THE REGULATORY NETWORK OF PITX1 DURING MOUSE LIMB AND CRANIOFACIAL DEVELOPMENT

by

JIALIANG WANG

(Under the Direction of Douglas B. Menke)

ABSTRACT

In multicellular organisms, cell fates are established through precise patterns of gene expression. This requires a complex set of interactions among *cis*-regulatory elements and transcription factors. Enhancers, in particular, play a central role in driving cell-type specific gene expression patterns and are capable of activating transcriptional target genes. Both the limb and the mandible are classic model systems for the study of developmental regulation. In mouse, the hindlimb and mandible share similar developmental processes, including specification of the prospective field, induction of tissue outgrowth, maintenance of the outgrowth and patterning tissue along different axes. Many major signaling pathways contribute to the development of both organs, including FGF, SHH, BMP and WNT. Beyond signaling pathways, there is a relatively narrow network of hindlimb-restricted or mandible-restricted regulators, such as PITX1, TBX4, ISL1, DLX5/6 and HAND1/2. PITX1 is a tissue-specific transcription factor, which is encoded by the paired-like homeodomain 1. Strong *Pitx1* expression is found in both the hindlimb and mandible mesenchyme. Haploinsufficiency for *Pitx1* in mice and humans can result in clubfoot, mandibular hypoplasia and other hindlimb- or jaw-related malformations. However, the hindlimb and mandible PITX1-transcriptional targets that are misregulated in the

absence of PITX1 remain largely unknown. In this work, I have applied a combination of comparative genomics, epigenetic profiles and functional analyses to study the regulatory network of PITX1 transcriptional targets that direct mouse hindlimb and mandible development. I performed a genome-wide ChIP-seq analysis to identify the location of *cis*-regulatory elements bound by PITX1 during embryonic development of the mouse hindlimb and mandible. The transcriptional targets of PITX1 in developing hindlimbs and mandibles are revealed by performing RNA-seq. PITX1-dependent enhancers and target genes of hindlimb and mandible are defined with the combination of sequence conservation, ChIP-seq, RNA-seq and functional analyses. The results support the conclusions that PITX1 can promote chondrogenic and myogenic differentiation in mouse hindlimb through conserved regulatory targets, and modulate both hindlimb and mandible development through the shared regulatory network.

INDEX WORDS: PITX1, Hindlimb, Mandible, Chromatin Immunoprecipitation sequencing (ChIP-seq), RNA sequencing (RNA-seq), Chondrogenesis, Myogenesis

THE REGULATORY NETWORK OF PITX1 DURING MOUSE LIMB AND CRANIOFACIAL DEVELOPMENT

by

JIALIANG WANG

B.S., China Agricultural University, China, 2012

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2017

THE REGULATORY NETWORK OF PITX1 DURING MOUSE LIMB AND CRANIOFACIAL DEVELOPMENT

by

JIALIANG WANG

Major Professor:	Douglas B. Menke
Committee:	Brian G. Condie James D. Lauderdale Jonathan T. Eggenschwiler Jianfu Chen

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
December 2017

DEDICATION

I would like to dedicate this work to my loving parents Yanping Wang and Baijing Liang. They are the first people who believe in me and inspire me to pursue my dream in academia. They never lose faith in me and love me throughout my whole life. Specially, I would like to thank my husband Jundi Liu. He is always there during my PhD life and so supportive of every decision I have made at turning points of my life. I wouldn't be the same without him standing by my side.

ACKNOWLEDGEMENTS

I would like to thank my mentor and major professor Dr. Douglas B. Menke for all of the guidance and attentions that he has provided throughout my entire graduate life. I am greatly inspired by his passion and commitment to science. He is an amazing mentor who balances so well between guidance and independence during my training. I can't remember how many times that I knocked on Doug's door when I got stuck. He would always set aside what he was working on and discuss the experimental design with me. He not only encourages communication between us to convey very clear ideas but also supports me to think and work independently. Without his patience and encouragement, I would not have made thus far and made up my mind to become a scientist.

I would like to express my sincere gratitude for the guidance from all my committee members, Dr. Brian G. Condie, Dr. Jonathan T. Eggenswhiler, Dr. James D. Lauderdale and Dr. Jianfu Chen. They have given me critical and constructive advice for my projects and have witnessed every progress that I have made. They have provided valuable ideas for my dissertation research and genuine career advices along the way.

I would like to thank current and previous members of the Menke lab, including Sungdae Park, Shana Pau, Ashley Rasys, Sergio Minchey, Aaron Alcala, Carlos Infante and Alexandra Mihala. You all have made my PhD life filled with happiness and laughter. Especially Sungdae, he is always available to troubleshoot with me and help me out with my experiments.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Limb Development and Major Genetic Mechanisms

Vertebrate limbs, including forelimbs and hindlimbs, are paired appendages that rise from lateral plate mesoderm (LPM) at fixed positions along the rostrocaudal axis of the body (Fig. 1.5). The genes of the *Hox* clusters have restricted expression patterns along the LPM and are thought to specify the level of which limb buds will appear (Cohn et al., 1996; Cohn and Tickle, 1999). *HoxC4* and *HoxC5* are expressed in the presumptive forelimb region and *HoxC9, 10,* and *11* are expressed in the presumptive hindlimb area (Duboc and Logan, 2011).

Fgf10 serves as an early marker in the prospective limb fields, where it exhibits restricted expression in the lateral plate mesoderm (Ohuchi et al., 1997). *Fgf10* expression acts to induce *Fgf8* expression in the apical ectodermal ridge (AER) (Fig. 1.1), a specialized epithelial surface along the dorsoventral (DV) boundary (Niswander, 2003). *Fgf8* expression in the AER, in return, maintains *Fgf10* expression in the mesenchyme of the growing limb. The mesenchymal-epithelial feedback loop plays an important role in limb outgrowth (Saunders, 1948; Sekine et al., 1999). In addition to controlling proper limb outgrowth, AER-FGF is required for proximodistal (PD) patterning in the limb (Niswander et al., 1993). Retinoid acid (RA), which is produced in the proximal mesenchyme, and AER-FGF have opposing activities. RA signal specifies the proximal fates and AER-FGF signal specifies the distal fates in the PD axis at early developmental stages (Mercader et al., 2000). AER-FGF keeps distal mesenchymal cells in an undifferentiated state. Once proximal cells are no longer under the exposure of AER-FGF, they

are determined and fated to give rise to proximal structures. This is the "differentiation-front model" in the PD patterning (Tabin and Wolpert, 2007).

The zone of polarizing activity (ZPA) is an organizer located in the posterior limb mesenchyme. The ZPA patterns the anteroposterior (AP) axis of the limb by producing sonic hedgehog (SHH) (Riddle et al., 1993). The determination of AP patterning has been proposed as the "spatial and temporal gradient model" (Harfe et al., 2004). In mouse, digit 2 and anterior digit 3 are specified by long-range SHH signaling. Posterior digit 3, digit 4, digit 5 and the ulna (or fibula in the hindlimb) are derivatives of mesenchymal cells that previously expressed *Shh*. The AP patterning of the humerus (femur), radius (tibia) and digit 1 are SHH-independent. AER-FGF and ZPA-SHH are reciprocal. Expression of *Shh* is maintained by FGFs in the AER, and SHH, the other way round, feeds back to maintain AER-FGF signal (Sun et al., 2002).

The WNT signaling pathway is necessary for the dorsoventral patterning (Kengaku et al., 1998). The canonical WNT pathway (*Wnt3, β-catenin*) acts to establish ventral limb identity by acting upstream of *Engrailed 1* (*En1*) (Barrow et al., 2003). *En1* is essential for the specification of the ventral limb (Loomis et al., 1996). Another WNT ligand, *Wnt7a,* is necessary for the specification of dorsal limb through the β-catenin independent pathway (Logan et al., 1997). The activity of *Wnt7a* is restricted in the dorsal ectoderm by the presence of *En1*.

Mandible Development and Major Genetic Mechanisms

The head ectoderm is subdivided into neural ectoderm and non-neural ectoderm. At the boundary of neural and non-neural ectoderm, a small subset of epithelial cells adopts a mesenchymal character and gives rise to neural crest cells (NCCs) (Trainer et al., 2003). Several signaling centers are involved in and are crucial to this process, such as BMP, WNT signals and

Sox transcription factors (Szabo-Rogers et al., 2010). After the formation of the neural tube, NCCs lie sandwiched between neural ectoderm and facial ectoderm.

Cranial NCCs follow stereotypical migratory pathways that are conserved among vertebrates. The hindbrain is segregated into rhombomeres (r). NCCs derived from posterior mesencephalon, r1 and r2 fill in the 1st pharyngeal arch (PA1) (Kimmel et al, 2001). PA1 will finally develop into maxillary and mandibular in mammals. The homeodomain transcription factors of the *Hox* gene family provide a nested and combinatorial expression along the AP axis of PAs (Minoux and Rijli, 2010). *Hox* gene expression controls the NCC pre-migrating identity along the AP axis (Krumlauf, 1993). PA1 is devoid of *Hox* gene expression (Prince and Lumsden, 1994). *Hoxa2* is expressed in the 2nd pharyngeal arch (PA2). When *Hoxa2* is inactivated in PA2, it will result in the homeotic transformation of PA2 to PA1-like elements (Rijli et al., 1993). The *Hox*-free ground pattern is the "default" state and is shared in all pharyngeal arches. It is concluded that the distal PA1 structures are specified from anterior to posterior in a *Hox*-independent pattern system. The patterns in other PAs are modified by expression of *Hox* genes. In PA2, *Hoxa2* can repress several downstream genes, such as *Pitx1, Lhx6, Alx4* and *Bapx1*. These genes are normally expressed in PA1 (Bobola et al., 2003; Kanzler et al., 1998; Minoux et al., 2009; Santagati et al., 2005).

The mouse has six *Dlx* genes (*Dlx1/2, Dlx3/4, Dlx5/6*). They exhibit nested expression patterns. In PA1, *Dlx1/2* is expressed in maxillary and mandibular regions. *Dlx5/6* is only restricted to the distal region of mandible (Stock et al., 1996). Two expression domains (*Dlx1/2* and *Dlx1/2/5/6*) actually extend the boundary to the hinge of the upper jaw and the lower jaw. Mouse *Dlx1/2* $^{-/-}$ double mutants present with malformations of the upper jaw and upper components of the hinge (Depew et al., 2005). *Dlx5/6* $^{-/-}$ double mutants have homeotic

transformations of the lower jaw to the upper jaw (Depew et al., 2002). These phenotypes led to the conclusion that *Dlx* code specifies spatial identity along the DV axis in PA1. *Dlx1/2* is the default and *Dlx5/6* is the selector for the mandibular process (Jeong et al., 2008). There are about 20 genes downstream of *Dlx5/6* in the mandibular region, and they are classified into three groups. *Dlx5/6* induce and maintain the expressions of Group A genes (*Dlx3/4, Hand1/2, Gbx2, Gsc, Alx3/4* and *Pitx1*) in the mandibular region, while repressing Group B (*Pou3f3*) and Group C (*Foxl2*) genes so that their expressions are confined to the maxillary region (Funato et al., 2016). *Dlx1/2,* on the other hand, induce and maintain Group B genes in the maxillary region. There are currently two models related to the *Dlx* code that support the DV patterning of mammalian jaw. The qualitative hypothesis supports the unique activities of DLX1/2 versus DLX5/6 proteins in specifying each domain of PA1. The quantitative hypothesis proposes that the higher level of total DLX proteins in the mandible differentiates it from the maxillary. It is still unclear which hypothesis is correct.

Endothelin 1 is expressed in the ventral surface ectoderm of pharyngeal arches, as well as endoderm and mesoderm. Experiments on *Edn1*$^{-/-}$ and *Ednra*$^{-/-}$ mice show that the loss of these genes results in a homeotic transformation of a mandibular to a maxillary state, which is similar to phenotype in *Dlx5/6*$^{-/-}$ mice (Ruest et al., 2004). Ectopic expression of *Ednra* or *Hand1/2* results in the homeotic transformation of the maxilla to a mandibular state. The above results led to the proposal Edn1-ventralizing activity acts through *Edn1/ Ednra --- Dlx5/6 --- Hand1/2* signaling pathway (Sato et al., 2008). *Edn1* has a very important role in the development of intermediate arches (primary articulation, jaw joint). EDN1 can regulate the restricted expression of *Bapx1* at an intermediate DV position by inducing the expression of *Hand1/2* (Miller et al., 2003). This *Bapx1* expression defines where the jaw joint will form. Jagged-Notch signaling has

a dorsal specification ability. *Edn1* can repress dorsal fate in ventral skeletal precursors by inhibiting Jagged-Notch signaling (Zuniga et al., 2010).

Shh is another environmental cue in PA1 formation. *Shh* expression has several sources. However, the foregut endoderm-expressed *Shh* is the key source in PA1 development (Couly et al., 2002). The activation of *Shh* in foregut endoderm can result in the death of facial skeletogenic neural crest cells (FSNC) and then leads to phenotypes that are similar to holoprosencephaly (Cordero et al., 2004). One phenotype is the complete absence of the lower jaw. In the normal development, *Shh* expression in the foregut endoderm results in the survival of FSNC cells. FSNC cells will migrate to their destination and contribute to the proper PA1 formation (Hu and Helms, 1999; Jeong et al., 2004).

Ffg8 expression is also found in different regions. *Ffg8* expressed from the PA ectoderm is a NCC survival factor (Abu-Issa et al., 2002). The loss of *Ffg8* function can cause the apoptosis of NCC cells (Trumpp et al., 1999). *Ffg8* expression in mid- and hindbrain directs pharyngeal ectodermal cells to migrate laterally (Creuzet et al., 2004). This is essential for the formation of pharyngeal pouch and the generation of PA cartilages. *Ffg8* expressed in the oral epithelia participates in the AP polarizing of PA1. High level of *Ffg8* expression causes the expression of *Lhx6/7* in the oral domain and restricts *Gsc* expression to the aboral domain (Grigoriou et al., 1998). *Ffg8* can also restrict the expression of *Bapx1* to the aboral domain through induction of *Edn1* expression (Wilson and Tucker, 2004). *Bmp4* is expressed in the PA1 ectoderm. *Msx1* is expressed in the underlying mesenchyme in response to the higher level of *Bmp4. Msx1* expression can repress and restrict *Bapx1* expression in PA1 as well (Shigetani et al., 2000).

The molecular network that controls PA1 patterning is very complex and dynamic. As mentioned earlier, the epithelial-mesenchymal interactions are crucial for specifying the identity

of the pre-mandibular and maxillary-mandibular regions. The positional identity of PA1 forming NCCs can be specified at pre-migratory stages by intrinsic programs. Positional identity is plastic and can be altered by environmental signals. Both the intrinsic transcription factors and the environmental cues are involved in the development of 1st pharyngeal arch (Minoux and Rijli, 2010; Abbasi, 2001; Meideros and Grump, 2012).

Chromatin Landscape and Epigenetic Regulation

Gene expression can be regulated through the binding of transcription factors to *cis*-regulatory elements. There are several types of *cis*-regulatory elements within the genome: promoters, promoter-proximal elements, silencers, enhancers and insulators (Spitz and Furlong, 2012). Transcription factors can recognize small 6-12bps degenerate DNA sequence motifs. This intrinsic DNA affinity has relatively low specificity and can be altered by the control of DNA sequence accessibility. The most important mechanism that controls the transcription factor-binding site is the chromatin state. Eukaryotic DNA is packed into chromatin. The nucleosome, as the basic unit, usually becomes the barrier for transcription factor binding (Adams and Workman, 1995; Mirny, 2010). Euchromatin and heterochromatin are the most common chromatin states. Euchromatin is also called “open chromatin”. It has lower DNA density, high accessibility and higher rate of gene transcription. On the other hand, heterochromatin is associated with high DNA density, increased resistance and lower rate of gene transcription (Meister et al., 2011). The chromatin state holds the key to permit or repress gene expression in the right cell type and at the right time point.

There are different processes involved in the changing chromatin landscape during development and differentiation. The first step is to open the chromatin and create space for

transcription factors to bind. One of the most well known models for establishing an open chromatin state is the action of pioneer factors (Cirillo et al., 2002). Pioneer factors can directly bind at the DNA sites in the compacted nucleosomes while other factors are inaccessible at these closed chromatin sites. This distinguishes pioneer factors from regular transcription factors (Iwafuchi-Doi and Zaret., 2014). The characteristics of pioneer factors are well studied in FOXA proteins (Cirillo et al., 2002). The pioneer function of FOXA was first discovered through its binding activity at a silent liver enhancer (Gualdi et al., 1996). This binding interaction allows the activation of the nearby liver gene and induces the commitment of a liver cell fate. Genome-wide ChIP-seq analysis has shown that FOXA binds mostly in regulatory enhancer regions and to a lesser extent at promoters (Motallebipour et al., 2009). Through the binding at enhancer regions, FOXA proteins can open up the targeted silent chromatin and enable the occupancy of these regions by other transcription factors. The recruitment of other transcription factors always requires additional chromatin remodeling and chromatin modification. The initiated regulatory events can then induce either gene activation or repression (Zaret and Mango, 2016).

Other than pioneer factors, there are at least two more ways of triggering the chromatin accessibility at the enhancer regions: cooperative binding and incorporation of histone variants. Many transcription factors cannot bind their sites because of the nucleosome barriers. However, cooperative interactions among multiple factors can allow their binding events (Adam and Workman, 1995). The binding events and activation of enhancers are often associated with protein-protein interactions between transcription factors bound to adjacent sites. Transcription factors (TFs) can also bind through a cofactor, which may lead to an increase in the affinity of each transcription factor for their binding sites (Lin et al., 1990). Many studies have shown that the histone variants H2A.Z and H3.3 are present at both active enhancers and promoters (Jin and

Felsenfeld, 2007). In most cases, canonical histones are replaced by H2A.Z and H3.3 with the help of TIP60 and HIRA (Goldberg et al., 2010; Jin et al., 2009). H2A.Z or H3.3-containing nucleosomes are less stable. It creates the nucleosomal hypermobility to facilitate the initial transcription factor binding events (Svotelis et al, 2009).

Once the chromatin is accessible, the next step is to recruit nonhistone proteins (remodeling ATPases and other transcription factors) through the function of histone modification. Chromatin remodeling can enzymatically change the distribution of nucleosomes. There are basically four subfamilies of ATP-dependent remodeling complexes: ISWI, CHD, SWI/SNF and INO80 (Clapier and Cairns, 2009). Depending on the composition of remodelers, these complexes can uncover enhancers by sliding or relocating nucleosomes to activate transcription, as well as hide or close the sites to repress transcription (Clapier et al., 2017). Histone modifications can affect chromatin structure by altering the association between different histones in nucleosomes or the interaction of histones with DNA (Kouzarides, 2007). Acetylation is strongly associated with the activation of transcription. For example, H3K27ac usually marks the nucleosome that directly flanks the transcription factor-binding region (Fig. 1.3). This modification has been widely used as a readout of active enhancers within the whole genome (Creyghton et al., 2010). Other than H3K27ac, H3K4me1 is also enriched at flanking regions around transcription factor binding sites and is absent from the nucleosomal depleted regions within enhancers. Bivalent domains are identified with coexisting methylation sites that have opposing modification functions (Bernstein et al., 2005). One of the classical examples is the coexistance of H3K4me1 and H3K27me3 on a subset of enhancers (Fig. 1.3). The H3K27me3 modification is generally enriched on silent chromatin. Regions marked by both H3K4me1 and H3K27me3 are though to represent poised enhancers that have not yet been activated. Another important function of histone modification is

to maintain the domains of euchromatin and heterochromatin. Euchromatin has higher level of acetylation, while heterochramatin has lower level of acetylation and higher level of methylation. The boundary between euchromation and heterochromatin is marked by the methylation at H3K4 and H3K9, where H3K4me1/H3K4me3 antagonize *de novo* DNA methylation and H3K9me3 contributes to heterochromatin compaction (Kouzarides, 2007; Rose and Klose, 2014; Becker et al., 2017).

PITX1 in Organ Development and Human Malformations

PITX1 is a paired-like homeodomain transcription factor. It was first discovered when analyzing the mechanism for pituitary-specific transcription of the POMC gene (Lamonerie et al., 1996). PITX1 acts as a transcriptional activator and binds the core motif TAA(T/G)C(C/T). The PITX1 homeodomain contains a lysine at position 9, which defines it as a member of the *bicoid* subfamily (Hanes and Brent, 1989). There are two domains (posterior lateral mesoderm and stomodeum) that have early *Pitx1* expression. For the posterior lateral mesoderm, in mouse, the expression starts at E6.5. Its expression is maintained in all derivatives including hindlimbs and external genitalia (Lanctôt et al., 1997). At E9.5, *Pitx1* is expressed in the lateral plate mesoderm at the level at which hindlimbs will emerge. From E10.5-E13.5, *Pitx1* is robustly expressed in hindlimb buds. At later stages, *Pitx1* is absent from the center of chondrogenesis and is restricted in the perichondral region and soft tissues in hindlimbs, and the downregulation of *Pitx1* expression in the hindlimbs occurs in a proximal to distal fashion.

In order to study the function of *Pitx1* in hindlimb development in mice, the whole exon 3 of *Pitx1* was deleted to generate a *Pitx1* null allele (Szeto et al., 1999). In *Pitx1* null mutants, the ilium of the pelvis fails to form and the femur is much shorter. The patella is also missing, and

the knee joint morphology more closely resembles that of the elbow joint in the forelimb. In addition, the tibia is reduced in size and fibula is enlarged with the morphology of these bones becoming more similar to the radius and ulna, respectively. These phenotypes suggest a partial hindlimb-forelimb (HL-FL) transformation and the loss of hindlimb features (Szeto et al., 1999; Lanctôt et al., 1999). When *Pitx1* expression is ectopically induced in chick wings with the *Pitx1*-expressing virus, the infected wings develop a hindlimb-like morphology, including changes in the flexure, the absence of the patagium and reduced featherbuds (Logan and Tabin, 1999). However, the most obvious change is in the skeletal pattern induced by *Pitx1* misexpression. The infected wing forms digit II, III and IV with similar lengths. The fourth digit is usually formed by the bifurcation of the digit III. The infected wing morphologically resembles the wild-type leg. Both loss- and gain-of-function studies show that *Pitx1* directs normal *Tbx4* expression in the hindlimb and is necessary for hindlimb outgrowth (Duboc and Logan, 2011). TBX4 is another hindlimb-specific transcriptional factor (Fig. 1.4). On the other hand, *Pitx1* can shape and determine hindlimb morphology independent of TBX4, including HL buds, muscles, tendons and bones (DeLaurier., 2006). When *Pitx1* expression is induced in embryonic mouse forelimbs using a transgene driven by the *Prx1* limb enhancer, it leads to a partial transformation of forelimb structures towards hindlimb morphology. The wrist bones in the *Prx1-Pitx1* mouse transgenic embryos are transformed to resemble the bones in the foot and ankle. The forelimb *extensor carpi radialis* muscle is transformed and the *extensor indicis proprius* muscle is translocated to reflect the *tibialis* muscle and the *extensor digitorum brevis* muscle in the hindlimb.

Pitx2 is a paralog of *Pitx1*. Mutation of human PITX2 results in the Rieger syndrome (Semina et al., 1996). *Pitx2* has a left-side specific expression and is involved in the asymmetric

development of internal organs including the heart and stomach (Yoshioka et al., 1998). The first asymmetric expression is detected in left lateral plate mesoderm at E8.0 in mouse. This expression pattern only lasts from E8.0 to E9.5. The relationship between *Pitx1* and *Pitx2* in hindlimb development is partially redundant (Marcil et al., 2003). In *Pitx1*$^{-/-}$ mouse embryos, the ossification process of the hindlimb bones is much delayed and more affected than the left side of the embryo due to partial compensation by *Pitx2* (Fig. 1.4). When *Pitx1* and *Pitx2* are both knocked out, it results in a more extreme phenotype in pelvis and hindlimb than the phenotypes observed in the single mutant.

Clubfoot is a congenital birth defect in the lower limbs of humans. It affects 1 in 1000 live births and is characterized by the misalignment of bones and joints in the foot and ankle (Brewer et al., 1998). Though not common, *Pitx1* is one of the causative genes for clubfoot (Gurnett et al., 2008). A missense mutation (E130K) was identified at the conserved homeodomain region of *PITX1* locus (Gurnett et al., 2008) within a multigenerational family exhibiting lower extremity abnormalities including clubfoot. A separate novel microdeletion (chromosome 5q31) within the *PITX1* locus was identified in a family with bilateral clubfoot and short stature (Alvarado et al., 2011). The *PITX1* haploinsufficiency can cause defects in lower legs and produce morphological defects resembling clubfoot. Liebenberg syndrome is another human limb syndrome caused by sequence alterations at the *PITX1* locus (Spielman et al., 2012). In this syndrome, the patients' arms undergo a partial transformation to legs. In these patients, the distal humerus is broadened at the elbow to resemble the femur and the proximal ulna is enlarged to resemble a tibia. The boundary between the PITX1 topologically associating domain (TAD) and the nearby TAD is deleted, so the two TADs are fused into one single TAD (Petit et al., 2017). Due to deletions and structural rearrangements, three enhancers with limb activity are relocated and surround the

PITX1 locus. This leads to the misexpression of *PITX1* in the forelimb, and results in the acquisition of hindlimb morphological characteristics and the arm-to-leg homeotic transformation.

Outside the hindlimb, *Pitx1* is also expressed in the first branchial arch in E9 mouse embryos (Lanctôt et al., 1997). During later stages, *Pitx1* is expressed in the derivatives of the first branchial arch, including portions of Meckel's cartilage and the mandible mesenchyme. *Pitx1*$^{-/-}$ mouse embryos exhibit severe mandible malformations, such as much reduced mandible bones, which resembles human mandibular hypoplasia. There are other defects, including cleft palate and bifurcate tongue (Lanctôt et al., 1999). The bHLH transcription factor HAND2 is required for the maxilla-to-mandible transformation (Funato et al., 2016). When *Hand2* expression is induced in the neural crest cells of the branchial arches, both *Hand2* and *Pitx1* are ectopically expressed in the maxilla. On the other hand, when *Hand2* expression is disrupted, *Pitx1* expression is also down-regulated in the mandible. In HAND2 ChIP-qPCR analysis, HAND2-binding complexes are enriched in the regulatory domain of PITX1 (Osterwalder et al., 2014). There is a possibility that *Hand2* could directly regulate PITX1 through HAND2-mediated signaling pathways. It has been shown that the expression of *Pitx1* depends on the presence of *Fgfs* in the epithelium (Bobola et al., 2003). *Hoxa2* is involved in the specification of 2nd pharyngeal arch (PA2). It can repress the FGF-dependent *Pitx1* expression in PA2. In *Hoxa2*$^{-/-}$ mice, *Pitx1* expression is not only maintained in PA1 but also expanded in the PA2 mesenchyme. This provides an explanation of why *Pitx1* expression is restricted in PA1.

Figures

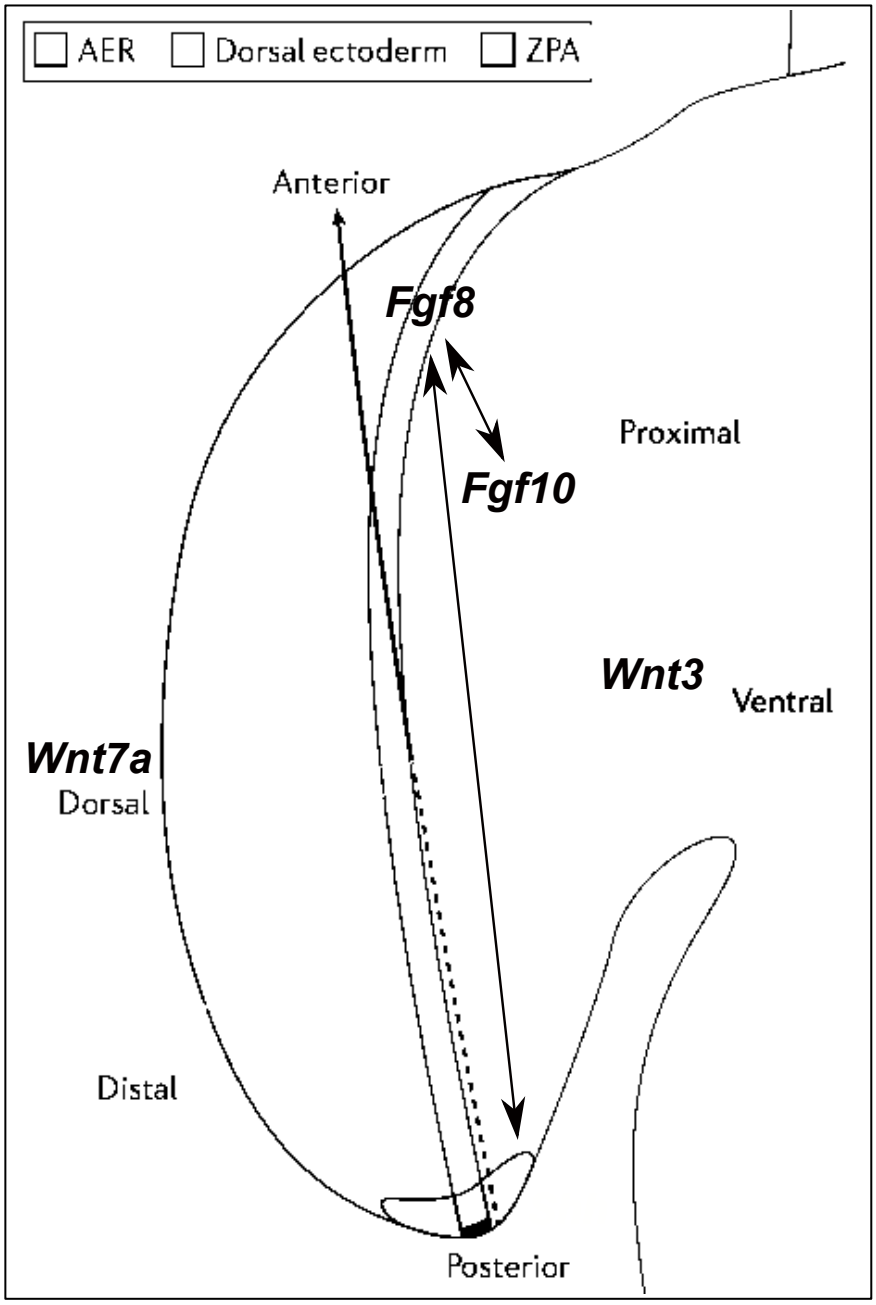


Figure 1.1. Three axes involved in limb patterning. Adapted from (Petit et al., *Nature Genetics* 2017). The limb is patterned along three different axes (anterior-posterior, proximal-distal, dorsal-ventral). AER: apical ectodermal ridge; ZPA: zone of polarizing activity.

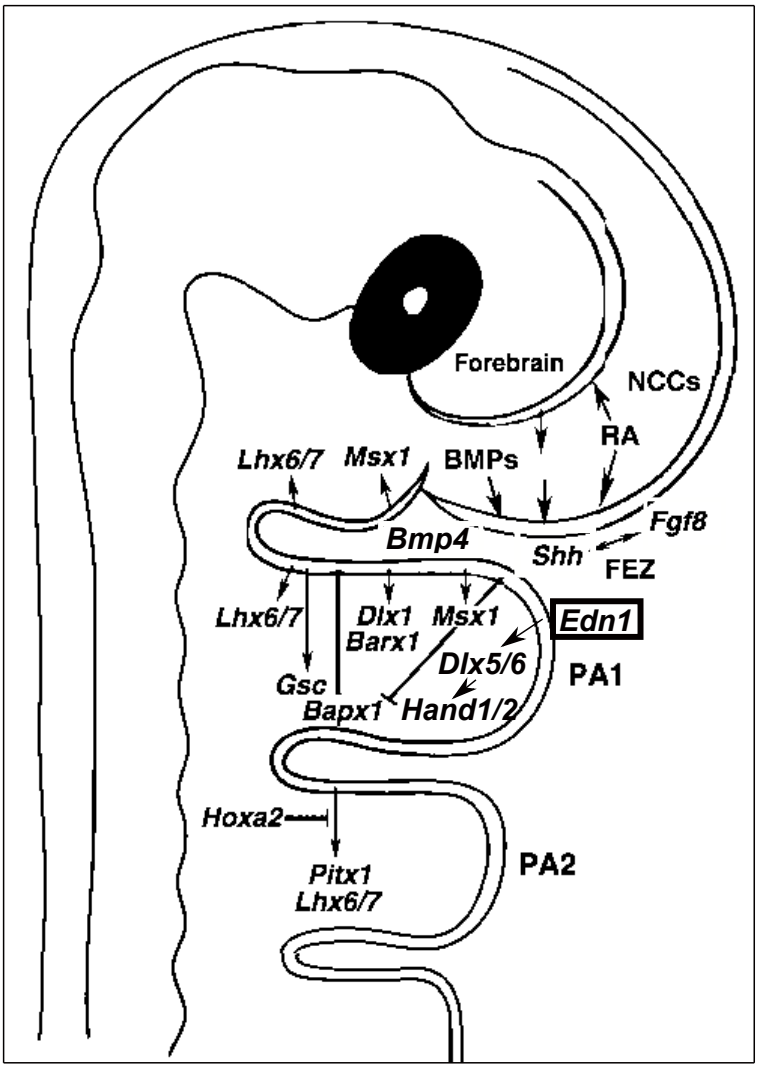


Figure 1.2. Signaling and patterning during craniofacial development. Adapted from (Minous and Rijli, *Development* 2010). A sagittal section through the vertebrate embryo showing the frontonasal process, the first and second pharyngeal arches. Genes and signals that are involved in the craniofacial development are listed.

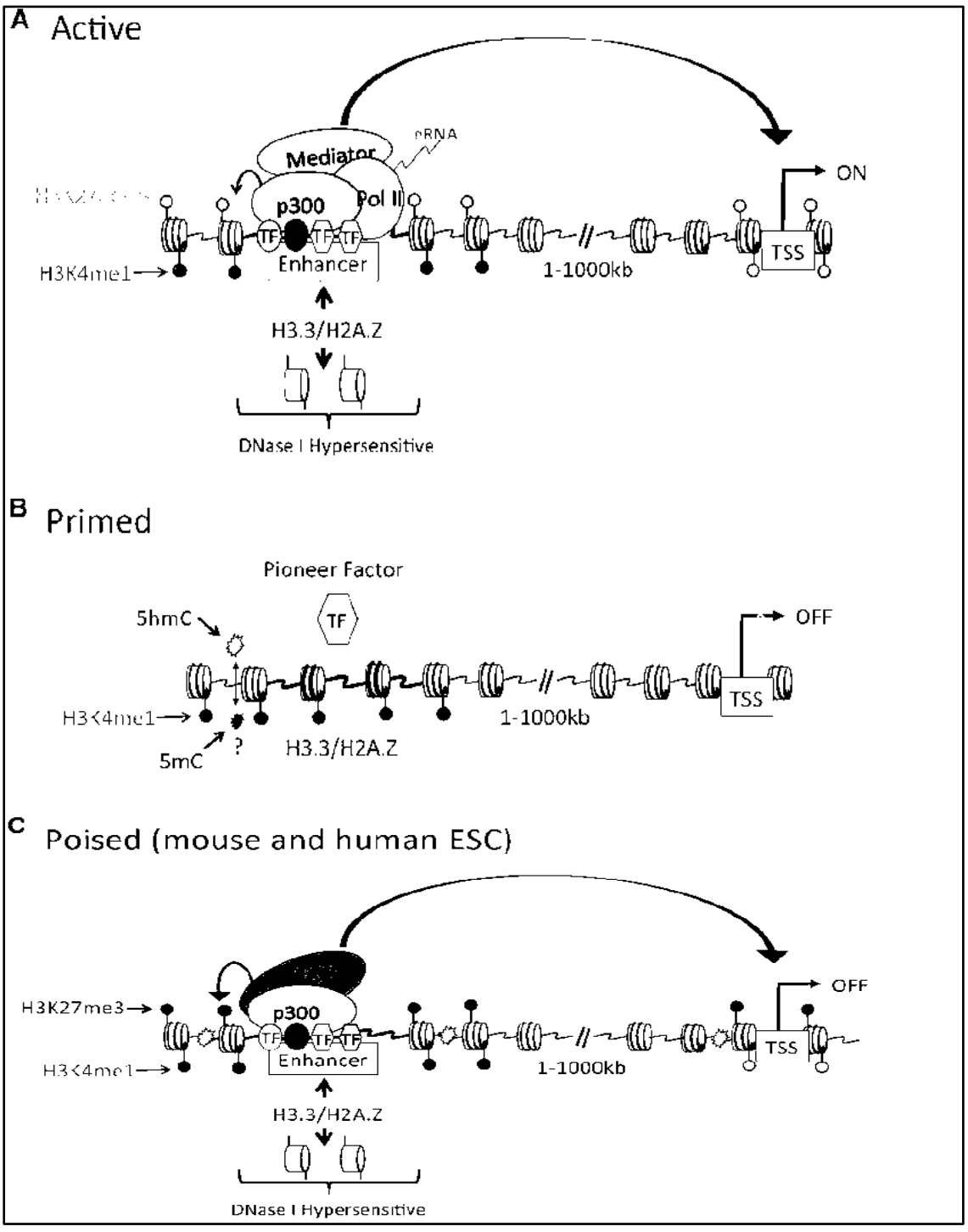


Figure 1.3. Chromatin landscape and epigenetic regulation. Adapted from (Calo and Wysocka, *Mol. Cell* 2013). The major chromatin features found at active, primed and poised enhancers.

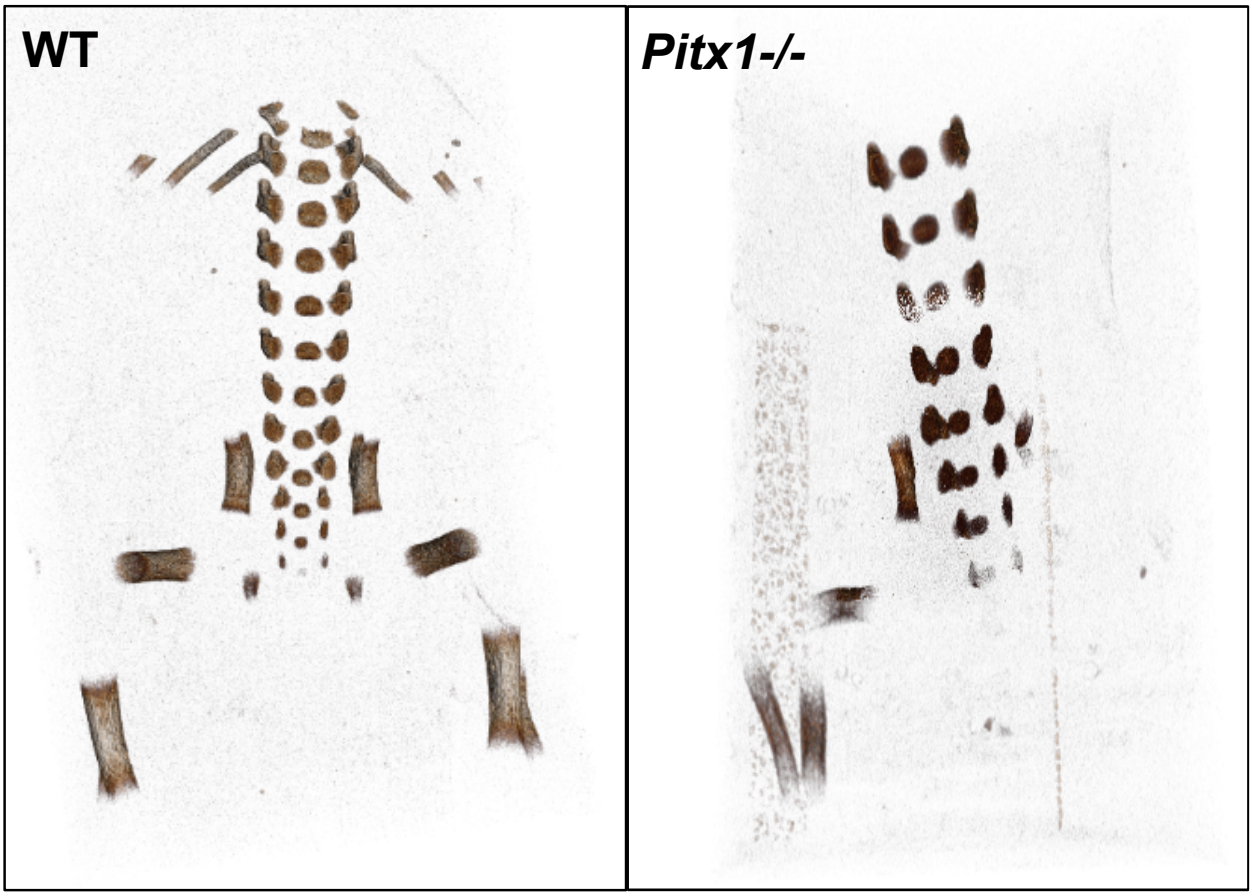


Figure 1.4. 3D X-ray microscopy in mouse embryos. E17.5 mouse wild-type and *Pitx1*$^{-/-}$ embryos were examined. Much delayed ossification process is identified on the right side of the *Pitx1*$^{-/-}$ embryo.

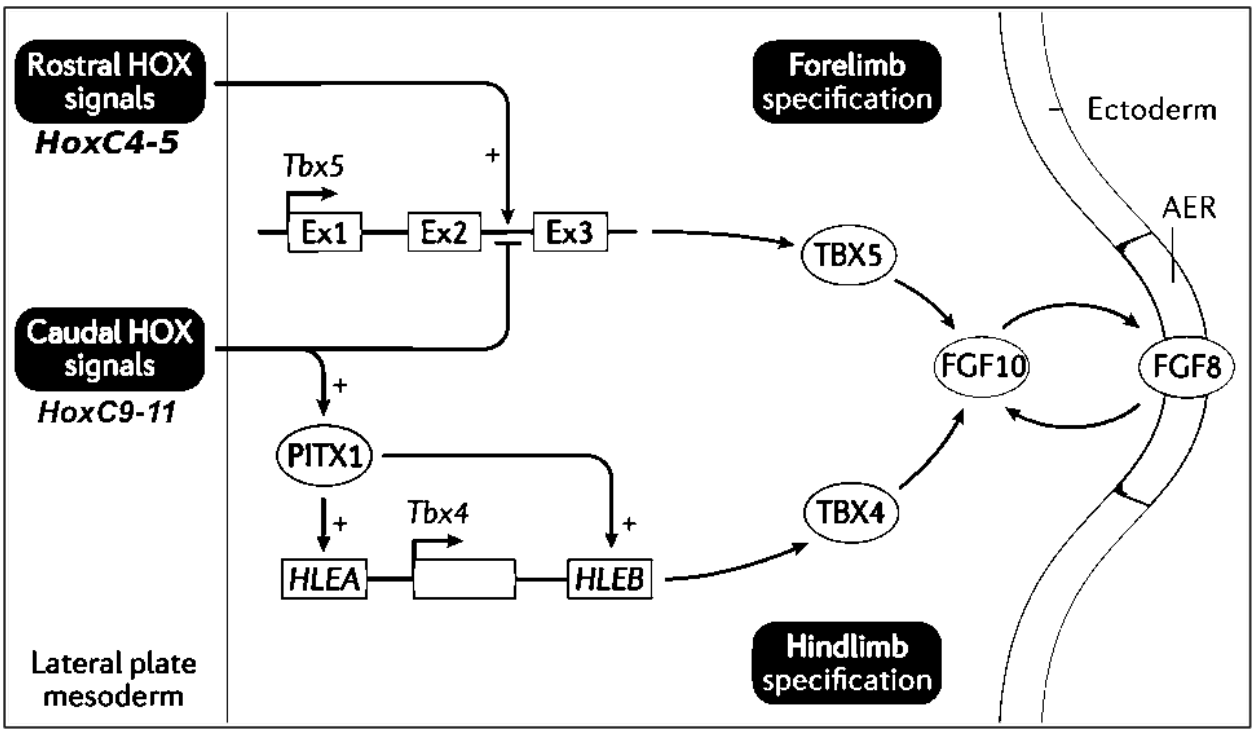


Figure 1.5. Specification of forelimb and hindlimb. Adapted from (Petit et al., *Nature Genetics* 2017). Genes and signaling pathways that are involved in the specification of forelimbs and hindlimbs.

CHAPTER 2

PITX1 PROMOTES CHONDROGENESIS AND MYOGENESIS IN MOUSE HINDLIMBS THROUGH CONSERVED REGULATORY TARGETS[1]

[1]Wang, J.S., Infante, C.R., Park, S., Menke, D.B., submitted to *Developmental Biology*, 10/01/17.

Abstract

The PITX1 transcription factor is expressed during hindlimb development, where it plays a critical role in directing hindlimb growth and the specification of hindlimb morphology. While it is known that PITX1 regulates hindlimb formation, in part, through activation of the *Tbx4* gene, other transcriptional targets remain to be elucidated. We have used a combination of ChIP-seq and RNA-seq to investigate enhancer regions and target genes that are directly regulated by PITX1 in embryonic mouse hindlimbs. In addition, we have analyzed PITX1 binding sites in hindlimbs of *Anolis* lizards to identify ancient PITX1 regulatory targets. We find that PITX1-bound regions in both mouse and *Anolis* hindlimbs are strongly associated with genes implicated in limb and skeletal system development. Gene expression analyses reveal a large number of misexpressed genes in the hindlimbs of *Pitx1*$^{-/-}$ mouse embryos. By intersecting misexpressed genes with genes that have neighboring mouse PITX1 binding sites, we identified 440 candidate targets of PITX1. Of these candidates, 68 exhibit ultra-conserved PITX1 binding events that are shared between mouse and *Anolis* hindlimbs. Among the ancient targets of PITX1 are important regulators of cartilage and skeletal muscle development, including *Sox9* and *Six1*. Our data suggest that PITX1 promotes chondrogenesis and myogenesis in the hindlimb by direct regulation of several key members of the cartilage and muscle transcriptional networks.

Keywords: Hindlimb, Anolis, Mouse, Chondrogenesis, Myogenesis

Introduction

The *Pitx1* gene encodes a bicoid-class homeodomain transcription factor that plays a central role in growth and patterning of the vertebrate hindlimb (Lanctôt et al., 1999; Szeto et al., 1999). The complete ablation of *Pitx1* function in mice results in reduced hindlimb size and the loss of hindlimb-specific features, as well as developmental defects of the mandible, teeth, palate, and pituitary gland. Haploinsufficiency for *Pitx1* in mice and humans can result in clubfoot and other malformations of the leg, demonstrating that hindlimb development is sensitive to *Pitx1* dosage (Alvarado et al., 2011; Gurnett et al., 2008). Since ectopic expression of *Pitx1* in the developing forelimbs of chickens, mice, and humans results in the forelimb developing a more hindlimb-like morphology (Delaurier et al., 2006; Logan and Tabin, 1999; Spielmann et al., 2012), it is apparent that the role of *Pitx1* in hindlimb formation extends to the control of limb-type identity. Furthermore, reductions in *Pitx1* expression have been linked to the evolution of pelvic fin loss in natural populations of threespine sticklebacks and to the formation of wing-like feathers on the hindlimbs of certain breeds of domesticated pigeon (Chan et al., 2010; Domyan et al., 2016; Shapiro et al., 2004). Thus, *Pitx1* is not only important for the formation of the hindlimb, but changes in *Pitx1* hindlimb expression have contributed to the evolution of differences in hindlimb morphology.

In *Pitx1* knockout mice, impaired hindlimb development is apparent by embryonic day 10.5 (E10.5), by which point there is a clear reduction in hindlimb bud size relative to wild-type embryos (Marcil et al., 2003). In *Pitx1* null embryos, this reduction in hindlimb size is more severe in the right hindlimb than the left due to partial compensation by *Pitx2*, a paralog of *Pitx1* which is expressed in the left lateral plate mesoderm during early embryogenesis. At later stages of hindlimb development, dramatic morphological alterations to the hindlimb skeleton are

evident in *Pitx1* null embryos. For instance, the ilium of the pelvis is either completely missing or present as a small rudiment, femur and tibia dimensions are greatly reduced, the patella is absent, and the relative proportions of the fibula and tibia are shifted to become more similar in size (Lanctôt et al., 1999; Szeto et al., 1999). Moreover, alterations to the hindlimb skeleton of *Pitx1* knockout mice are not limited to changes in the position, shape, and size of the hindlimb bones, but also include a decrease in mineralization and reduced formation of extracellular matrix. Finally, the development of soft tissues is also impacted in the *Pitx1* knockout embryos with patterning alterations that include shifts in the size and position or even the complete loss of certain hindlimb muscles and tendons (Delaurier et al., 2006).

Initial studies of *Pitx1* mouse mutants determined that hindlimb expression of the *Tbx4* gene is significantly reduced in the absence of *Pitx1* function (Lanctôt et al., 1999; Szeto et al., 1999). Additional work showed that the ectopic expression of *Pitx1* in the embryonic forelimb is sufficient to induce *Tbx4* expression (Delaurier et al., 2006; Logan and Tabin, 1999). Our further investigations revealed that PITX1 binds to and regulates hindlimb enhancers of the *Tbx4* gene, demonstrating that *Tbx4* is a direct transcriptional target of PITX1 (Infante et al., 2013). Since *Tbx4* encodes a T-box transcription factor that is critical for hindlimb bud outgrowth (Naiche and Papaioannou, 2003), it was proposed that reduced *Tbx4* expression contributes to the hindlimb phenotypes found in *Pitx1* knockout mice. Consistent with this hypothesis, subsequent work demonstrated that restoration of *Tbx4* expression in *Pitx1* null embryos is sufficient to rescue hindlimb growth defects and restore femur length, tibia length, and formation of the ilium (Duboc and Logan, 2011; Ouimette et al., 2010). However, restoring *Tbx4* expression does not rescue other hindlimb patterning defects of the skeleton, muscles, and tendons. Therefore, the misregulation of additional PITX1 transcriptional targets likely contribute to *Pitx1* mutant

phenotypes. While we have shown that the PITX1 transcription factor binds to a large number of limb *cis*-regulatory elements and may promote hindlimb formation by regulating many different genes, *Tbx4* remains the only well-validated direct, regulatory target of PITX1 (Infante et al., 2013).

In this study, we apply a comparative ChIP-seq approach to discover deeply conserved PITX1 binding events that occur in the embryonic hindlimbs of mice and *Anolis* lizards. We find that ancient PITX1 binding sites are enriched near many limb patterning genes and near components of the Hedgehog, BMP, and WNT signaling pathways. We also find that PITX1 binding events are enriched near genes that are misregulated in the hindlimbs of *Pitx1* knockout mice. We compare the location of PITX1 binding events to the position of genes that exhibit PITX1-dependent expression to reveal putative direct transcriptional targets of PITX1. Our results suggest that PITX1 promotes hindlimb development through ancient regulatory interactions with several key members of the chondrocyte and muscle transcriptional networks.

Materials and Methods

Animals

Pitx1 knockout mice were previously described (Szeto et al., 1999). The *Pitx1* knockout allele was maintained on a *129/Sv* background but outcrossed onto an outbred ICR background (Envigo) for the generation of embryos for RNA-seq and *in situ* hybridization experiments. Adult *Anolis carolinensis* were purchased from Candy's Quality Reptiles (Reserve, LA, USA), housed at the University of Georgia, and bred to produce embryos. All procedures involving animals were performed in accordance with guidelines issued by the Institutional Animal Care

and Use Committees (IACUC) at the University of Georgia under approved Animal Use Protocols (mouse protocol A2014 06-019; *Anolis* protocol A2015 02-020).

ChIP-seq

PITX1 ChIP-seq data for E11.5 mouse hindlimbs was previously described (GEO accession GSE41591; (Infante et al., 2013)). For PITX1 ChIP-seq on *Anolis carolinensis* hindlimbs, embryos were staged according to Sanger (Sanger et al., 2008) and hindlimbs from stages 6 to 7 were collected. Two independent *Anolis* ChIP-seq replicates and input control libraries were generated, using two separate pools of chromatin collected from embryonic lizard hindlimbs (50 pairs of hindlimbs per replicate for a total of 700 μg of chromatin in each ChIP). PureProteome™ Protein G Magnetic Beads (Millipore) were pre-incubated with PITX1 antibody (Santa Cruz Biotechnology, sc-18922) before incubating overnight with chromatin. All ChIP and input chromatin control libraries were produced using the NEBNext Ultra DNA Library Prep Kit for Illumina as previously reported (Infante et al., 2015). Libraries were submitted to the Georgia Genomics Facility and sequenced on the NextSeq 500 platform. *Anolis* PITX1 ChIP-seq data generated for this work have been deposited in the Gene Expression Omnibus (GSE104460) (Edgar et al., 2002).

ChIP-seq Data Analysis

Sequencing read quality was evaluated using FastQC (version 0.5.1, https://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/). ChIP-seq reads were aligned to the mouse genome (mm9) or *Anolis carolinensis* genome (anoCar2) using bowtie v1.1.0 (Langmead et al., 2009) with the parameters described previously (Infante et al., 2013). PITX1 peaks were

identified using MACS2 with default parameters except for the effective genome size, which was set for either mouse (mm9) or lizard (anoCar2). Peaks were associated with Gene Ontology (GO) terms using the Genomic Regions Enrichment of Annotation Tool (GREAT) (Mclean et al., 2010). The assignment of target genes was performed by associating PITX1 peaks with neighboring genes using GREAT. *Anolis* enhancer coordinates (anoCar2) were translated to the mouse genome (mm9) using the UCSC liftOver Tool. Only the *Anolis* enhancer regions that were successfully lifted-over to the mouse genome were used to compare PITX1 signal at orthologous enhancers between two species. The significance of PITX1 peak enrichment near putative targets was evaluated using Pearson's Chi-Square Test. A permutation test was used to determine significant overlap between enhancer datasets. A distribution was created by randomly reshuffling genomic region coordinates 1000 times and performing overlaps using BEDTools v2.26.0 (Quinlan and Hall, 2010). A p-value was calculated by ranking the observed number of overlaps within the number of overlaps from the random distribution.

RNA-seq

In order to minimize variation in RNA-seq data collected from E9.5 embryos we collected the entire hindlimb field (ectoderm and underlying lateral plate mesoderm) together with the paraxial and axial mesoderm from both wild-type and *Pitx1*$^{-/-}$ embryos. For E11.5 and E12.5 RNA-seq, only the hindlimb buds on the right side were collected from wild-type and *Pitx1* null embryos since expression of *Pitx2* on the left side of embryos can partially compensate for the loss of *Pitx1*. Embryos were sexed via PCR and only XY embryos were used. All embryos used in RNA-seq were staged based on the forelimb morphology as described previously (Wanek et al., 1989). For each developmental stage and each genotype, RNA-seq libraries were made from

three separate embryos. Total RNA was isolated using the *mir*Vana RNA Isolation Kit (ThermoFisher Scientific). RNA quality was evaluated on an Agilent 2100 Bioanalyzer, and only samples with RIN values >8 were used. Libraries were constructed with TruSeq Stranded mRNA Sample Prep Kit for Illumina. For construction of E9.5 libraries 250 ng of total RNA was used. 500 ng of total RNA was used to construct E11.5 and E12.5 hindlimb libraries. Libraries were submitted to the Georgia Genomics Facility and sequenced on the Illumina NextSeq platform to produce approximately 50 million, single-end, 75bp reads per library. Mouse RNA-seq data generated for this work have been deposited in the Gene Expression Omnibus (GSE104460).

RNA-seq Data Analysis

Sequenced libraries were evaluated and trimmed with FastQC. R package (https://www.r-project.org/) was used to calculate the correlations among libraries. Sequenced reads were aligned to mm9 using Tophat2 on the Galaxy platform (Kim et al., 2013). Cuffdiff was used to generate qualified RPKM and to identify differentially expressed genes and transcripts (Trapnell et al., 2012). GO Analysis was performed with DAVID Bioinformatics Resources 6.7 (Huang et al., 2009). DAVID annotation categories that were applied to identify functional annotation clusters include “Functional Categories”, “Literature”, “Pathways” and “Tissue Expression”.

Comparison of PITX1 Binding to PITX1-depedent Transcription

To examine the correlation of mouse PITX1-enriched regions with the transcriptional regulation of neighboring genes, we compared the genomic distribution of PITX1 peaks with the location of the misregulated genes that we identified through RNA-seq analysis. We defined the regulatory domains of 976 misexpressed genes from E11.5 *Pitx1*$^{-/-}$ hindlimbs based upon the

"Basal plus extension" approach using GREAT (Mclean et al., 2010). Each gene was assigned a basal regulatory domain with the minimum distance both upstream (5 kb) and downstream (1 kb) of the gene's transcription start site (TSS). The whole gene regulatory domain was then extended in both directions to the nearest gene's basal domain but no more than the maximum extension (1000 kb), and intersections between mouse PITX1 ChIP-Seq peaks and gene regulatory domains were performed. We used the same approach to search for overlapping regions between mouse/*Anolis* conserved PITX1-bound regions and the regulatory domains of misexpressed genes.

Whole-mount In Situ Hybridization and qRT-PCR

Whole-mount mRNA *in situ* hybridization was performed as described previously (Wilkinson, 1992). A minimum of three wild-type and three *Pitx1*$^{-/-}$ embryos were stained for each gene, and all replicates showed similar expression patterns. Embryos were staged based on the forelimb morphology. Riboprobes were generated by synthesizing a 500-640bp template using gBlocks Gene Fragments (Integrated DNA Technologies), amplifying the template by PCR, and transcribing with T3 RNA polymerase (Jarvis and Condie, 2017). Quantitative real time-PCR was used to verify expression differences between wide-type and *Pitx1*$^{-/-}$ hindlimbs that were detected by RNA-seq. E11.5 and E12.5 right hindlimb total RNA was isolated with the *mir*Vana RNA Isolation Kit (ThermoFisher Scientific) and cDNA was synthesized using the ProtoScript II First Strand cDNA Synthesis Kit (New England Biolabs). qPCR assays were performed on a LightCycler® 480 using SYBR Green I Master reagent (Roche). *Gapdh* was used as the internal control for normalization. The $2^{-\Delta\Delta Ct}$ method was used to detect expression fold change for each target gene with three biological replicates. The primers used in qPCR were as

follows: *Alx3*-F: 5'-CTATGACATCTCCGTACTGCC-3'; *Alx3*-R: 5'-TCTGGAGACATGAGACAGGG-3'; *Pax9*-F: 5'-GTGAATGGATTGGAGAAGGGAG-3'; *Pax9*-R: 5'-CATGTAGGGTGACACTTGGG-3'; *Cxcl12*-F: 5'-CGCTCTGCATCAGTGACG-3'; *Cxcl12*-R: 5'-GTTTGGAGTGTTGAGGATTTTCAG-3'; *Sox9*-F: 5'-GCCGACTCCCCACATTC-3'; *Sox9*-R: 5'-CGCTTCAGATCAACTTTGCC-3'; *Col2a1*-F: 5'-GGTTCACATACACTGCCCTG-3'; *Col2a1*-R: 5'-AAATTCCTGTTCAGCCCCTC-3'; *Runx2*-F: 5'-GTAGCCAGGTTCAACGATCTG-3'; *Runx2*-R: 5'-CCGTCCACTGTCACTTTAATAGC-3'; *Six1*-F: 5'-CAGGTCAGCAACTGGTTTAAG-3'; *Six1*-R: 5'-CAGAGGAGAGAGTTGATTCTGC-3'; *Mef2c*-F: 5'-CCAGATCTCCGCGTTCTTATC-3'; *Mef2c*-R: 5'-CCTCCCATTCCTTGTCCTG-3'; *Myod1*-F: 5'-CAGAATGGCTACGACACCG-3'; *Myod1*-R: 5'-ATGCGCTCCACTATGCTG-3'; *Gapdh*-F: 5'-AAGGTCGGTGTGAACGGATTTG-3'; *Gapdh*-R: 5'-GTCGTTGATGGCAACAATCTCC-3

Results

Conserved PITX1 Binding Events are Enriched Near Limb Genes

In previous work, we performed PITX1 ChIP-seq on mouse E11.5 hindlimbs (Infante et al., 2013). These experiments revealed the genome-wide distribution of mouse PITX1 binding sites during hindlimb development. While many of these PITX1-associated regions fall within well-conserved non-coding sequences, DNA sequence conservation is not sufficient to conclude that PITX1 is actually bound to orthologous *cis*-regulatory elements that are conserved in other species. In order to identify deeply conserved PITX1-binding events, we performed PITX1 ChIP-seq on embryonic hindlimb chromatin from the green anole lizard, *Anolis carolinensis*. The common ancestor of mammals and lizards lived more than 300 million years ago (Pyron, 2010). Therefore, PITX1-binding events that are shared between these species are likely to represent ancient binding interactions that are common to diverse species of limbed amniotes.

Anolis PITX1 ChIP-seq was carried out on chromatin isolated from the hindlimbs of stage 6 to 7 lizard embryos (Sanger et al., 2008). At these embryonic stages, *Anolis* hindlimbs are roughly comparable in development to E11.5 mouse hindlimbs. We determined that the PITX1-specific antibody that we previously used for mouse PITX1 ChIP-seq, also cross-reacts with *Anolis* PITX1 (Fig. S2.1). Therefore, we were able to use the same PITX1 antibody for mouse and *Anolis* ChIP-seq experiments. Analysis of our *Anolis* hindlimb PITX1 ChIP-seq replicates resulted in the identification of 11,090 peaks, a number similar to the 10,625 PITX1 ChIP-seq peaks that we identified in mouse hindlimbs (Table S2.2) (Infante et al., 2013). *De novo* motif searches performed within +/- 50 bps of PITX1 peak summits revealed that TAATCC, the core PITX1 binding motif, was significantly enriched in both *Anolis* (p-value=1 x 10^{-986}; 61% of target sequences) and mouse (p-value=1 x 10^{-1044}; 69% of target sequences) PITX1 peaks (Fig.

2.1A and Table S2.1). Moreover, we found that both mouse and *Anolis* PITX1 peaks are significantly enriched near genes associated with limb development (Fig. 2.1D, G, and J and Fig. S2.2).

To identify deeply conserved enhancers that are bound by PITX1 in the developing hindlimbs of mouse and *Anolis* embryos, we began by determining the fraction of *Anolis* PITX1 peaks that have sequence orthologs in the mouse genome. Of 11,090 *Anolis* peaks, 2479 (22.4%) exhibit DNA sequence conservation between the *Anolis* and mouse genomes (Table S2.2). We intersected the mouse genome coordinates of these 2479 conserved sequences with the coordinates of the 10,625 mouse PITX1 ChIP-seq peaks to reveal 574 overlapping regions. These 574 regions represent the subset of putative *cis*-regulatory elements that exhibit sequence conservation between mouse and *Anolis* and that are bound by PITX1 in the developing hindlimbs of both species. The degree of overlap between mouse peaks and *Anolis* peaks is significantly higher than expected by chance ($p < 0.001$). Prior work has demonstrated that binding site turnover is a common feature of enhancer evolution, and individual binding sites for a transcription factor can be lost and then compensated for by the gain of new binding sites for the same transcription factor (Hare et al., 2008). Therefore, we extended our search for putative *cis*-regulatory elements that are bound by PITX1 in mouse and *Anolis* by identifying regions where the sequence underlying an *Anolis* PITX1 peak is conserved in mouse and occurs within 500 base pairs of a neighboring mouse PITX1 peak. This led to the identification of an additional 87 putative *cis*-regulatory regions that are conserved between mouse and *Anolis* but where the location of the PITX1 binding event within the region differs between the mouse and *Anolis* orthologs. Together, our analyses revealed a total of 661 putative *cis*-regulatory elements that

exhibit primary sequence conservation between mouse and *Anolis* and that are bound by PITX1 in both species (Table S2.2).

The majority of the mouse/*Anolis* PITX1-bound regions are located within 100 kb of transcription start sites with a distribution that is similar to the larger set of all mouse PITX1 peaks (Fig. 2.1A and 2.1B). As expected, the top ranked *de novo* motif found within the mouse/*Anolis* PITX1 regions matches the known PITX1 motif (Fig. 2.1B and Table S2.1). Notably, this motif is essentially indistinguishable from the top motif found in "mouse-specific" PITX1 peaks, the subset of regions that are bound by PITX1 in mice but that are not bound by PITX1 in *Anolis* or that do not have detectable sequence conservation in *Anolis* (Fig. 2.1C). To determine whether conserved PITX1 binding events are enriched near particular gene categories, we used GREAT (Mclean et al., 2010). This analysis demonstrated that mouse/*Anolis* PITX1 bound regions are significantly enriched near genes that are expressed in developing limbs, with 5 of the top 10 ranked gene expression GO terms associated with embryonic limb expression (Fig. 2.1E). Examination of signaling pathway GO terms showed that genes associated with the Hedgehog, WNT, and BMP signaling pathways are also enriched near conserved PITX1 bound regions (Fig. 2.1K). For PITX1 peaks that are mouse-specific, enrichments for similar sets of limb-related and signaling pathway GO terms were observed (Fig. 2.1F, I and L). Thus, mouse/*Anolis* conserved and mouse-specific PITX1 binding events display similar gene associations.

Genes Related to Patterning, Chondrogenesis, and Myogenesis are Misregulated in Pitx1$^{-/-}$ Hindlimbs

In order to identify genes with PITX1-dependent expression, we performed global gene expression comparisons between the developing hindlimbs of *Pitx1*$^{-/-}$ and wild-type mouse embryos. For our analyses, RNA-seq was carried out separately on E9.5 hindlimb fields, E11.5 hindlimbs, and E12.5 hindlimbs. Given the small size of the E9.5 hindlimb field, we collected the prospective hindlimb region (ectoderm and underlying lateral plate mesoderm) together the associated paraxial and axial mesoderm. This helped to decrease variation in the tissue samples that we collected. Because the E9.5 samples contained a substantial amount of tissue outside of the prospective hindlimb, we deeply sequenced the RNA-seq libraries that we generated to improve our ability to detect differentially expressed genes (~50 million reads were sequenced per library). Since *Pitx2* expression on the left side of embryos can partially compensate for the loss of *Pitx1* (Marcil et al., 2003), we exclusively used right hindlimb buds to prepare RNA-seq libraries at E11.5 and E12.5 as the hindlimbs can be precisely dissected at these stages. In the early hindlimb field, we found that a relatively modest number of genes (55 in total) are significantly misregulated in *Pitx1* mutants (Fig. 2.2A, Table S2.3). In contrast, several hundred genes are misregulated at E11.5 and E12.5. At these later stages, we also observed that a greater portion of misregulated genes are down-regulated than up-regulated, consistent with PITX1 acting predominantly as a transcriptional activator.

Previous studies have demonstrated that *Tbx4* expression is reduced in the hindlimb buds of *Pitx1* knockout embryos and that *Tbx4* is a direct regulatory target of PITX1 (Infante et al., 2013; Lanctôt et al., 1999; Szeto et al., 1999). Consistent with these earlier findings, our RNA-seq analyses demonstrate that *Tbx4* expression is significantly reduced in the absence of PITX1

(expressed at 34% and 56% of wild-type at E11.5 and E12.5, respectively; Table S2.3). Prior work also established that the expression of hindlimb-specific *HoxC10* and *HoxC11* genes can be induced in the forelimb through ectopic expression of *Pitx1* (Delaurier et al., 2006; Logan and Tabin, 1999; Park et al., 2014). However, we found that the expression levels of *HoxC10* and *HoxC11* are not significantly reduced during hindlimb development of *Pitx1*$^{-/-}$ embryos, indicating that PITX1 function is not essential for expression of these genes. Finally, we examined the expression of *Isl1*, a gene that is essential for hindlimb bud formation but plays no role in forelimb development (Kawakami et al., 2011). In wild-type embryos, *Isl1* is expressed transiently in the hindlimb field and early hindlimb bud but is down-regulated by E10.5. In E9.5 *Pitx1*$^{-/-}$ embryos, we find that *Isl1* is expressed at significantly higher levels than wild-type (1.5x higher; Table S2.3). Therefore, of the known transcription factors with robust, hindlimb-biased expression, only *Tbx4* exhibits reduced expression in the absence of PITX1.

We next used the DAVID functional annotation tool to determine whether specific pathways or gene categories are enriched among genes that are misexpressed in *Pitx1* mutants (Huang et al., 2009). Among the categories of genes that are misregulated in *Pitx1*$^{-/-}$ hindlimb field at E9.5, the top ranked cluster relates to limb morphogenesis and pattern specification (Fig. 2.2B). By E11.5, the top ranked terms that are associated with misregulated genes include extracellular matrix, cartilage and bone development, and skeletal muscle development (Fig. 2.2C). Similar enrichments are observed at E12.5, with significant associations with genes related to extracellular matrix, cell adhesion, collagen, and muscle function (Fig. 2.2D). Thus, in *Pitx1*$^{-/-}$ embryos we find evidence of altered transcription of limb patterning genes in the early hindlimb field by E9.5, just prior to hindlimb bud outgrowth. By E11.5 large-scale alterations in

gene expression are apparent with genes related to extracellular matrix, cartilage and skeletal muscle being enriched among misregulated genes.

PITX1 Binding Events are Enriched Near Misregulated Genes

To examine the relationship between PITX1-bound regions and the transcriptional regulation of neighboring genes, we compared the genomic distribution of PITX1 peaks with the location of the misregulated genes identified through RNA-seq on *Pitx1*$^{-/-}$ hindlimbs. Since our PITX1 ChIP-seq data was collected from E11.5 hindlimb chromatin, we compared PITX1 binding sites with our list of E11.5 misregulated genes. We began by associating the 10,625 mouse PITX1 ChIP-seq peaks with gene locations using GREAT (Mclean et al., 2010). We found that PITX1 binding sites are significantly enriched near genes that are misexpressed in *Pitx1*$^{-/-}$ hindlimbs (2.45-fold enriched; p-value < 2.2×10^{-16}) (Fig. 2.3C). In addition, the association between PITX1 binding events is stronger with down-regulated genes (2.98-fold enriched; p-value < 2.2×10^{-16}) than up-regulated genes (1.65-fold enriched; p-value = 8.97×10^{-6}). When we limited our analysis to the PITX1 binding events that are shared between mouse and *Anolis*, we observed that these sites are also enriched near genes that are misregulated in the absence of PITX1 (2.82-fold enriched; p = 2.8×10^{-16}).

In order to identify direct transcriptional targets of PITX1, we next compared the genomic positions of misexpressed genes to the location of PITX1 bound regions. Since precise maps of gene regulatory domains are not currently available for embryonic limbs, we associated PITX1 peaks with specific genes by using GREAT (Mclean et al., 2010). GREAT operates by assigning each gene a basal regulatory domain around the gene's transcriptional start site and then extends the regulatory domain to the nearest upstream and downstream genes to a maximum of 1

megabase. Therefore, each PITX1 binding site has the potential to be assigned to more than one gene. Using this approach, we identified 440 genes as candidate, direct transcriptional targets of PITX1 (Table S2.4). Among the candidate targets of PITX1, the expression levels of 121 are increased and the levels of 319 are decreased in hindlimbs of *Pitx1* knockout embryos. Of the 440 putative transcriptional targets of PITX1, 68 are associated with one or more neighboring mouse/*Anolis* conserved PITX1 binding event. The PITX1 dependent expression of these 68 genes coupled with the presence of conserved PITX1 binding events suggests that these genes are ancient transcriptional targets of PITX1 (Table S2.4).

PITX1 Directly Influences Hindlimb Development through Essential Regulators of Cartilage and Muscle Development

Our initial DAVID analysis of genes that are misregulated in *Pitx1* null embryos included both direct transcriptional targets of PITX1 and genes that are misregulated as a secondary consequence of *Pitx1* inactivation (see section 3.2). We revisited this analysis by focusing exclusively on our list of putative direct transcriptional targets of PITX1 and by performing separate DAVID analyses on up-regulated and down-regulated PITX1 targets. This revealed that up-regulated targets are enriched for terms related to cell cycle control, limb morphogenesis/patterning, homeobox genes, and components of the BMP and Hedgehog signaling pathways (Fig. 2.3A). The up-regulated genes underlying these enriched GO terms include a large number of factors with well-established roles in the initiation of limb bud formation and early patterning of the limb buds and include *Alx3*, *Bmp2*, *Bmp4*, *Bmp7*, *Fgf10*, *Grem1*, *HoxA9*, *HoxD10*, *Irx3*, *Irx5*, *Meis2*, and *Tbx2* (Table S2.4). In contrast, the PITX1 target genes that are down-regulated in *Pitx1* mutants are strongly enriched for terms related to

extracellular matrix, cell adhesion, cartilage/bone development, and skeletal muscle development (Fig. 2.3B). The cartilage and bone related genes include *FoxP4*, *Snai1*, *Sox6*, *Sox8*, and *Sox9* (Table S2.4), which are all transcription factors that regulate chondrogenesis (Bi et al., 1999; Chen and Gridley, 2013; Smits et al., 2001; Zhao et al., 2015). In addition, multiple collagens and other extracellular structural protein important for cartilage and bone formation have reduced expression in the absence of PITX1. Further examination of down-regulated PITX1 targets revealed multiple regulators of myogenesis including *Cxcl12*, *Eya1*, *Mef2c*, *Nfatc2*, and *Six1* (Table S2.4; (Daou et al., 2013; Grifone et al., 2007; 2005; Molkentin et al., 1995; Vasyutina et al., 2005)). We also found that the *Dcn*, *Lum*, and *Mkx* genes, which are all important for tendon development (Danielson et al., 1997; Ito et al., 2010; Jepsen et al., 2002), were among the putative PITX1 transcriptional targets that we identified. In order to gain a better understanding of the *Pitx1*$^{-/-}$ expression defects we detected, we generated time-course charts from our RNA-seq data. This allowed us to examine the expression of misregulated genes at different developmental stages. We found that *HoxD10* expression is significantly elevated in both E11.5 and E12.5 *Pitx1*$^{-/-}$ hindlimbs (Fig. 2.4). Similar patterns of elevated expression were observed for *Alx3* and *Meis2* (Table S2.4). These genes were up-regulated by 1.3x-2x in *Pitx1*$^{-/-}$ hindlimbs relative to wild-type at E11.5. Other early patterning genes, including *Grem1*, *Bmp2*, *Bmp4*, *Bmp7*, *Fgf10,* and *Tbx2* (Fig. 2.4 and Table S2.4), exhibited transient up-regulation in *Pitx1* mutants hindlimbs at E11.5 before returning to normal or near normal expression levels at E12.5. In contrast, chondrogenesis and myogenesis PITX1 target genes exhibited a delay in their up-regulation with decreased expression apparent at E11.5 (Fig. 2.4 and Table S2.4).

We characterized the expression patterns of several misregulated genes by whole-mount *in situ* hybridization (Fig. 2.5). In general, the overall pattern of up-regulated genes was very

similar between the hindlimbs of *Pitx1*$^{-/-}$ and wild-type embryos, though we did find that *Alx3* expression was expanded towards the center of *Pitx1*$^{-/-}$ hindlimb buds (Fig. 2.5A-B). *In situ* hybridization for genes associated with cartilage and muscle development produced more dramatic differences. For instance, *Sox9* displayed altered expression in the proximal portion of the hindlimb bud at E12 (Fig. 2.5G-H), and *Mef2c* and other markers of myogenesis had qualitatively less staining by *in situ* hybridization at E11.5-E12.5 (Fig. 2.5M-R).

PITX1 Binding Profiles at Conserved Transcriptional Targets

To explore the ancient transcriptional targets of PITX1 in hindlimbs, we performed a final DAVID analysis on the 68 putative conserved targets of PITX1. This analysis shows that skeletal- and muscle-related GO terms ("Mesenchyme/Bone development", "Skeletal system development" and "Muscle organ development") are enriched among these 68 genes (Fig. S2.2A). Patterning-related terms ("Transcription factor activity" and "Pattern specification") are also significantly enriched in the conserved target gene set (Fig. S2.2A). Thus, the functional terms associated with these 68 conserved PITX1 targets are similar to the terms enriched in the larger set of 440 putative PITX1 targets.

One of the predicted ancient regulatory targets of PITX1 is *Tbx4*. In mice, we previously demonstrated that PITX1 regulates *Tbx4*, at least in part, through binding of the HLEA *cis*-regulatory element (Infante et al., 2013; Menke et al., 2008). While HLEA is not conserved in *Anolis*, we find that there are two conserved PITX1 binding events downstream of *Tbx4* (Fig. 2.6C). Therefore, in the mouse hindlimb there are both conserved and mouse-specific binding events occur at the *Tbx4* locus. A further examination of the 68 predicted ancient targets of PITX1 showed that this group includes *Sox9* and *Six1*, important regulators of cartilage and

skeletal muscle development, respectively. In the case of *Sox9*, a significant PITX1 peak is identified 232 kb downstream of its transcription start site (Fig. 2.6A). The sequence of this binding region is conserved in the *Anolis* genome, and a PITX1 peak is also located at the orthologous sequence located near the *Anolis sox9* gene. Similarly, there is a strongly enriched PITX1 peak located just upstream of *Six1* in mouse (Fig. 2.6B). This region is conserved in the *Anolis* genome and is bound by PITX1 in embryonic *Anolis* hindlimbs. These results are consistent with the idea that PITX1 promotes chondrocyte and myoblast differentiation in the hindlimb by direct regulation of key factors involved in the chondrogenic and myogenic regulatory networks. Moreover, some of these factors are likely to be ancient regulatory targets of PITX1.

Discussion

Nearly two decades ago functional studies of PITX1 demonstrated that this transcription factor performs pivotal roles in promoting hindlimb growth and specifying hindlimb morphology (Delaurier et al., 2006; Lanctôt et al., 1999; Logan and Tabin, 1999; Szeto et al., 1999). Yet the direct regulatory targets of PITX1 have remained largely unknown. Our previous work demonstrated that PITX1 is bound to thousands of different sites during mouse hindlimb development, including many known limb enhancers. However, the presence of a PITX1 binding interaction does not itself prove that the interaction directly influences gene expression or enhancer activity. Since it is not practical to perform direct functional tests on each of the thousands of different PITX1 binding sites that occur in developing mouse hindlimbs, we have used ChIP-seq combined with expression analyses to identify putative regulatory targets of PITX1.

Our RNA-seq analyses of early and later stages of hindlimb development revealed that many genes important for the initial stages of limb growth and patterning are up-regulated in the hindlimbs of *Pitx1*$^{-/-}$ embryos. The expression of several of these genes, including *Bmp7*, *Fgf10*, *Meis2*, and *Tbx2*, have previously been analyzed in *Pitx1*$^{-/-}$;*Pitx2*$^{-/+}$ embryos by whole-mount *in situ* hybridization and were found to have grossly normal patterns of expression (Marcil et al., 2003). However, our quantitative RNA-seq analyses show that the expression of these genes is elevated in the absence of PITX1. All four of these genes have neighboring PITX1 ChIP-seq peaks and are putative direct transcriptional targets of PITX1. Prior work has suggested that PITX1 can act as a transcriptional repressor in some contexts (Qi et al., 2011), but additional studies will be required to determine whether the binding of PITX1 is directly acting to reduce gene expression of certain limb-patterning genes.

Previous anatomical and histological analysis of *Pitx1* knockout mice showed that the hindlimbs are reduced in overall size and the relative size and shape of hindlimb skeletal elements are altered. Lanctôt and colleagues also demonstrated that skeletal elements of *Pitx1* knockout hindlimbs exhibit impaired ossification as evidenced by abnormal mineralization of the femur and tibia at E16.5 (Lanctôt et al., 1999). Therefore, the reduced and delayed expression of chondrogenesis genes that we have detected in *Pitx1*$^{-/-}$ hindlimbs aligns well with earlier histological studies. However, it has remained unclear whether the ossification defects of *Pitx1* knockout embryos are directly due to a role of PITX1 in cartilage and bone differentiation or are a secondary consequence of growth and patterning defects that occur earlier in hindlimb development. Our mouse PITX1 ChIP-seq data strongly suggest that PITX1 directly influences the differentiation of skeletal tissues by regulating several factors that are involved in cartilage and bone development (Fig. 2.7A). For instance, we found that *Sox9*, an early chondrogenic

marker necessary for mesenchymal condensation and early chondrocyte differentiation (Akiyama et al., 2002; Bi et al., 1999), has significantly reduced expression at E11.5 in the absence of PITX1. We also identified a strong PITX1 binding site downstream of *Sox9*'s transcription start site (Fig. 2.6A). This PITX1-bound element is a known limb enhancer (VISTA enhancer mm636; (Visel et al., 2007)) that is conserved among amniotes, and we found that PITX1 is bound to the orthologous *Anolis* sequence during embryonic development of lizard hindlimbs. Additionally, our ChIP-seq data revealed a series of PITX1 binding sites located within a large (>1MB; mm9 chr11:111,424,585-112,518,902) region upstream of *Sox9*. This region contains 12 mouse/*Anolis* conserved PITX1 binding events, and this genomic domain has previously been identified as an extended transcriptional control region for *Sox9* (Gordon et al., 2009; Pfeifer et al., 1999; Wunderle et al., 1998). A number of distinct, tissue-specific enhancers have been found in this interval, including three limb enhancers (Gordon et al., 2014; Visel et al., 2007), and we have found that each of these limb enhancers overlaps with a conserved PITX1 binding event. We therefore hypothesize that that the *Sox9* gene is an ancient, direct transcriptional target of PITX1.

In addition to *Sox9*, our ChIP-seq and RNA-seq data suggest that *FoxP4*, *Mef2c*, *Snai1*, and *Sox6* are all direct targets of PITX1 during hindlimb development. Each of these genes has been implicated in chondrogenesis or osteogenesis (Arnold et al., 2007; Chen and Gridley, 2013; Smits et al., 2001; Zhao et al., 2015), and two of these loci (*FoxP4* and *Sox6*) have conserved non-coding sequences that are bound by PITX1 in mouse and *Anolis*. Although we detected no conserved PITX1 binding events near *Mef2c* or *Snai1*, the mouse and *Anolis* orthologs of these genes both have species-specific PITX1 peaks (Fig. S2.2D and data not shown). Therefore, despite the apparent absence of conserved PITX1 bound sequences, *Mef2c* and *Snai1* may also

be ancient regulatory targets of PITX1. We also note that *Runx2* and *Osterix* have reduced expression in *Pitx1*$^{-/-}$ hindlimbs by E12.5 (>2x reduced; Table S2.3). Both of these genes encode transcription factors that have well-established functions in promoting chondrocyte hypertrophy and osteoblast differentiation, and these genes are normally up-regulated shortly after *Sox9* (Komori et al., 1997; Nakashima et al., 2002; Otto et al., 1997; Yoshida et al., 2004). Though we detected conserved PITX1 binding events adjacent to *Runx2* and *Osterix* genes, our ChIP-seq data was collected at E11.5 rather than at E12.5, the stage at which we detected significant misregulation of the *Runx2* and *Osterix* genes in *Pitx1* knockout mice. Therefore, while we have not included these genes in our list of putative PITX1 targets, it remains possible that *Runx2* and *Osterix* are directly regulated by Pitx1.

Prior studies established that the *Pitx2* gene is expressed in muscle progenitors throughout the developing embryo, including those that contribute to the limb musculature (L'Honor et al., 2007; Shih et al., 2007). Further investigations by L'Honoré demonstrated that PITX2 is required for the onset of *MyoD* expression in muscle precursors of the forelimbs and hindlimbs and showed that PITX2 directly binds to the *MyoD* core enhancer to activate *MyoD* expression (L'Honore et al., 2010). In contrast to *Pitx2*, the *Pitx1* gene is expressed throughout the hindlimb mesenchyme and can be detected in both muscle progenitors and cells that give rise to the hindlimb skeleton (Marcil et al., 2003). Though *Pitx1* is expressed in hindlimb muscle progenitors, a direct role for PITX1 in promoting myogenesis has not previously been demonstrated. Our expression analyses reveal that *MyoD* expression is delayed and decreased in *Pitx1*$^{-/-}$ hindlimbs in a manner similar to the *MyoD* expression defects observed in *Pitx2* mutants (Fig. 2.5Q-R; Table S2.3). However, unlike PITX2 we find no evidence of PITX1 binding at the *MyoD* core enhancer or any other location within 600kb upstream or downstream of the *MyoD*

locus. Instead, we have identified direct myogenic targets of PITX1 that include *Six1*, *Eya1* and *Mef2c* (Fig. 2.7B). The SIX1/EYA1 transcriptional complex functions as an upstream activator of myogenic determination genes, including *MyoD*, as evidenced by the loss of all hypaxial dermomyotome derived muscles in *Six1/4* and *Eya1/2* double mutants (Grifone et al., 2007; 2005). We found that *Six*1 is significantly down-regulated in E11.5 *Pitx1* null hindlimbs (Fig. 2.4; Fig. S2.2B), and conserved PITX1 peaks are located in its regulatory domain (Fig. 2.6B). *Eya1* also shows significantly reduced expression in E11.5 *Pitx1*$^{-/-}$ hindlimbs (Table S2.3) and has species-specific PITX1 binding events in both mouse and *Anolis* hindlimbs (Table S2.2). In addition to the function of *Mef2c* in skeletal development (see above), this gene also promotes myogenic differentiation (Molkentin et al., 1995). Thus, we believe the reduced *MyoD* expression in *Pitx1* knockout hindlimbs is a secondary consequence of the absence of PITX1.

Beyond myogenesis genes that encode transcription factors, we also identified *Cxcl12* as a candidate PITX1 regulatory target. CXCL12 (also known as SDF1) is a secreted chemokine that acts as a ligand for the chemokine receptor CXCR4 (Vasyutina et al., 2005). CXCR4/CXCL12 signaling plays an important role in the migration of muscle progenitor cells into the limb buds with the *Cxcr4* gene expressed in migrating muscle progenitors and *Cxcl12* in the developing limbs. Taken together, we find strong evidence that PITX1 promotes myogenesis through direct regulation of multiple myogenesis genes that include transcriptional regulators and the *Cxcl12* chemokine.

In conclusion, we have found significant delays in the expression of important cartilage, bone, and muscle genes in the absence of PITX1. While the relative importance of individual PITX1-binding sites remains to be tested, the presence of deeply conserved binding events that

are shared between mammals and reptiles strongly suggests that PITX1 directly regulates the transcription of key components of the chondrogenesis and myogenesis regulatory networks.

Acknowledgements

This work was supported by grants awarded to D.B.M. from the NSF (#1149453) and NIH (HD081034) and by two Junior Faculty Research Grants (#1847 and #2503) from the Office of the Vice President of Research at the University of Georgia. This study was supported in part by resources from the Georgia Advanced Computing Resource Center.

Figures

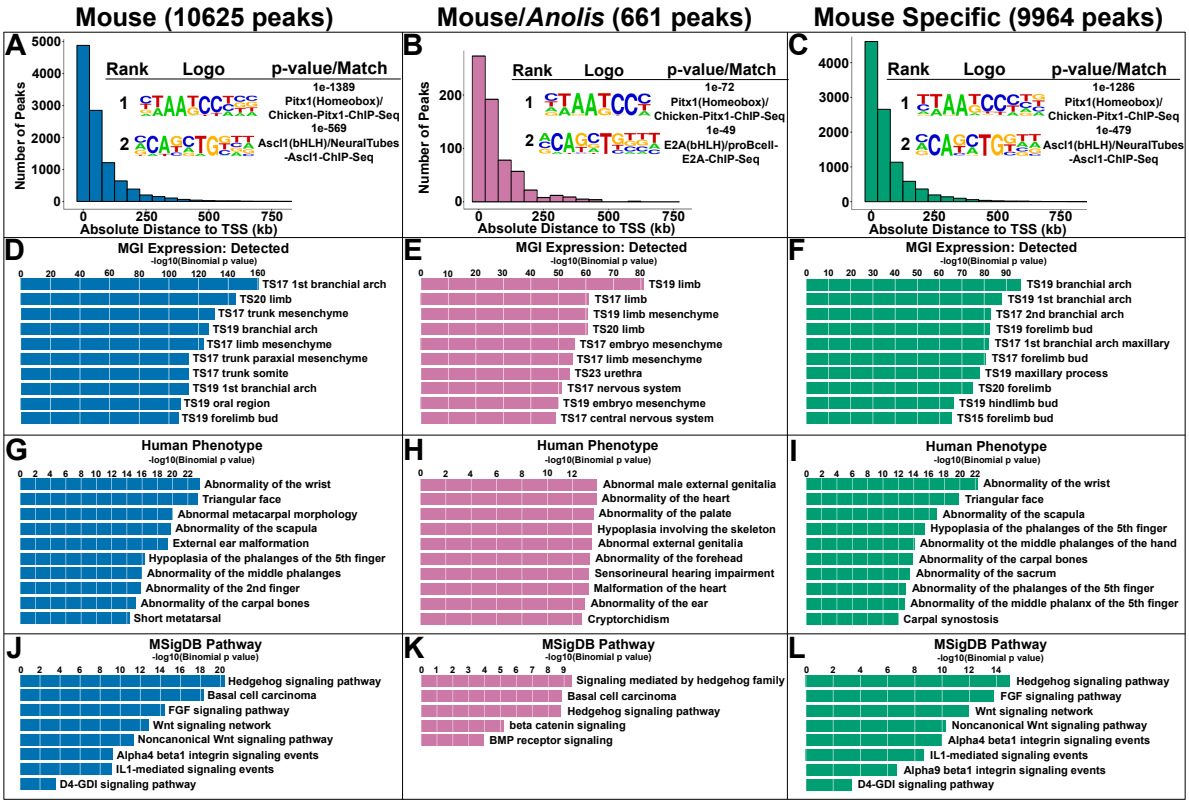

Mouse (10625 peaks)
Mouse/Anolis (661 peaks)
Mouse Specific (9964 peaks)
A
B
C
Number of Peaks
Absolute Distance to TSS (kb)
Rank
Logo
p-value/Match
1e-1389
Pitx1(Homeobox)/
Chicken-Pitx1-ChIP-Seq
1e-569
Ascl1(bHLH)/NeuralTubes
-Ascl1-ChIP-Seq
1e-72
Pitx1(Homeobox)/
Chicken-Pitx1-ChIP-Seq
1e-49
E2A(bHLH)/proBcell-
E2A-ChIP-Seq
1e-1286
Pitx1(Homeobox)/
Chicken-Pitx1-ChIP-Seq
1e-479
Ascl1(bHLH)/NeuralTubes
-Ascl1-ChIP-Seq
D
E
F
MGI Expression: Detected
-log10(Binomial p value)
TS17 1st branchial arch
TS20 limb
TS17 trunk mesenchyme
TS19 branchial arch
TS17 limb mesenchyme
TS17 trunk paraxial mesenchyme
TS17 trunk somite
TS19 1st branchial arch
TS19 oral region
TS19 forelimb bud
TS19 limb
TS17 limb
TS19 limb mesenchyme
TS20 limb
TS17 embryo mesenchyme
TS17 limb mesenchyme
TS23 urethra
TS17 nervous system
TS19 embryo mesenchyme
TS17 central nervous system
TS19 branchial arch
TS19 1st branchial arch
TS17 2nd branchial arch
TS19 forelimb bud
TS17 1st branchial arch maxillary
TS17 forelimb bud
TS19 maxillary process
TS20 forelimb
TS19 hindlimb bud
TS15 forelimb bud
G
H
I
Human Phenotype
-log10(Binomial p value)
Abnormality of the wrist
Triangular face
Abnormal metacarpal morphology
Abnormality of the scapula
External ear malformation
Hypoplasia of the phalanges of the 5th finger
Abnormality of the middle phalanges
Abnormality of the 2nd finger
Abnormality of the carpal bones
Short metatarsal
Abnormal male external genitalia
Abnormality of the heart
Abnormality of the palate
Hypoplasia involving the skeleton
Abnormal external genitalia
Abnormality of the forehead
Sensorineural hearing impairment
Malformation of the heart
Abnormality of the ear
Cryptorchidism
Abnormality of the wrist
Triangular face
Abnormality of the scapula
Hypoplasia of the phalanges of the 5th finger
Abnormality ot the middle phalanges of the hand
Abnormality of the carpal bones
Abnormality of the sacrum
Abnormality of the phalanges of the 5th finger
Abnormality of the middle phalanx of the 5th finger
Carpal synostosis
J
K
L
MSigDB Pathway
-log10(Binomial p value)
Hedgehog signaling pathway
Basal cell carcinoma
FGF signaling pathway
Wnt signaling network
Noncanonical Wnt signaling pathway
Alpha4 beta1 integrin signaling events
IL1-mediated signaling events
D4-GDI signaling pathway
Signaling mediated by hedgehog family
Basal cell carcinoma
Hedgehog signaling pathway
beta catenin signaling
BMP receptor signaling
Hedgehog signaling pathway
FGF signaling pathway
Wnt signaling network
Noncanonical Wnt signaling pathway
Alpha4 beta1 integrin signaling events
IL1-mediated signaling events
Alpha9 beta1 integrin signaling events
D4-GDI signaling pathway

Figure 2.1. Genome-wide enrichment of PITX1 binding in mouse and *Anolis* hindlimbs. **A - C)** Distribution of PITX1 peaks relative to transcription start sites (TSSs) and top two *de novo* motifs enriched within +/- 50 bps of **A)** mouse, **B)** mouse/*Anolis* conserved and **C)** mouse-specific PITX1 peak summits and their best matches to known motifs using HOMER. **D - F)** The top 10 MGI (Mouse Genome Informatics) Expression terms associated with **D)** mouse, **E)** mouse/*Anolis* conserved and **F)** mouse-specific PITX1 peaks. **G - I)** The top 10 Human Phenotype terms associated with **G)** mouse, **H)** mouse/*Anolis* conserved and **I)** mouse-specific PITX1 peaks. **J - L)** The top 10 MSigDB (Molecular Signature Database) Pathway terms associated with **J)** mouse, **K)** mouse/*Anolis* conserved and **L)** mouse-specific PITX1 peaks.

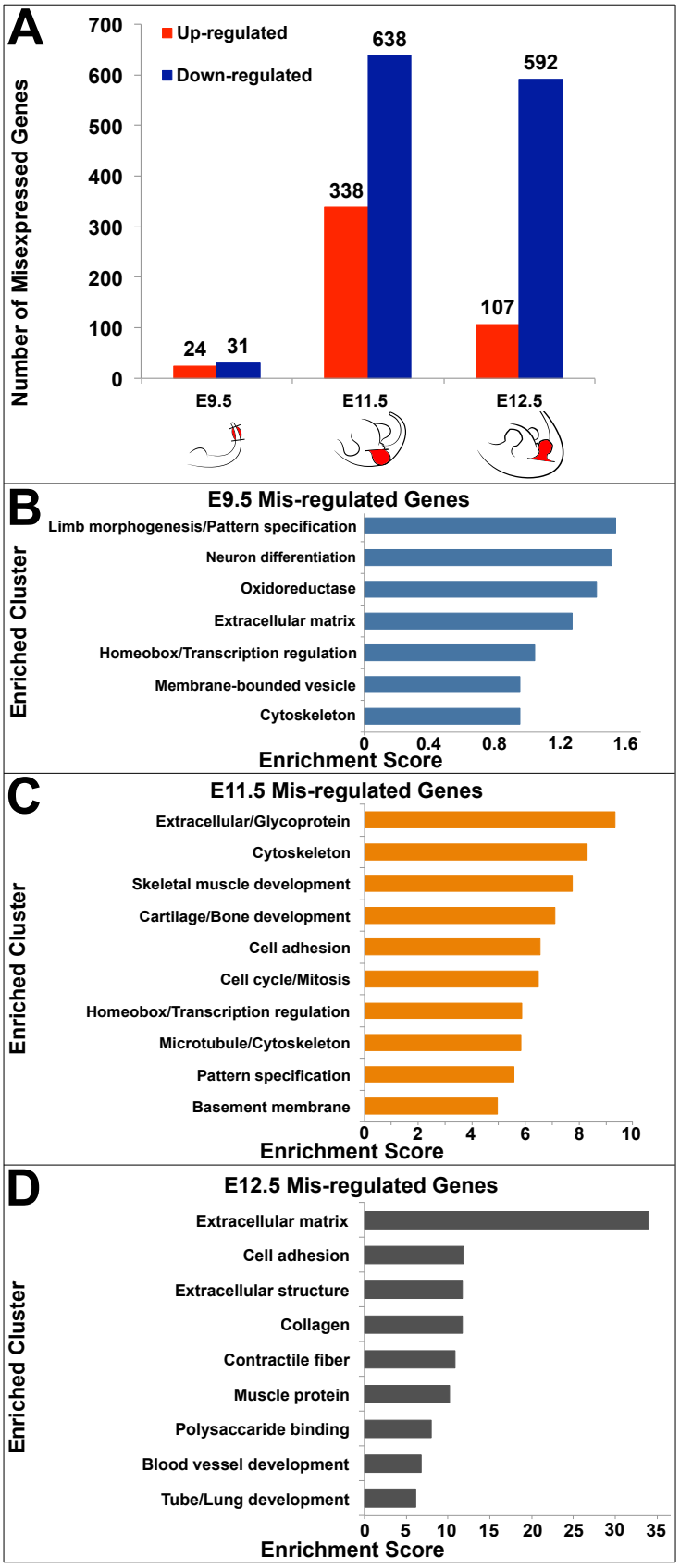
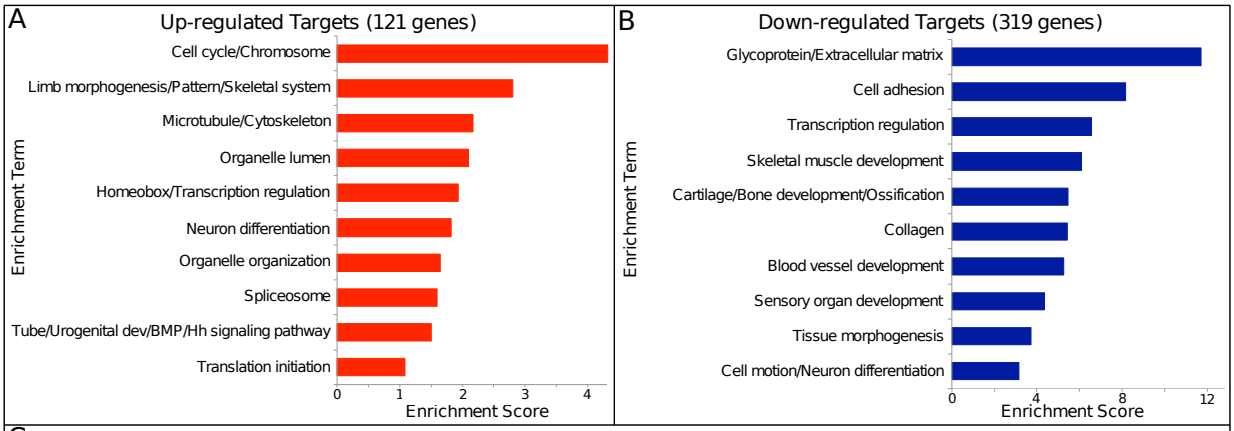

A
Number of Misexpressed Genes
700
600
500
400
300
200
100
0
Up-regulated
Down-regulated
24
31
338
638
107
592
E9.5
E11.5
E12.5
B
E9.5 Mis-regulated Genes
Enriched Cluster
Limb morphogenesis/Pattern specification
Neuron differentiation
Oxidoreductase
Extracellular matrix
Homeobox/Transcription regulation
Membrane-bounded vesicle
Cytoskeleton
0
0.4
0.8
1.2
1.6
Enrichment Score
C
E11.5 Mis-regulated Genes
Enriched Cluster
Extracellular/Glycoprotein
Cytoskeleton
Skeletal muscle development
Cartilage/Bone development
Cell adhesion
Cell cycle/Mitosis
Homeobox/Transcription regulation
Microtubule/Cytoskeleton
Pattern specification
Basement membrane
0
2
4
6
8
10
Enrichment Score
D
E12.5 Mis-regulated Genes
Enriched Cluster
Extracellular matrix
Cell adhesion
Extracellular structure
Collagen
Contractile fiber
Muscle protein
Polysaccaride binding
Blood vessel development
Tube/Lung development
0
5
10
15
20
25
30
35
Enrichment Score

Figure 2.2. Misregulated genes in the hindlimbs of *Pitx1*$^{-/-}$ mice at different stages of embryonic development. **A)** The number of misexpressed genes in *Pitx1*$^{-/-}$ compared to wild-type mouse hindlimbs at E9.5, E11.5 and E12.5. Red shading in cartoons represents the *Pitx1* expression domain in the hindlimbs. Solid lines indicate where cuts were made for collection of tissue for RNA-Seq. **B - D)** The top scoring gene clusters associated with misexpressed genes in *Pitx1*$^{-/-}$ hindlimbs at **B)** E9.5, **C)** E11.5, and **D)** E12.5.



C

PITX1 peaks	Mis-regulated		Up-regulated		Down-regulated	
	p-value	Fold difference	p-value	Fold difference	p-value	Fold difference
Mouse PITX1 peaks	$<2.20 \times 10^{-16}$	2.45	8.97×10^{-6}	1.65	$<2.20 \times 10^{-16}$	2.98
Mouse/*Anolis* PITX1 peaks	2.75×10^{-16}	2.82	7.89×10^{-6}	2.63	1.24×10^{-12}	2.93

Figure 2.3. Putative direct targets of PITX1 are strongly associated with limb patterning, chondrogenesis and myogenesis in mouse hindlimbs. **A-B)** The top 10 enriched clusters associated with **A)** up- and **B)** down-regulated candidate targets of PITX1 in mouse hindlimbs. **C)** Mouse and mouse/*Anolis* conserved PITX1 binding sites are both significantly enriched near genes that are misexpressed in *Pitx1*$^{-/-}$ mouse hindlimbs.

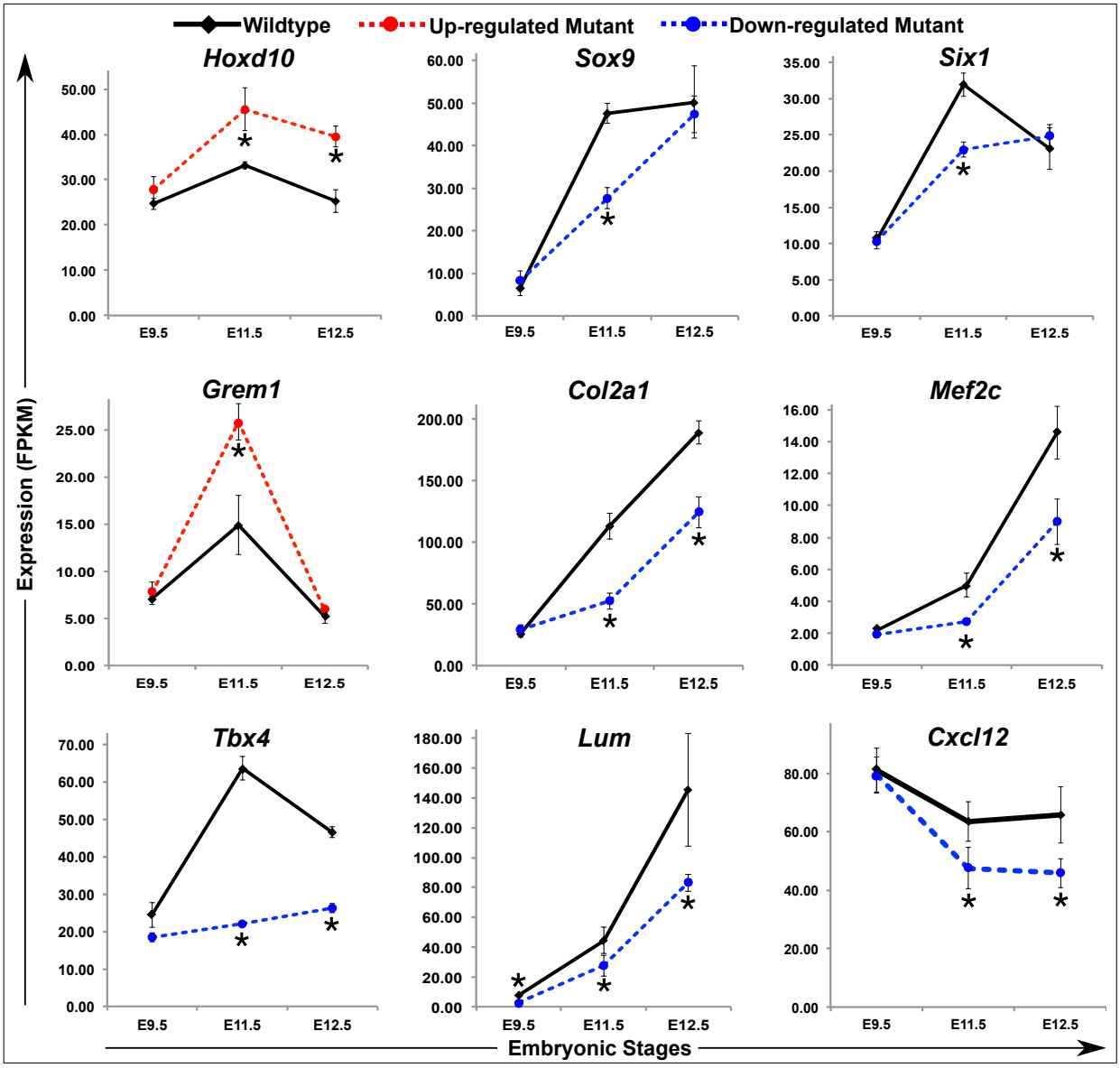

Wildtype
Up-regulated Mutant
Down-regulated Mutant
Hoxd10
Sox9
Six1
Grem1
Col2a1
Mef2c
Tbx4
Lum
Cxcl12
Expression (FPKM)
E9.5
E11.5
E12.5
Embryonic Stages

Figure 2.4. Expression levels of putative PITX1 transcriptional targets in wild-type and *Pitx1*$^{-/-}$ hindlimbs at different developmental stages. Time charts display the average RNA-Seq expression level from three independent biological replicates. Asterisks represent developmental stages with significantly different expression in *Pitx1*$^{-/-}$ compared to wild-type hindlimbs (FDR-adjusted p-value < 0.05). Error bars represent the standard error of the mean.

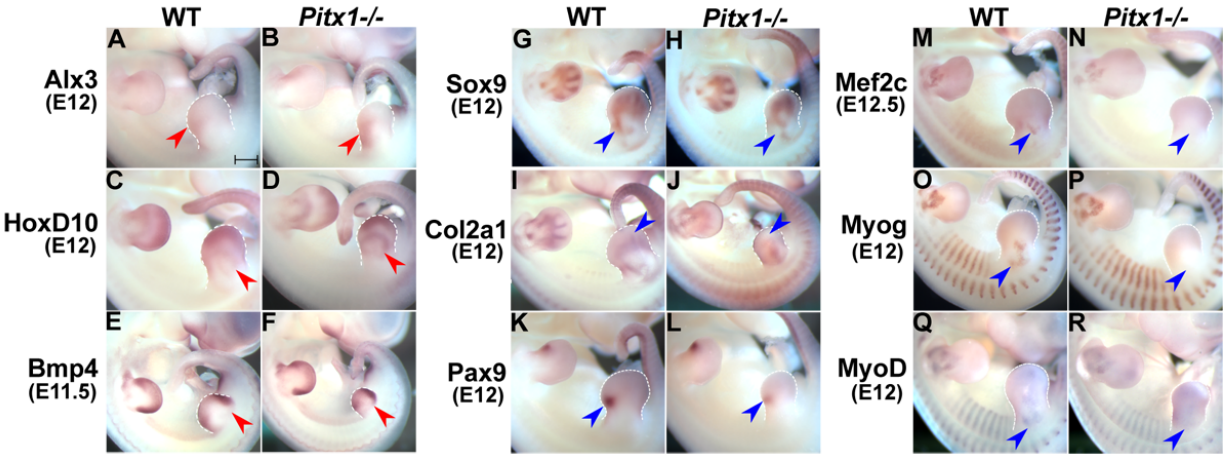

WT
Pitx1-/-
WT
Pitx1-/-
WT
Pitx1-/-
A
B
G
H
M
N
Alx3
(E12)
Sox9
(E12)
Mef2c
(E12.5)
C
D
I
J
O
P
HoxD10
(E12)
Col2a1
(E12)
Myog
(E12)
E
F
K
L
Q
R
Bmp4
(E11.5)
Pax9
(E12)
MyoD
(E12)

Figure 2.5. Whole-mount *in situ* hybridization of PITX1-dependent genes in wide-type and *Pitx1*$^{-/-}$ mouse embryos. **A-F)** The expression pattern of limb patterning genes (*Alx3, HoxD10* and *Bmp4*). **G-L)** The expression pattern of genes related to cartilage development (*Sox9, Col2a1* and *Pax9*). **M-R)** The expression pattern of genes related to muscle development (*Mef2c, MyoD* and *Myogenin*). Red arrowheads point to regions with increased expression while blue ones point to regions with decreased expression in *Pitx1*$^{-/-}$ mouse embryos. All images were taken at the same magnification. Scale bar = 500 μm.

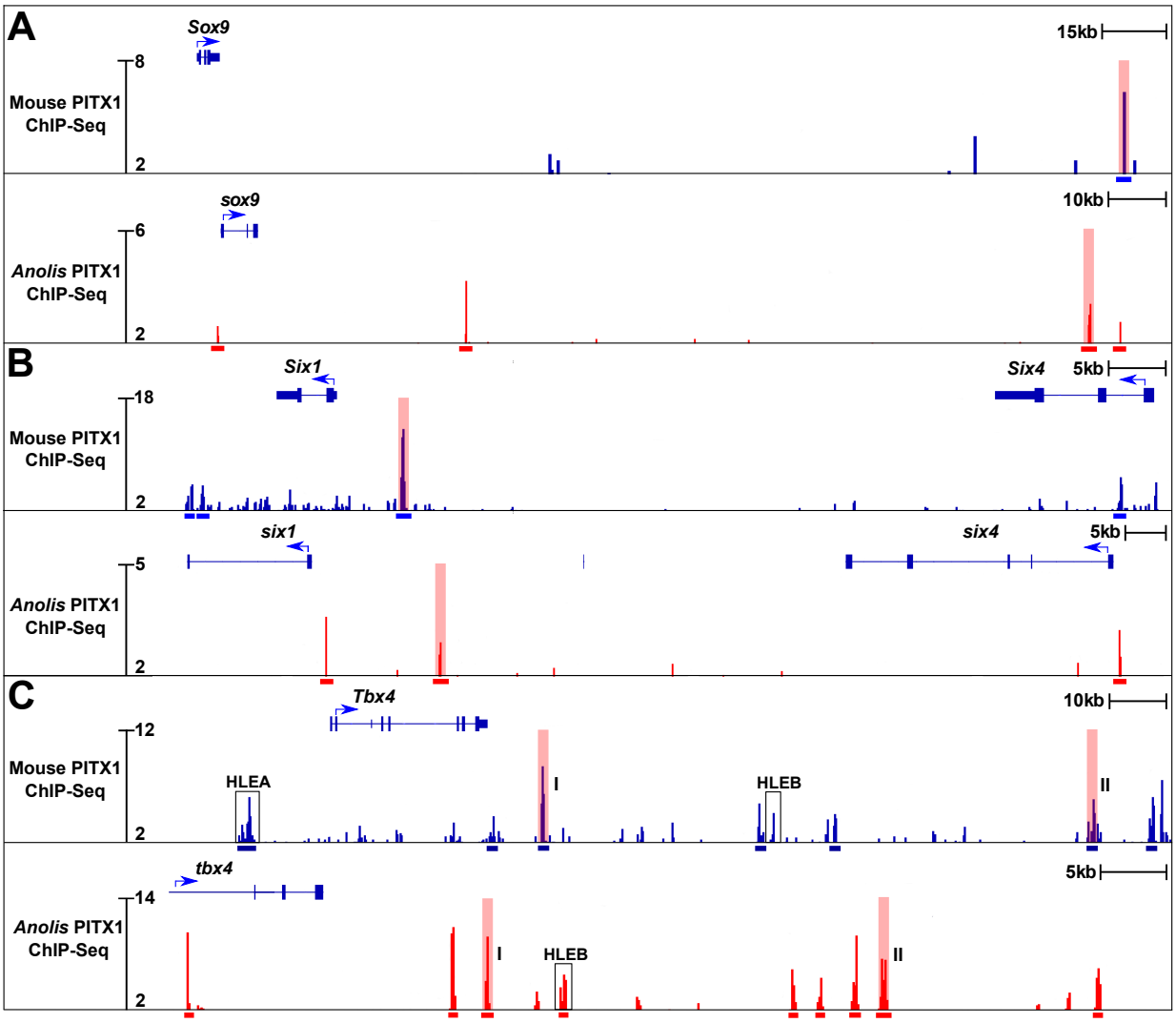

A
Sox9
15kb
8
Mouse PITX1
ChIP-Seq
2
sox9
10kb
6
Anolis PITX1
ChIP-Seq
2
B
Six1
Six4
5kb
18
Mouse PITX1
ChIP-Seq
2
six1
six4
5kb
5
Anolis PITX1
ChIP-Seq
2
C
Tbx4
10kb
12
Mouse PITX1
ChIP-Seq
HLEA
I
HLEB
II
2
tbx4
5kb
14
Anolis PITX1
ChIP-Seq
I
HLEB
II
2

Figure 2.6. PITX1 ChIP-seq profiles in mouse and *Anolis* hindlimbs at putative PITX1 transcriptional targets. **A)** *Sox9,* **B)** *Six1* and **C)** *Tbx4*. Blue bars denote mouse PITX1 peaks and red bars denote *Anolis* PITX1 peaks. Conserved binding events identified in both species are highlighted in pink. I and II are used to distinguish two different pairs of conserved peaks at the *Tbx4* locus. The location of known *Tbx4* hindlimb-specific enhancers, HLEA and HLEB, are outlined by black boxes.

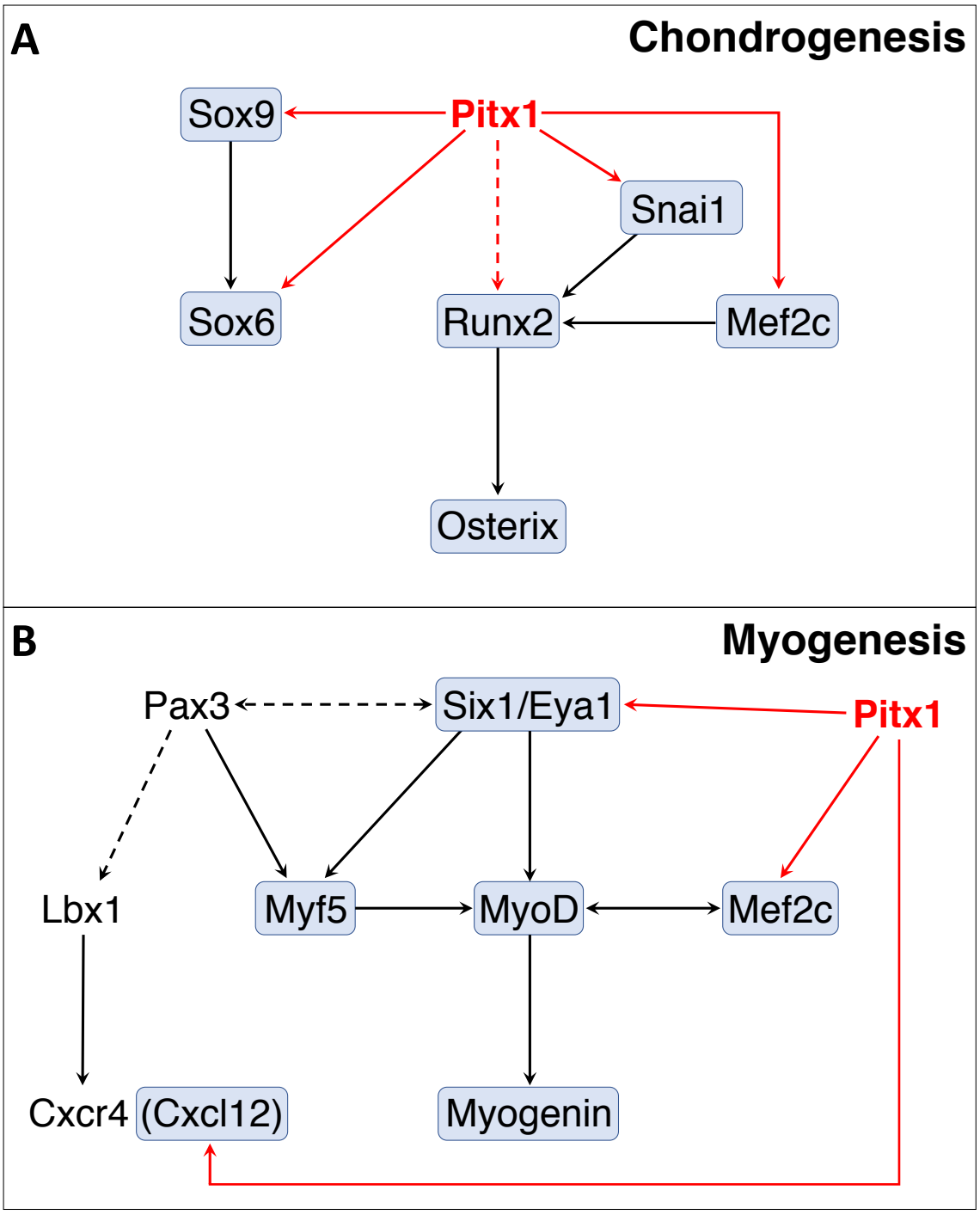

A
Chondrogenesis
Sox9
Pitx1
Snai1
Sox6
Runx2
Mef2c
Osterix
B
Myogenesis
Pax3
Six1/Eya1
Pitx1
Lbx1
Myf5
MyoD
Mef2c
Cxcr4
(Cxcl12)
Myogenin

Figure 2.7. Proposed PITX1 regulatory interactions with components of the chondrogenesis and myogenesis regulatory networks. Solid arrows represent direct regulatory interactions and dashed arrows represent indirect or unproven regulatory interactions. Red arrows highlight PITX1 regulation of downstream targets identified from our data and black arrows represent interactions identified from the published literature. Genes in shaded boxes are down-regulated in *Pitx1*$^{-/-}$ hindlimbs.

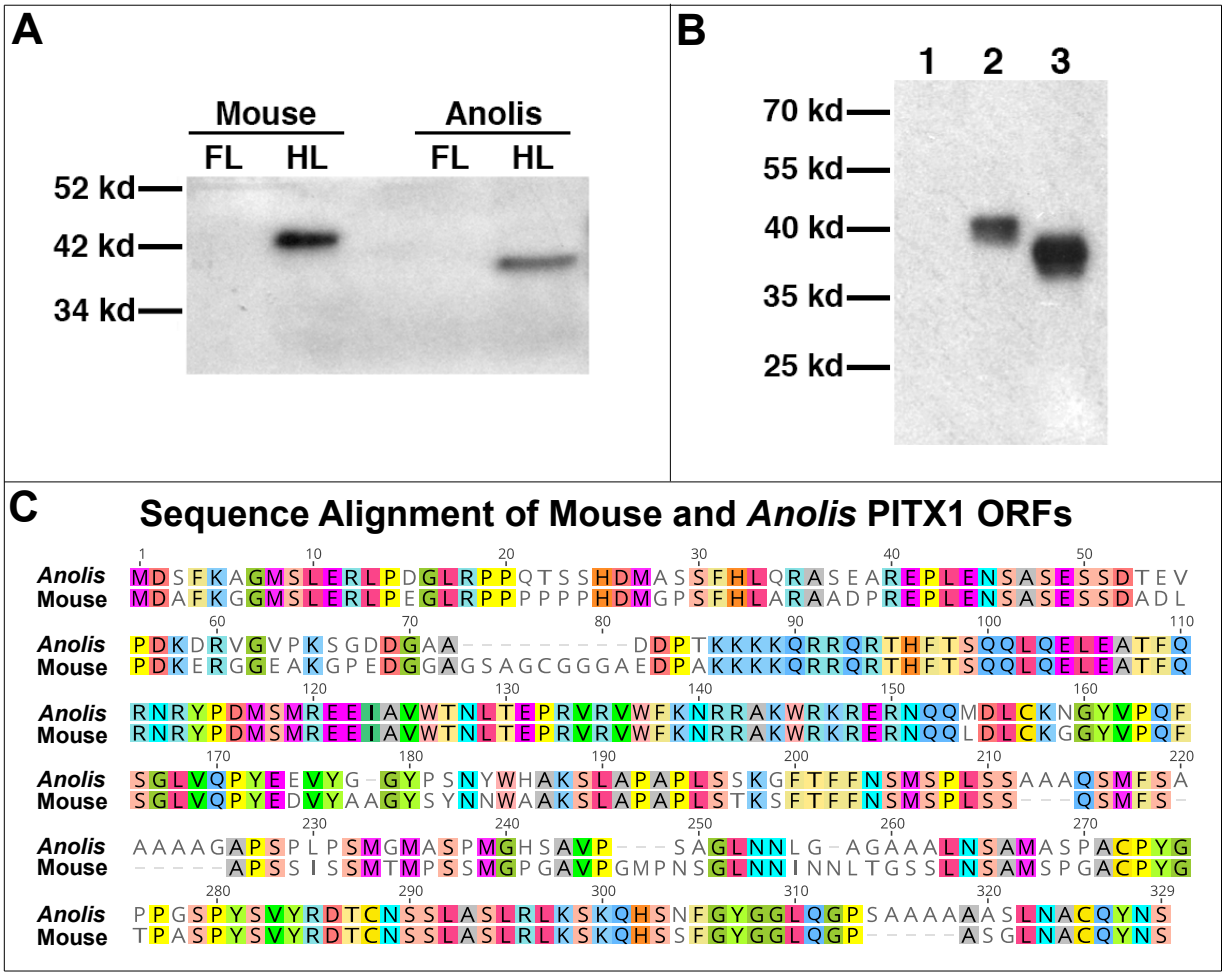

A
Mouse
Anolis
FL
HL
FL
HL
52 kd
42 kd
34 kd
B
1
2
3
70 kd
55 kd
40 kd
35 kd
25 kd
C
Sequence Alignment of Mouse and Anolis PITX1 ORFs
Anolis
Mouse

Figure S2.1. PITX1 antibody specificity and sequence alignment of mouse and *Anolis* PITX1. A) Western blot analysis of PITX1-specific antibody against protein extracts from mouse and Anolis hindlimbs. The antibody recognizes a hindlimb-specific protein in the embryonic hindlimbs of both species. 20 micrograms of protein were loaded onto each lane. B) Western blot using PITX1 antibody on HEK293T cells transformed with expression constructs containing GFP (lane 1) or the full ORF of Mouse PITX1 (lane 2) or Anolis PITX1 (lane 3). Mouse and Anolis PITX1 proteins exhibit different migration sizes despite having similar molecular weights. 10 micrograms of protein were loaded onto each lane. C) Sequence alignment of predicted mouse and Anolis PITX1 open reading frames (ORFs). The overall amino acid identify is 73%. Identical amino acids are highlighted in color.

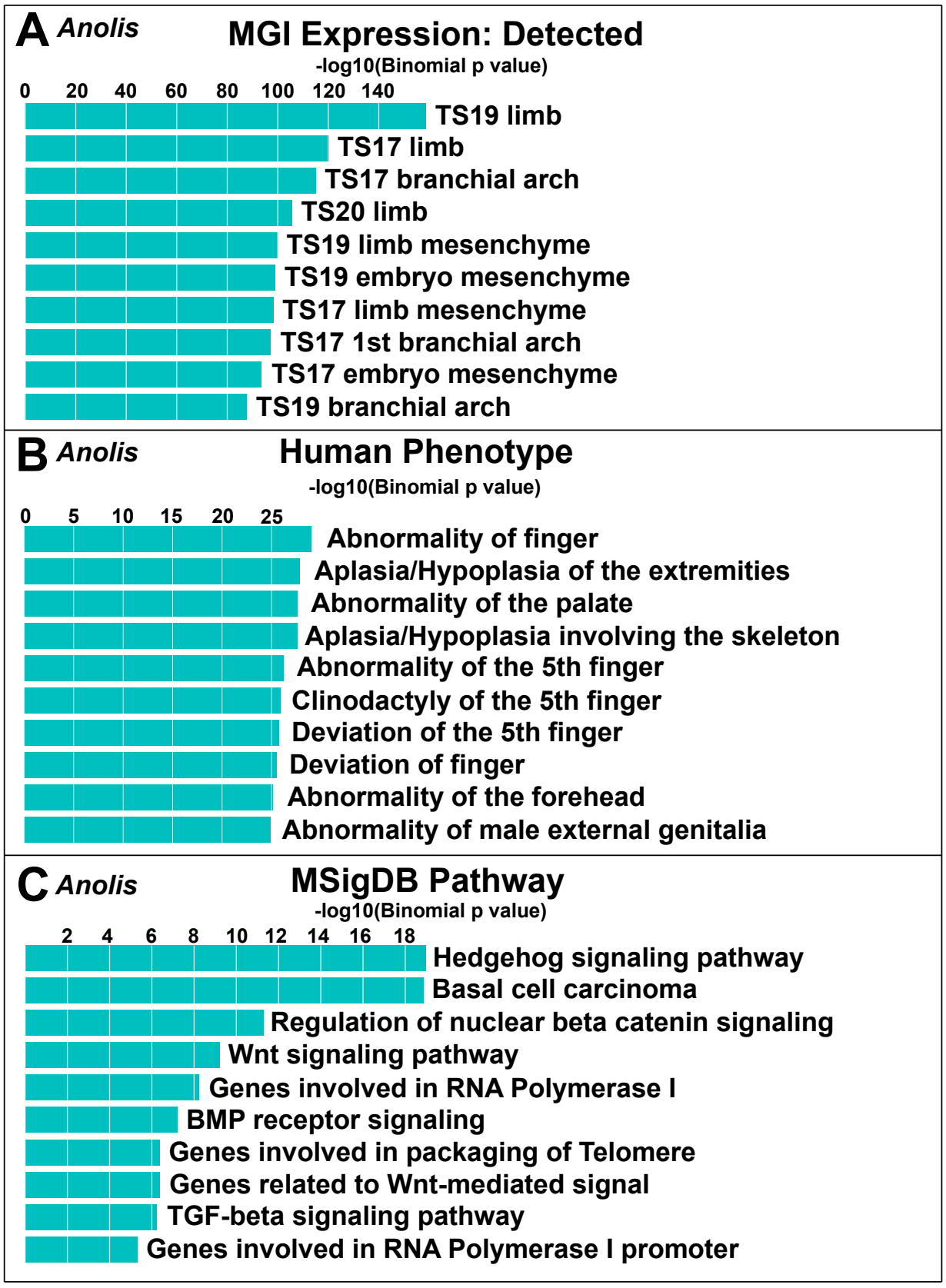

A Anolis
MGI Expression: Detected
-log10(Binomial p value)
0 20 40 60 80 100 120 140
TS19 limb
TS17 limb
TS17 branchial arch
TS20 limb
TS19 limb mesenchyme
TS19 embryo mesenchyme
TS17 limb mesenchyme
TS17 1st branchial arch
TS17 embryo mesenchyme
TS19 branchial arch
B Anolis
Human Phenotype
-log10(Binomial p value)
0 5 10 15 20 25
Abnormality of finger
Aplasia/Hypoplasia of the extremities
Abnormality of the palate
Aplasia/Hypoplasia involving the skeleton
Abnormality of the 5th finger
Clinodactyly of the 5th finger
Deviation of the 5th finger
Deviation of finger
Abnormality of the forehead
Abnormality of male external genitalia
C Anolis
MSigDB Pathway
-log10(Binomial p value)
2 4 6 8 10 12 14 16 18
Hedgehog signaling pathway
Basal cell carcinoma
Regulation of nuclear beta catenin signaling
Wnt signaling pathway
Genes involved in RNA Polymerase I
BMP receptor signaling
Genes involved in packaging of Telomere
Genes related to Wnt-mediated signal
TGF-beta signaling pathway
Genes involved in RNA Polymerase I promoter

Figure S2.2. Gene ontology associations of PITX1 binding sites in *Anolis* hindlimbs. The top 10 GO terms associated with the subset of *Anolis* PITX1 peaks that have detectable sequence conservation in the mouse genome for the following gene ontology databases: A) Mouse Genome Informatics (MGI) Expression, B) Human Phenotype and C) MSigDB (Molecular Signature Database) Pathway. Analysis was performed using GREAT (McLean et al., 2010).

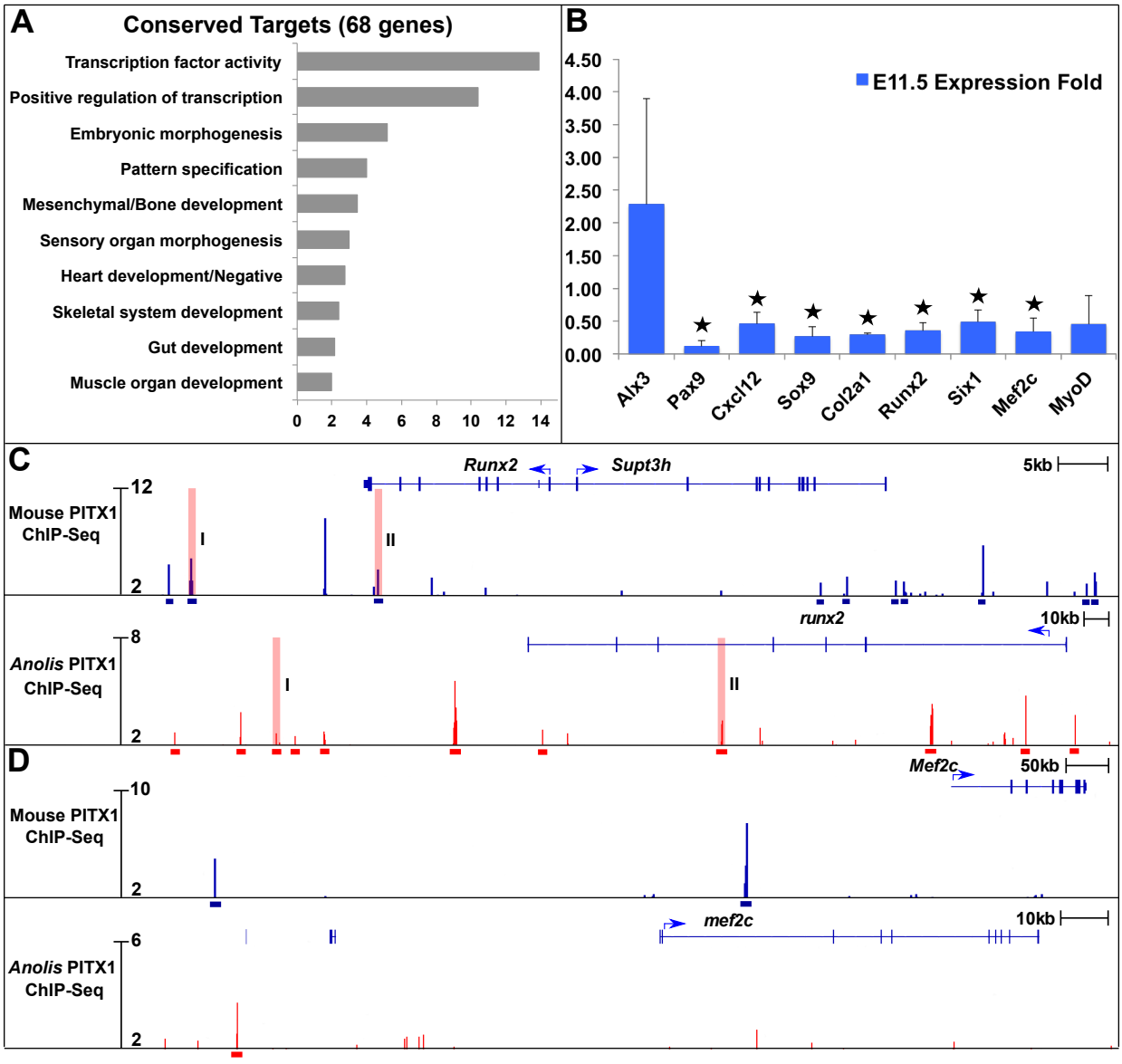

A
Conserved Targets (68 genes)
Transcription factor activity
Positive regulation of transcription
Embryonic morphogenesis
Pattern specification
Mesenchymal/Bone development
Sensory organ morphogenesis
Heart development/Negative
Skeletal system development
Gut development
Muscle organ development
0 2 4 6 8 10 12 14
B
E11.5 Expression Fold
4.50
4.00
3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
Alx3
Pax9
Cxcl12
Sox9
Col2a1
Runx2
Six1
Mef2c
MyoD
C
Runx2
Supt3h
5kb
Mouse PITX1 ChIP-Seq
12
2
I
II
runx2
10kb
Anolis PITX1 ChIP-Seq
8
2
I
II
D
Mef2c
50kb
Mouse PITX1 ChIP-Seq
10
2
mef2c
10kb
Anolis PITX1 ChIP-Seq
6
2

Figure S2.3. The binding activities and expression analyses of PITX1 putative direct targets. A) The top 10 enriched clusters associated with conserved targets of PITX1 in developing mouse hindlimbs. B) qRT-PCR of limb patterning, chondrogenesis- and myogenesis-related genes in wide-type versus Pitx1-/- mouse hindlimbs at E11.5. Asterisks indicate $p < 0.05$ (Two-tailed t test). C-D) Comparison of PITX1 binding profiles in mouse and Anolis hindlimbs showing the positions of mouse, mouse/*Anolis* conserved and *Anolis* PITX1 peaks in the neighboring regions of putative direct targets C) *Runx2* and D) *Mef2c*. Blue bars underline identified mouse PITX1 peaks and red bars underline identified Anolis PITX1 peaks. Conserved peaks identified in both species are highlighted in pink. I and II are used to distinguish two different pairs of conserved peaks at the *Runx2* and *Mef2c* loci.

Table S2.1. Top 10 enriched motifs identified in mouse, *Anolis*, conserved, and mouse-specific PITX1 peaks.

Rank	Mouse De Novo Motif	P-value	Log P-value	% of Targets	% of Background	STD(Bg STD)	Best Match
1		1e-1389	-3.2E+03	58.79%	18.52%	22.4bp (44.5bp)	Pitx1(Homeobox)/Chicken-Pitx1-ChIP-Seq(GSE38910)/Homer(0.948)
2		1e-569	-1.3E+03	31.02%	10.05%	27bp (46.3bp)	Ascl1(bHLH)/NeuralTubes-Ascl1-ChIP-Seq(GSE55840)/(0.957)
3		1e-452	-1.1E+03	43.76%	21.28%	27.2bp (45.6bp)	ERG(ETS)/VCaP-ERG-ChIP-Seq(GSE14097)/(0.842)
4		1e-371	-8.6E+02	45.98%	25.01%	27.5bp (45.9bp)	ZFX(Zf)/mES-Zfx-ChIP-Seq(GSE11431)/Homer(0.665)
5		1e-342	-7.9E+02	62.54%	41.15%	28.2bp (46.7bp)	CDC5(MYB)/Arabidopsis thaliana/AthaMap(0.729)
6		1e-340	-7.8E+02	42.7%	23.05%	26.6bp (42.8bp)	HOXD13(Homeobox)/Chicken-Hoxd13-ChIP-Seq(GSE38910)/Homer(0.822)
7		1e-280	-6.5E+02	28.34%	13.2%	27.1bp (45.0bp)	ADR1/Literature(Harbison)/Yeast(0.627)
8		1e-207	-4.8E+02	24.4%	12%	28.4bp (46.0bp)	RIM101/Literature(Harbison)/Yeast(0.746)
9		1e-146	-3.4E+02	6.77%	1.68%	28.6bp (39.3bp)	CG7056/dmmpmm(Noyes_hd)/fly(0.779)
10		1e-142	-3.3E+02	14.15%	6.22%	27.3bp (46.1bp)	EBF1(EBF)/Near-E2A-ChIP-Seq(GSE21512)/(0.873)

Rank	*Anolis* De Novo Motif	P-value	Log P-value	% of Targets	% of Background	STD(Bg STD)	Best Match
1		1e-1434	-3.3E+03	46.17%	16.9%	22.7bp (42.7bp)	Pitx1(Homeobox)/Chicken-Pitx1-ChIP-Seq(GSE38910)/Homer(0.920)
2		1e-683	-1.6E+03	45.62%	24.39%	26.6bp (45.4bp)	SCL(bHLH)/HPC7-Scl-ChIP-Seq(GSE13511)/Homer(0.919)
3		1e-573	-1.3E+03	32.53%	15.38%	27.2bp (44.1bp)	EWS:ERG-fusion(ETS)/CADO_ES1-EWS:ERG-ChIP-Seq(SRA014231)/Homer(0.847)
4		1e-499	-1.2E+03	52.41%	33.35%	27.8bp (44.4bp)	RIM101/Literature(Harbison)/Yeast(0.729)
5		1e-491	-1.1E+03	36.25%	19.42%	27.7bp (44.6bp)	Smad3(MAD)/NPC-Smad3-ChIP-Seq(GSE36673)/(0.689)
6		1e-401	-9.2E+02	31.49%	16.98%	27.3bp (43.9bp)	EBF1(EBF)/Near-E2A-ChIP-Seq(GSE21512)/Homer(0.801)
7		1e-377	-8.7E+02	36.17%	21.21%	27.3bp (42.5bp)	CDX1/MA0878.1/Jaspar(0.888)
8		1e-319	-7.4E+02	24.77%	13.01%	26.8bp (45.8bp)	POL010.1_DCE_S_III/Jaspar(0.732)
9		1e-225	-5.2E+02	31.32%	20.12%	28.6bp (44.9bp)	CST6(MacIsaac)/Yeast(0.961)
10		1e-222	-5.1E+02	20.88%	11.61%	27.8bp (50.4bp)	SeqBias: G/A bias(0.968)

Rank	Conserved De Novo Motif	P-value	Log P-value	% of Targets	% of Background	Best Match
1		1e-72	-1.7E+02	53.56%	17.71%	Pitx1(Homeobox)/Chicken-Pitx1-ChIP-Seq(GSE38910)/Homer(0.962)
2		1e-49	-1.1E+03	34.34%	9.63%	E2A(bHLH)/proBcell-E2A-ChIP-Seq(GSE21978)/Homer(0.914)
3		1e-35	-8.2E+01	49.17%	23.59%	EWS:FLI1-fusion(ETS)/SK_N_MC-EWS:FLI1-ChIP-Seq(SRA014231)/Homer(0.813)
4		1e-28	-6.6E+01	35.85%	15.47%	Meis1(Homeobox)/MastCells-Meis1-ChIP-Seq(GSE48085)/Homer(0.744)
5		1e-23	-5.4E+01	41.3%	21.5%	Abd-B/dmmpmm(Bergman)/fly(0.709)
6		1e-20	-4.7E+01	6.35%	0.41%	CDX2/MA0465.1/Jaspar(0.848)
7		1e-17	-4.1E+01	15.28%	4.78%	POL009.1_DCE_S_II/Jaspar(0.800)
8		1e-16	-3.9E+01	11.8%	3.07%	PHA-4(Forkhead)/cElegans-Embryos-PHA4-ChIP-Seq(modEncode)/(0.787)
9		1e-15	-3.6E+01	3.78%	0.1%	SeqBias: polyC-repeat(0.849)
10		1e-13	-3.1E+01	4.99%	0.53%	REI1/MA0364.1/Jaspar(0.802)

Rank	Mouse-specific De Novo Motif	P-value	Log P-value	% of Targets	% of Background	STD(Bg STD)	Best Match
1		1e-1286	-3.0E+03	52.61%	14.42%	22.0bp (44.3bp)	Pitx1(Homeobox)/Chicken-Pitx1-ChIP-Seq(GSE38910)/Homer(0.956)
2		1e-479	-1.1E+03	28.71%	9.41%	26.9bp (45.7bp)	Ascl1(bHLH)/NeuralTubes-Ascl1-ChIP-Seq(GSE55840)/Homer(0.946)
3		1e-465	-1.1E+03	58.95%	33.47%	27.6bp (46.4bp)	EWS:FLI1-fusion(ETS)/SK_N_MC-EWS:FLI1-ChIP-Seq(SRA014231)/Homer(0.872)
4		1e-398	-9.2E+02	51.95%	28.88%	28.2bp (44.8bp)	CDC5(MYB)/Arabidopsis thaliana/AthaMap(0.622)
5		1e-310	-7.2E+02	47.71%	27.6%	27.2bp (44.3bp)	LIN54/MA0619.1/Jaspar(0.791)
6		1e-208	-4.8E+02	18.48%	7.5%	27.2bp (41.7bp)	Abd-B/MA0165.1/Jaspar(0.929)
7		1e-199	-4.6E+02	43.71%	27.75%	26.3bp (43.0bp)	MET28/MA0332.1/Jaspar(0.732)
8		1e-146	-3.4E+02	20.34%	10.3%	27.8bp (46.8bp)	Tbx20(T-box)/Heart-Tbx20-ChIP-Seq(GSE29636)/Homer(0.766)
9		1e-142	-3.3E+02	15.15%	6.7%	27.4bp (41.5bp)	Kr/dmmpmm(SeSiMCMC)/fly(0.710)
10		1e-129	-3.0E+02	9.46%	3.29%	27.8bp (47.8bp)	Tbx20(T-box)/Heart-Tbx20-ChIP-Seq(GSE29636)/Homer(0.710)

Table S2.2. Representatives of PITX1 peaks identified in mouse and *Anolis* hindlimbs. Coordinates of PITX1 mouse peaks (mm9), *Anolis* peaks (anoCar2), *Anolis* lifted-over peaks (mm9) and mouse/*Anolis* conserved peaks (mm9) are listed.

Mouse

Chromosome	start	end
chr1	3641251	3641880
chr1	4833459	4834433
chr1	4847289	4847891
chr1	4900698	4901098
chr1	7300850	7301432
chr1	7686479	7686824
chr1	9573461	9574408
chr1	10000552	10001016
chr1	10074353	10074946
chr1	11017007	11017583
chr1	11320090	11320743
chr1	11734890	11735431
chr1	11990451	11990985
chr1	12041970	12042369
chr1	12446916	12447405
chr1	12500159	12500862
chr1	12511078	12511778
chr1	12561859	12562554
chr1	12605244	12606069
chr1	12680516	12681224
chr1	12682323	12682838
chr1	12804908	12805367
chr1	13362533	13362974
chr1	13592921	13593334
chr1	13719989	13720552
chr1	13761132	13761645
chr1	13772607	13772903
chr1	13836622	13837073
chr1	13849779	13850460
chr1	13850891	13851938
chr1	13917842	13918614
chr1	13921229	13921903
chr1	14026183	14026692
chr1	14135938	14136565
chr1	14297021	14297744
chr1	14334446	14334968
chr1	14347592	14348321
chr1	14412502	14413053
chr1	14419325	14419831
chr1	14480148	14480591
chr1	14557084	14557682
chr1	14625538	14626422
chr1	14790420	14790849
chr1	14831376	14833009
chr1	15024070	15024612
chr1	16238873	16239722
chr1	16283183	16283973
chr1	16596716	16597267
chr1	16999499	17000031
chr1	18096151	18096674

Anolis

Chromosome	Start	End
chr1	592459	592744
chr1	620893	621112
chr1	665027	665480
chr1	683795	684413
chr1	707810	708028
chr1	906229	906651
chr1	933517	934253
chr1	937943	938796
chr1	1150909	1151116
chr1	1327621	1327790
chr1	2376710	2377119
chr1	2694855	2695253
chr1	2810959	2811460
chr1	3549579	3549984
chr1	3590607	3590898
chr1	3882897	3883109
chr1	4357540	4357718
chr1	4640235	4640483
chr1	4707999	4708406
chr1	4717027	4717371
chr1	4717437	4718074
chr1	4720218	4720393
chr1	4758441	4758704
chr1	4760258	4760678
chr1	5042037	5042352
chr1	5093193	5093688
chr1	5143272	5143761
chr1	5148523	5149056
chr1	5164365	5164552
chr1	5230271	5230510
chr1	5266374	5266694
chr1	5289840	5290685
chr1	5469268	5469680
chr1	5505953	5506442
chr1	5509158	5509591
chr1	5537037	5537254
chr1	5914715	5914963
chr1	6347879	6348361
chr1	6535921	6536424
chr1	6631103	6631668
chr1	6671678	6672577
chr1	6760836	6761121
chr1	6769742	6770473
chr1	6786785	6787398
chr1	6824388	6825177
chr1	6833477	6834441
chr1	6853257	6853699
chr1	6868783	6868976
chr1	6882888	6883214

Conserved

Chromosome	Start	End
chr1	12500159	12500862
chr1	14831376	14833009
chr1	48316517	48317019
chr1	56890186	56890661
chr1	61168474	61169309
chr1	62678366	62678981
chr1	76968936	76969520
chr1	76978526	76979093
chr1	77137302	77138063
chr1	81740730	81741882
chr1	82199283	82200204
chr1	92015584	92016285
chr1	140988641	140989102
chr1	140989229	140989833
chr1	164146757	164147313
chr1	164147333	164151771
chr1	169732556	169733495
chr1	170226367	170226756
chr1	186128577	186129211
chr1	186158575	186159607
chr1	186177765	186178336
chr1	186236466	186236864
chr1	186236866	186237191
chr1	187657960	187658621
chr1	190550577	190551449
chr1	194625523	194626453
chr10	17446626	17448163
chr10	23281953	23282570
chr10	23430326	23432276
chr10	29392337	29392841
chr10	37841749	37842166
chr10	42890445	42891517
chr10	50483166	50483901
chr10	52811249	52811867
chr10	56759953	56760818
chr10	59342033	59342497
chr10	66648107	66649225
chr10	67670509	67671465
chr10	67671508	67671999
chr10	67817170	67817790
chr10	83867918	83868201
chr10	95942360	95942854
chr10	98725158	98726347
chr10	98727435	98728889
chr10	98731082	98732182
chr10	115309429	115310058
chr10	119854933	119856684
chr10	119884867	119885438
chr10	119885451	119886126

Table S2.3. Transcriptome analyses (RPKM) between the developing hindlimbs of wild-type and *Pitx1*$^{-/-}$ in mouse embryos. E9.5, top E11.5 and E12.5 DE genes are listed.

E9.5_DE_Genes	WT_RPKM	Mutant_RPKM
Cwc22	30.5897	88.6103
Ntrk2	0.459703	1.16248
Hoxd12	3.03072	7.53455
Cxcl14	8.90267	21.3072
Gnpda1	2.07765	4.89893
Lrrn3	2.08189	4.72683
Msx3	1.93099	4.31589
Gpr137b	2.20342	4.58121
Ckb	31.205	63.9681
Hdhd3	5.79958	11.4809
Cabp1	6.17091	12.046
Aldh1b1	4.26914	8.27561
Col23a1	8.75142	15.9479
Miat	1.95565	3.53924
Rps15a	15.2481	24.1864
Rabl2	13.7092	21.7302
Hba-a1,Hba-a2	865.124	1368.94
Hoxd11	10.1825	15.3615
Isl1	15.0243	22.6326
Alad	26.6519	39.4395
Fam181b	32.3672	46.4324
Ldhb	47.2719	66.9844
Aldh1a2	77.1618	107.497
Efcab9	0	0.743611
1700017J07Rik	0.753254	0
Gm7244	0.773053	0
Prtg	29.0388	21.2068
Col1a2	34.2969	24.1526
Dusp6	68.6996	48.0477
Sgce	39.4617	27.0543
Amot	29.2119	19.8739
Hapln1	29.1789	19.7094
Acta2	75.0798	46.4861
Mbnl3	3.11254	1.83359
Steap2	2.44875	1.42998
Hoxc8	16.4609	9.53272
Cbl	9.30667	5.262
Phlda2	54.4413	29.708
Tubb2a	28.8684	15.2509
Postn	43.8099	23.0097
Gm15772	393.19	203.979
Cdx4	19.7112	9.42677
Lhx2	6.15045	2.8231
Nfam1	1.80229	0.772171
Rgs9bp	1.46873	0.621175
Pitx1	54.7924	22.5044
Dcn	4.17128	1.70544
Afp	5.17829	2.03612
Dcaf12l1	4.24132	1.6654
Gm7120	20.5291	7.35507
Lum	7.55861	2.63739
Ahcy	6.89368	2.27906
Rgn	2.97905	0.962457
Amph	5.02575	1.26837
Gpr137b-ps	3.66223	0.572625

E11.5_DE_Genes	WT_RPKM	Mutant_RPKM
Gpx6	0.123836	2.50368
AA388235	0.899564	4.38258
Zfp42	0.185373	0.898901
Xlr3a	0.41979	1.95116
Gbx2	0.438629	1.90423
Bco2	0.121492	0.506394
Lbx2	0.498235	2.04662
Myl4	1.77474	0.438993
Mylpf	23.0776	5.64899
Vgll2	4.06698	0.994768
Tmem8c	2.3782	0.575498
Hrc	0.735877	0.175308
Zic1	0.662055	0.155466
Dok7	0.316239	0.0709495
Gmnc	0.818283	0.182208
Tnc	5.29093	1.16917
Myog	20.3572	4.40807
T	0.775334	0.167593
Atp2a1	3.35543	0.691147
Fitm1	2.81556	0.545998
Rapsn	1.14683	0.219753
Klhl41	2.72152	0.520168
Arpp21	1.36336	0.2562
Fam155a	0.338274	0.062701
Tnni1	4.39654	0.8131
Col11a2	3.28854	0.606641
Zfp429	3.25466	0.593001
Chrng	1.2889	0.227618
Trim55	0.75871	0.132487
Cntnap3	0.283745	0.04887
Pitx1	119.705	20.1692
Acta2	16.965	2.80056
Alpk2	0.460704	0.0741868
Mylk4	0.57292	0.0918443
Tnnc1	8.43364	1.29744
Crhbp	0.918598	0.139248
Hand1	0.883962	0.133735
Asb2	0.278211	0.0407787
Eci3	0.666236	0.0961995
Gm7325	3.93141	0.53095
Apobec2	0.95319	0.127463
Actg2	0.675778	0.0882698
Pax7	0.414831	0.0516825
Actn2	1.75892	0.215423
E130114P18Rik	6.76123	0.819635
Smyd1	1.72708	0.207884
Lect1	8.70069	1.02733
Mip	1.16066	0.134733
Neb	0.35508	0.0395189
Ablim3	0.532497	0.0584637
Myl1	3.13834	0.338199
Unc45b	0.999196	0.103853
Tnnt2	3.19537	0.313995
Tnnc2	2.77563	0.247112
Actc1	44.5338	3.89511
Matn4	14.9946	1.06035
Casq2	1.06624	0.070556
B3galt5	0.266473	0.0169028
Afp	4.23262	0.260381
Acan	0.782567	0.0438972
Erv3	0.532137	0.0198609

E12.5_DE_Genes	WT_RPKM	Mutant_RPKM
Kirrel2	0.0675347	0.564369
A530016L24Rik	0.0660723	0.433277
Anxa8	0.671731	3.77352
Mpo	0.0908115	0.395144
Sbspon	0.912329	0.220942
Oprk1	0.375426	0.0905448
Chodl	15.997	3.84472
Lypd1	4.29085	1.01786
Slc26a7	6.37609	1.4657
Tcf24	0.536491	0.122608
Aqp3	1.7787	0.405177
C3	1.52637	0.343915
Cytl1	22.0629	4.92166
Slc27a2	0.466744	0.100845
Lgals3	2.52703	0.540976
Napsa	1.47568	0.311875
Acan	6.40981	1.35113
Steap4	0.942286	0.198461
Adra1b	0.597122	0.122387
Tac2	8.38631	1.69331
Prap1	72.8807	14.705
Cpa3	2.90645	0.584632
Cfb	0.388025	0.0751552
Dhrs7c	1.75346	0.331964
F2	0.793165	0.143844
Htr7	0.374673	0.066299
Smpd3	0.732436	0.127654
Havcr1	0.605897	0.105594
Calca	1.5341	0.267082
Lrat	0.425908	0.0736345
Myh8	6.00316	1.03731
Mfap5	0.904956	0.152463
H60c	0.510959	0.0808335
Ankrd2	0.983329	0.155158
Lect1	22.1906	3.38413
1500015O10Rik	2.75425	0.401714
Apom	3.86361	0.560745
Trf	27.5728	3.91002
Scrg1	7.06889	0.998335
Apoa1	20.9207	2.90693
Il1rn	0.435793	0.0586958
Apoc2	2.25905	0.301127
Pitx1	90.4245	11.9324
Slc17a6	0.396657	0.043761
Dpt	2.58004	0.271616
Fgg	1.55527	0.163481
Alb	2.04718	0.208035
Plg	0.499157	0.0502891
Ano5	0.464633	0.0453527
Ttr	8.76092	0.84508
Tnmd	7.54371	0.635057
Apoa2	7.40444	0.620402
Calb1	1.98668	0.153524
Comp	2.36525	0.179473
Afp	98.3084	7.03809
Apoa4	8.18427	0.565502
Fgb	1.7473	0.116907
Fga	1.3852	0.0857986
Matn1	12.9749	0.696121
Apob	0.785626	0.0333467
Rps3a1	96.3858	2.58277
Zfp429	3.66996	0.0740512

Table S2.4. Top putative and conserved targets of PITX1 in mouse and *Anolis* hindlimbs.

Putative_Targets	WT_RPKM	Mutant_RPKM
Zfp42	0.185373	0.898901
Ahcy	2.91864	8.9065
Zfp804a	0.384559	0.966646
Alx3	5.92269	13.6907
Rps3a1	109.153	225.71
Pmaip1	57.5102	118.244
Cbln4	5.64022	2.43223
Myof	1.18361	0.482221
Wt1	0.967013	0.387018
Cybrd1	1.22337	0.489297
Grik1	0.539423	0.215697
Rprm	25.7932	10.2257
Prelp	0.929198	0.364542
Sfmbt2	0.474757	0.184857
Rxrg	1.0034	0.389757
Hist1h4i	13.1762	5.10603
Colgalt2	1.12239	0.432804
Fam101a	4.6263	1.78378
Ntrk2	0.559348	0.214402
Rgs8	0.50563	0.193335
Msc	8.65239	3.29362
Myo18b	0.314682	0.119648
Rbfox1	0.588566	0.222817
Galnt13	0.346786	0.129878
Luzp2	0.749047	0.279505
Creb3l1	1.02929	0.375363
Dll3	1.15619	0.409497
Tbx4	63.5552	21.9857
Pax9	5.14146	1.75997
Osmr	0.465421	0.159083
Sytl2	1.83745	0.590728
Tbx5	0.896295	0.277732
Pde8b	0.516244	0.158868
Hfe2	1.29535	0.396371
Nkx3-2	1.43415	0.425299
Rgs11	0.484756	0.142691
Nfatc2	10.3929	3.03548
Lypd1	0.591135	0.172484
Syt6	0.220996	0.0606483
Tlr1	0.511427	0.137316
Cntn4	0.774592	0.197705
Tmem8c	2.3782	0.575498
Gmnc	0.818283	0.182208
Tnc	5.29093	1.16917
Arpp21	1.36336	0.2562
Fam155a	0.338274	0.062701
Acta2	16.965	2.80056
Alpk2	0.460704	0.0741868
Mylk4	0.57292	0.0918443
Hand1	0.883962	0.133735
Apobec2	0.95319	0.127463
E130114P18Rik	6.76123	0.819635
Lect1	8.70069	1.02733
B3galt5	0.266473	0.0169028
Erv3	0.532137	0.0198609

Conserved Targets	WT RPKM	Mutant RPKM
Rps3a1	109.153	225.71
Tfap2c	2.15604	3.80622
Elavl2	1.8794	2.95496
Irx5	5.79053	8.85749
Meis2	19.7131	29.1234
Emx2	6.45033	9.47438
Chordc1	62.7232	87.5323
Bmp2	5.21626	7.16325
Etaa1	12.1443	16.6486
Bmp7	18.6723	25.4532
Irx3	10.2082	13.8434
Tbca	251.19	340.532
Fam107b	28.628	38.5512
Top1	122.516	163.28
Id2	108.463	143.812
Prickle1	14.8676	19.5063
Hoxa9	36.9536	48.069
Bmp4	35.6604	46.0218
Tbx2	20.3265	26.2276
Zfhx4	27.44	35.2491
Dab2	23.9038	18.0753
Tshz2	8.05828	6.06511
Nfatc1	8.16233	6.11076
Sulf1	38.6293	28.8892
Pknox2	11.0126	8.10715
Tcf7l1	29.9719	21.9642
Lrig1	18.4208	13.2909
Six1	31.9197	22.9497
Sema3d	6.47787	4.54272
Rgmb	10.2904	7.21622
Mpped2	35.7919	25.0055
Ebf1	15.1716	10.4679
Meox1	10.5321	7.24331
Foxp4	33.9251	23.3248
Hic1	10.894	7.41835
Nav2	3.38355	2.29433
Nfia	3.79449	2.56491
Neurl1a	4.89519	3.28814
Sobp	3.81109	2.54894
Foxc1	8.28319	5.50017
Tcf7l2	16.0114	10.472
Sema6a	10.1826	6.57234
Tbx18	43.4143	27.8003
Hs3st3b1	7.97933	4.89371
Plce1	3.36703	2.05769
Dctd	36.7541	22.4506
Col25a1	15.1507	9.24137
Mkx	6.74459	3.94767
Sox9	47.6136	27.5962
Slc35f1	1.91773	1.11074
Kirrel3	1.74951	0.959135
Creb5	2.19827	1.19778
Hlx	3.99852	2.10187
Sox6	6.09323	3.03038
Dmrt2	6.45683	3.18506
Glis3	1.06264	0.518861
Jph2	0.805717	0.380846
Pmp22	18.5116	8.14517
Car8	0.342092	0.139856
Msc	8.65239	3.29362
Tbx4	63.5552	21.9857
Pax9	5.14146	1.75997
Nkx3-2	1.43415	0.425299
E130114P18Rik	6.76123	0.819635

CHAPTER 3

PITX1 MODULATES MOUSE HINDLIMB AND MANDIBLE DEVELOPMENT THROUGH SHARED REGULATORY NETWORK[2]

[2] Wang, J.S., Menke, D.B., to be submitted to *Development Biology*.

Abstract

Both the forelimb-hindlimb (FL-HL) and maxilla-mandible (MX-MD) have been considered to be two series of morphologically and developmentally homologous structures. The forelimb and maxilla have sometimes been regarded as the default state, while the hindlimb and mandible are diverged anatomically and functionally. *Pitx1*, a transcription factor gene, has a hindlimb- and mandible-restricted expression patterns. The genomic mechanisms underlying how PITX1 regulates hindlimb- or mandible-specific development remain poorly understood. We first applied a combination of ChIP-seq and RNA-seq to investigate enhancer regions and target genes that are directly regulated by PITX1 in embryonic mouse hindlimb and mandible separately. We further compared the role of PITX1 in embryonic hindlimb and mandible development from the transcriptome and genome-wide binding levels. We found that PITX1-dependent expressed genes shared between hindlimb and mandible are strongly associated with limb and craniofacial development. PITX-bound enhancer regions that are identified both in hindlimb and mandible are significantly enriched near genes implicated in hindlimb, craniofacial and skeletal system development. Both *Sox9* and *Gtf2ird1* are among the putative direct targets of PITX1 in the developing hindlimb and mandible. Our data suggest that PITX1 modulates mouse limb and craniofacial development through shared transcriptional targets, including enhancer regions and target genes.

Keywords: PITX1, Mandible, Hindlimb, ChIP-seq, RNA-seq, Shared *cis*-regulatory network

Introduction

In vertebrates, both the limb and jaw are great model systems to study evolution and development (evo-devo). From the aspect of evo-devo, both forelimb-hindlimb (FL-HL) and maxilla-mandible (MX-MD) belong to the model of "homologous structure followed by functional and anatomical divergence" (Young and Hallgrimsson, 2005; Beverdam et al., 2002). The hypothesis is that both forelimb and maxilla are the default states during organ development, where hindlimb and mandible are shaped via the transcription of HL- and MD-specific genes (Duboc and Logan, 2011; Depew et al., 2002). Similar developmental processes are involved in the development of limb and jaw, including initiation, outgrowth, patterning and differentiation. Both organs develop complex musculoskeletal structures within their common frameworks. It all supports the parallel comparison between the limb and jaw development in the genomic and molecular mechanisms.

The *Pitx1* gene encodes a bicoid-class homeodomain transcription factor that plays a central role in growth and patterning of the vertebrate hindlimb (Lanctôt et al., 1999; Szeto et al., 1999). The complete ablation of *Pitx1* function in mice results in reduced hindlimb size, the loss of hindlimb-specific features, as well as developmental defects in the mandible, teeth, palate, and pituitary gland. Since ectopic expression of *Pitx1* in the developing forelimbs of chickens, mice, and humans results in the forelimb developing a hindlimb-like morphology (Delaurier et al., 2006; Logan and Tabin, 1999; Spielmann et al., 2012), it is apparent that the role of PITX1 in hindlimb formation extends to the control of limb-type identity. We further identified the core regulatory network of PITX1 that underlies hindlimb development, and proved that PITX1 can promote chondrogenic and myogenic differentiation through direct regulation of the transcription of key conserved genes (Wang et al., under review).

Other than expressed in the derivatives of the posterior lateral mesoderm, *Pitx1* is expressed in the first branchial arch mesenchyme and its derivatives, including mandible (Lanctôt et al., 1997). When knocked out in mice, the *Pitx1*$^{-/-}$ embryos exhibit severe mandible malformations, which resembles human mandibular hypoplasia. The bHLH transcription factor HAND2 is essential for the maxilla-to-mandible transformation (Funato et al., 2016). When *Hand2* expression is induced in the neural crest cells of the branchial arches, both *Hand2* and *Pitx1* are ectopically expressed in the maxilla. On the other hand, when *Hand2* expression is disrupted, *Pitx1* expression is also down-regulated in the mandible. It supports that PITX1 could regulate mandible outgrowth and direct mandible-specific identity.

The regulatory networks of PITX1-directed hindlimb and mandible programs still remain unclear. Here, we performed the global gene expression analyses in embryonic hindlimbs and mandibles in both wild-type and *Pitx1*$^{-/-}$ mice. We also incorporated the chromatin profiles of PITX1 and p300 of mouse hindlimb and mandible. Both putative enhancers and target genes that are shared between hindlimb and mandible are isolated from the combination of ChIP-seq and RNA-seq datasets. Based on the strong enrichment of PITX1 binding activity and PITX1-dependent genes, it suggests that PITX1 can direct hindlimb and mandible development by regulating the transcription of shared target genes through the binding interactions in shared enhancer regions.

Materials and Methods

Mice

Pitx1 knockout mice were previously reported (Szeto et al., 1999). The *Pitx1* knockout allele was maintained on a *129/Sv* background but outcrossed onto an outbred ICR background (Envigo) to generate embryos for mandible RNA-seq. ICR mice were purchased to generate embryos for mandible ChIP-seq. All procedures involving animals were performed in accordance with guidelines issued by the Institutional Animal Care and Use Committees (IACUC) at the University of Georgia under approved Animal Use Protocols (mouse protocol A2014 06-019).

ChIP-seq

PITX1 ChIP-seq data for E11.5 mouse hindlimbs was performed as previously described (GEO accession GSE41591; (Infante et al., 2013)). For ChIP-seq on mouse E11.5 mandible, two independent ChIP-seq replicates and input control libraries were generated. Two separate pools of chromatin collected from embryonic mouse mandibles were used (70 mandibles per replicate for a total of 500 µg of chromatin in each ChIP). PureProteome™ Protein G Magnetic Beads (Millipore) were pre-incubated with PITX1 antibody (Santa Cruz Biotechnology, sc-18922) before incubating overnight with chromatin. All ChIP and input chromatin control libraries were produced using the NEBNext Ultra DNA Library Prep Kit for Illumina as previously reported (Infante et al., 2015). Libraries were submitted to the Georgia Genomics Facility and sequenced on the NextSeq 500 platform. The mouse mandible PITX1 ChIP-seq data generated for this work will be deposited in the Gene Expression Omnibus (Edgar et al., 2002).

ChIP-seq Data Analysis

ChIP-seq data for E11.5 hindlimb was analyzed and described previously (Infante et al., 2013). For E11.5 mandible, sequencing read quality was evaluated using FastQC (version 0.11.2, http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). ChIP-seq reads were aligned to the mouse genome (mm9) using bowtie v2 (Langmead and Salzberg, 2012) with the default parameters. PITX1 peaks were identified using MACS2 with the effective genome size set for mouse (mm9) (Zhang et al., 2008). Peaks were associated with Gene Ontology (GO) terms using the Genomic Regions Enrichment of Annotation Tool (GREAT) (Mclean et al., 2010). The genes that are associated with PITX1 binding activities are assigned by searching PITX1 peaks in the *cis*-regulatory domains of neighboring genes using GREAT.

Additional Mouse ChIP-seq Data

Aligned reads from p300 ChIP-seq experiments in mouse E11.5 limb was downloaded from the Gene Expression Omnibus (GEO) database (GEO: GSM348066; Visel et al., 2009). p300 ChIP-seq aligned read dataset for E11.5 facial tissue was downloaded from the GEO database (GEO: GSM1199037; Attanasio et al., 2013).

RNA-seq

RNA-seq data for E11.5 mouse hindlimbs have been described previously and deposited in the Gene Expression Omnibus (GSE104460; (Wang et al., under review)). For E11.5 mandible, embryos were sexed via PCR and only XY embryos were used. All embryos used in RNA-seq were staged based on the forelimb morphology as described previously (Wanek et al., 1989). For each genotype, RNA-seq libraries were made from three separate embryos. Total RNA was

isolated using the *mir*Vana RNA Isolation Kit (ThermoFisher Scientific). Libraries were constructed with TruSeq Stranded mRNA Sample Prep Kit for Illumina. 500 ng of total RNA was used to construct E11.5 mandible libraries. Libraries were submitted to the Georgia Genomics Facility and sequenced on the Illumina NextSeq platform to produce approximately 50 million, single-end, 75bp reads per library. Mouse RNA-seq data generated for this work will be deposited in the Gene Expression Omnibus (Edgar et al., 2002).

RNA-seq Data Analysis

RNA-seq dataset of E11.5 mouse hindlimb was analyzed and described previously (Wang et al., under review). For E11.5 mandible data, sequenced libraries were evaluated and trimmed with FastQC (version 0.11.2, http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). R package (https://www.r-project.org/) was used to calculate the correlations among libraries. Sequenced reads were aligned to the mouse genome (mm9) using Tophat2 on the Galaxy platform (Kim et al., 2013). Both Cuffdiff and DESeq2 were used to detect differentially expressed genes. Cuffdiff was used to generate qualified RPKM and to identify differentially expressed genes and transcripts (Trapnell et al., 2012). htseq-count was used to calculate the number of reads that maps to each feature (gene) (Anders et al., 2014). Followed by DESeq2, differentially expressed genes were determined from the count tables generated from htseq-count (Love et al., 2014). GO Analysis was performed with DAVID Bioinformatics Resources 6.7 (Huang et al., 2009). DAVID annotation categories that were applied to identify functional annotation clusters include “Functional Categories”, “Literature”, “Pathways” and “Tissue Expression”.

Results

Genes Related to Musculoskeletal Development are Misregulated in Pitx1$^{-/-}$ Mandibles

In order to identify genes with PITX1-depedent expression, we performed global gene expression analyses between E11.5 *Pitx1*$^{-/-}$ and wild-type mouse mandibles. 113 genes are significantly mis-regulated in *Pitx1*$^{-/-}$ mandible (Table S3.1). Among those, 57 of them are up-regulated and 56 are down-regulated. We then used the DAVID functional annotation tool to determine whether specific gene categories are enriched among genes that are misexpressed in *Pitx1* mutants (Huang et al., 2009). Among the categories that are enriched, the top ranked clusters relate to chromatin assembly, skeletal system development, muscle organ development and Homeobox (Fig. 3.1). Thus, in *Pitx1*$^{-/-}$ embryonic mandible, we find clear evidence of altered expressions in the genes that are related to organ bone and muscle development.

Published literature has demonstrated that the *Dlx* code provides the spatial identity along dorsal-ventral axis in the first branchial arch. *Dlx1/2* are factors that are involved in the default setting and their expressions are consistent with maxillary development. In the meantime, *Dlx5/6* are the selectors for the determination and development of mandibular process (Jeong et al., 2008). Our RNA-seq analyses demonstrate that *Dlx2* expression is significantly up-regulated in *Pitx1*$^{-/-}$ mandible (Table S3.1). It shows that the expression of maxilla-related genes has increased when *Pitx1* expression is disrupted. We also examined the expression of *Gtf2ird1*, a gene that is required for mammalian mandible development (Tassabehji et al., 2005). In wild-type embryos, *Gtf2ird1* is expressed in the mandible mesenchyme. In *Pitx1*$^{-/-}$ embryos, we find that *Gtf2ird1* expression level is significantly reduced (Table S3.1). When *Gtf2ird1* is knocked out in mouse embryos, the null mice exhibit abnormal skeletons both in the skull and jaws, which resembles human disorder Williams-Beuren syndrome (WBS).

Mouse PITX1 Binding Events are Involved in Craniofacial and Skeletal Development

To examine the binding activity of PITX1, two ChIP-seq replicates and input control libraries were constructed in E11.5 mouse mandible. Significant PITX1 peaks were identified using MACS2 peakcalling (Zhang et al., 2008). 2,518 highly reproducible PITX1 binding sites were yielded from two replicates using BEDTools v2.26.0. (Table S3.2; Quinlan and Hall, 2010). *De novo* motif searches was performed within +/- 50 bps of PITX1 peak summits using HOMER and revealed that TAATCC, the core PITX1 binding motif, is significantly enriched in mouse mandible (p-value = 1×10^{-251}; 50.63% of target sequences) PITX1 peaks (Fig. 3.2A; Table S3.3; Heinz et al., 2010). Other than the core PITX1 binding motif, the known binding motif of MSX1 is also enriched among the top *de novo* motifs of PITX1 mandible peaks and could be mapped to 11.65% of targeted sequences (Table S3.3). *Msx1* is expressed in the mesenchyme of the 1st branchial arch (BA1) in response of *Bmp4* signal in the BA1 ectoderm (Shigetani et al., 2000). PITX1 and MSX1 core motifs share three same bases AAT. This could be the evidence that PITX1 and MSX1 are co-factors or somehow related in regulating mouse mandible development. In Mouse Genome Informatics, we found that mandible PITX1 peaks are significantly enriched near genes that are expressed in limb and 1st branchial arch (Fig. 3.2B). Moreover, these PITX1 binding events are also strongly enriched near genes that are functioning not only in abnormal limb development, but also craniofacial and skeletal development (Fig. 3.2C). The binding activity of PITX1 in mandible suggests that PITX1 is required for craniofacial development, especially in skeletal development.

Transcriptome Analyses Identify Shared PITX1-dependent Genes in Hindlimb and Mandible

In order to identify shared regulatory network of PITX1, we started by comparing the gene expression profiles from different tissues to examine the shared function of PITX1 during organ development. For the hierarchical clustering, we included the whole transcriptome data of wild-type E11.5 hindlimbs, E11.5 mandibles and E12.5 genital tubercles. It can calculate the distances between different samples based on the similarity of the whole transcriptome. To begin with, wild-type replicates are closely related and they were clustered within each tissue. The mandible transcriptome is more similar and closer to the hindlimb transcriptome, which is consistent with that hindlimb and mandible share more similarity during organ patterning and differentiation (Fig. 3.3A). In contrast, the mouse genital tubercle transcriptome is a more different and distinct from the hindlimb and mandible transcriptomes, highlighting a possible separate developmental process of genitalia at this stage. To assess the genes that are differentially expressed in *Pitx1*$^{-/-}$ in both hindlimb and mandible, we further re-analyzed our RNA-seq data from E11.5 hindlimb and mandible. Of the 1044 genes that are differentially misexpressed either in mutant hindlimb or mandible, three distinct groups were clustered based on log2 fold change (Fig. 3.3B). Genes within the first cluster have hindlimb-specific expression and they are significantly mis-regulated only in hindlimb. Though not significant, most of these genes have a trend of down regulation in *Pitx1*$^{-/-}$ mandible. 45 genes are significantly misexpressed in both *Pitx1*$^{-/-}$ hindlimb and mandible (Table S3.4). For most of these genes, their expression patterns are altered in the same direction in both organs. 14 genes are up-regulated both in *Pitx1*$^{-/-}$ hindlimb and mandible, and 25 are down-regulated in both mutant tissues. It suggests that PITX1 could regulate the expressions of target genes through consistent activation or repression in the developing hindlimb and mandible.

Shared PITX1 Binding Activities are Revealed in Mouse Hindlimb and Mandible

In order to identify PITX1 binding sites that are shared between developing hindlimb and mandible, we intersected PITX1 peaks from hindlimb against PITX1 peaks from mandible. We were able to identify 1160 PITX1 peaks that are shared between hindlimb and mandible (Table S3.5). The shared binding sites are significantly enriched near genes that are expressed in the limb and genes that are involved in the craniofacial and limb development (Data not shown). Since the binding activity identified from PITX1 ChIP-seq is the combination of promoters and enhancer regions, we decided to extract specifically the enhancers that are bound by PITX1 in both organs. We examined the pattern of enhancer activity in embryonic mouse hindlimb and mandible using p300, which deposits histone acetylation and marks the enhancer region (Visel et al., 2009). p300 ChIP-seq aligned read datasets from mouse E11.5 limb and facial tissue were downloaded from the Gene Expression Omnibus (GEO) database (Visel et al., 2009; GEO: GSM348066; Attanasio et al., 2013; GEO: GSM1199037). PITX1 binding sites that were identified either from embryonic hindlimb or mandible were further grouped by the relative p300 signal using k-mean clustering (Ye et al., 2011). Five major groups of loci could be distinguished (Fig. 3.4A). The first and second clusters have p300 signal in both limb and mandible, with the first cluster has a relative stronger p300 signal. It means that the enhancer regions from the first two clusters are bound by PITX1 in both hindlimb and mandible. The third group contains loci that are hindlimb-specific, while the fourth group contains loci that are mandible-specific. These two groups represent the enhancers that are bound by PITX1 either in hindlimb or in mandible.

We were able to further extract 847 enhancers from the first cluster (Table S3.6). These enhancers are strongly bound by PITX1 in both organs. The *de novo* motif search performed with +/- 50bp of the peak summit of these enhancers reveals that the top *de novo* motif (p-value

= 1 x 10^{-63}; 12.63% of target sequences) matches to the known PITX1 binding motif, TAATCC (Fig. 3.4B). These organ-shared enhancers are strongly enriched near genes that are involved in two major clusters: appendage/limb development and craniofacial development (Fig. 3.4C). We then searched the genes that are associated with shared enhancers using GREAT "Basal plus extension" approach (Mclean et al., 2010). Each gene was assigned a basal regulatory domain with the minimum distance both upstream (5 kb) and downstream (1 kb) of the gene's transcription start site (TSS). The whole gene regulatory domain was then extended in both directions to the nearest gene's basal domain but no more than the maximum extension (1000 kb). 927 genes have shared enhancer regions that fall in their *cis*-regulatory domains (data not shown). These associated genes are more specific and precisely enriched in clusters like skeletal development, hindlimb morphogenesis and cranial skeletal morphogenesis (Fig. 3.4D). In summary, the comparisons performed in ChIP-seq and RNA-seq between hindlimb and mandible support our conclusion that PITX1 can modulate hindlimb and mandible development through shared *cis*-regulatory elements and target genes.

Mouse PITX1 Binding Profiles at Shared Putative Transcriptional Targets

Within the shared regulatory network of PITX1, we further examined PITX1 binding profiles at several putative direct targets that are involved in the development of both hindlimb and mandible, including *Gtf2ird1* and *Sox9*. *Gtf2ird1* is significantly down-regulated in both *Pitx1*$^{-/-}$ hindlimb and mandible (Table S3.4). A significant PITX1 peak is identified at the promoter region of *Gtf2ird1* at E11.5 mouse hindlimbs (Fig. 3.5A). Significant PITX1 peaks are also identified at the *Gtf2ird1* promoter in both E11.5 mouse mandible and E12.5 genital tubercle. There is strong evidence that PITX1 could regulate *Gtf2ird1* expression in both

hindlimb and mandible via direct binding events at the *Gtf2ird1* promoter. *Sox9* expression is significantly reduced in both *Pitx1*$^{-/-}$ hindlimb and mandible (Table S3.4). In our previous work, there is a strongly enriched PITX1 peak located about 232 kb downstream of *Sox9* transcription start site (TSS) in mouse (Wang et al., under review). This region has sequence conservation in the *Anolis* lizard genome and is also bound by PITX1 in *Anolis* embryonic hindlimbs. At the same region, a significant PITX1 peak was identified in mouse mandible (Fig. 3.5B). This PITX1-bound region has been experimentally verified with gene enhancer activity in VISTA Enhancer Browser (Visel et al., 2007). Other than strong staining in forelimbs and hindlimbs, it can also drive lacZ staining in the mandible in transgenic mouse embryos. PITX1 could direct *Sox9* expression in both hindlimb and mandible through this conserved and shared enhancer region. These results strongly suggest that PITX1 can directly regulate mouse hindlimb and mandible development through the shared regulatory network, including enhancer regions and target genes.

Discussion

There are a few reasons of comparing PITX1 function between the mouse hindlimb and mandible. *Pitx1* is robustly expressed in the hindlimb and mandible mesenchyme (Lanctôt et al., 1997), and its expression is absent from the forelimb and maxilla. In *Pitx1*$^{-/-}$ mouse embryos, there is a series of complex skeletal defects in the hindlimb and mandible, including shorter femur, reduced pelvis, absent of patella, alterations in tibia and fibula, and undersized mandible bone (Szeto et al., 1999; Lanctôt et al., 1999). In mouse, both hindlimb and mandible share similar developmental processes, including specification of the prospective field, induction of initiation, maintenance of the outgrowth and patterning (Zeller et al., 2009; Minoux and Rijli, 2010). Many major signaling pathways contribute to the development of both organs, including FGF, SHH, BMP and WNT. For example, the FGF signaling in the hindlimb and mandible shares both similarities and differences. In hindlimb, *Fgf8* is expressed in the apical ectodermal ridge (AER) and the AER-FGF signaling is required for both limb outgrowth and proximodistal (PD) patterning (Niswander, 2003; Sekine et al., 1999; Niswander et al., 1993). In mandible, *Ffg8* is expressed in the ectoderm of the 1st pharyngeal arch (PA1) and its expression is essential for the formation of PA1, which is similar to hindlimb (Abu-Issa et al., 2002). *Ffg8* is also expressed in a different region, the oral epithelia, to participate in the anteroposterior (AP) polarizing of PA1 (Grigoriou et al., 1998).

The combination of ChIP-seq and RNA-seq has provided ways to identify PITX1-related enhancer regions and target genes at the whole-genome level. From the RNA-seq analyses performed in wild-type and *Pitx1*$^{-/-}$ mandible, several genes related to jaw development have significantly altered expressions in the mutant, including up-regulated *Dlx1as*, *Dlx2*, *Dlx6os1* and down-regulated *Six2* (Table S3.1). *Dlx1as* is the antisense RNA of *Dlx1* and can modulate

Dlx1 transcript level and stability (Kraus et al., 2013). *Dlx1/2* is expressed in maxilla and mandible, and is the default setting for both organs to develop (Depew et al., 2005; Jeong et al., 2008). *Dlx6os1* is also known as *Evf1*. It is a non-coding RNA transcribed upstream of *Dlx6* gene and has an important role in ventral forebrain and craniofacial development (Kohtz and Fishell, 2004). The comparison between hindlimb and mandible transcriptomes shows that many genes involved in skeletal development are significantly misexpressed in both tissues, including *Sox9*, *Sox10*, *Lect1*, *Matn4* and *Tbx22* (Table S3.4). With the incorporation of PITX1 ChIP-seq datasets, we were able to identify putative direct targets of PITX1 during hindlimb and mandible development, such as *Sox9* and *Gtf2ird1* (Fig. 3.5). *Sox9* is the early marker of chondrogenic differentiation and is strongly tied to skeletal phenotypes and morphogenesis (Akiyama et al., 2002; Gordan et al., 2014). *Gtf2ird1* is required for mammalian mandible development. When *Gtf2ird1* is knocked out in mouse embryos, the null mice exhibit skeletal abnormalities both in the skull and jaws, which resembles the human disorder Williams-Beuren syndrome (WBS) (Tassabehji et al., 2005). Our work has provided an integrative method to study the role of PITX1 in limb and craniofacial development by introducing the whole-genome level comparative approaches.

Acknowledgements

This work was supported by grants awarded to D.B.M. from the NIH (HD081034) and by two Junior Faculty Research Grants (#1847 and #2503) from the Office of the Vice President of Research at the University of Georgia. This study was supported in part by resources from the Georgia Advanced Computing Resource Center.

Figures

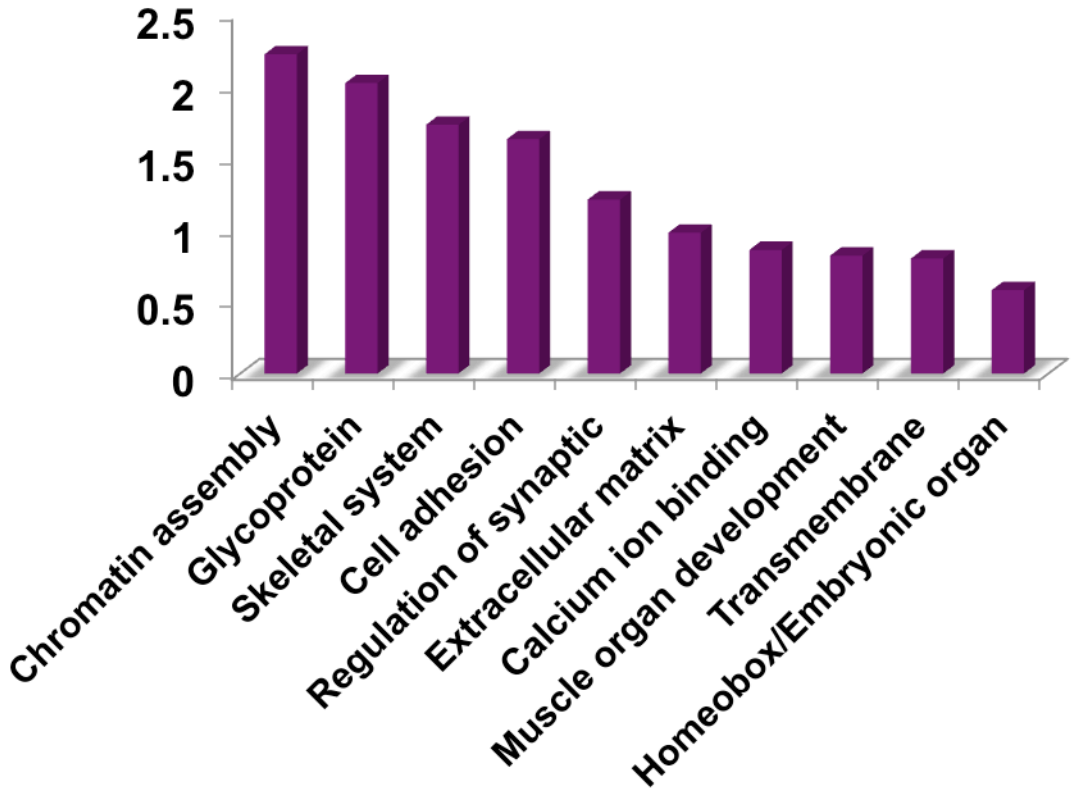


Figure 3.1. Misregulated genes in mouse *Pitx1*$^{-/-}$ mandible. The top 10 scoring gene clusters associated with misexpressed genes in *Pitx1*$^{-/-}$ mandibles at E11.5.

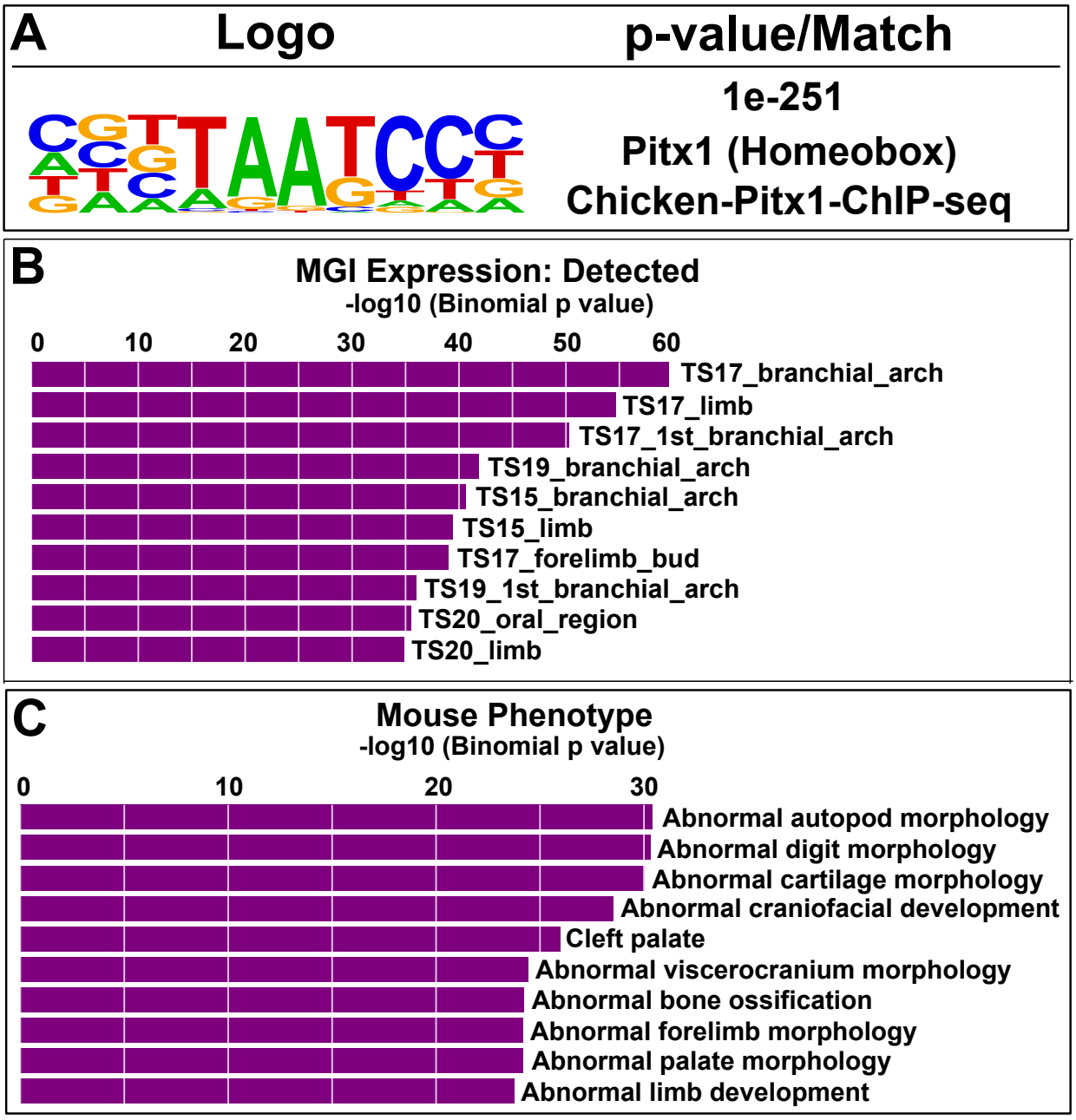


Figure 3.2. Genome-wide enrichment of PITX1 binding activity in mouse mandible. **A)** Top *de novo* motif enriched within +/- 50 bps of mouse PITX1 peak summits and its best match to the known motif using HOMER. **B)** The top 10 MGI (Mouse Genome Informatics) Expression terms associated with mouse mandible PITX1 peaks. **C)** The top 10 Mouse Phenotype terms associated with mouse mandible PITX1 peaks.

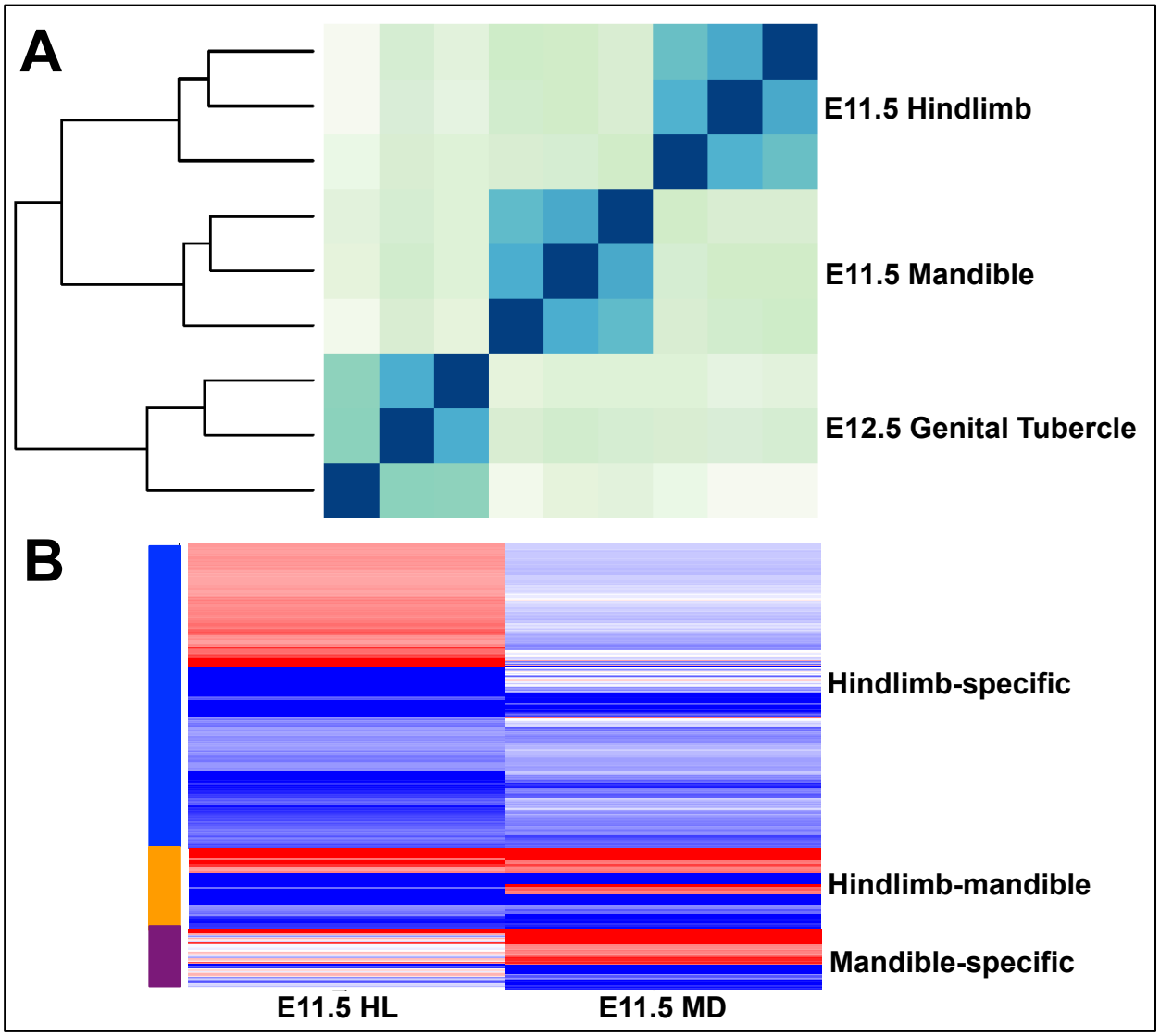


Figure 3.3. Transcriptome analyses between mouse hindlimb, mandible and genital tubercle. **A)** Hierarchical clustering between E11.5 hindlimb, mandible and E12.5 genital tubercle based on wild-type RNA-seq datasets. **B)** PITX1-dependent genes are clustered into hindlimb-specific, hindlimb-mandible and mandible-specific groups based on log2 fold change of RNA-seq data.

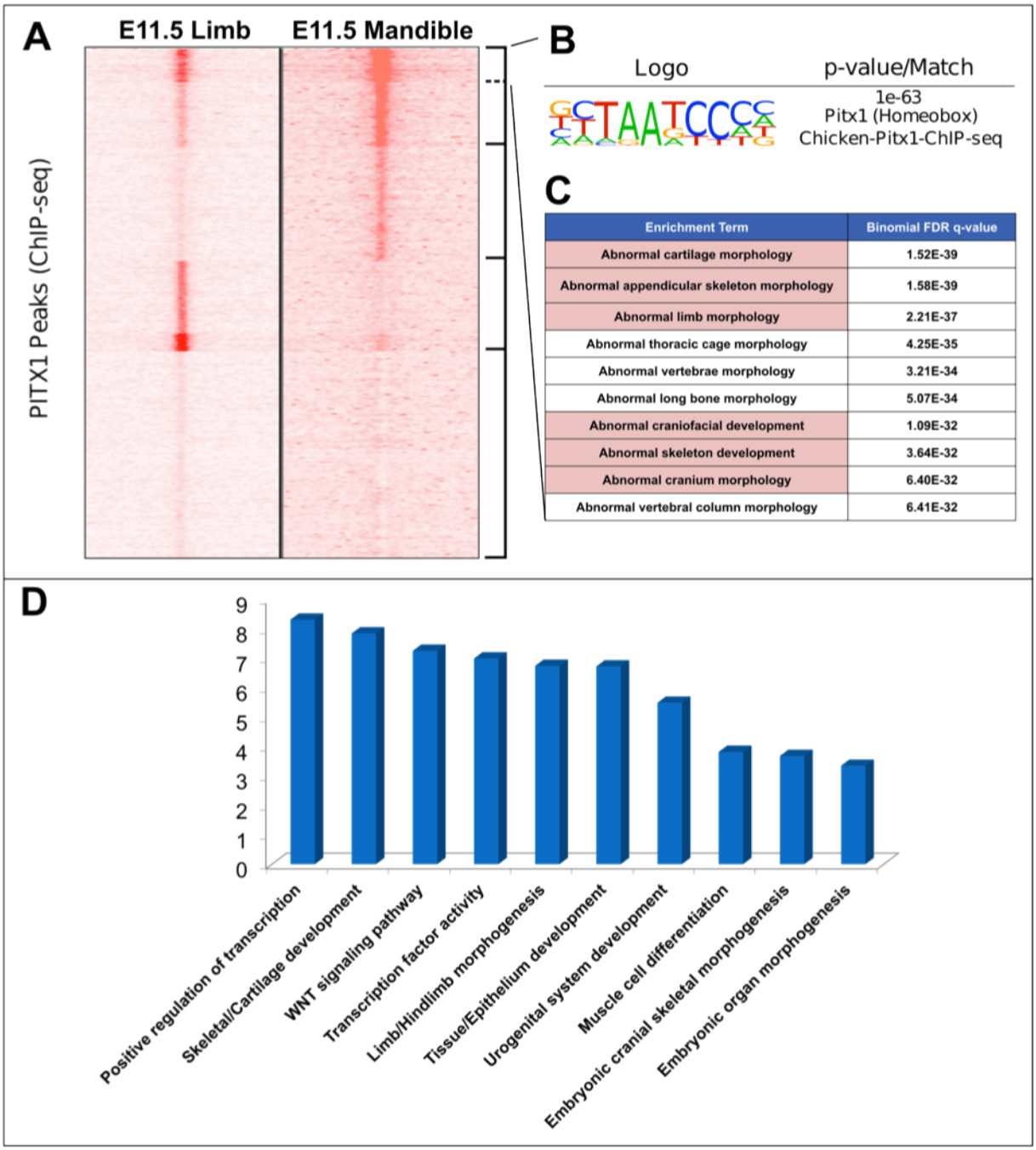


Figure 3.4. Enrichment analyses of shared enhancers of PITX1 in developing hindlimb and mandible. **A)** K-mean clustering of PITX1 peaks based on p300 ChIP-seq signal. **B)** Top *de novo* motif enriched within shared PITX1 enhancers and its best match to the known motif. **C)** The top 10 Mouse Phenotype terms associated with shared PITX1 enhancers. **D)** The top 10 enriched clusters of genes that are associated with shared PITX1 enhancers.

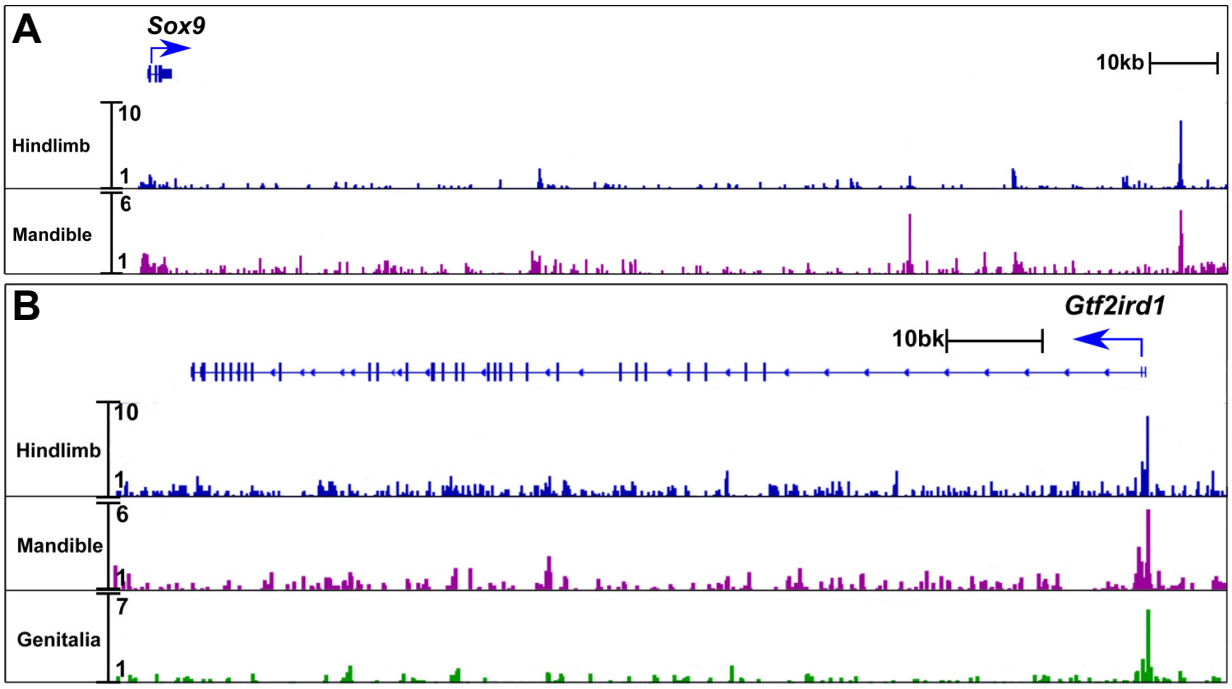


Figure 3.5. PITX1 ChIP-seq profiles in mouse hindlimbs and mandible at putative PITX1 transcriptional targets. **A)** *Sox9* and **B)** *Gtf2ird1*.

Table S3.1. Transcriptome analyses (RPKM) between the developing mandible of wild-type and *Pitx1*$^{-/-}$ in mouse embryos. E11.5 differentially expressed genes are listed.

E11.5_DE_Gene	WT_RPKM	Mut_RPKM
Klk1	0	1.2587
Rmrp	20.7626	242.982
Hist4h4	0.770436	6.42085
Rpph1	13.5325	111.919
Hist1h4k	0.988043	6.33522
Hist2h2bb	1.63355	9.84551
Lars2	81.873	427.162
Prap1	7.65701	37.9904
3110099E03Rik	0.589917	2.90013
Hist1h1a	1.17093	5.35808
Hist1h3e	2.27077	8.50814
Ahcy	2.24902	8.39224
Hist1h1b	1.06326	3.95146
Sim2	3.18343	10.1804
Sstr1	0.558825	1.78142
Hist1h1e	1.36813	4.33551
Galnt13	0.216026	0.613719
Dlx1as	13.563	37.2338
Lyrm7	0.958872	2.49024
Grik1	0.663668	1.52575
AA388235	1.16419	2.66565
Rps2	31.5085	69.9673
Rps18	52.6766	115.602
Ppp1cc	3.23724	6.80145
Stmn2	3.32828	6.21257
Zfp708	3.19817	5.89209
Avpr1a	2.14966	3.93093
Dlx6os1	13.0555	23.8585
Slitrk5	1.13061	2.04426
Tpt1	17.4278	31.3071
Rps7	19.5913	35.0919
Cap2	3.79422	6.42227
Bbs2	4.98261	8.30197
Ptx3	10.3159	16.7315
Ebf2	6.7333	10.7917
4833420G17Rik	13.8926	21.6916
Tbx22	9.911	15.3113
Hic1	11.3052	17.0543
Efemp1	12.1754	18.2243
Coq2	17.4478	25.8405
Myh3	4.43931	6.51218
Glo1	60.39	87.001
Pcdh10	15.8067	22.7697
Cox7b	175.603	251.739
Actc1	61.4351	87.3569
Crabp1	329.45	465.162
Myog	26.8851	37.8263
Dlx2	12.1255	17.0356
Mpzl2	11.1864	15.7049
Rgs2	16.395	22.9924
Sepp1	27.2635	38.1919
Cdk2ap1	56.4965	78.5794
1810011O10Rik	35.4314	48.6552
Nr2f1	35.3384	48.1732
Cited1	143.496	195.236
Bod1	62.3951	82.7952

Table S3.2. Representatives of PITX1 peaks identified in mouse mandibles. Coordinates of PITX1 mouse peaks (mm9) are listed.

Chromosome	Start	End
chr1	7649998	7650114
chr1	9573871	9574084
chr1	10272215	10272319
chr1	11320410	11320596
chr1	12275628	12276088
chr1	12562237	12562382
chr1	12804844	12805391
chr1	13485273	13485384
chr1	14026568	14026687
chr1	14297342	14297611
chr1	14480194	14480433
chr1	14611210	14611351
chr1	15168421	15168572
chr1	17708128	17708326
chr1	19166358	19166447
chr1	21511677	21511987
chr1	21810355	21810443
chr1	22311294	22311384
chr1	22562726	22562826
chr1	24791828	24792056
chr1	30594252	30594753
chr1	36360384	36360500
chr1	36458575	36458690
chr1	38745094	38745383
chr1	38821245	38821408
chr1	41949359	41949549
chr1	42828998	42829144
chr1	44258081	44258191
chr1	45100176	45100465
chr1	45817551	45817818
chr1	45817930	45818632
chr1	46789391	46789501
chr1	46959110	46959345
chr1	47564755	47564979
chr1	52628734	52628819
chr1	54275521	54275660
chr1	56258208	56258321
chr1	56307441	56307627
chr1	56596588	56596837
chr1	56839133	56839238
chr1	56863027	56863244
chr1	56922265	56922401
chr1	57137382	57137492
chr1	57545997	57546481
chr1	59327392	59327522
chr1	59367151	59367351
chr1	59391497	59391923
chr1	59944846	59945002
chr1	61290297	61290845
chr1	61388848	61388996

Table S3.3. *De novo* motifs enriched in mouse mandible PITX1 peaks.

Rank	Conserved De Novo Motif	P-value	Log P-value	% of Targets	% of Background	STD(Bg STD)	Best Match
1		1e-251	-5.7E+02	41.83%	12.12%	11.9bp (17.9bp)	Pitx1(Homeobox)/Chicken-Pitx1-ChIP-Seq(GSE38910)/Homer(0.968)
2		1e-50	-1.2E+02	12.19%	4.16%	12.8bp (14.6bp)	Bhlha15/MA0607.1/Jaspar(0.876)
3		1e-23	-5.4E+01	0.92%	0.02%	10.7bp (10.6bp)	CHR(?)/Hela-CellCycle-Expression/Homer(0.696)
4		1e-20	-4.7E+01	7.01%	2.94%	12.7bp (14.8bp)	NFATC2/MA0152.1/Jaspar(0.709)
5		1e-19	-4.6E+01	16.59%	9.98%	12.7bp (15.3bp)	POL009.1_DCE_S_II/Jaspar(0.655)
6		1e-17	-4.0E+01	0.97%	0.06%	12.1bp (11.6bp)	PH0152.1_Pou6f1_2/Jaspar(0.695)
7		1e-14	-3.4E+01	11.65%	6.84%	12.8bp (17.1bp)	MSX1/MA0666.1/Jaspar(0.922)
8		1e-14	-3.4E+01	0.44%	0.01%	8.9bp (4.9bp)	RUNX(Runt)/HPC7-Runx1-ChIP-Seq(GSE22178)/Homer(0.829)
9		1e-14	-3.3E+01	0.53%	0.01%	8.3bp (7.7bp)	MF0008.1_MADS_class/Jaspar(0.727)
10		1e-13	-3.2E+01	0.63%	0.03%	11.0bp (10.7bp)	ETS:E-box(ETS,bHLH)/HPC7-Scl-ChIP-Seq(GSE22178)/Homer(0.656)

Table S3.4. Genes with PITX1-dependent expression in both mouse hindlimb and mandible.

Gene_ID	HL_log2(FC)	MD_log2(FC)
2610203C20Rik	-0.545282	-0.553868
4833420G17Rik	0.798858	0.642856
AA388235	2.28493	1.1952
Acan	-4.15171	-2.72257
Actc1	-3.46846	0.507893
Adamtsl3	-1.1062	-1.3804
Ahcy	1.61177	1.89979
Alpk1	-0.965198	-1.05472
Aph1b	-0.464836	-1.01191
Bbs2	0.789873	0.736585
C130046K22Rik	-1.37301	-1.3957
Cdk2ap1	0.399188	0.476031
Col9a3	-1.51623	-1.23906
Coq2	0.475178	0.566638
Cox7b	0.994276	0.51941
Crabp1	0.366629	0.497713
Crispld2	-0.805215	-0.885958
Galnt13	-1.41565	1.50642
Glo1	1.07481	0.526986
Grik1	-1.3025	1.20103
Gtf2ird1	-0.465413	-0.471046
H2-Q7	1.24961	1.47527
Hic1	-0.551907	0.593188
Igfbp5	-0.431954	-0.452795
Kif27	-1.04409	-1.79342
Lars2	0.537165	2.38334
Lect1	-3.06869	-2.31126
Malat1	-0.478372	-0.669214
Matn4	-3.81821	-3.05599
Miat	-0.944688	-0.938831
Myog	-2.20236	0.492614
Nfatc2	-1.77459	-1.00308
Pitx1	-2.5669	-2.23662
Postn	-0.59825	-0.553372
Ppp1cc	1.40165	1.07112
Prrt2	-0.794778	-0.789668
Rps18	1.89479	1.13396
Rps2	1.072	1.15056
Slc15a2	-0.738264	-1.15644
Sox10	-0.975112	-1.95262
Sox9	-0.784984	-0.973009
Susd5	-1.8467	-2.24749
Tbx22	-1.1511	0.627534
Zfp429	-2.4277	-2.04196
Zfp459	-0.820253	-1.35959

Table S3.5. Representatives of PITX1 peaks identified in both hindlimb and mandible.

Coordinates of PITX1 mouse peaks (mm9) are listed.

Chromosome	Start	End
chr1	9573461	9574408
chr1	11320090	11320743
chr1	12561859	12562554
chr1	12804908	12805367
chr1	14026183	14026692
chr1	14297021	14297744
chr1	14480148	14480591
chr1	21511506	21512150
chr1	22310934	22311517
chr1	24791758	24792143
chr1	30594064	30594875
chr1	38744927	38745427
chr1	41949182	41949688
chr1	42828588	42829288
chr1	45817292	45818654
chr1	52628573	52628930
chr1	54275404	54275756
chr1	56307356	56307770
chr1	56922066	56922546
chr1	59326533	59327690
chr1	59391118	59391903
chr1	59944460	59945269
chr1	61290179	61291449
chr1	62062840	62063514
chr1	63609907	63610337
chr1	64034180	64034766
chr1	65081912	65082621
chr1	66377707	66378084
chr1	72048982	72049689
chr1	72098473	72098979
chr1	72534117	72534600
chr1	75454390	75454817
chr1	75555904	75556534
chr1	75829028	75830407
chr1	75982267	75982722
chr1	77103970	77104537
chr1	77223438	77224157
chr1	77224461	77224803
chr1	78936324	78936825
chr1	90716955	90717361
chr1	93849458	93850004
chr1	122611641	122612526
chr1	122902618	122903057
chr1	128579985	128580815
chr1	129339293	129340018
chr1	134521537	134522199
chr1	135016307	135016828
chr1	135492737	135493635
chr1	136413148	136413668
chr1	137926854	137927286

Table S3.6. Representatives of PITX1-bound enhancers that are shared in both mouse hindlimb and mandible. Coordinates of PITX1 mouse peaks (mm9) are listed.

Chromosome	Start	End
chr11	4208627	4209470
chr11	32825228	32826375
chr11	44950204	44951278
chr11	51203153	51204391
chr11	62780161	62781061
chr11	64464099	64464956
chr11	68567578	68568630
chr11	85445329	85446720
chr11	108808502	108809608
chr11	112273482	112274359
chr11	112875451	112876351
chr16	34935745	34936440
chr11	11708588	11709881
chr11	11754840	11755821
chr11	11778714	11779127
chr11	11837136	11837762
chr16	44531067	44531849
chr11	19333840	19334365
chr16	44615863	44616962
chr16	44632640	44633330
chr16	50753890	50754306
chr11	28635007	28635552
chr6	91185837	91186889
chr16	63451908	63452708
chr6	93862429	93863440
chr2	152564093	152565376
chr16	67251436	67252542
chr16	74092936	74094447
chr2	165974117	165974948
chr2	166005082	166006235
chr2	167994201	167995562
chr16	77329019	77330742
chr16	77636269	77637346
chr11	47192554	47193225
chr6	98643091	98643963
chr6	100371756	100372521
chr11	61298528	61299074
chr6	100926803	100927876
chr11	64781936	64783098
chr11	67876168	67876852
chr11	68451275	68452262
chr2	57481884	57482146
chr2	66027265	66027512
chr2	91933758	91933918
chr11	78410163	78410908
chr11	81262592	81263536
chr6	124690608	124691827
chr6	124869659	124870607
chr2	181747448	181747990
chr6	127000716	127001278

CHAPTER 4

CONCLUSIONS

In this dissertation I investigated on the regulatory network of PITX1 during limb and craniofacial development. The majority of previous publications have been focused on understanding the function of PITX1 from the traditional way, which is by elucidating the regulation between PITX1 and another particular gene or a specific pathway. This could narrow our thoughts and findings since it is much more complex to orchestrate organogenesis. Therefore, my dissertation aimed to take an integrative view to study the role of PITX1 in limb and craniofacial development by introducing the whole-genome level comparative approaches.

In Chapter 2, I started by applying PITX1 ChIP-seq in embryonic mouse and *Anolis* hindlimbs. Since the common ancestor of mammals and reptiles lived more than 300 million years ago, I was able to identify the ancient binding interactions, which are the common binding sites that are shared between mouse and green anole lizard. The conserved PITX1 binding events are strongly enriched near genes that are involved in limb morphogenesis. I further performed the globel gene experssion comparisons and identified a large number of misexpressed genes in *Pitx1-/-* hindlimbs. I revealed 440 putative targets of PITX1. They are significant misexpressed in the mutant hindlimb and have mouse PITX1 peaks that fall into their *cis*-regulatory domains. Up-regulated putative targets are enriched in limb patterning while down-regulated putative targets are enriched in cartilage and muscle development. 68 putative targets are ultra-conserved. They have conserved PITX1 binding sites in both mouse and *Anolis* hindlimbs. Our findings

conclude that PITX1 could promote chondrogenesis and myogenesis in the hindlimb by direct regulation of several key members of the cartilage and muscle transcriptional networks.

Chapter 3 of this dissertation provided a direction of comparing PITX1 function between developing hindlimb and mandible. I first performed ChIP-seq and RNA-seq in mouse mandibles. Both PITX1 binding events and PITX1-dependent expressed genes are strongly involved in craniofacial and skeletal development. I then compared the function of PITX1 in developing hindlimb and mandible at the transcriptome and the genome-wide binding levels. I successfully extracted both shared enhancer regions and PITX1-dependent genes that are significantly enriched in limb and craniofacial morphogenesis. These initial analyses suggest that PITX1 could modulate mouse hindlimb and mandible development through shared *cis*-reguatory elements and downstream targets.

Collectively the datasets presented in Chapter 2 and 3 provide for the importance of applying transscriptome and epigenetic analyses in studying the function of transcription factors and their associated *cis*-regulatory networks during embryonic development. Future directions of studying *cis*-regulatory elements and transcriptional complexes will be mainly focus on the application of high-throughput approaches during functional analyses, such as Genome-scale CRISPR-Cas9 knockout and transcriptinal activation screening.

REFERENCES

Abbasi, A.A., 2011. Evolution of vertebrate appendicular structures: Insight from genetic and palaeontological data. Dev. Dyn. 240, 1005-1016.

Abu-Issa, R., Smyth, G., Smoak, I., Yamamura, K., Meyers, E.N., 2002. Fgf8 is required for pharyngeal arch and cardiovascular development in the mouse. Development 129, 4613-4625.

Adams, C.C., Workman, J.L., 1995. Binding of disparate transcriptional activators to nucleosomal DNA is inherently cooperative. Mol. Cell. Biol. 15, 1405–1421.

Akiyama, H., Chaboissier, M.C., Martin, J.F., Schedl, A., de Crombrugghe, B., 2002. The transcription factor Sox9 has essential roles in successive steps of the chondrocyte differentiation pathway and is required for expression of Sox5 and Sox6. Genes Dev. 16, 2813–2828. doi:10.1101/gad.1017802

Alvarado, D.M., McCall, K., Aferol, H., Silva, M.J., Garbow, J.R., Spees, W.M., Patel, T., Siegel, M., Dobbs, M.B., Gurnett, C.A., 2011. Pitx1 haploinsufficiency causes clubfoot in humans and a clubfoot-like phenotype in mice. Hum. Mol. Genet. 20, 3943–3952. doi:10.1093/hmg/ddr313

Anders, S., Pyl, P.T., Huber, W., 2014. HTSeq: a Python framework to work with high-throughput sequencing data. Bioinformatics, 31 (2), 166–169. doi:10.1093/bioinformatics/btu638

Arnold, M.A., Kim, Y., Czubryt, M.P., Phan, D., McAnally, J., Qi, X., Shelton, J.M., Richardson, J.A., Bassel-Duby, R., Olson, E.N., 2007. MEF2C transcription factor controls

chondrocyte hypertrophy and bone development. Dev. Cell 12, 377–389. doi:10.1016/j.devcel.2007.02.004

Attanasio, C., Nord, A.S., Zhu, Y., Blow, M.J., Li, Z., Liberton, D.K., et al., 2013. Fine tuning of craniofacial morphology by distant-acting enhancers. Science 342(6157), 1241006–1241006. doi.org/10.1126/science.1241006

Barrow, J.R., Thomas, K.R., Boussadia-Zahui, O., Moore, R., Kemler, R., Capecchi, M.R., McMahon, A.P., 2003. Ectodermal Wnt3/beta-catenin signaling is required for the establishment and maintenance of the apical ectodermal ridge. Genes Dev. 17(3), 394–409. doi.org/10.1101/gad.1044903

Becker, J.S., Nicetto, D., Zaret, K.S., 2016. H3K9me3-Dependent Heterochromatin: Barrier to Cell Fate Changes. Trends in Genetics, 32(1), 29–41. doi.org/10.1016/j.tig.2015.11.001

Bernstein, B.E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., Bailey, D.K., Huebert, D.J., McMahon, S., Karlsson, E.K., Kulbokas, E.J., Gingeras, T.R., et al., 2005. Genomic maps and comparative analysis of histone modifications in human and mouse. Cell 120, 169–181.

Bi, W., Deng, J.M., Zhang, Z., Behringer, R.R., de Crombrugghe, B., 1999. Sox9 is required for cartilage formation. Nat. Genet. 22, 85–89. doi:10.1038/8792

Bobola, N., Carapuco, M., Ohnemus, S., Kanzler, B., Leibbrandt, A., Neubuser, A., Drouin, J., Mallo, M., 2003. Mesenchymal patterning by Hoxa2 requires blocking Fgf-dependent activation of Ptx1. Development 130, 3403-3414.

Cale, E., Wysocka, J., 2013. Modification of enhancer chromatin: what, how, and why? Mol. Cell 49(5):825-37.

Chan, Y.F., Marks, M.E., Jones, F.C., Villarreal, G., Shapiro, M.D., Brady, S.D., Southwick, A.M., Absher, D.M., Grimwood, J., Schmutz, J., Myers, R.M., Petrov, D., Jónsson, B.,

Schluter, D., Bell, M.A., Kingsley, D.M., 2010. Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. Science 327, 302–305. doi:10.1126/science.1182213

Chen, Y., Gridley, T., 2013. Compensatory regulation of the Snai1 and Snai2 genes during chondrogenesis. J. Bone Miner. Res. 28, 1412–1421. doi:10.1002/jbmr.1871

Cirillo, L.A., Lin, F.R., Cuesta, I., Friedman, D., Jarnik, M., Zaret, K.S., 2002. Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. Mol. Cell 9, 279–289.

Clapier, C.R., Cairns, B.R., 2009. The biology of chromatin remodeling complexes. Annu Rev Biochem. 78:273–304.

Clapier, C.R., Iwasa, J., Cairns, B.R., Peterson, C.L., 2017. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. Nat. Rev. Mol. Cell Biol. 18(7), 407–422. doi.org/10.1038/nrm.2017.26

Cohn, M.J., Tickle, C., 1996. Limbs: a model for pattern formation within the vertebrate body plan. Trends Genet 12, 253-257.

Cohn, M.J., Tickle, C., 1999. Developmental basis of limblessness and axial patterning in snakes. Nature 399, 474-479.

Cordero, D., Marcucio, R., Hu, D., Gaffield, W., Tapadia, M., Helms, J.A., 2004. Temporal perturbations in sonic hedgehog signaling elicit the spectrum of holoprosencephaly phenotypes. J. Clin. Invest. 114, 485-494.

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

Creuzet, S., Schuler, B., Couly, G., Le Douarin, N.M., 2004. Reciprocal relationships between Fgf8 and neural crest cells in facial and forebrain development. Proc. Natl. Acad. Sci. USA 101, 4843-4847.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., et al., 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc. Natl. Acad. Sci. USA 107, 21931–21936.

Danielson, K.G., Baribault, H., Holmes, D.F., Graham, H., Kadler, K.E., Iozzo, R.V., 1997. Targeted disruption of decorin leads to abnormal collagen fibril morphology and skin fragility. J. Cell Biol. 136, 729–743.

Daou, N., Lecolle, S., Lefebvre, S., Gaspera, B.D., Charbonnier, F., Chanoine, C., Armand, A.S., 2013. A new role for the calcineurin/NFAT pathway in neonatal myosin heavy chain expression via the NFATc2/MyoD complex during mouse myogenesis. Development 140, 4914–4925. doi:10.1242/dev.097428

Delaurier, A., Schweitzer, R., Logan, M.P.O., 2006. Pitx1 determines the morphology of muscle, tendon, and bones of the hindlimb. Dev. Biol. 299, 22–34. doi:10.1016/j.ydbio.2006.06.055

Depew, M.J., Lufkin, T., Rubenstein, J.L., 2002. Specification of jaw subdivisions by Dlx genes. Science 298, 381-385.

Depew, M.J., Simpson, C.A., Morasso, M., Rubenstein, J.L., 2005. Reassessing the Dlx code: the genetic regulation of branchial arch skeletal pattern and development. J. Anat. 207, 501-561.

Domyan, E.T., Kronenberg, Z., Infante, C.R., Vickrey, A.I., Stringham, S.A., Bruders, R., Guernsey, M.W., Park, S., Payne, J., Beckstead, R.B., Kardon, G., Menke, D.B., Yandell,

M., Shapiro, M.D., 2016. Molecular shifts in limb identity underlie development of feathered feet in two domestic avian species. eLife 5, e12115. doi:10.7554/eLife.12115

Duboc, V., Logan, M.P.O., 2011. Pitx1 is necessary for normal initiation of hindlimb outgrowth through regulation of Tbx4 expression and shapes hindlimb morphologies via targeted growth control. Development 138, 5301–5309. doi:10.1242/dev.074153

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

Funato, N., Kokubo, H., Nakamura, M., Yanagisawa, H., Saga, Y., 2016. Specification of jaw identity by the Hand2 transcription factor. Sci. Rep. 6, 28405. doi.org/10.1038/srep28405

Goldberg, A.D., Banaszynski, L.A., Noh, K.M., Lewis, P.W., Elsaesser, S.J., Stadler, S., Dewell, S., Law, M., Guo, X., Li, X., et al. 2010. Distinct factors control histone variant H3.3 localization at specific genomic regions. Cell 140, 678–691.

Gordon, C.T., Tan, T.Y., Benko, S., FitzPatrick, D., Lyonnet, S., Farlie, P.G., 2009. Long-range regulation at the SOX9 locus in development and disease. J. Med. Genet. 46, 649–656. doi:10.1136/jmg.2009.068361

Gordon, C.T., Attanasio, C., Bhatia, S., Benko, S., Ansari, M., Tan, T.Y., Munnich, A., Pennacchio, L.A., Abadie, V.R., Temple, I.K., Goldenberg, A., van Heyningen, V., Amiel, J., FitzPatrick, D., Kleinjan, D.A., Visel, A., Lyonnet, S., 2014. Identification of Novel Craniofacial Regulatory Domains Located far Upstream of SOX9and Disrupted in Pierre Robin Sequence. Hum. Mutat. 35, 1011–1020. doi:10.1002/humu.22606

Grifone, R., Demignon, J., Houbron, C., Souil, E., Niro, C., Seller, M.J., Hamard, G., Maire, P., 2005. Six1 and Six4 homeoproteins are required for Pax3 and Mrf expression during myogenesis in the mouse embryo. Development 132, 2235–2249. doi:10.1242/dev.01773

Grifone, R., Demignon, J., Giordani, J., Niro, C., Souil, E., Bertin, F., Laclef, C., Xu, P.-X., Maire, P., 2007. Eya1 and Eya2 proteins are required for hypaxial somitic myogenesis in the mouse embryo. Dev. Biol 302, 602–616. doi:10.1016/j.ydbio.2006.08.059

Grigoriou, M., Tucker, A.S., Sharpe, P.T., Pachnis, V., 1998. Expression and regulation of Lhx6 and Lhx7, a novel subfamily of LIM homeodomain encoding genes, suggests a role in mammalian head development. Development 125, 2063-2074.

Gualdi, R., Bossard, P., Zheng, M., Hamada, Y., Coleman, J.R., Zaret, K.S., 1996. Hepatic specification of the gut endoderm in vitro: cell signaling and transcriptional control. Genes Dev. 10(13), 1670–1682.

Gurnett, C.A., Alaee, F., Kruse, L.M., Desruisseau, D.M., Hecht, J.T., Wise, C.A., Bowcock, A.M., Dobbs, M.B., 2008. Asymmetric lower-limb malformations in individuals with homeobox PITX1 gene mutation. Am. J. Hum. Genet. 83, 616–622. doi:10.1016/j.ajhg.2008.10.004

Hanes, S.D., Brent, R., 1989. DNA specificity of the bicoid activator protein is determined by homeodomain recognition helix residue 9. Cell 57, 1275-1283.

Hare, E.E., Peterson, B.K., Iyer, V.N., Meier, R., Eisen, M.B., 2008. Sepsid even-skipped Enhancers Are Functionally Conserved in Drosophila Despite Lack of Sequence Conservation. PLoS Genet. 4, e1000106–13. doi:10.1371/journal.pgen.1000106

Harfe, B.D., Scherz, P.J., Nissim, S., Tian, H., McMahon, A.P., Tabin, C.J., 2004. Evidence for an expansion-based temporal Shh gradient in specifying vertebrate digit identities. Cell 118, 517-528.

Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A., et al., 2007. Distinct and predictive chromatin signatures of

transcriptional promoters and enhancers in the human genome. Nat. Genet. 39, 311–318.

Heinz, S., Benner, C., Spann, N., Bertolino, E et al., 2010. Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. Mol. Cell 38(4):576-589.

Hu, D., Helms, J.A., 1999. The role of sonic hedgehog in normal and abnormal craniofacial morphogenesis. Development 126, 4873-4884.

Huang, D.W., Sherman, B.T., Lempicki, R.A., 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 4, 44–57. doi:10.1038/nprot.2008.211

Infante, C.R., Park, S., Mihala, A.G., Kingsley, D.M., Menke, D.B., 2013. Pitx1 broadly associates with limb enhancers and is enriched on hindlimb cis-regulatory elements. Dev. Biol. 374, 234–244. doi:10.1016/j.ydbio.2012.11.017

Infante, C.R., Mihala, A.G., Park, S., Wang, J.S., Johnson, K.K., Lauderdale, J.D., Menke, D.B., 2015. Shared Enhancer Activity in the Limbs and Phallus and Functional Divergence of a Limb-Genital cis-Regulatory Element in Snakes. Dev. Cell 35, 107–119. doi:10.1016/j.devcel.2015.09.003

Ito, Y., Toriuchi, N., Yoshitaka, T., Ueno-Kudoh, H., Sato, T., Yokoyama, S., Nishida, K., Akimoto, T., Takahashi, M., Miyaki, S., Asahara, H., 2010. The Mohawk homeobox gene is a critical regulator of tendon differentiation. Proc. Natl. Acad. Sci. U.S.A. 107, 10538–10542. doi:10.1073/pnas.1000525107

Iwafuchi-Doi, M., Zaret, K.S., 2014. Pioneer transcription factors in cell reprogramming. Genes Dev. 28, 2679–2692.

Jarvis, B.R., Condie, B.G., 2017. Synthetic DNA templates for the production of in situ

hybridization probes. bioRxiv 108613. doi:10.1101/108613

Jeong, J., Li, X., McEvilly, R.J., Rosenfeld, M.G., Lufkin, T., Rubenstein, J.L., 2008. Dlx genes pattern mammalian jaw primordium by regulating both lower jaw-specific and upper jaw-specific genetic programs. Development 135,2905-2916.

Jepsen, K.J., Wu, F., Peragallo, J.H., Paul, J., Roberts, L., Ezura, Y., Oldberg, A., Birk, D.E., Chakravarti, S., 2002. A Syndrome of Joint Laxity and Impaired Tendon Integrity in Lumican- and Fibromodulin-deficient Mice. J. Biol. Chem. 277, 35532–35540. doi:10.1074/jbc.M205398200

Jin, C., Felsenfeld, G., 2007. Nucleosome stability mediated by histone variants H3.3 and H2A.Z. Genes Dev. 21, 1519–1529.

Jin, C., Zang, C., Wei, G., Cui, K., Peng, W., Zhao, K., and Felsenfeld, G., 2009. H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. Nat. Genet. 41,941–945.

Kanzler,B., Kuschert, S. J., Liu, Y.H., Mallo, M., 1998. Hoxa-2 restricts the chondrogenic domain and inhibits bone formation during development of the branchial area. Development 125, 2587-2597.

Kawakami, Y., Marti, M., Kawakami, H., Itou, J., Quach, T., Johnson, A., Sahara, S., O'Leary, D.D.M., Nakagawa, Y., Lewandoski, M., Pfaff, S., Evans, S.M., Belmonte, J.C.I., 2011. Islet1-mediated activation of the β-catenin pathway is necessary for hindlimb initiation in mice. Development 138, 4465–4473. doi:10.1242/dev.065359

Kengaku, M., Capdevila, J., Rodriguez Esteban, C., La Pena, De, J., Johnson, R. L., Izpisua Belmonte, J. C., Tabin, C. J., 1998. Distinct WNT pathways regulating AER formation and dorsoventral polarity in the chick limb bud. Science 280(5367), 1274–1277.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., Salzberg, S.L., 2013. TopHat2: accurate alignment of transcriptomes inthe presence of insertions, deletions and genefusions. Genome Biol. 14, R36. doi:10.1186/gb-2013-14-4-r36

Kimmel, C.B., Miller, C.T., Keynes, R.J., 2001. Neural crest patterning and the evolution of the jaw. J Anat. 199: 105-20.

Kohtz, J.D., Fishell, G., 2004. Developmental regulation of EVF-1, a novel non-coding RNA transcribed upstream of the mouse Dlx6 gene. Gene Expr. Patterns 4(4): 407-412.

Komori, T., Yagi, H., Nomura, S., Yamaguchi, A., Sasaki, K., Deguchi, K., Shimizu, Y., Bronson, R.T., Gao, Y.H., Inada, M., Sato, M., Okamoto, R., Kitamura, Y., Yoshiki, S., Kishimoto, T., 1997. Targeted disruption of Cbfa1 results in a complete lack of bone formation owing to maturational arrest of osteoblasts. Cell 89, 755–764.

Kouzarides, T., 2007. Chromatin modifications and their function. Cell 128(4): 693-705.

Kraus, P., Sivakamasundari, V., Lim, S.L., Xing, X., Lipovish, L., Lufkin, T., 2013. Making sense of Dlx1 antisense RNA. Dev. Biol. 376(2): 224-235.

Krumlauf, R., 1993. Hox genes and pattern formation in the branchial region of the vertebrate head. Trends Genet. 9, 106-112.

Lamonerie, T., Tremblay, J.J., Lanctôt, C., Therrien, M., Gauthier, Y., Drouin, J., 1996. Ptx1, a bicoid-related homeo box transcription factor involved in transcription of the pro-opiomelanocortin gene. Genes Dev. 10, 1284-1295.

Lanctôt, C., Lamolet, B., Drouin, J., 1997. The bicoid-related homeoprotein Ptx1 defines the most anterior domain of the embryo and differentiates posterior from anterior lateral mesoderm. Development 124, 2807-2817.

Lanctôt, C., Moreau, A., Chamberland, M., Tremblay, M.L., Drouin, J., 1999. Hindlimb patterning and mandible development require the Ptx1 gene. Development 126, 1805-1810.

Langmead, B., Trapnell, C., Pop, M., Salzberg, S.L., 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10, R25. doi:10.1186/gb-2009-10-3-r25

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nature Methods 9(4), *357–359.* doi:10.1038/nmeth.1923

Lin, Y.S., Carey, M., Ptashne, M., Green, M.R. 1990. How different eukaryotic transcriptional activators can cooperate promiscuously. Nature 345, 359–361.

Logan, C., Hornbruch, A., Campbell, I., Lumsden, A., 1997. The role of Engrailed in establishing the dorsoventral axis of the chick limb. Development, 124(12), 2317–2324.

Logan, M., Tabin, C.J., 1999. Role of Pitx1 upstream of Tbx4 in specification of hindlimb identity. Science 283, 1736-1739.

Loomis, C.A., Harris, E., Michaud, J., Wurst, W., Hanks, M., Joyner, A.L., 1996. The mouse Engrailed-1 gene and ventral limb patterning. Nature 382(6589), 360–363. doi.org/10.1038/382360a0

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology, 15 (12). doi:10.1186/s13059-014-0550-8

L'Honor, A., Coulon, V., Marcil, A., Lebel, M.L., Lafrance-Vanasse, J., Gage, P., Camper, S., Drouin, J., 2007. Sequential expression and redundancy of Pitx2 and Pitx3 genes during muscle development. Dev. Biol. 307, 421–433. doi:10.1016/j.ydbio.2007.04.034

L'Honore, A., Ouimette, J.F., Lavertu-Jolin, M., Drouin, J., 2010. Pitx2 defines alternate pathways acting through MyoD during limb and somitic myogenesis. Development 137,

3847–3856. doi:10.1242/dev.053421

Marcil, A., Dumonitier, E., Chamberland, M., Camper, S.A., Drouin, J., 2003. Pitx1 and Pitx2 are required for development of hindlimb buds. Development 130, 45-55.

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., Bejerano, G., 2010. GREAT improves functional interpretation of cis-regulatory regions. Nat. Biotechnol. 28, 495–501. doi:10.1038/nbt.1630

Medeiros, D.M., Grump, J.G., 2012. New perspectives on pharyngeal dorsoventral patterning in development and evolution of the vertebrate jaw. Dev. Biol 371, 121-135.

Meister, P., Mango, S.E., Gasser, S.M., 2011. Locking the genome: nuclear organization and cell fate. Curr. Opin. Genet. Dev. 21(2), 167–174. doi.org/10.1016/j.gde.2011.01.023

Menke, D.B., Guenther, C., Kingsley, D.M., 2008. Dual hindlimb control elements in the Tbx4 gene and region-specific control of bone size in vertebrate limbs. Development 135, 2543–2553. doi:10.1242/dev.017384

Mercader, N., Leonardo, E., Piedra, M.E., Martinez-A, C., Ros, M.A., Torres, M., 2000. Opposing RA and FGF signals contral proximodistal vertebrate limb development through regulation of Meis genes. Development 127, 3961-3970.

Miller, C.T., Yelon, D., Stainier, D.Y., Kimmel, C.B., 2003. Two endothelin 1 effectors, hand2 and bapx1, pattern ventral pharyngeal cartilage and the jaw joint. Development 130, 1353-1365.

Minoux, M., Antonarakis, G.S., Kmita, M., Duboule, D., Rijli, F.M., 2009. Rostral and caudal pharyngeal arches share a common neural crest ground pattern. Development 136, 637-645.

Minoux, M., Rijli, F.M., 2010. Molecular mechanisms of cranial neural crest cell migration and patterning in craniofacial development. Development, 137(16), 2605–2621. doi.org/10.1242/dev.040048

Mirny, L.A., 2010. Nucleosome-mediated cooperativity between transcription factors. Proc. Natl. Acad. Sci. USA 107, 22534–22539.

Motallebipour, M., Ameur, A., Reddy Bysani, M.S., Patra, K., Wallerman, O., Mangion, J., et al., 2009. Differential binding and co-binding pattern of FOXA1 and FOXA3 and their relation to H3K4me3 in HepG2 cells revealed by ChIP-seq. Genome Biology, 10(11), R129. doi.org/10.1186/gb-2009-10-11-r129

Molkentin, J.D., Black, B.L., Martin, J.F., Olson, E.N., 1995. Cooperative activation of muscle gene expression by MEF2 and myogenic bHLH proteins. Cell 83, 1125–1136. doi:10.1016/0092-8674(95)90139-6

Naiche, L.A., Papaioannou, V.E., 2003. Loss of Tbx4 blocks hindlimb development and affects vascularization and fusion of the allantois. Development 130, 2681–2693.

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

Niswander, L., Tickle, C., Vogel, A., Booth, I., Martin, G.R., 1993. FGF-4 replaces the apical ectodermal ridge and directs outgrowth and patterning of the limb. Cell 75, 579-587.

Niswander, L., 2003. Pattern formation: old models out on a limb. Nat Rev Genet. 4, 133-143.

Ohuchi, H., Nakagawa, T., Yamamoto, A., Araga, A., Ohata, T., Ishimaru, Y., Yoshioka, H., Kuwana, T., Nohno, T., Yamasaki, M., Itoh, N., Noji, S., 1997. The mesenchymal factor,

FGF10, initiates and maintains the outgrowth of the chick limb bud through interaction with FGF8, an apical ectodermal factor. Development 124, 2235-2244.

Osterwalder, M., Speziale, D., Shoukry, M., Mohan, R., Ivanek, R., Kohler, M., Beisel, C., Wen, X., Scales, S.J et al., 2014. HAND2 Targets Define a Network of Transcriptional Regulators that Compartmentalize the Early Limb Bud Mesenchyme. Dev. Cell 31, 345–357.

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., Selby, P.B., Owen, M.J., 1997. Cbfa1, a candidate gene for cleidocranial dysplasia syndrome, is essential for osteoblast differentiation and bone development. Cell 89, 765–771.

Ouimette, J.F., Jolin, M.L., L'honore, A., Gifuni, A., Drouin, J., 2010. Divergent transcriptional activities determine limb identity. Nat. Commun. 1, 1–9. doi:10.1038/ncomms1036

Park, S., Infante, C.R., Rivera-Davila, L.C., Menke, D.B., 2014. Conserved regulation of hoxc11 by pitx1 in Anolis lizards. J. Exp. Zool. B Mol. Dev. Evol. 322, 156–165. doi:10.1002/jez.b.22554

Petit, F., Sears, K. E., Ahituv, N., 2017. Limb development: a paradigm of gene regulation. Nature Rev. Genet. 18(4), 245–258. doi.org/10.1038/nrg.2016.167

Pfeifer, D., Kist, R., Dewar, K., Devon, K., Lander, E.S., Birren, B., Korniszewski, L., Back, E., Scherer, G., 1999. Campomelic dysplasia translocation breakpoints are scattered over 1 Mb proximal to SOX9: evidence for an extended control region. Am. J. Hum. Genet. 65, 111–124. doi:10.1086/302455

Prince, V., Lumsden, A., 1994. Hoxa-2 expression in normal and transposed rhombomeres: independent regulation in the neural tube and neural crest. Development 120, 911-923.

Pyron, R.A., 2010. A likelihood method for assessing molecular divergence time estimates and

the placement of fossil calibrations. Syst. Biol. 59, 185–194. doi:10.1093/sysbio/syp090

Qi, D.L., Ohhira, T., Fujisaki, C., Inoue, T., Ohta, T., Osaki, M., Ohshiro, E., Seko, Tomomi., Aoki, Shinsuke., Oshimura, Mitsuo., Kugoh, H., 2011. Identification of PITX1 as a TERT suppressor gene located on human chromosome 5. Mol. Cell. Biol. 31, 1624–1636. doi.org/10.1128/MCB.00470-10

Quinlan, A.R., Hall, I.M., 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842.

Riddle, R.D., Johnson, R.L., Laufer, E., Tabin, C., 1993. Sonic hedgehod mediates the polarizing activity of the ZPA. Cell 75, 1401-1416.

Rijli, F.M., Mark, M., Lakkaraju, S., Dierich, A., Dolle, P., Chambon, P., 1993. A homeotic transformation is generated in the rostral branchial region of the head by disruption of Hoxa-2, which acts as a selector gene. Cell 75, 1333-1349.

Rose, N.R., Klose, R.J., 2014. Understanding the relationship between DNA methylation and histone lysine methylation. Biochimica Et Biophysica Acta 1839(12), 1362–1372. doi.org/10.1016/j.bbagrm.2014.02.007

Ruest, L.B., Xiang, X., Lim, K.C., Levi, G., Clouthier, D.E., 2004. Endothelin-A receptor-dependent and -independent signaling pathways in establishing mandibular identity. Development 131, 4413-4423.

Sanger, T.J., Losos, J.B., Gibson-Brown, J.J., 2008. A developmental staging series for the lizard genus Anolis: a new system for the integration of evolution, development, and ecology. J. Morphol. 269, 129–137. doi:10.1002/jmor.10563

Santagati, F., Rijli, F.M., 2003. Cranial neural crest and the building of the vertebrate head. Nat. Rev. Neurosci. 4,806-818.

Sato, T., Kurihara, Y., Asai, R., Kawamura, Y., Tonami, K., Uchijima, Y., Heude, E., Ekker, M., Levi, G., Kurihara, H., 2008. An endothelin-1 switch specifies maxillomandibular identity. Proc. Natl. Acad. Sci. USA 105, 18806-18811.

Semina, E.V., Reiter, R., Leysens, N.J., Alward, W.L., Small, K.W., Datson, N.A., Siegel-Bartelt, J., Bierke-Nelson, D., Bitoun, P., Zabel, B.U., Carey, J.C., Murray, J.C., 1996. Cloning and characterization of a novel bicoid-related homeobox transcription factor gene, RIEG, involved in Rieger syndrome. Nat Genet. 14, 392-399.

Shapiro, M.D., Marks, M.E., Peichel, C.L., Blackman, B.K., Nereng, K.S., Jónsson, B., Schluter, D., Kingsley, D.M., 2004. Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. Nature 428, 717–723. doi:10.1038/nature02415

Shigetani, Y., Nobusada, Y., Kuratani, S., 2000. Ectodermally derived FGF8 defines the maxillomandibular region in the early chick embryo: epithelial-mesenchymal interactions in the specification of the craniofacial ectomesenchyme. Dev. Biol. 228, 73-85.

Shih, H.P., Gross, M.K., Kioussi, C., 2007. Expression pattern of the homeodomain transcription factor Pitx2 during muscle development. Gene Expr. Patterns 7, 441–451. doi:10.1016/j.modgep.2006.11.004

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

Spielmann, M., Brancati, F., Krawitz, P.M., Robinson, P.N., Ibrahim, D.M., Franke, M., Hecht, J., Lohan, S., Dathe, K., Nardone, A.M., Ferrari, P., Landi, A., Wittler, L., Timmermann, B., Chan, D., Mennen, U., Klopocki, E., Mundlos, S., 2012. Homeotic Arm-to-Leg Transformation Associated with Genomic Rearrangements at the PITX1 Locus. Am. J. Hum.

Genet. 91, 629–635. doi:10.1016/j.ajhg.2012.08.014

Spitz, F., Furlong, E.M., 2012. Transcription factors: from enhancer binding to developmental control. Nature Rev. Genet. 13, 613-626.

Stock, D.W., Ellies, D.L., Zhao, Z., Ekker, M., Ruddle, F. H., Weiss, K.M., 1996. The evolution of the vertebrate Dlx gene family. Proc. Natl. Acad. Sci. USA 93, 10858-10863.

Sun, X., Mariani, F.V., Martin, G.R., 2002. Function of FGF signaling from the apical ectodermal ridge in limb development. Nature 410, 501-508.

Svotelis, A., Gevry, N., Gaudreau, L., 2009. Regulation of gene expression and cellular proliferation by histone H2A.Z. Biochem. Cell Biol. 87, 179–188.

Szabo-Rogers, H.L., Geetha-Loganathan, P., Nimmagadda, S., Fu, K.K., Richman, J.M., 2008. FGF signals from the nasal pit are necessary for normal facial morphogenesis. Dev. Biol. 318, 289-302.

Szeto, D.P., Rodriguez-Esteban, C., Ryan, A.K., O'Connell, S.M., Liu, F., Kioussi, C., Gleiberman, A.S., Izpisúa-Belmonte, J.C., Rosenfeld, M.G., 1999. Role of the Bicoid-related homeodomain factor Pitx1 in specifying hindlimb morphogenesis and pituitary development. Genes Dev. 13, 484–494.

Tabin, C., Wolpert, L., 2007. Rethinking the proximodistal axis of the vertebrate limb in the molecular era. Genes Dev. 21, 1433-1442.

Tassabehji, M., Hammond, P., Karmiloff-Smith, A., Thompson, P., Thorgeirsson, S. S., Durkin, M. E., et al., 2005. GTF2IRD1 in craniofacial development of humans and mice. Science 310(5751), 1184–1187. doi:/10.1126/science.1116142

Trainor, P.A., Melton, K.R., Manzanares, M., 2003. Origins and plasticity of neural crest cells and their roles in jaw and craniofacial evolution. Int J Dev Biol. 47(7-8), 541–553.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., Pachter, L., 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat. Protoc. 7, 562–578. doi:10.1038/nprot.2012.016

Trumpp, A., Depew, M.J., Rubenstein, J.L., Bishop, J.M., Martin, G.R., 1999. Cre-mediated gene inactivation demonstrates that FGF8 is required for cell survival and patterning of the first branchial arch. Genes Dev. 13, 3136-3148.

Vasyutina, E., Stebler, J., Brand-Saberi, B., Schulz, S., Raz, E., Birchmeier, C., 2005. CXCR4 and Gab1 cooperate to control the development of migrating muscle progenitor cells. Genes Dev. 19, 2187–2198. doi:10.1101/gad.346205

Visel, A., Minovitsky, S., Dubchak, I., Pennacchio, L.A., 2007. VISTA Enhancer Browser--a database of tissue-specific human enhancers. Nucleic Acids Res. 35, D88–92. doi:10.1093/nar/gkl822

Visel, A., Blow, M. J., Li, Z., Zhang, T., Akiyama, J. A., Holt, A., et al., 2009. ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature 457(7231), 854–858. doi.org/10.1038/nature07730

Wanek, N., Muneoka, K., Holler-Dinsmore, G., Burton, R., Bryant, S.V., 1989. A staging system for mouse limb development. J. Exp. Zool. 249, 41–49. doi.org/10.1002/jez.1402490109

Wang, J.S., Infante, C.R., Park, S., Menke, D.B., 2017. PITX1 promotes chondrogenesis and myogenesis in mouse hindlimbs through conserved regulatory targets. Under review.

Wilkinson, D.G., 1992. In situ hybridization: a practical approach. Oxford University Press.

Wilson, J., Tucker, A.S., 2004. Fgf and Bmp signals repress the expression of Bapx1 in the mandibular mesenchyme and control the position of the developing jaw joint. Dev. Biol. 266, 138-150.

Wunderle, V.M., Critcher, R., Hastie, N., Goodfellow, P.N., Schedl, A., 1998. Deletion of long-range regulatory elements upstream of SOX9 causes campomelic dysplasia. Proc. Natl. Acad. Sci. U.S.A. 95, 10649–10654.

Ye, T., Krebs, A.R., Choukrallah, M.A., Keime, C., Plewniak, F., Davidson, I., Tora, L., 2011. seMINER: an integrated ChIP-seq data interpretation platform. Nucleic Acids Res. 39(6): e35. Doi:10.1093/nar/gkq1287.

Yoshida, C.A., Yamamoto, H., Fujita, T., Furuichi, T., Ito, K., Inoue, K.-I., Yamana, K., Zanma, A., Takada, K., Ito, Y., Komori, T., 2004. Runx2 and Runx3 are essential for chondrocyte maturation, and Runx2 regulates limb growth through induction of Indian hedgehog. Genes Dev. 18, 952–963. doi:10.1101/gad.1174704

Yoshioka, H., Meno, C., Koshiba, K., Sugihara, M., Itoh, H., Ishimaru, Y., Inoue, T., Ohuchi, H., Semina, E.V., Murray, J.C., Hamada, H., Noji, S., 1998. Pitx2, a bicoid-type homeobox gene, is involved in a lefty-signaling pathway in determination of left-right asymmetry. Cell 94, 299-305.

Young, N. M., Hallgrímsson, B., 2005. Serial homology and the evolution of mammalian limb covariation structure. Evolution 59(12), 2691–2704.

Zaret, K.S., Mango, S.E., 2016. Pioneer transcription factors, chromatin dynamics, and cell fate control. Curr. Opi. Genet. Dev. *37*, 76–81. doi:10.1016/j.gde.2015.12.003

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

Zhao, H., Zhou, W., Yao, Z., Wan, Y., Cao, J., Zhang, L., Zhao, J. PITX1 promotes chondrocyte and myoblast differentiation through conserved regulatory targets, Li, H., Zhou, R., Li, B., Wei, G., Zhang, Z., French, C.A., Dekker, J.D., Yang, Y., Fisher, S.E., Tucker, H.O., Guo, X., 2015. Foxp1/2/4 regulate endochondral ossification as a suppresser complex. Dev. Biol. 398, 242–254. doi:10.1016/j.ydbio.2014.12.007

Zuniga, E., Stellabotte, F., Crump, J.G., 2010. Jagged-Notch signaling ensures dorsal skeletal identity in the vertebrate face. Development 137, 1843-1852.