Komori, T. (2011). Early onset of Runx2 expression caused craniosynostosis, ectopic bone formation, and limb defects. Bone *49*, 673-682.

Marchesini, N., Luberto, C., and Hannun, Y.A. (2003). Biochemical properties of mammalian neutral sphingomyelinase 2 and its role in sphingolipid metabolism. J Biol Chem *278*, 13775-13783.

Mardis, E.R. (2007). ChIP-seq: welcome to the new frontier. Nat Methods *4*, 613-614.

Martin, J.W., Zielenska, M., Stein, G.S., van Wijnen, A.J., and Squire, J.A. (2011). The Role of RUNX2 in Osteosarcoma Oncogenesis. Sarcoma *2011*, 282745.

Maruyama, Z., Yoshida, C.A., Furuichi, T., Amizuka, N., Ito, M., Fukuyama, R., Miyazaki, T., Kitaura, H., Nakamura, K., Fujita, T.*, et al.* (2007). 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.

Masters, B.A., Quaife, C.J., Erickson, J.C., Kelly, E.J., Froelick, G.J., Zambrowicz, B.P., Brinster, R.L., and Palmiter, R.D. (1994). Metallothionein III is expressed in neurons that sequester zinc in synaptic vesicles. J Neurosci *14*, 5844-5857.

Matsubara, T., Kida, K., Yamaguchi, A., Hata, K., Ichida, F., Meguro, H., Aburatani, H., Nishimura, R., and Yoneda, T. (2008). BMP2 regulates Osterix through Msx2 and Runx2 during osteoblast differentiation. J Biol Chem *283*, 29119-29125.

McCarthy, T.L., Ji, C., Chen, Y., Kim, K.K., Imagawa, M., Ito, Y., and Centrella, M. (2000). Runt domain factor (Runx)-dependent effects on CCAAT/ enhancer-binding protein delta expression and activity in osteoblasts. J Biol Chem *275*, 21746-21753.

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol *28*, 495-501.

Meech, R., Edelman, D.B., Jones, F.S., and Makarenkova, H.P. (2005). The homeobox transcription factor Barx2 regulates chondrogenesis during limb development. Development *132*, 2135-2146.

Mikasa, M., Rokutanda, S., Komori, H., Ito, K., Tsang, Y.S., Date, Y., Yoshida, C.A., and Komori, T. (2011). Regulation of Tcf7 by Runx2 in chondrocyte maturation and proliferation. J Bone Miner Metab *29*, 291-299.

Miyoshi, H., Ohira, M., Shimizu, K., Mitani, K., Hirai, H., Imai, T., Yokoyama, K., Soeda, E., and Ohki, M. (1995). Alternative splicing and genomic structure of the AML1 gene involved in acute myeloid leukemia. Nucleic Acids Res *23*, 2762-2769.

Moffatt, P., Gaumond, M.H., Salois, P., Sellin, K., Bessette, M.C., Godin, E., de Oliveira, P.T., Atkins, G.J., Nanci, A., and Thomas, G. (2008). Bril: a novel bone-specific modulator of mineralization. J Bone Miner Res *23*, 1497-1508.

Mundlos, S., Otto, F., Mundlos, C., Mulliken, J.B., Aylsworth, A.S., Albright, S., Lindhout, D., Cole, W.G., Henn, W., Knoll, J.H.*, et al.* (1997). Mutations involving the transcription factor CBFA1 cause cleidocranial dysplasia. Cell *89*, 773-779.

Munro, S., and Pelham, H.R. (1987). A C-terminal signal prevents secretion of luminal ER proteins. Cell *48*, 899-907.

Nakashima, K., Zhou, X., Kunkel, G., Zhang, Z., Deng, J.M., Behringer, R.R., and de Crombrugghe, B. (2002). The novel zinc finger-containing transcription factor osterix is required for osteoblast differentiation and bone formation. Cell *108*, 17-29.

Nordberg, M. (1998). Metallothioneins: historical review and state of knowledge. Talanta *46*, 243-254.

O'Riordan, M., and Grosschedl, R. (2000). Transcriptional regulation of early B-lymphocyte differentiation. Immunol Rev *175*, 94-103.

Ogawa, S., Harada, H., Fujiwara, M., Tagashira, S., Katsumata, T., and Takada, H. (2000). Cbfa1, an essential transcription factor for bone formation, is expressed in testis from the same promoter used in bone. DNA Res *7*, 181-185.

Ohba, S., Kawaguchi, H., Kugimiya, F., Ogasawara, T., Kawamura, N., Saito, T., Ikeda, T., Fujii, K., Miyajima, T., Kuramochi, A.*, et al.* (2008). Patched1 haploinsufficiency increases adult bone mass and modulates Gli3 repressor activity. Dev Cell *14*, 689-699.

Okuda, T., van Deursen, J., Hiebert, S.W., Grosveld, G., and Downing, J.R. (1996). AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. Cell *84*, 321-330.

Otto, F., Lubbert, M., and Stock, M. (2003). Upstream and downstream targets of RUNX proteins. J Cell Biochem *89*, 9-18.

Otto, F., Thornell, A.P., Crompton, T., Denzel, A., Gilmour, K.C., Rosewell, I.R., Stamp, G.W., Beddington, R.S., Mundlos, S., Olsen, B.R.*, et al.* (1997). Cbfa1, a candidate gene for cleidocranial dysplasia syndrome, is essential for osteoblast differentiation and bone development. Cell *89*, 765-771.

Palmiter, R.D., Findley, S.D., Whitmore, T.E., and Durnam, D.M. (1992). MT-III, a brain-specific member of the metallothionein gene family. Proc Natl Acad Sci U S A *89*, 6333-6337.

Park, M.H., Shin, H.I., Choi, J.Y., Nam, S.H., Kim, Y.J., Kim, H.J., and Ryoo, H.M. (2001). Differential expression patterns of Runx2 isoforms in cranial suture morphogenesis. J Bone Miner Res *16*, 885-892.

Peters, H., Wilm, B., Sakai, N., Imai, K., Maas, R., and Balling, R. (1999). Pax1 and Pax9 synergistically regulate vertebral column development. Development *126*, 5399-5408.

Porte, D., Tuckermann, J., Becker, M., Baumann, B., Teurich, S., Higgins, T., Owen, M.J., Schorpp-Kistner, M., and Angel, P. (1999). Both AP-1 and Cbfa1-like factors are required for the induction of interstitial collagenase by parathyroid hormone. Oncogene *18*, 667-678.

Pratap, J., Lian, J.B., and Stein, G.S. (2011). Metastatic bone disease: role of transcription factors and future targets. Bone *48*, 30-36.

Ptashne, M. (1986). Gene regulation by proteins acting nearby and at a distance. Nature *322*, 697-701.

Reinhold, M.I., and Naski, M.C. (2007). Direct interactions of Runx2 and canonical Wnt signaling induce FGF18. J Biol Chem *282*, 3653-3663.

Ririe, T.O., Fernandes, J.S., and Sternberg, P.W. (2008). The Caenorhabditis elegans vulva: a post-embryonic gene regulatory network controlling organogenesis. Proc Natl Acad Sci U S A *105*, 20095-20099.

Robertson, G., Hirst, M., Bainbridge, M., Bilenky, M., Zhao, Y., Zeng, T., Euskirchen, G., Bernier, B., Varhol, R., Delaney, A.*, et al.* (2007). Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. Nat Methods *4*, 651-657.

Roca, H., Phimphilai, M., Gopalakrishnan, R., Xiao, G., and Franceschi, R.T. (2005). Cooperative interactions between RUNX2 and homeodomain protein-binding sites are critical for the osteoblast-specific expression of the bone sialoprotein gene. J Biol Chem *280*, 30845-30855.

Rodrigo, I., Hill, R.E., Balling, R., Munsterberg, A., and Imai, K. (2003). Pax1 and Pax9 activate Bapx1 to induce chondrogenic differentiation in the sclerotome. Development *130*, 473-482.

Ruiz-Perez, V.L., Blair, H.J., Rodriguez-Andres, M.E., Blanco, M.J., Wilson, A., Liu, Y.N., Miles, C., Peters, H., and Goodship, J.A. (2007). Evc is a positive mediator of Ihh-regulated bone growth that localises at the base of chondrocyte cilia. Development *134*, 2903-2912.

Saito, T., Ikeda, T., Nakamura, K., Chung, U.I., and Kawaguchi, H. (2007). S100A1 and S100B, transcriptional targets of SOX trio, inhibit terminal differentiation of chondrocytes. EMBO Rep *8*, 504-509.

Salsi, V., Vigano, M.A., Cocchiarella, F., Mantovani, R., and Zappavigna, V. (2008). Hoxd13 binds in vivo and regulates the expression of genes acting in key pathways for early limb and skeletal patterning. Dev Biol *317*, 497-507.

Sato, M., Morii, E., Komori, T., Kawahata, H., Sugimoto, M., Terai, K., Shimizu, H., Yasui, T., Ogihara, H., Yasui, N.*, et al.* (1998). Transcriptional regulation of osteopontin gene in vivo by PEBP2alphaA/CBFA1 and ETS1 in the skeletal tissues. Oncogene *17*, 1517-1525.

Schipani, E., Ryan, H.E., Didrickson, S., Kobayashi, T., Knight, M., and Johnson, R.S. (2001). Hypoxia in cartilage: HIF-1alpha is essential for chondrocyte growth arrest and survival. Genes Dev *15*, 2865-2876.

Schmidt, K., Schinke, T., Haberland, M., Priemel, M., Schilling, A.F., Mueldner, C., Rueger, J.M., Sock, E., Wegner, M., and Amling, M. (2005). The high mobility group transcription factor Sox8 is a negative regulator of osteoblast differentiation. J Cell Biol *168*, 899-910.

Schoor, M., Schuster-Gossler, K., Roopenian, D., and Gossler, A. (1999). Skeletal dysplasias, growth retardation, reduced postnatal survival, and impaired fertility in mice lacking the SNF2/SWI2 family member ETL1. Mech Dev *85*, 73-83.

Schroeder, A., Mueller, O., Stocker, S., Salowsky, R., Leiber, M., Gassmann, M., Lightfoot, S., Menzel, W., Granzow, M., and Ragg, T. (2006). The RIN: an RNA integrity number for assigning integrity values to RNA measurements. BMC Mol Biol *7*, 3.

Selvamurugan, N., Pulumati, M.R., Tyson, D.R., and Partridge, N.C. (2000). Parathyroid hormone regulation of the rat collagenase-3 promoter by protein kinase A-dependent transactivation of core binding factor alpha1. J Biol Chem *275*, 5037-5042.

Shimoyama, A., Wada, M., Ikeda, F., Hata, K., Matsubara, T., Nifuji, A., Noda, M., Amano, K., Yamaguchi, A., Nishimura, R.*, et al.* (2007). Ihh/Gli2 signaling promotes osteoblast differentiation by regulating Runx2 expression and function. Mol Biol Cell *18*, 2411-2418.

Shirakabe, K., Terasawa, K., Miyama, K., Shibuya, H., and Nishida, E. (2001). Regulation of the activity of the transcription factor Runx2 by two homeobox proteins, Msx2 and Dlx5. Genes Cells *6*, 851-856.

Silver, D.L., Geisbrecht, E.R., and Montell, D.J. (2005). Requirement for JAK/STAT signaling throughout border cell migration in Drosophila. Development *132*, 3483-3492.

Singh, M.K., Petry, M., Haenig, B., Lescher, B., Leitges, M., and Kispert, A. (2005). The T-box transcription factor Tbx15 is required for skeletal development. Mech Dev *122*, 131-144.

Smits, P., Dy, P., Mitra, S., and Lefebvre, V. (2004). Sox5 and Sox6 are needed to develop and maintain source, columnar, and hypertrophic chondrocytes in the cartilage growth plate. J Cell Biol *164*, 747-758.

Smits, P., Li, P., Mandel, J., Zhang, Z., Deng, J.M., Behringer, R.R., de Crombrugghe, B., and Lefebvre, V. (2001). The transcription factors L-Sox5 and Sox6 are essential for cartilage formation. Dev Cell *1*, 277-290.

Sohaskey, M.L., Jiang, Y., Zhao, J.J., Mohr, A., Roemer, F., and Harland, R.M. (2010). Osteopotentia regulates osteoblast maturation, bone formation, and skeletal integrity in mice. J Cell Biol *189*, 511-525.

Song, W.J., Sullivan, M.G., Legare, R.D., Hutchings, S., Tan, X., Kufrin, D., Ratajczak, J., Resende, I.C., Haworth, C., Hock, R.*, et al.* (1999). Haploinsufficiency of CBFA2 causes familial thrombocytopenia with propensity to develop acute myelogenous leukaemia. Nat Genet *23*, 166-175.

Soung do, Y., Dong, Y., Wang, Y., Zuscik, M.J., Schwarz, E.M., O'Keefe, R.J., and Drissi, H. (2007). Runx3/AML2/Cbfa3 regulates early and late chondrocyte differentiation. J Bone Miner Res *22*, 1260-1270.

St-Jacques, B., Hammerschmidt, M., and McMahon, A.P. (1999). Indian hedgehog signaling regulates proliferation and differentiation of chondrocytes and is essential for bone formation. Genes Dev *13*, 2072-2086.

Stathopoulos, A., and Levine, M. (2005). Localized repressors delineate the neurogenic ectoderm in the early Drosophila embryo. Dev Biol *280*, 482-493.

Stein, G.S., Lian, J.B., van Wijnen, A.J., Stein, J.L., Montecino, M., Javed, A., Zaidi, S.K., Young, D.W., Choi, J.Y., and Pockwinse, S.M. (2004). Runx2 control of organization, assembly and activity of the regulatory machinery for skeletal gene expression. Oncogene *23*, 4315-4329.

Stewart, M., Terry, A., Hu, M., O'Hara, M., Blyth, K., Baxter, E., Cameron, E., Onions, D.E., and Neil, J.C. (1997). Proviral insertions induce the expression of bone-specific isoforms of PEBP2alphaA (CBFA1): evidence for a new myc collaborating oncogene. Proc Natl Acad Sci U S A *94*, 8646-8651.

Stickens, D., Behonick, D.J., Ortega, N., Heyer, B., Hartenstein, B., Yu, Y., Fosang, A.J., Schorpp-Kistner, M., Angel, P., and Werb, Z. (2004). Altered endochondral bone development in matrix metalloproteinase 13-deficient mice. Development *131*, 5883-5895.

Stoffel, W., Jenke, B., Holz, B., Binczek, E., Gunter, R.H., Knifka, J., Koebke, J., and Niehoff, A. (2007). Neutral sphingomyelinase (SMPD3) deficiency causes a novel form of chondrodysplasia and dwarfism that is rescued by Col2A1-driven smpd3 transgene expression. Am J Pathol *171*, 153-161.

Stricker, S., Fundele, R., Vortkamp, A., and Mundlos, S. (2002). Role of Runx genes in chondrocyte differentiation. Dev Biol *245*, 95-108.

Takeda, S., Bonnamy, J.P., Owen, M.J., Ducy, P., and Karsenty, G. (2001). Continuous expression of Cbfa1 in nonhypertrophic chondrocytes uncovers its ability to induce hypertrophic chondrocyte differentiation and partially rescues Cbfa1-deficient mice. Genes Dev *15*, 467-481.

Tang, F., Barbacioru, C., Nordman, E., Li, B., Xu, N., Bashkirov, V.I., Lao, K., and Surani, M.A. (2010). RNA-Seq analysis to capture the transcriptome landscape of a single cell. Nat Protoc *5*, 516-535.

Taniuchi, I., Osato, M., Egawa, T., Sunshine, M.J., Bae, S.C., Komori, T., Ito, Y., and Littman, D.R. (2002). Differential requirements for Runx proteins in CD4 repression and epigenetic silencing during T lymphocyte development. Cell *111*, 621-633.

Thirunavukkarasu, K., Mahajan, M., McLarren, K.W., Stifani, S., and Karsenty, G. (1998). Two domains unique to osteoblast-specific transcription factor Osf2/Cbfa1 contribute to its transactivation function and its inability to heterodimerize with Cbfbeta. Mol Cell Biol *18*, 4197-4208.

Thomas, D.M., Carty, S.A., Piscopo, D.M., Lee, J.S., Wang, W.F., Forrester, W.C., and Hinds, P.W. (2001). The retinoblastoma protein acts as a transcriptional coactivator required for osteogenic differentiation. Mol Cell *8*, 303-316.

Tohmonda, T., Miyauchi, Y., Ghosh, R., Yoda, M., Uchikawa, S., Takito, J., Morioka, H., Nakamura, M., Iwawaki, T., Chiba, K., *et al.* (2011). The IRE1alpha-XBP1 pathway is essential for osteoblast differentiation through promoting transcription of Osterix. EMBO Rep *12*, 451-457.

Tribioli, C., and Lufkin, T. (1999). The murine Bapx1 homeobox gene plays a critical role in embryonic development of the axial skeleton and spleen. Development *126*, 5699-5711.

Ueta, C., Iwamoto, M., Kanatani, N., Yoshida, C., Liu, Y., Enomoto-Iwamoto, M., Ohmori, T., Enomoto, H., Nakata, K., Takada, K., *et al.* (2001). Skeletal malformations caused by overexpression of Cbfa1 or its dominant negative form in chondrocytes. J Cell Biol *153*, 87-100.

Vaes, B.L., Ducy, P., Sijbers, A.M., Hendriks, J.M., van Someren, E.P., de Jong, N.G., van den Heuvel, E.R., Olijve, W., van Zoelen, E.J., and Dechering, K.J. (2006). Microarray analysis on Runx2-deficient mouse embryos reveals novel Runx2 functions and target genes during intramembranous and endochondral bone formation. Bone *39*, 724-738.

van Wijnen, A.J., Stein, G.S., Gergen, J.P., Groner, Y., Hiebert, S.W., Ito, Y., Liu, P., Neil, J.C., Ohki, M., and Speck, N. (2004). Nomenclature for Runt-related (RUNX) proteins. Oncogene *23*, 4209-4210.

Vega, R.B., Matsuda, K., Oh, J., Barbosa, A.C., Yang, X., Meadows, E., McAnally, J., Pomajzl, C., Shelton, J.M., Richardson, J.A., *et al.* (2004). Histone deacetylase 4 controls chondrocyte hypertrophy during skeletogenesis. Cell *119*, 555-566.

Vortkamp, A., Lee, K., Lanske, B., Segre, G.V., Kronenberg, H.M., and Tabin, C.J. (1996). Regulation of rate of cartilage differentiation by Indian hedgehog and PTH-related protein. Science *273*, 613-622.

Vu, T.H., Shipley, J.M., Bergers, G., Berger, J.E., Helms, J.A., Hanahan, D., Shapiro, S.D., Senior, R.M., and Werb, Z. (1998). MMP-9/gelatinase B is a key regulator of growth plate angiogenesis and apoptosis of hypertrophic chondrocytes. Cell *93*, 411-422.

Wada, M., Yazumi, S., Takaishi, S., Hasegawa, K., Sawada, M., Tanaka, H., Ida, H., Sakakura, C., Ito, K., Ito, Y., *et al.* (2004). Frequent loss of RUNX3 gene expression in human bile duct and pancreatic cancer cell lines. Oncogene *23*, 2401-2407.

Wagner, T., Wirth, J., Meyer, J., Zabel, B., Held, M., Zimmer, J., Pasantes, J., Bricarelli, F.D., Keutel, J., Hustert, E., *et al.* (1994). Autosomal sex reversal and campomelic dysplasia are caused by mutations in and around the SRY-related gene SOX9. Cell *79*, 1111-1120.

Wang, W., Wang, Y.G., Reginato, A.M., Glotzer, D.J., Fukai, N., Plotkina, S., Karsenty, G., and Olsen, B.R. (2004). Groucho homologue Grg5 interacts with the transcription factor Runx2-Cbfa1 and modulates its activity during postnatal growth in mice. Dev Biol *270*, 364-381.

Westendorf, J.J. (2006). Transcriptional co-repressors of Runx2. J Cell Biochem *98*, 54-64.

Wilczynski, B., and Furlong, E.E. (2010). Dynamic CRM occupancy reflects a temporal map of developmental progression. Mol Syst Biol *6*, 383.

Woolf, E., Xiao, C., Fainaru, O., Lotem, J., Rosen, D., Negreanu, V., Bernstein, Y., Goldenberg, D., Brenner, O., Berke, G.*, et al.* (2003). Runx3 and Runx1 are required for CD8 T cell development during thymopoiesis. Proc Natl Acad Sci U S A *100*, 7731-7736.

Wright, E., Hargrave, M.R., Christiansen, J., Cooper, L., Kun, J., Evans, T., Gangadharan, U., Greenfield, A., and Koopman, P. (1995). The Sry-related gene Sox9 is expressed during chondrogenesis in mouse embryos. Nat Genet *9*, 15-20.

Wu, B., Shao, Y., Chen, B., Liu, C., Xue, Z., Wu, P., and Li, H. (2010). Identification of a novel mouse brachyury (T) allele causing a short tail mutation in mice. Cell Biochem Biophys *58*, 129-135.

Xiao, G., Jiang, D., Ge, C., Zhao, Z., Lai, Y., Boules, H., Phimphilai, M., Yang, X., Karsenty, G., and Franceschi, R.T. (2005). Cooperative interactions between activating transcription factor 4 and Runx2/Cbfa1 stimulate osteoblast-specific osteocalcin gene expression. J Biol Chem *280*, 30689-30696.

Xiao, W.H., and Liu, W.W. (2004). Hemizygous deletion and hypermethylation of RUNX3 gene in hepatocellular carcinoma. World J Gastroenterol *10*, 376-380.

Xiao, Z., Zhang, S., Magenheimer, B.S., Luo, J., and Quarles, L.D. (2008). Polycystin-1 regulates skeletogenesis through stimulation of the osteoblast-specific transcription factor RUNX2-II. J Biol Chem *283*, 12624-12634.

Xiao, Z.S., Hinson, T.K., and Quarles, L.D. (1999). Cbfa1 isoform overexpression upregulates osteocalcin gene expression in non-osteoblastic and pre-osteoblastic cells. J Cell Biochem *74*, 596-605.

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

Xiao, Z.S., Thomas, R., Hinson, T.K., and Quarles, L.D. (1998). Genomic structure and isoform expression of the mouse, rat and human Cbfa1/Osf2 transcription factor. Gene *214*, 187-197.

Yamashita, S., Andoh, M., Ueno-Kudoh, H., Sato, T., Miyaki, S., and Asahara, H. (2009). Sox9 directly promotes Bapx1 gene expression to repress Runx2 in chondrocytes. Exp Cell Res *315*, 2231-2240.

Yang, X., and Karsenty, G. (2004). ATF4, the osteoblast accumulation of which is determined post-translationally, can induce osteoblast-specific gene expression in non-osteoblastic cells. J Biol Chem *279*, 47109-47114.

Yang, X., Matsuda, K., Bialek, P., Jacquot, S., Masuoka, H.C., Schinke, T., Li, L., Brancorsini, S., Sassone-Corsi, P., Townes, T.M.*, et al.* (2004). ATF4 is a substrate of RSK2 and an essential regulator of osteoblast biology; implication for Coffin-Lowry Syndrome. Cell *117*, 387-398.

Yano, T., Ito, K., Fukamachi, H., Chi, X.Z., Wee, H.J., Inoue, K., Ida, H., Bouillet, P., Strasser, A., Bae, S.C.*, et al.* (2006). The RUNX3 tumor suppressor upregulates Bim in gastric epithelial cells undergoing transforming growth factor beta-induced apoptosis. Mol Cell Biol *26*, 4474-4488.

Yoshida, C., Tokumasu, F., Hohmura, K.I., Bungert, J., Hayashi, N., Nagasawa, T., Engel, J.D., Yamamoto, M., Takeyasu, K., and Igarashi, K. (1999). Long range interaction of cis-DNA elements mediated by architectural transcription factor Bach1. Genes Cells *4*, 643-655.

Yoshida, C.A., Yamamoto, H., Fujita, T., Furuichi, T., Ito, K., Inoue, K., Yamana, K., Zanma, A., Takada, K., Ito, Y.*, et al.* (2004). Runx2 and Runx3 are essential for chondrocyte maturation, and Runx2 regulates limb growth through induction of Indian hedgehog. Genes Dev *18*, 952-963.

Yu, V.W., Akhouayri, O., and St-Arnaud, R. (2009). FIAT is co-expressed with its dimerization target ATF4 in early osteoblasts, but not in osteocytes. Gene Expr Patterns *9*, 335-340.

Zelzer, E., Mamluk, R., Ferrara, N., Johnson, R.S., Schipani, E., and Olsen, B.R. (2004). VEGFA is necessary for chondrocyte survival during bone development. Development *131*, 2161-2171.

Zhang, Y., Kong, L., Carlson, C.S., and Liu, C.J. (2008a). Cbfa1-dependent expression of an interferon-inducible p204 protein is required for chondrocyte differentiation. Cell Death Differ *15*, 1760-1771.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W.*, et al.* (2008b). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.

Zhang, Y.W., Yasui, N., Ito, K., Huang, G., Fujii, M., Hanai, J., Nogami, H., Ochi, T., Miyazono, K., and Ito, Y. (2000). A RUNX2/PEBP2alpha A/CBFA1 mutation displaying impaired transactivation and Smad interaction in cleidocranial dysplasia. Proc Natl Acad Sci U S A *97*, 10549-10554.

Zheng, Q., Zhou, G., Morello, R., Chen, Y., Garcia-Rojas, X., and Lee, B. (2003). Type X collagen gene regulation by Runx2 contributes directly to its hypertrophic chondrocyte-specific expression in vivo. J Cell Biol *162*, 833-842.

Zheng, Y.J., Chung, H.J., Min, H., Kang, M., Kim, S.H., Gadi, J., and Kim, M.H. (2009). In vitro osteoblast differentiation is negatively regulated by Hoxc8. Appl Biochem Biotechnol *160*, 891-900.

Zhong, J., Pevny, L., and Snider, W.D. (2006). "Runx"ing towards sensory differentiation. Neuron *49*, 325-327.