

UCSF

UC San Francisco Electronic Theses and Dissertations

Title

Cis-regulatory elements in limb development and human congenital malformations

Permalink

<https://escholarship.org/uc/item/97z7q8px>

Author

VanderMeer, Julia

Publication Date

2013

Peer reviewed|Thesis/dissertation

Cis-regulatory elements in limb development and human congenital malformations

by

Julia E. VanderMeer

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biomedical Sciences

in the

Copyright (2013)

By

Julia E VanderMeer

Acknowledgements

This work would not have been possible without the support and contributions of many individuals. First, I want to thank my mentor, Nadav. His ideas, creativity, dedication and fearlessness have shaped me into a better scientist. I am proud to have been one of the first members of the Ahituv lab and it's been fun to watch the lab grow and mature.

Members of the Ahituv lab have not only made many contributions to my work, but have made working in the lab a lot of fun. Arriving to work with a new PI with Loan and Meej to set up a functional lab and zebrafish facility was an adventure. The scientific mentorship and friendship of Robin, Ramon, Walter and Taka have been instrumental in creating a scientifically challenging environment. Traveling through grad school along with Meej and Nir has been like having a built-in support group for help with reagents, cloning, spelling, fashion and music. Karl has been the lab's grumpy technician who can always be relied on to give straight answers, know how to fix things that are broken and identify California wildlife. Numerous other members of the lab, whether they were there for a few months or for the long haul, have all contributed to the lab culture and enriched our group and I'm glad I got the chance to work with them.

Beyond any other person, Yien has been a mentor and a role model for science and life. As a lab neighbor, I could always go to her for protocols and troubleshooting advice and I knew that she would be able to help. As a friend, I could always go to her for recipes, stories, encouragement and life-troubleshooting advice and she was there to help in those situations as well. I will be forever grateful for her friendship and for teaching me to really think like a scientist.

I'd like to thank the members of my thesis committee, Matthias and Jun, for their much-appreciated input and advice along the way. The expertise they brought to different aspects of my projects was very helpful in shaping this work.

I also want to thank people outside my lab at UCSF that provided help and inspiration along the way. Group meetings with the Wynshaw-Boris, Klein, Hsaio, Lomvardas labs and others were a source of scientific conversations and interesting challenges to the way I viewed my data. The staff of various UCSF core facilities lent me their expertise in techniques and access to equipment that let me do experiments that would not have been possible without their support.

Thank you to members of the BMS office, especially Lisa and Monique, for taking care of so many things for me that I didn't even know about.

I feel very grateful to have made such good friends in San Francisco that have brought joy to the years I spent on this. Thank you to my friends in the BMS program, particularly Justin, Ashley, Robyn, Kasia and Karen. Thank you to Mike for everything from protocols to friendship. Being able to share successes and good times and to commiserate with you about the rest of graduate school has been essential to getting through this process. I've been incredibly lucky to have friends from home in Toledo who also moved to San Francisco: Amirah, Priya, Reena – thank you for being here for laughter and advice, for dinners and hikes, and for friends I met through you particularly Dave, Arun and Cooper.

My mom and dad have been supportive even when they had no idea what I was doing or why I wanted to do it. Karen and Brian have always been a source of encouragement and bad puns. Thank you for everything, I love you all.

Contributions of co-authors to the presented work

This dissertation contains excerpts of previously published material. Permission for this material has been obtained from the publishers.

Chapters 1-3 contain material from VanderMeer, J.E., and Ahituv, N. (2012). “cis-Regulatory Enhancer Mutations are a Cause of Human Limb Malformations.” In Gene Regulatory Sequences and Human Disease, pp. 73–93. JEV is the author of this material and NA provided comments and editorial assistance. (VanderMeer and Ahituv, 2012)

Chapter 3 contains material from: VanderMeer, J.E., Afzal, M., Alyas, S., Haque, S., Ahituv, N., and Malik, S. (2012). *A novel ZRS mutation in a Balochi tribal family with triphalangeal thumb, pre-axial polydactyly, post-axial polydactyly, and syndactyly*. Am. J. Med. Genet. Part A 158A, 2031–2035 and Laurell, T., VanderMeer, J.E., Wenger, A.M., Grigelioniene, G., Nordenskjöld, A., Arner, M., Ekblom, A.G., Bejerano, G., Ahituv, N., and Nordgren, A. (2012). *A novel 13 base pair insertion in the sonic hedgehog ZRS limb enhancer (ZRS/LMBR1) causes preaxial polydactyly with triphalangeal thumb*. Hum. Mutat. 33, 1063–1066. JEV did all molecular experiments and wrote the manuscripts. MA, SA, SH, SM, TL, GG, AN, MA, AGE identified and examined subjects. AMW and GB assisted with computational analysis of TFBS. NA, SM and AN provided comments on the experiments and manuscripts.

Chapter 5 contains material from VanderMeer, J.E., and Ahituv, N. (2011). *cis-Regulatory Mutations are a Genetic Cause of Human Limb Malformations*. Dev. Dyn. 240, 920–930. JEV is the author of this material and NA provided comments and editorial assistance.

The role of *cis*-regulatory enhancers in tetrapod limb development

Julia E VanderMeer

Abstract

Regulatory elements provide information necessary for the spatial, temporal and dosage appropriate expression of genes. Developmental genes in particular rely on *cis*-regulatory enhancers to control expression in the various tissues where they are active and cause defects in development. Congenital limb malformations are the second most common class of human birth defects and can be caused both by environmental and genetic factors, and identifying the causal mutation in a patient with an isolated limb malformation is often difficult. The difficulty of identification may be due, in part, to the growing number of cases with isolated limb malformations that are shown to be the result of nucleotide changes in regulatory elements. These regulatory mutations affect gene expression in the developing limb and can cause dramatic changes to patterning, leading to congenital limb malformations. There are multiple examples of mutations in an enhancer known as the zone of polarizing activity regulatory sequence (ZRS) that cause preaxial polydactyly and other malformation phenotypes. The identification of further ZRS mutations along with changes they have to transcription factor binding and resulting

phenotypes can help elucidate the mechanisms by which it controls gene expression. While the genes and pathways that determine specific limb signaling centers have been described, the identification of enhancers that determine these centers has been limited. It is possible to identify enhancers that are specific to the zone of polarizing activity (ZPA) and apical ectodermal ridge (AER) signaling centers by isolating these regions. Using H3K27ac ChIP-seq on mouse E11.5 ZPA and AER fluorescently sorted cells, I identified thousands of specific signaling center enhancers. Mouse transgenic assays confirmed that several of them function as ZPA and AER enhancers. Combined, these results provide novel ZPA and AER enhancers that may play important roles in limb development. Because the ZPA and AER have critical roles in establishing the three axes and patterning the limb, changes in enhancer function can result in malformations of limb structures.

Table of Contents

Chapter 1 – Introduction

1.1 – <i>Cis</i> -regulatory enhancers and gene regulation.....	1
1.2 – Identifying and studying enhancers	4
1.3 – Tetrapod limb development as a model.....	7

Chapter 2 – Models of enhancer activity in limb development

2.1 – Limb development; tissue patterning along three axes.....	10
2.2 – Studying enhancers in models of the tetrapod limb.....	13
2.3 – Modeling limb enhancer activity in zebrafish	15
2.4 – Modeling limb enhancer activity in mouse.....	20
2.5 – Studying congenital limb malformations in human patients	21

Chapter 3 – The ZRS and human malformations

3.1 – Background of the ZRS enhancer in limb development.....	24
3.2 – Patient Sequencing Results	33
3.3 – ZRS 278 C>A	34
3.4 – ZRS 603ins13	40
3.5 – ZRS 402 C>T Homozygous Mutation.....	46
3.6 – Conclusions.....	54
3.7 – Methods.....	56

Chapter 4 – ChIP-seq identifies limb signaling center enhancers

4.1 – ChIP-seq introduction	59
4.2 – Genome-wide identification of enhancers in signaling centers	61
4.3 – Results and discussion	62
4.4 – Methods.....	68

Chapter 5 – Conclusions and future directions

5.1 – Summary	71
5.2 – Evidence for additional <i>cis</i> -regulatory mutations	71
5.3 – Interpreting regulatory mutations	74
5.4 – Outlook	75

References	77
-------------------------	----

Appendix I	91
-------------------------	----

Appendix II	101
--------------------------	-----

Appendix III	115
---------------------------	-----

List of Tables

Chapter 2

Table 1. Isolated limb malformation sample database summary.....23

Chapter 3

Table 2. Mutations in the ZRS and associated human phenotypes..... 30
Table 3. Phenotypes in the ZRS 287C>A family.38
Table 4. Phenotypes in the ZRS 402 C>T families.49
Table 5. Primers for human ZRS sequencing.57

List of Figures

Chapter 1

Figure 1. <i>Cis</i> -regulatory elements control gene expression.....	1
Figure 2. Methods for identifying enhancers.....	5
Figure 3. A model of enhancer activity in development.....	8

Chapter 2

Figure 4. The limb bud is patterned along three axes.....	11
Figure 5. The <i>in vivo</i> enhancer assay.....	15
Figure 6. Testing limb enhancers in zebrafish.....	19
Figure 7. Phenotypes of limb malformations.....	22

Chapter 3

Figure 8. ZRS expression and preaxial polydactyly.....	26
Figure 9. Pedigree and chromatograms of the ZRS 287C>A.....	35
Figure 10. Phenotypic presentation of ZRS 287C>A.....	37
Figure 11. Pedigree and phenotypic presentation of ZRS 603ins13.....	41
Figure 12. Chromatogram and TFBS of ZRS 603ins13.....	42
Figure 13. Pedigree and phenotypic presentation of ZRS 402C>T.....	48
Figure 14. Chromatogram and mouse transgenic assay of ZRS 402C>T.....	51

Chapter 4

Figure 15. ChIP-seq method.....	59
Figure 16. ZPA and AER ChIP-seq results.....	63
Figure 17. Signaling center genes are marked by H3K27ac.....	65
Figure 18. Active enhancers in the ZPA and AER.....	66

Chapter 1 – Introduction

1.1 – *Cis*-regulatory enhancers and gene regulation

With the analysis of the complete human genome sequence, it quickly became apparent that it contains fewer protein-coding genes than expected and a lot more DNA that does not code for protein. The human genome consists of ~3.2 billion base pairs and contains around 20,500 genes (Clamp et al., 2007) which make up only ~1.6% of its content.

Repetitive sequences make up an additional ~45% and the remaining 53% is made up of DNA sequences of unknown function. This relatively small number of genes reinforced the idea that many genes serve multiple functions, a fact that is especially evident in genes with roles in tissue patterning and embryonic development. The regulation of these genes is very important in order for these roles to be executed at the proper time, place and expression levels. It appears that some of the regulation for many developmental genes occurs through the actions of *cis*-regulatory elements - sequences of noncoding DNA that control the expression of nearby genes. There are multiple types of *cis*-regulatory elements that can affect gene expression (Figure 1).

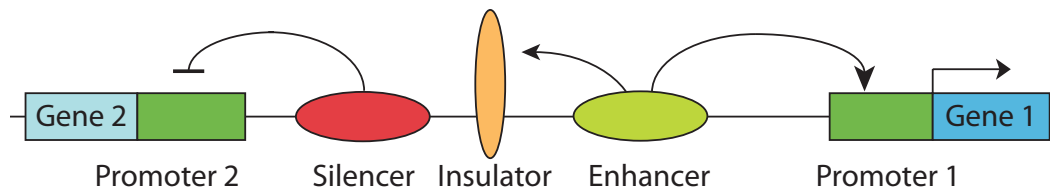


Figure 1. *Cis*-regulatory elements control gene expression. In this example scenario, the enhancer acts on the promoter of Gene 1 to activate expression. The insulator prevents the enhancer from acting on Gene 2 and the silencer prevents expression of Gene 2.

The best-characterized *cis*-regulatory element is the promoter. The promoter is a region at the beginning of the gene that serves as the site where transcription machinery assembles and transcription of the gene begins. The core promoter is the minimal stretch of DNA that allows the RNA polymerase II machinery to initiate transcription (Butler and Kadonaga, 2002; Smale and Kadonaga, 2003). It is typically only 35-40 base pairs (bp) long and often includes sequence motifs like the TATA box, TFIIB recognition element (BRE), initiator element (Inr) and the downstream core promoter element (DPE). The promoter region can also include a nearby *cis*-regulatory region known as the proximal promoter that extends up to a few hundred base pairs upstream of the transcription start site and includes many transcription factor binding sites (TFBS).

While promoters are critical for gene expression, the majority of transcription control lies beyond the promoter (Gerstein et al., 2012; Thurman et al., 2012). The *cis*-regulatory element that is the focus of this dissertation is the enhancer. Enhancers are DNA sequences that up-regulate gene expression in a specific temporal and spatial manner. They often have modular expression patterns, and a gene that is active in many tissues is likely influenced by multiple enhancers (Visel et al., 2009a). Enhancer function is usually considered to be independent of location or orientation relative to the gene. Enhancers are not located at fixed positions relative to the genes that they regulate. They can be 3' or 5' of the gene and have been found farther than 1 megabase away from their target gene promoter. Enhancers can be found in intergenic DNA or within introns of the regulated gene or unrelated genes. They are thought to function through the recruitment of transcription factors (TFs) and subsequent physical interactions with the gene promoter.

Developmental enhancers are often conserved through evolution, but this is not always the case (Blow et al., 2010). In the tissue context where they are active, enhancers usually show a characteristic pattern of histone modifications including monomethylation at both H3K4 and H2BK5, low H3K4 tri-methylation, H3K27 acetylation and enrichment of the H2A.Z histone variant (Barski et al., 2007; Heintzman et al., 2009). Enhancers are thought to function through the recruitment of transcription factors (TFs) and subsequent physical interactions with the gene promoter (Levine, 2010; Maston et al., 2006).

A related *cis*-regulatory element is the silencer, which acts to repress gene expression. Like enhancers, silencers are thought to be orientation-independent and can be located almost anywhere with regard to the genes that they regulate. Silencers are thought to act by recruiting repressor proteins, which can influence the local chromatin or by inhibiting the recruitment of activating factors (Cecchini et al., 2009; Maston et al., 2006; Perissi et al., 2010). Some sequences may function as silencers in one context and enhancers in another, depending on their chromatin state and the presence of particular transcription factors.

Insulators, also known as boundary elements, are *cis*-regulatory regions that prevent the transcriptional activity of one gene from affecting neighboring genes. The mechanism by which insulators delineate regions of active gene transcription is unknown, but involves limiting the range of permissive chromatin modifications. One protein that was shown to be important for insulator activity is the CCCTC-binding factor (CTCF), which is thought to insulate by modifying DNA methylation (Maston et al., 2006; West et al., 2002).

1.2 – Identifying and studying enhancers

One major challenge in the study of *cis*-regulatory elements is our relatively poor ability to detect and characterize them. Currently, other than the promoter, the most studied element is the enhancer. Enhancers can be identified using various techniques. One such method is the enhancer trap; the integration of a reporter gene along with a minimal promoter (a promoter that is not sufficient to drive reporter expression without the presence of a functional enhancer) randomly in the genome followed by a screen for the reporter gene expression (Bellen, 1999; Korn et al., 1992; Parinov et al., 2004). Two drawbacks of this method are that it does not pinpoint the exact location of the enhancer or the identity of the gene that it regulates.

A more efficient way to identify enhancers is by using comparative genomics. By aligning genomes of different species, it is possible to identify noncoding regions that have a high degree of conservation. Evolutionary conservation can imply that a sequence is functional and it is thought that many noncoding conserved regions may have *cis*-regulatory roles (Figure 2) (Dermitzakis et al., 2005; Pennacchio et al., 2006). Depending on the degree of conservation and the species used for the alignments, between 10% and 60% of these conserved noncoding elements have been shown to function as enhancers (Pennacchio et al., 2006; Woolfe et al., 2005). In general, comparisons between distantly related species yield fewer conserved sequences, but the sequences that are conserved between these species are likely to be functional. Though this approach can identify some enhancers, there is also evidence that not all enhancers can be detected by evolutionary conservation (Blow et al., 2010). Relying on conservation is also problematic because

enhancers that have undergone recent selection due to changes in their functional roles will not be detected even though they could well be some of the most important enhancers dictating changes in morphology between species.

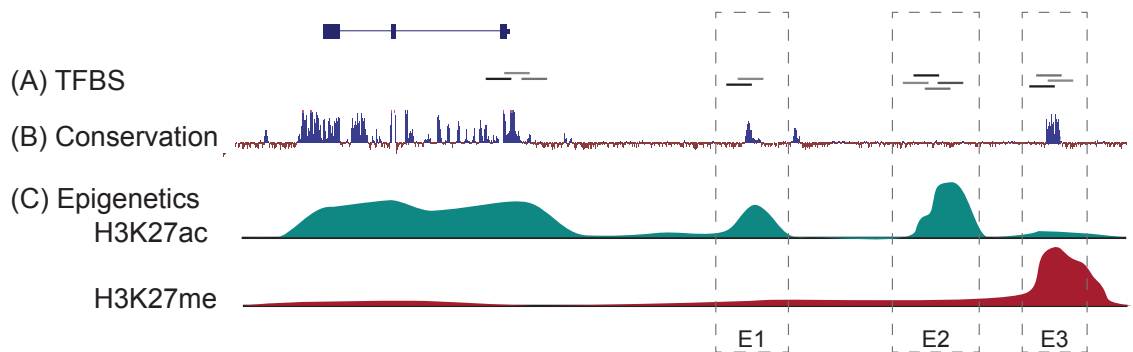


Figure 2. Different characteristics of enhancers have been used to identify them in the genome. Enhancers can be identified based on (A) the presence and pattern of TFBS present in the sequence, as depicted by gray lines (B) the conservation of genomic sequence between species (blue peaks signify evolutionary conserved regions), and (C) epigenetic state – with active marks like H3K27ac or repressive marks like H3K27me3 (peaks signify marked regions). In this cartoon, the gene is expressed (as indicated by the epigenetic profiles). There are three potential enhancers labeled E1, E2 and E3. E1 shows marks of enhancer status based on TFBS predictions, conservation and epigenetic state. E2 shows marks of enhancer status based on TFBS predictions and epigenetic state, but is not conserved. E3 shows conservation and TFBS predictions, but does not have the epigenetic signature of an active enhancer. E3 could function as an enhancer in a different tissue other than the one profiled and would have the active epigenetic signature in those cells.

Another approach used to detect enhancers focuses on the role of TFs in enhancer activity by computationally identifying regions where the TFBS sequences appear at some frequency in specific combinations, (Gotea et al., 2010; Vokes et al., 2008). Many efforts have been made to discover a ‘code’ of TFBS that determines the enhancer potential of a given sequence (Smith et al., 2013a, 2013b).

In addition to computational predictions for TF binding, biochemical methods to measure actual interactions between TFs and DNA can be used to identify enhancers. Enhancers that are involved in a particular genetic pathway have been identified by using chromatin immunoprecipitation (ChIP) in combination with microarrays (ChIP-chip) or next-generation sequencing technologies (ChIP-seq) to locate regions bound by specific transcription factors or enhancer-associated proteins like E1A binding protein p300 (EP300/p300) (Gotea et al., 2010; Visel et al., 2009b; Vokes et al., 2008). ChIP uses antibodies to specific DNA-interacting proteins to selectively pull down only the fragments of sequence that are bound to those proteins. These fragments can then be realigned to the genome in order to identify the location where the protein was bound.

A final approach is to use more general properties of *cis*-regulatory elements to detect them in a genome-wide manner. The most common epigenetic marks used for identifying enhancers are histone modifications. DNA is packaged with histone proteins where 147bp of DNA is wrapped around 8 histone proteins in a complex composed of H2A, H2B, H3 and H4 protein pairs to make up a nucleosome. These histone proteins have tails that extend from the core and can have various chemical modifications including methylation, acetylation, phosphorylation, ubiquination and sumoylation. These modifications impact the interaction between the protein and DNA and can determine the state of the chromatin (Rando, 2012; Wang et al., 2008). Enhancers are located in open chromatin, which is associated with a single methyl group on the 4th lysine of histone H3 (H3K4me1) among others. Closed chromatin, indicative of silencing, can be identified by methyl groups on the lysine in position 27 of H3 (H3K27me) among others. Antibodies

developed against these modifications can allow the identification of chromatin state and regulatory potential (active, silenced etc.) of specific sequences using ChIP-seq (Figure 2) (Barski et al., 2007; Heintzman et al., 2009).

Within the broad category of *cis*-regulatory enhancers, it is possible to use epigenetic signatures to distinguish between enhancers that are active and those that are “poised” but not active yet in a particular tissue. While both types of enhancers have the H3K4me1 modification and are associated with p300, active enhancers are also marked by acetylation of the 27th lysine of H3 (H3K27ac) while poised enhancers do not have this mark (Creyghton et al., 2010; Rada-Iglesias et al., 2012). This kind of distinction between classes of enhancer allows a better understanding of the mechanisms of regulation in development. ChIP-based methods have the advantage of identifying enhancers that are specific to the tissue used for the assay and they have the ability to detect enhancers that do not show significant evolutionary conservation (Blow et al., 2010). By looking at cells in different stages of differentiation or tissues at different developmental stages, it is possible to begin to reconstruct dynamic changes in enhancer activity.

1.3 – Tetrapod limb development as a model

Many developmental genes play key roles in multiple tissues and at different developmental stages. A coding mutation in one of these genes would cause a syndromic phenotype that consists of the sum of the effects to these different tissues and stages. On the other hand, a mutation in a *cis*-regulatory element that controls one aspect of the gene’s expression would only cause a phenotype due to the change to the particular tissue

where the *cis*-regulatory element is active (Figure 3). This model implies that malformations which occur in both syndromic and isolated forms could represent the results of mutations in coding genes and in the *cis*-regulatory elements that control these genes, respectively. Human limb malformations occur in both syndromic and isolated forms, making this class of malformation an informative field for studying the effect of *cis*-regulatory mutations on development. In order to study the effect of these mutations, we can use the rich knowledge of tetrapod limb development. Many years of research on limb development has led to a detailed characterization of gene expression patterns and phenotypes that result from changes to gene expression.

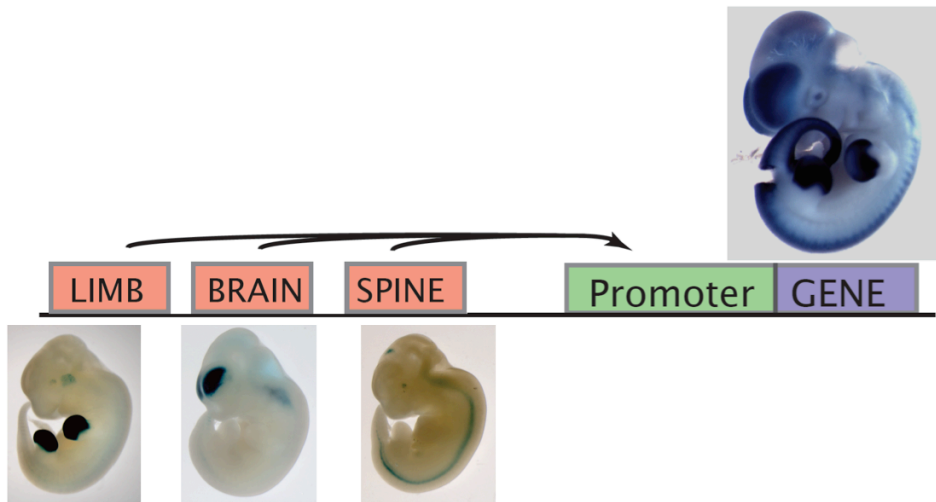


Figure 3. The model of enhancer activity in development focuses on the modular expression patterns of different enhancers for the same gene. In the example above, there are three enhancers with each controlling expression in a specific tissue. The combined profile of enhancer expression makes the complete gene expression profile. While a mutation in the gene would be expected to affect the limb, brain and spine tissues, a mutation in just the limb enhancer would leave the other tissues unaffected and cause only a limb phenotype.

Human congenital limb malformations occur in as many as 1 in 500 births and, collectively, represent the second most common form of congenital defect (Moore and Persaud, 1998). These malformations display a wide range of severities from small changes to digit morphology that do not impair function to more detrimental malformations such as digit fusion or limb truncations that cause severe functional impairment. In some cases, surgery is necessary to restore or improve function. Some types of limb malformations are caused by environmental factors such as exposure to teratogens or by physical constraints like amniotic bands and vascular disruptions in developing embryos, but there can also be genetic causes. The genetic etiology of congenital limb malformations is important to understand for purposes of genetic counseling. Discovering the various genetic mechanisms that cause human limb malformations also provides insight to genes and pathways that are important for tetrapod limb development.

While human limb malformations have been studied since the early 19th century (Farabee, 1903), the identification of causal mutations is difficult. Though gene mutations have been shown to be the cause of some limb malformations, these are often in the context of a syndrome with phenotypes that affect other tissues and organs (Schwabe and Mundlos, 2004). The genetic causes of isolated limb malformations – those that occur in a patient who has no other tissues or organs affected – have proven more difficult to discover. This may be because some of the mutations that cause congenital limb malformations are in regulatory DNA regions that affect the expression of genes important in limb development.

Chapter 2 – Models of enhancer activity in limb development

2.1 – Limb development; tissue patterning along three axes

Many limb malformations can arise from changes in patterning that occur in the early stages of limb development. Normal limb development requires the coordinated establishment of the anterior-posterior (AP), proximal-distal (PD), and dorsal-ventral (DV) axes (Figure 4). Through classical developmental biology studies on developing mouse and chicken limbs, a lot is known about the genes that control these axes and the networks through which they interact.

Early development and axis specification

The limb bud begins as a thickening of mesenchyme cells from the lateral plate mesoderm and somites at the flank of the embryo. This bulge of cells is called the limb bud. Limb bud outgrowth begins with expression of fibroblast growth factor 10 (*FGF10*) in the lateral plate mesoderm which signals through Wnt proteins to induce fibroblast growth factor 8 (*FGF8*) in the ectoderm. *FGF8* in turn stimulates further expression of *FGF10* in the mesoderm, creating a feedback loop that causes proliferation. This signaling also induces the overlying ectoderm cells to form a structure called the apical ectodermal ridge (AER; Figure 4) at the edge of the limb bud. The AER becomes the primary signaling center that determines the outgrowth of the limb and has a role in establishing the PD axis. Signals from the AER sustain mitotic proliferation in the underlying cells and if AER signaling is removed, physically or through genetic manipulations, further development of the distal limb ceases (Summerbell, 1974).

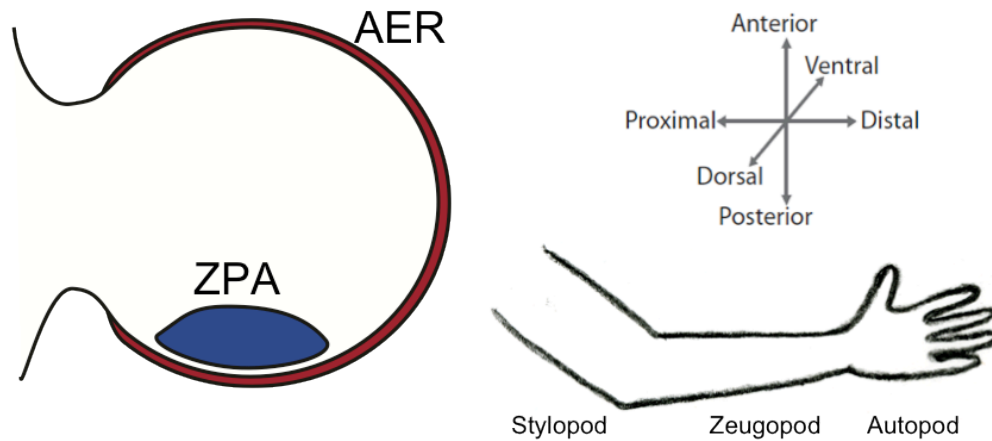


Figure 4. The limb bud is patterned along three axes: anterior-posterior, proximal-distal and dorsal-ventral. The two primary signaling centers of the limb bud are the AER (in red) and the ZPA (in blue). As the limb bud grows out and develops, the limb is divided into three parts: stylopod, zeugopod and autopod.

The AP axis is also controlled by a signaling center called the zone of polarizing activity (ZPA; Figure 4). The ZPA is composed of a small region of mesodermal cells at the posterior of the limb bud and is required for determining the AP polarity of the limb. Many experiments have shown that the ZPA defines the AP axis (Saunders and Gasseling, 1968) and that it does so by expressing Sonic Hedgehog (*SHH*). The SHH protein undergoes an autocatalytic cleavage to generate an active N-terminal fragment that is covalently bound to cholesterol and acts as a morphogen to directly signal to other cells. *SHH* also acts indirectly through bone morphogenetic protein 2 (*BMP2*) and GLI family zinc finger 3 (*GLI3*) in establishing AP gradients and regulating digit identity. The expression of *SHH* defines the posterior portion of the limb and grafting a secondary ZPA or an alternative source of *SHH* signal on the anterior side of a chick limb bud causes the

development of supernumerary preaxial digits that develop as a mirror image to the normal digits (Riddle et al., 1993).

Crosstalk between the ZPA and AER coordinates limb growth. *SHH* from the ZPA induces gremlin 1 (*GREM1*), which in turn induces fibroblast growth factor 4 (*FGF4*) in the posterior AER, a gene that is required for the maintenance of the ZPA. These interactions control the growth and patterning of the limb and illustrate how interference with these pathways could affect development on multiple axes.

Limb structures

The mature tetrapod limb is divided into three parts: the stylopod, zeugopod and autopod (Figure 4). The stylopod is the most proximal part of the limb and is composed of one long bone (femur/radius) that is attached to the pelvic/pectoral girdle. The middle part of the limb is the zeugopod, which is composed of two long bones (tibia/radius and fibula/ulna). The most distal part of the limb is the autopod (foot/hand), which is composed of many bones. The autopod can be further divided into the tarsal/carpal bones of the wrist, metatarsal/metacarpal bones in the hand and phalanges that make up the digits. After the three axes have been established in the limb bud, the next phase of limb development consists of the development of the components of the limb such as the muscles, tendons, skeleton and nerves. The skeletal elements of the limb form through chondrogenic differentiation where some cells of the limb bud turn into chondrocytes and begin to produce cartilage. This process involves the bone morphogenic protein (BMP) pathway as well as signals from Hox and Sox family transcription factors. Mesenchymal

condensations form in the distal limb bud and eventually develop into the bones of the autopod.

Studies from classical embryology show that early disruption of genes expressed by the ZPA and AER can cause changes in limb morphology and that disruption of signals at later time points can cause problems with bone development. Given our understanding of the basic roles of these signaling centers in normal limb development, it is clear that the disruption of the primary patterning of limb axes causes these limb malformations.

Because many of the genes involved in these pathways are regulated by *cis*-regulatory enhancers, mutations that affect these enhancers could change gene expression patterns and lead to problems in development that result in congenital limb malformations. This has been shown to be the case with mutations that change enhancers that control genes involved in early limb patterning and later morphological changes like bone and cartilage development.

2.2 – Studying enhancers in models of the tetrapod limb

A major reason that enhancers are well studied compared to other *cis*-regulatory elements is that the functional assays to characterize them are straightforward and highly reproducible. These assays are based on a common design: the putative enhancer is placed in a vector that contains a reporter gene controlled by a minimal promoter. This construct allows the enhancer to direct the expression of the reporter gene. This type of assay is commonly used in cell culture systems with a luciferase reporter gene (Heintzman et al., 2007), in mouse embryos with a LacZ reporter (Kothary et al., 1989;

Pennacchio et al., 2006) and in zebrafish with a GFP reporter (Fisher et al., 2006). Each system has its advantages – the high throughput capacity of cell culture assays, the similarity to human development of the mouse model, and the relative ease and temporal tracking of reporter expression in zebrafish. Many attempts are being made to study other regulatory elements using similar principles. As is the case for enhancers, silencers and insulators are not readily identifiable from sequence data alone. Insulators have been identified through ChIP-chip and ChIP-seq studies looking for CTCF binding regions in specific cell types (Cuddapah et al., 2009; Kim et al., 2007). Functional assays for *in vivo* silencer or insulator activity have proven difficult to design and while some assays have been developed (Bessa et al., 2009; Petrykowska et al., 2008), there are currently no robust *in vivo* functional assays for these elements in widespread use.

To study enhancers in limb development, I used multiple systems. To functionally validate enhancers, I started with a zebrafish system and eventually moved to a mouse system. In each system, the assay vector is injected into a single-cell embryo, which is examined for reporter expression at specific time points (Figure 5). The reporter gene is only active in the tissues of the embryo where the enhancer is active. Mutations in enhancer sequence can produce patterns of reporter expression that are different from the pattern driven by the normal enhancer, similar to what might happen to the gene normally regulated by the enhancer. The primary drawback to these systems is that the expression patterns are qualitative and can only show changes to the expression domains rather than more subtle changes to the degree of expression because of the different number of

enhancer construct integrations and variation in the genomic site of integration in both mouse and zebrafish.

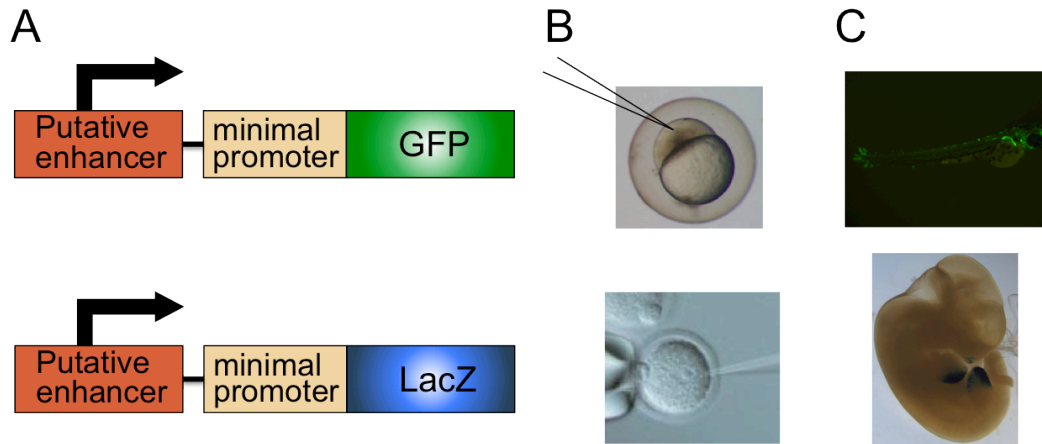


Figure 5. The *in vivo* enhancer assay uses a reporter gene to show the expression pattern controlled by an enhancer. (A) The assay vector uses the enhancer in conjunction with a minimal promoter – a promoter that is not active without the presence of a functional enhancer – to control a reporter gene; GFP in the zebrafish assay and LacZ in the mouse assay. (B) The assay construct is injected into the single-cell embryo. (C) Representative limb enhancers in zebrafish and mouse are shown. For the zebrafish, the GFP is observed in the live embryos allowing for tracking of expression at multiple developmental stages.

2.3 – Modeling limb enhancer activity in zebrafish

Despite the fact that zebrafish clearly do not have limbs, they can be used as a model for the early stages of tetrapod limb development because many of the same genes and pathways are used in the zebrafish pectoral fin.

The mature fin does not have all of the same structures that are seen in the tetrapod limb. The tetrapod limb is composed of the stylopod, zeugopod and autopod. In contrast, the pectoral fins of the zebrafish are made up of proximal radials, distal radials and fin rays (Grandel and Schulte-Merker, 1998). The proximal radials are columnar bones that are

attached to the pelvic/pectoral girdle. The distal radials are small bones at the end of the proximal radials. These bones are formed in a manner similar to the bones of the tetrapod limb, but they make up only a small portion of the fin (Yano and Tamura, 2013). The largest portion of the fin is composed of fin rays in the distal part of the fin. Fin rays are made up of thin rod-like bones that develop as exoskeleton directly from mesenchymal cells (Géraudie and Landis, 1982; Landis and Géraudie, 1990).

Despite these differences, the tetrapod limb and the fish pectoral fin are homologous organs because they share a common evolutionary ancestor and develop from the same tissues (Mercader, 2007; Yano and Tamura, 2013). These fins and limbs originate from a locomotive structure in a shared common ancestral vertebrate. Like the limb bud, the fin bud is patterned along three axes (AP, PD and DV) to develop into a three-dimensional structure. In the zebrafish fin, the AP axis is controlled by the expression of genes that are also seen in the tetrapod limb AP axis including the zone of polarizing activity (ZPA) defined by signals such as *Shh* (Dahn and Fallon, 2000; Neumann et al., 1999; Yonei-Tamura et al., 2008). The PD axis is controlled by an apical ectodermal ridge (AER) that later develops into an apical fold (AF) and expresses genes seen in the tetrapod AER like *Fgf8* and *Wnt3a* (Fernandez-Teran and Ros, 2008; Grandel and Brand, 2011; Kengaku et al., 1998; Lu et al., 2008)

These shared developmental pathways mean that changes in enhancer activity that affect gene expression in the tetrapod limb might have a similar impact on the zebrafish fin. The two tissues are only comparable at the early stages of development, but this is the stage

where the final pattern of the limb is being established and likely a time of dynamic enhancer activity. Many genes that are shown to cause early developmental defects of the limb have similar effects on the fin (Mercader, 2007) so it is reasonable to assume that some of the same gene regulatory elements controlling these genes may function in tetrapods and in fish. It has been shown that human-specific enhancer sequences that are not conserved to zebrafish can act as enhancers in a zebrafish assay, showing that at least some of the regulation of limb development is conserved in fins (Ritter et al., 2010). The number of enhancers that are positive in a zebrafish assay ranges between 17% and 30% depending on the study (Ariza-Cosano et al., 2012; Ritter et al., 2010).

To better understand how enhancers can control genes in different species that have very divergent body plans and morphology, I collaborated with Gill Bejerano's group from Stanford University that was studying enhancers with very deep evolutionary conservation. Two enhancers that show conservation between deuterostomes (including humans and sea urchins) and protostomes (including flies and worms) were identified. These represent *cis*-regulatory elements that were the first shared enhancer elements between these two clades. The homologous sequences for these enhancers from different species retain the same enhancer activity. This suggests that they are responding to the same signaling pathways in these distantly related organisms and highlights the power of enhancer assays even when the model organism is not the source of the enhancer sequence (Clarke et al., 2012).

I created zebrafish assay vectors with these two ancient conserved enhancers from human, zebrafish, sea urchin, tick or limpet genomes. The first enhancer is upstream of a homolog of the highly conserved inhibitor of DNA binding (*Id*) gene. This “Id element” is a short ~100 base pair sequence that is only weakly conserved in these genomes with 60-65% pairwise identity. In the zebrafish assay, all of the enhancers showed expression throughout the embryo at 24 hours post fertilization (hpf) and in the nervous system with the highest expression at 48 hpf (see Appendix I). This pattern is similar to the expression of the zebrafish homolog *idl* at these time points. The second enhancer is upstream of a homolog of *Znf503*. This “Znf503 element” from each species drives expression of the reporter in the hindbrain in a pattern consistent with zebrafish *Znf503* expression and the known role of the gene in vertebrate hindbrain development (see Appendix I).

Having established the feasibility of the *in vivo* zebrafish assay as a tool for enhancer activity with non-zebrafish enhancer sequences, I turned to validating the limb model in the zebrafish pectoral fin. For this, I selected a group of six known enhancers that were shown to have limb expression in a comparable mouse assay (Figure 6A)(VISTA enhancer browser, <http://enhancer.lbl.gov>). These enhancers were selected to represent a range of limb expression patterns in enhancers that were conserved to zebrafish. While these mouse transgenic assays are limited to a single time point per embryo, zebrafish embryos allow for longitudinal studies over multiple developmental time points by non-invasive imaging of the reporter. The fin bud first appears around 26 hours post fertilization (hpf) and the early development that corresponds with tetrapod development is finished by 48 hpf.

To examine the expression patterns of these enhancers, the human homolog of each enhancer was cloned into two different zebrafish enhancer assay vectors with two different minimal promoters; e1b and cfos (Fisher et al., 2006; Li et al., 2010). These minimal promoters have slightly different patterns of nonspecific background expression based on work in the Ahituv lab. For example, the e1b minimal promoter has the highest background expression in the heart while cfos has the highest background expression in the midbrain (personal communication, Ahituv lab members). Transgenic embryos were examined at 24 and 48 hours for GFP expression. Results are summarized in Figure 6.

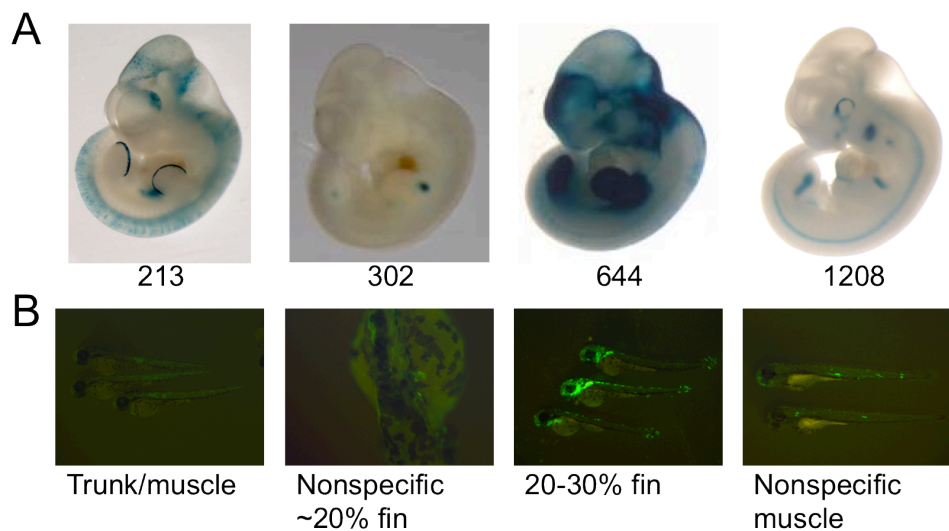


Figure 6. Known limb enhancers from mouse data were tested in zebrafish to compare expression patterns. (A) Four mouse enhancer sequences shown here were chosen from the VISTA Enhancer Browser to represent different expression patterns within the limb. (B) Corresponding zebrafish enhancer assays show a lack of fin-specific expression in three out of the four injected sequences. The most common expression pattern is listed below the fish along with an indication of any fin expression. Percent fin is calculated as fin-specific/total GFP-expressing.

Ultimately, only one enhancer had significant expression in the pectoral fin. This enhancer was expressed in the entire limb bud of the mouse and the entire fin bud of the zebrafish. The enhancers that were localized to the ZPA or AER in the mouse did not show expression in the zebrafish fin bud. Despite the potential benefits to a zebrafish model system, the reproducibility of known enhancer patterns was determined to be too low to pursue. Follow up work in the lab that looked at 22 known limb enhancers in the same zebrafish system concluded that only 45% of the enhancers had pectoral fin expression and that the specificity to the ZPA and AER was lost (Booker et al., 2013).

2.4 – Modeling limb enhancer activity in mouse

Another model organism that is commonly used to study the role of *cis*-regulatory elements in development is the mouse. Mice are a powerful tool for studying genetic pathways in the limb because of the homology to the human limb. The bones of the mouse limb all have homologous structures in the human limb and human congenital limb malformation phenotypes have homologous malformations in various mouse models and knockouts. Mouse transgenics can be used to examine the expression pattern of enhancers. Because of the homology of the mouse and human limbs, enhancer expression domains changes can easily be interpreted in the context of human congenital limb malformation phenotypes.

The mouse is also a good tool for molecular genetic studies of enhancers. The accessibility of transgenic mice allows for the collection of embryonic limb tissue from specific portions of the limb bud. Pooling of tissue from multiple embryos allows enough

tissue to be collected for ChIP-seq exclusively on cells from the signaling centers that control limb patterning. Chapter 4 focuses on mouse ChIP-seq studies that I have done to identify novel limb ZPA and AER enhancers.

2.5 – Studying congenital limb malformations in human patients

In addition to studying limb development in model organisms, we built a database of DNA samples from human patients with isolated congenital limb defects. Human limb malformations are the second most common form of birth defect with a frequency of nearly 1 in 500 live births (Moore and Persaud, 1998). Many of these are part of a syndrome that affects other tissues, but limb defects are also seen in isolation as the only phenotype in an individual. We looked specifically for those individuals who don't have any phenotypes outside of the limb, because those are the candidates for limb-specific enhancer mutations (Figure 3). To collect samples, we reached out to physicians and orthopedic surgeons around the United States and internationally through emails, posters at meetings and the creation of a brochure that they could distribute to patients (see Appendix I).

We accepted subjects with any type of limb malformations, but are particularly interested in those whose defect fits an aspect of limb patterning. These include phenotypes predicted to be caused by changes in the AP axis through ZPA mutations like preaxial polydactyly type II (PPD2), triphalangeal thumb (TPT), postaxial polydactyly (PAPA), syndactyly and Werner mesomelic syndrome (WMS) (tibial hypoplasia-polysyndactyly-triphalangeal thumb, THPSTPT). Other phenotypes that we were particularly interested

in are associated with changes in proximal-distal growth or maintenance of the AER. These include split hand/foot malformations type I-IV (SHFM), brachydactyly and clinodactyly. Additional phenotypes that are interesting for us include those that are limited to just the hands or just the feet like clubfoot. Figure 7 illustrates some of these phenotypes.

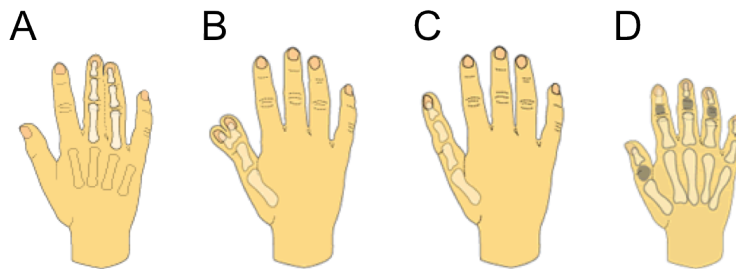


Figure 7. Phenotypes of limb malformations predicted to be the result of early patterning defects. (A) Syndactyly – fusion of digits. (B) Preaxial polydactyly – additional digits at the anterior of the limb. (C) Triphalangeal thumb – an additional phalanx in the first digit, transformation of the thumb to a finger-like identity. (D) Brachydactyly – shortening of the digits. (Images reproduced from http://www.peds.ufl.edu/divisions/genetics/teaching/hand_malformations.htm)

As of late 2013, our database consists of 896 affected subjects and 272 unaffected related controls covering a wide variety of limb malformation phenotypes (Table 1). While most of these affected subjects are from small families, there are 9 pedigrees with more than three generations of affected individuals in the database. We collect extensive phenotypic data on subjects and screened coding regions of genes where mutations have been shown to cause the subject's phenotype (Materna-Kiryluk et al., 2013)(see Appendix I for details). After excluding coding mutations, many of these samples have been sequenced for mutations in multiple candidate enhancers and 6 primary research articles with data

from sequencing work in this patient database have been published or submitted to date (Birnbaum et al., 2012a, 2012b; Laurell et al., 2012; Lu et al., 2012; VanderMeer et al., 2012; Vandermeer et al., 2014).

Table 1. Isolated limb malformation sample database summary.

Limb phenotype	Number
Preaxial polydactyly	122
Syndactyly	97
Clubfoot	142
Split hand/foot malformation	93
Brachydactyly	43

There are advantages to using different models to identify enhancers in limb development. Zebrafish offer a cost efficient and higher throughput model where enhancer expression can be tracked over time through non-invasive imaging of a GFP enhancer assay. A mouse model offers extensive genetic tools and a wealth of transgenic and knockout examples to study the impact of genes on limb development. Human subject sequences can be used to connect model organism studies to naturally occurring mutations. The combination of these three systems allows for multiple approaches to the study of tetrapod limb development and the role of enhancers in patterning the limb.

Chapter 3 – The ZRS and human malformations

3.1 – Background of the ZRS enhancer in limb development

One of the most studied developmental enhancers is the enhancer that controls the expression of SHH in the limb bud. Because this enhancer activates *SHH* in the posterior ZPA, it is known as the ZPA regulatory sequence (ZRS). This enhancer was discovered through a combination of mouse models and the study of human patients with preaxial polydactyly.

Developmental studies have shown that inducing ectopic *Shh* expression in the anterior side of the limb bud induces ectopic digits. There are multiple mouse models of preaxial polydactyly where the extra digits appear similar to what was seen in experiments where an ectopic ZPA was grafted to the anterior of a chick limb bud. Because these mouse models were generated or discovered by different screens, they have mutations in different genes, but many were found to have embryonic limb *Shh* expression that extended far beyond the normal posterior ZPA and, in some cases, was even considered a second “anterior ZPA” (Blanc et al., 2002; Chan et al., 1995; Masuya et al., 1995).

Some of these mice were discovered to have defects in genes that function upstream of *Shh* and would normally restrict its expression to the posterior ZPA (*xt* mutant; (Hui and Joyner, 1998)), *lst* mutant; (Qu et al., 1998)). Due to mutations in these genes, *Shh* could now be expressed in anterior tissues. In other mice, the phenotype was mapped to the *Shh*

locus, but no *Shh* coding mutations were found (Sharpe et al., 1999) and the ectopic *Shh* expression indicates that the gene is functional. The Sasquatch mutant (*ssq*) is an example where the insertion of a transgene caused preaxial polydactyly and the mutation was mapped to the *Shh* locus. This transgenic insertion created a 20 kilobase (kb) duplication within intron 5 of the limb region 1 (*Lmbr1*) gene, which is about 1 megabase away from *Shh* (Lettice et al., 2002). The *ssq* mutant showed not only ectopic *Shh* expression, but also expression of a similar pattern for the transgene in both the posterior ZPA and the anterior limb bud, suggesting that the transgene integrated into a region of the genome able to regulate the spatial expression of genes – that is to say, a region with an enhancer.

Identifying the ZRS

The homologous human region including *SHH* and *LMBR1* was also recognized to be important in limb patterning through studies of human patients with preaxial polydactyly (PPD). PPD patients have an extra digit on the anterior (thumb) side of the hand or foot, or triphalangeal thumb (TPT), a thumb with a third bone that has the appearance of a second index finger (Figure 8). This phenotype is relatively common and can occur either as an isolated phenotype or as part of a syndrome with other associated phenotypes outside of the limb. Through linkage analysis, isolated PPD was mapped in several families to a region of approximately 500 kb on chromosome 7 (Heus et al., 1999). A patient with isolated PPD that had a *de novo* chromosomal translocation in this region led to the fine-mapping of the PPD locus to a region within intron 5 of the gene *LMBR1*, the same region disrupted by the mouse *ssq* transgene insertion (Lettice et al., 2002). This intron contains a highly conserved sequence of about 800 base pairs (bp) that was found

to have enhancer activity in the posterior limb bud, where it normally controls *SHH* expression, and was called the ZRS. The ZRS has since been studied in many organisms where it was shown to be required for normal *Shh* expression and limb development (Sagai et al., 2005) and to harbor mutations that cause polydactyly in mice (Lettice et al., 2003; Masuya et al., 2007; Sagai et al., 2004), chickens (Maas and Fallon, 2004; Maas et al., 2011), dogs (Park et al., 2008) and cats (Lettice et al., 2008). Other human ZRS mutations that have been identified since the 2003 study have been named based on the patients where they were identified as well as the ZRS site where they are located, in accordance with the numbering system assigned by Lettice et al. (Lettice et al., 2003). Another conserved region next to the ZRS has also been identified as the proximal-ZRS (pZRS) and has been found to be the site of mutations that cause polydactyly in dogs, extending the length of the regulatory region where polydactyly mutations can reside to nearly 2 kb (Park et al., 2008).

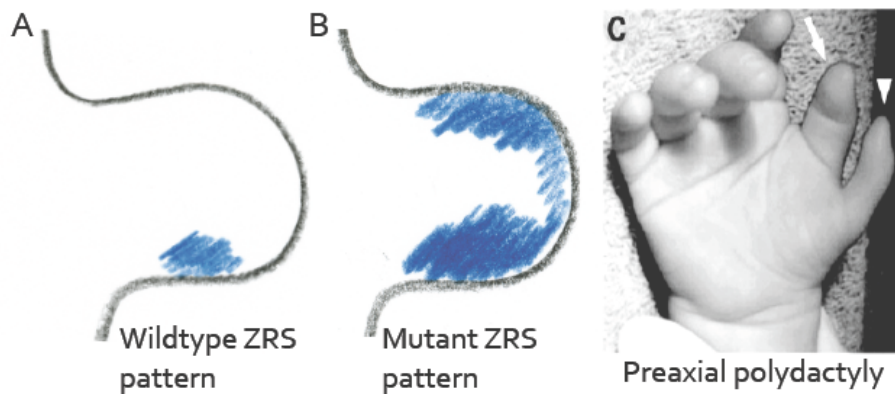


Figure 8. ZRS expression and preaxial polydactyly. (A) The expression of a LacZ reporter gene, indicated in blue, controlled by the normal ZRS enhancer is limited to the posterior mesenchyme that is the normal ZPA. (B) Mutations in the ZRS cause the reporter to be expressed more extensively in the posterior and in an ectopic region in the

anterior part of the limb bud. Mutations like this are thought to have the same impact on *SHH* expression and the patterning defect causes preaxial polydactyly (C) [reproduced with permission from Lettice et al. 2003].

Point Mutations within the ZRS region

Over a dozen single nucleotide mutations and one small (13 bp) insertion mutation within the ZRS have been shown to cause human limb malformations. These mutations result in preaxial polydactyly (PPD) and triphalangeal thumb (TPT) with or without supernumerary digits. The phenotype can affect the hands and or feet and can be unilateral or bilateral. The known human ZRS mutations are distributed throughout a 700 bp region (see Table 2 for a list of human ZRS mutations) within the conserved ZRS. These mutations are thought to change the enhancer activity of the ZRS. Support for this comes from the observation that when some of these mutations were tested in a mouse enhancer assay, the LacZ reporter was shown to be expressed in the anterior part of the limb bud (Figure 8A, B). These mutations don't all affect predicted transcription factor binding sites. In addition, the fact that ZRS-related phenotypes differ between the various reported mutations is further evidence that particular mutations in the ZRS can affect the enhancer's function in subtly different ways. Currently, it is not possible to predict phenotypic severity based on sequence alone.

The severity of the human phenotypes does appear to be related to the extent of *SHH* misexpression. While only a few ZRS mutations have been tested in mice for enhancer activity, some do show a correlation between a higher level of reporter gene expression – in the ectopic anterior region and/or the normal posterior region of the limb bud – and more severe human phenotypes. Two of the first reported mutations, referred to as Cuban

(ZRS 404G>A) and Belgian1 (ZRS 305A>T) illustrate this correlation. The more severe Cuban patient phenotype includes polydactyly and tibial malformations and the reporter assay showed a very strong anterior and posterior expression pattern (Lettice et al., 2003, 2008; Zguricas et al., 1999). The Belgian1 mutation that causes PPD2 (OMIM#174500), an extra thumb anterior to a triphalangeal thumb with no abnormalities in the long bones of the arms, shows only weak anterior reporter expression in the reporter assay (Lettice et al., 2003, 2008).

Most of the known ZRS mutations have complete penetrance within the affected family and are inherited in a dominant pattern (Al-Qattan et al., 2012). There is, however, one reported mutation that does not always cause polydactyly. The mutation at ZRS 295 (295 T>C) was first reported to be a neutral polymorphism because it was found in 10-30% of unaffected samples (Lettice et al., 2003). Later, this mutation was discovered to be associated with triphalangeal thumb in multiple English families (Furniss et al., 2008). Examination of this mutation in a mouse enhancer assay showed a weak anterior expression of the reporter, suggesting that this incompletely penetrant mutation might not always cause enough ectopic expression to lead to polydactyly (Furniss et al., 2008). The factors that determine whether the 295 T>C mutation causes a phenotype are not known and no other low-penetrance ZRS mutations have been identified to date.

While most point mutations in the ZRS cause preaxial polydactyly that is limited to the autopod, one particular site in the enhancer is thought to be associated with Werner mesomelic syndrome (OMIM # 188770), a limb phenotype that includes hypoplastic tibia

in addition to triphalangeal thumb polydactyly. Unrelated patients with Werner mesomelic syndrome were found to all have different mutations at ZRS 404, leading to the thought that there might be something special about this site that causes the phenotype to extend beyond the hands and feet (Cho et al., 2013; Wieczorek et al., 2010). Because the mechanisms that cause ZRS mutations to change gene expression are unknown, it is currently not possible to determine what makes this site different from the others identified throughout the ZRS.

Table 2. Mutations in the ZRS enhancer and associated phenotypes

Mutation Name	Mutation	Location (hg19)	Phenotype	Reference
<i>SHH</i> ZRS enhancer (chr7:156,583,562-156,584,711)				
739 A>G, Family A,C	point mutation	chr7:156,583,831	PPT, TPT	Gumett et al., 2007
621 C>G, Family B	point mutation	chr7:156,583,949	PPT, TPT	Gumett et al., 2007
619 C>T	point mutation	chr7:156,583,951	PPT, TPT, aplasia of the radius	Al-Qusban et al., 2012
603-604ins13	insertion	chr7:156,583,967	PPT, TPT	Laurell et al., 2012
463 T>G	point mutation	chr7:156,584,107	PPT, TPT	Farooq et al., 2010
404 G>C, Family 2	point mutation	chr7:156,584,166	WMS	Wieczorek et al., 2009
404 G>A, Family 1	point mutation	chr7:156,584,166	WMS	Wieczorek et al., 2009
404 G>A, Cuban	point mutation	chr7:156,584,166	PPD, WMS	Letfice et al., 2003
404 G>A	point mutation	chr7:156,584,166	WMS	Cho et al., 2013
402 C>T	point mutation	chr7:156,584,168	PPD, TPT, WMS	VanderMeer et al., 2014
396 C>T, Turkish 1	point mutation	chr7:156,584,174	PPD, TPT	Semerci et al., 2009
334 T>G, French 2	point mutation	chr7:156,584,236	PPD	Albuisson et al., 2010
329 T>C, Belgian 2	point mutation	chr7:156,584,241	PPD	Letfice et al., 2003
305 A>T, Belgian 1	point mutation	chr7:156,584,266	PPD	Letfice et al., 2003
297 G>A, French 1	point mutation	chr7:156,584,273	PPD	Albuisson et al., 2010
295 T>C	point mutation	chr7:156,584,275	TPT	Furniss et al., 2008
284 C>A	point mutation	chr7:156,584,283	PPD, TPT, PAP	VanderMeer et al., 2012
105 C>G, Dutch	point mutation	chr7:156,584,465	PPD	Letfice et al., 2003
Case, Letfice	translocation	9(5.7)(q11,q36)	PPD, TPT	Letfice et al., 2002
Family Klopocki	duplication	-chr7:156,143,386-156,732,204	TPTPS	Klopocki et al., 2008
Family 6, Sun	duplication	-chr7:156,241,020-156,699,998	TPTPS, SD4	Sun et al., 2008
Family 2, Sun	duplication	-chr7:156,241,020-156,677,739	TPTPS, SD4	Sun et al., 2008
Family 5, Sun	duplication	-chr7:156,241,020-156,619,399	TPTPS, SD4	Sun et al., 2008
Family 4, Sun	duplication	-chr7:156,254,085-156,687,613	TPTPS	Sun et al., 2008
Family 3, Sun	duplication	-chr7:156,354,085-156,619,399	TPTPS, SD4	Sun et al., 2008
Family 3, Wieczorek	duplication	-chr7:156,368,541-156,661,877	SD4, fibial hypoplasia	Wieczorek et al., 2009
Family Dai	duplication	-chr7:156,547,469-156,644,074	TPTPS, SD4, PAP of toe	Dai et al., 2013
Family 1, Sun	duplication	-chr7:156,539,605-156,699,998	TPTPS	Sun et al., 2008
Family Wu	duplication	-chr7:156,547,469-156,644,074	SD4, fibial hypoplasia	Wu et al., 2009
Family 4, Wieczorek	duplication	-chr7:156,572,751-156,661,877	TPTPS, SD4, PAP of toe	Wieczorek et al., 2009

PPD: preaxial polydactyly, TPT: triphalangeal thumb, WMS: Werner mesomelic syndrome, TPTPS: triphalangeal thumb poly/syndactyly, SD4: syndactyly type 4, PAP: postaxial polydactyly

ZRS duplications and complex polysyndactyly

Human limb malformations have also been attributed to duplications that encompass the ZRS and parts of the surrounding sequence. These duplications cause complex polysyndactyly phenotypes that entail fusion of soft tissue or bones of the autopod in addition to supernumerary digits including triphalangeal thumb polysyndactyly (TPTPS) and syndactyly type IV (Al-Qattan et al., 2013; Dai et al., 2013; Sun et al., 2008; Wu et al., 2009). Multiple ZRS duplications have been found in different families. These duplications do not have shared breakpoints and there is no discernable relationship between the size of the duplication and the severity of the phenotype. The smallest shared region between the various duplications is 47 kb, extends from intron 4 of *LBMR1* and continues into intron 5, ending past the 3' end of the ZRS. The human ZRS duplication phenotype is different than the mouse *ssq* phenotype which has a 20 kb duplication within intron 5 of *Lmbr1* that includes the ZRS (Sharpe et al., 1999), but shows only polydactyly with no fusion of the digits, suggesting either human-mouse phenotypic differences or that the duplicated sequence outside of the ZRS itself may have additional important limb regulatory elements.

ZRS and Acheiropodia

In addition to polydactyly and polysyndactyly, there is another human limb malformation phenotype that has been mapped to the region near the ZRS. Acheiropodia (OMIM #200500) is a severe limb malformation consisting of nearly complete truncations of all limbs and aplasia of the hands and feet. Acheiropodia is a very rare malformation caused by a homozygous deletion of a nearly 6 kb region that removes exon 4 from mRNA transcripts of *LMBR1*, but the deletion does not appear to extend as far as the ZRS in intron 5 (Ianakiev et al., 2001). A mouse model of

a ZRS knockout has a similar limb phenotype, but lacks only the ZRS and does not have disruptions in *Lmbr1* intron 4 or exon 4 (Sagai et al., 2004). So far, no additional *cis*-regulatory elements have been identified in the acheiropodia deletion.

Difficulties in linking ZRS mutations to phenotypes

Animal models of ZRS mutations appear to be of little use in predicting the phenotype caused by specific mutations. Human phenotypes from ZRS point mutations predominantly affect the hands, while mouse models of ZRS mutations tend to have a stronger phenotype in the hindlimbs (Knudsen and Kochhar, 1982; Lettice et al., 2003; Li et al., 2009). Furthermore, human patients homozygous for ZRS mutations show phenotypes no more severe than heterozygotes, unlike what has been seen in mice (Semerci et al., 2009). It is clear that not all cases of human isolated preaxial polydactyly are caused by ZRS mutations. There are even numerous families with preaxial polydactyly that is genetically linked to the ZRS that appear to have no mutation or duplications in either the ZRS or in any portion of the acheiropodia deletion (Gurnett et al., 2007; Lettice et al., 2003; Li et al., 2009). Whether other mechanisms are behind these patients' malformations or there is yet another *SHH* limb *cis*-regulatory element or elements in this locus remains to be seen.

ZRS looping

The ZRS has been shown to physically interact with the promoter of *Shh* in mouse limb tissue through DNA looping bringing these two regions into contact (Amano et al., 2009). While the exact looping mechanisms remain unclear, it appears to create specific DNA interactions in a location that correlates with gene expression (Kagey et al., 2010). This interaction occurs only in

regions of the limb where *Shh* is “poised” for activation – the posterior ZPA and the anterior mesenchyme region where ectopic *Shh* is observed in polydactylous mouse lines (Amano et al., 2009). In addition to this looping interaction, the same study showed that the looped *Shh*–ZRS complex moves out of its normal chromosome territory in the nucleus when *Shh* is transcribed. This chromosome territory shift normally happens only in the ZPA, suggesting that it is related to the activation of *Shh* expression. Other studies show a role for nuclear matrix proteins in the looping and physical interaction of *Shh* and the ZRS (Zhao et al., 2009). These mechanisms are not entirely clear, but show that there are multiple levels of control over *Shh* expression and potentially multiple ways this control could be disrupted by mutations.

3.2 – Patient Sequencing Results

From the LB human sample collection, we selected patients with preaxial polydactyly of the hands and/or feet with no record of any other congenital phenotypes. The full list of patients sequenced and sequencing results can be found in Appendix II. Of these patients, approximately 25% have a family history of preaxial polydactyly. From this group, we identified a number of novel ZRS mutations that are likely the cause of their preaxial polydactyly. Three individuals had mutations that have already been identified in the literature and three additional novel mutations were discovered. These novel mutations and their characterization are described in detail in the following three sections.

3.3 – ZRS 278 C>A

Clinical Report

A highly inbred Balochi tribal family was identified in a remote village of Southern Punjab, Pakistan. The pedigree comprised six generations with the trait segregating in three consecutive generations (IV–VI; Fig. 9A). The trait appeared *de novo* in subject IV-5 and was transmitted to 10 of his descendants in the next two generations (Figure 9A). The affected family subjects have isolated limb phenotypes primarily affecting the hands with minimal involvement of the feet. There were no symptoms in any other organ system. A total of 11 subjects (seven males, four females) were affected (Figure 9A). Photographs and X-rays of seven affected and one unaffected subjects were obtained.

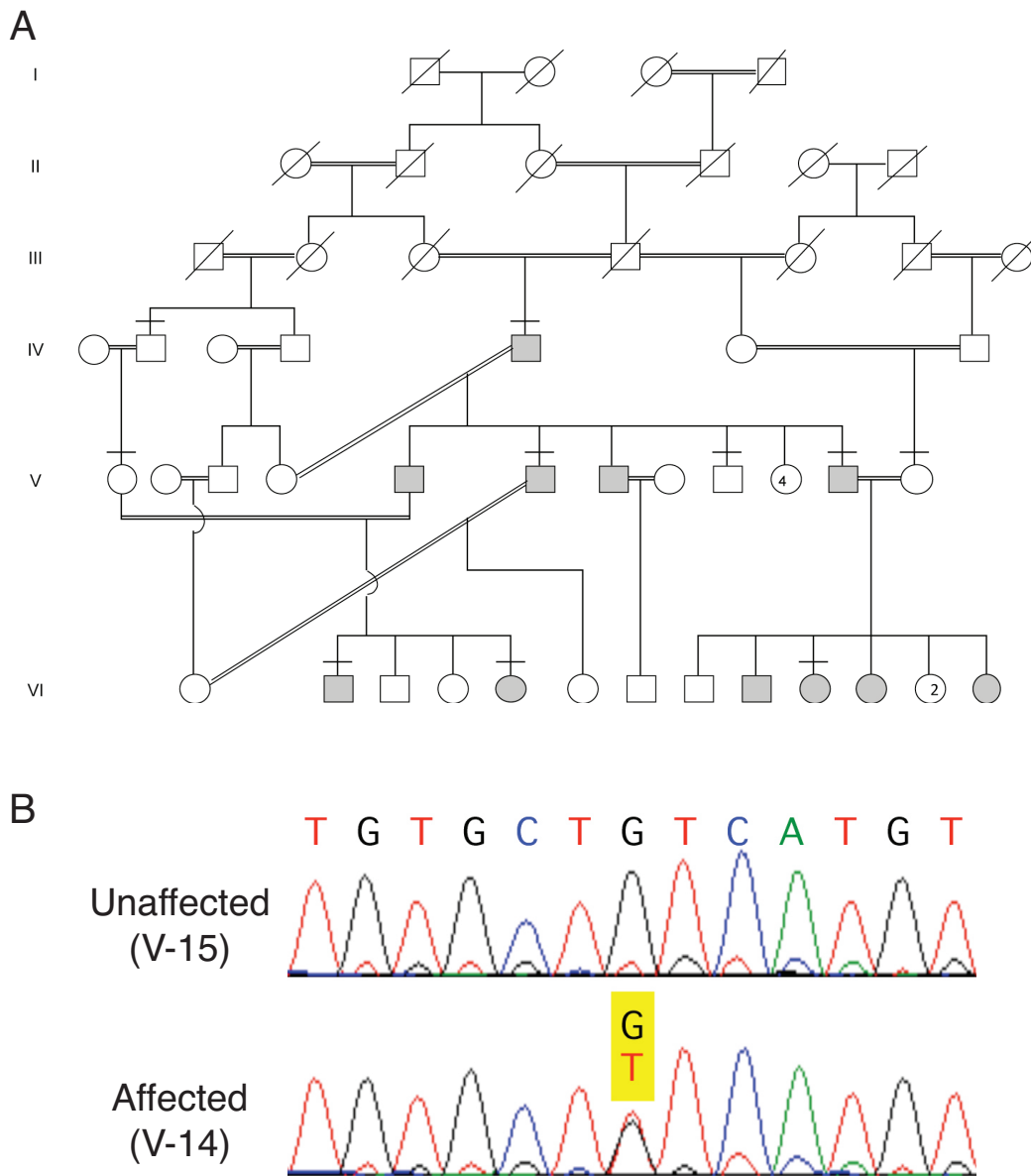


Figure 9. (A) Pedigree of the ZRS 287C>A family. Shaded symbols denote patients with hand malformations, details of phenotypes can be found in Table 3. Bars above symbols denote family members who were sequenced for ZRS mutations. (B) Representative chromatogram from an affected subject V-14 showing the heterozygous mutation at position ZRS 287 (bottom) compared to an unaffected subject V-15 wildtype sequence (top). The mutation name was designated in the opposite strand as C>A in accordance with previously published ZRS mutations.

The phenotype in this family is characterized by triphalangeal thumb, pre-axial polydactyly and postaxial polydactyly. Bilateral triphalangeal thumbs were observed in all of the affected subjects. The first digit was usually weak and had varus inclination (Figure 10). Three phalanges were observed in radiographs with a hypoplastic terminal phalanx visible on examination. Six of the nine affected subjects also had a supernumerary preaxial digit (Figure 10; Table 3). The additional digit was usually well-established with bony elements and a dorsal nail, but was non-functional.

Additional distinct clinical variants were identifiable in some patients. These include syndactyly of the postaxial digits (Figure 10D; Table 3). In these patients, syndactyly was bilateral and complete; nails of the syndactylous fingers were intimately fused, separable only by a median fissure. In the radiographs, terminal phalanges of the webbed digits depicted osseous fusion and symphalangism. Three patients had bilateral clinodactyly of the 5th fingers (Table 3). Another three patients had distal phalangeal hypoplasia of mesoaxial fingers (Figure 10B,D; Table 3). One patient had bifid halluces with fused nails and valgus deviation (V-5) (Figure 10C). Generally, patients showed crowding of carpals in the affected hands (Figure 10A, D). The distal heads of the radius and ulna were unremarkable.

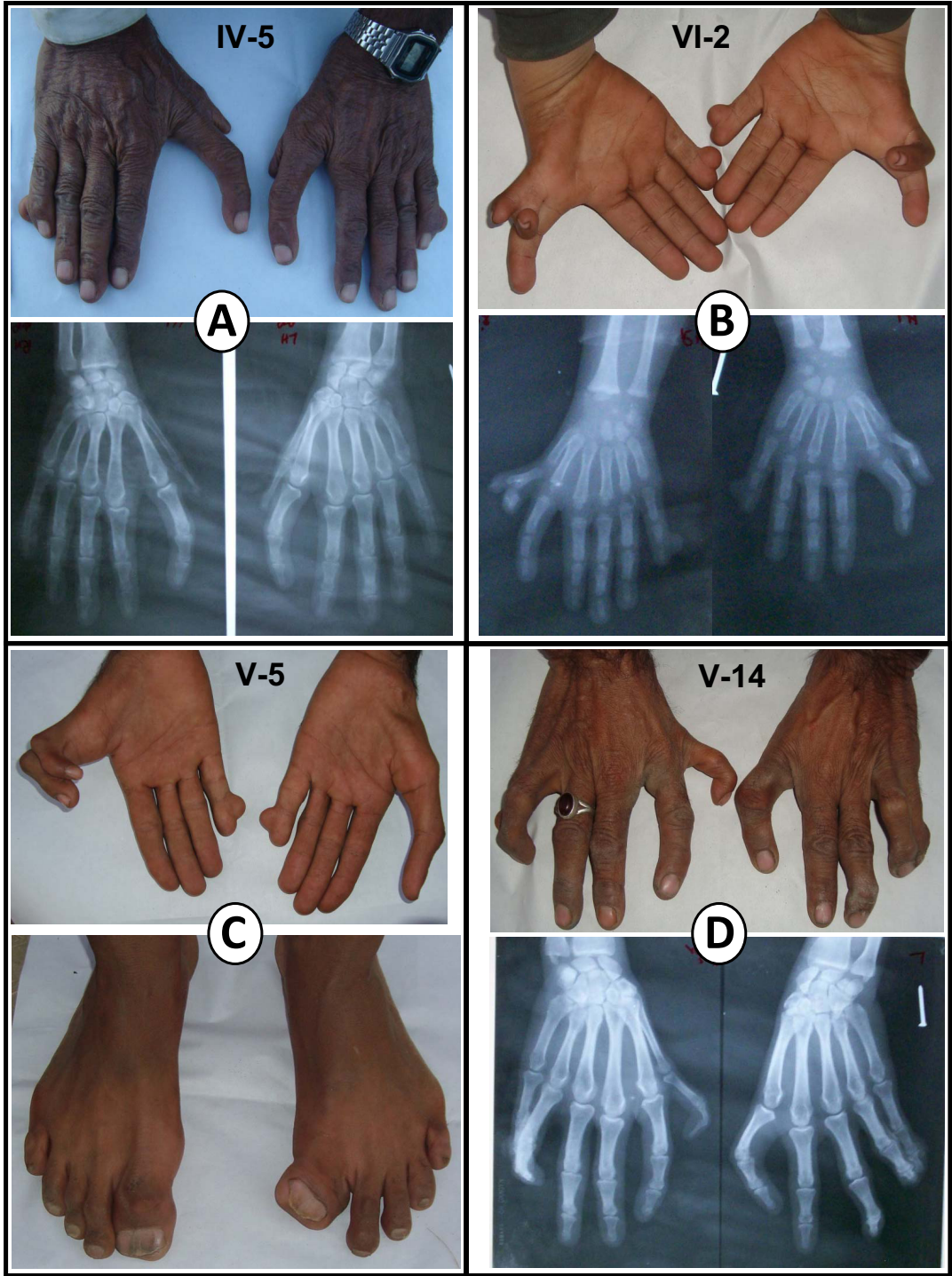


Figure 10. Phenotypic presentation in the affected family subjects. Triphalangeal thumb of essentially non-opposable nature is associated with duplication or triplication of preaxial digital ray (A, B, D). In addition to the pre- or post-axial polydactyly phenotypes, this family displays a cutaneous, knob-like tag on the 5th finger that contains no nail or bony element (A, B, C). Malformation in the feet was represented only by bifid halluces (C).

Roentgenograms revealed two or three phalangeal segments and symphalangism at the proximal inter-phalangeal (PIP) joint of the first digit (Figure 10). Some degree of postaxial polydactyly was observed in all affected patients. In most individuals, this presented as a small cutaneous tag without nail or any bony element juxtaposed to the ulnar aspect of the distal inter-phalangeal joint the 5th finger (Figure 10, Table 3). One individual had more severe postaxial polydactyly with a complete digit that was syndactylous with digits 4–5 (Table 3, subject V-7).

Table 3. Phenotypic spectrum in the affected subjects in the ZRS 287C>A family

Clinical variant	IV-5	V-5	VI-2	V-14	VI-9	VI-5	VI-10	V-6	V-7
Triphalangeal thumb	+	+	+	+	+	+	+	+	+
Varus curvature of 1 st digit	+	+	+	+	+	+	+	+	+
Pre-axial polydactyly	+	R	+, 1	R		+		R	
Extra digit	NF	NF	NF	F		NF		NF	
Nail on extra digit		+	+	+		+		+	
Post-axial polydactyly						+			+
Skin tag/lateral knob	+	+	+	+	+		L	+	+
Clinodactyly of 5 th digit	+					+	+		
Distal phalangeal hypoplasia of mesoaxial fingers		+	+	+					
Syndactyly/osseous fusion of 4-5 fingers				+, 2	+			+	+, 3
Bifid halluces		+							

(F = functional; L = in left hand only; NF = non-functional; R = in right hand only; 1 = seven digits in right hand; 2 = nails of syndactylous fingers fused; 3 = fusion of 4-5-6 fingers)

Sequence Analysis

In sequencing the proband (Patient IV-5, Figure 10A), we identified a heterozygous single base pair mutation in the ZRS. This mutation is a C>A substitution at position 287 according to

conventional ZRS mutation numbering (Lettice et al., 2002) and is located at chr7:156,584,283 (hg19) (Figure 10B). This mutation was present in all affected subjects (n=6) and absent in all unaffected subjects (n=4), and segregated perfectly with the limb anomaly in three generations. This mutation has not been reported in dbSNP (build 135; <http://www.ncbi.nlm.nih.gov/projects/SNP>) nor in the current available data from the 1,000 Genomes Project (<http://www.1000genomes.org>). Additionally, this mutation has not been observed in 42 unrelated subjects, including 19 unrelated Pakistani subjects, sequenced in our laboratory.

Transcription Factor Binding Site Analysis

UniProbe transcription factor binding site analysis indicates that the G>A mutation could result in reduced affinity of some transcription factors to the mutated allele (Newburger and Bulyk, 2009). In particular, MEIS1, MEIS2 and MEIS3 binding is predicted to be greatly reduced or absent. These proteins interact with HOXA13 in forelimb patterning and are known to be expressed in limb development, but are generally thought to be related to extension and proximal-distal limb growth (Mercader et al., 2009).

Discussion

There is strong evidence to suggest that the C>A mutation at ZRS287 could be the cause of some or all of the limb malformations in this extended family. This mutation is adjacent to the largest cluster of reported human ZRS mutations, which fall between ZRS295 and ZRS460 (VanderMeer and Ahituv, 2011). It is within a highly conserved portion of the ZRS enhancer and the C allele at this location is conserved in all of the 46 vertebrate genomes that are currently

available in the UCSC Genome Browser (<http://genome.ucsc.edu>). No other reported human sequences show any mutation at this site.

The phenotype of this family is striking compared to other examples of ZRS point mutations. Point mutations in the ZRS usually cause consistent and fully penetrant phenotypes of either triphalangeal thumb or triphalangeal thumb with pre-axial polydactyly (Albuisson et al., 2011; Farooq et al., 2010; Semerci et al., 2009), but some variation has also been seen among individuals (Gurnett et al., 2007; Lettice et al., 2003) and one family appears to show reduced penetrance with phenotypically normal carriers of the mutation (Gurnett et al., 2007). The family described here shows high phenotypic variability with regard to the severity of the anomalies and the limbs affected. Additionally, the clinical presentation of at least eight distinct phenotypic entities in this family expands the range of phenotypic variants associated with point mutations in the ZRS. While large duplications that include the ZRS have been shown to cause post-axial synpolydactyly (Klopocki et al., 2008; Sun et al., 2008; Wieczorek et al., 2010), this has not yet been reported for single base changes. The mechanism by which ZRS mutations cause digit malformations is not fully understood and it is possible that the post-axial polydactyly and syndactyly phenotypes in this family indicate that the site of their mutation within the ZRS is functionally distinct from other previously reported ZRS mutations.

3.4 – ZRS603ins13

Clinical Report

The proband (Figure 11A, individual V/14) has bilateral symmetrical triphalangeal thumbs with an extra hypoplastic radial thumb containing two phalanges and one metacarpal – preaxial

polydactyly type II (PPD2; MIM#174500) (Figure 11B, C). The triphalangeal thumb was not opposable and had a finger-like appearance. There was no involvement of the lower limbs in the proband or in any other affected family member. The proband belongs to a large 6 generation family with 78 individuals, 18 of whom are reported to be affected. DNA samples were available from 3 affected and 3 unaffected individuals in three generations and clinical examinations were performed on the proband's parents and unaffected sister.

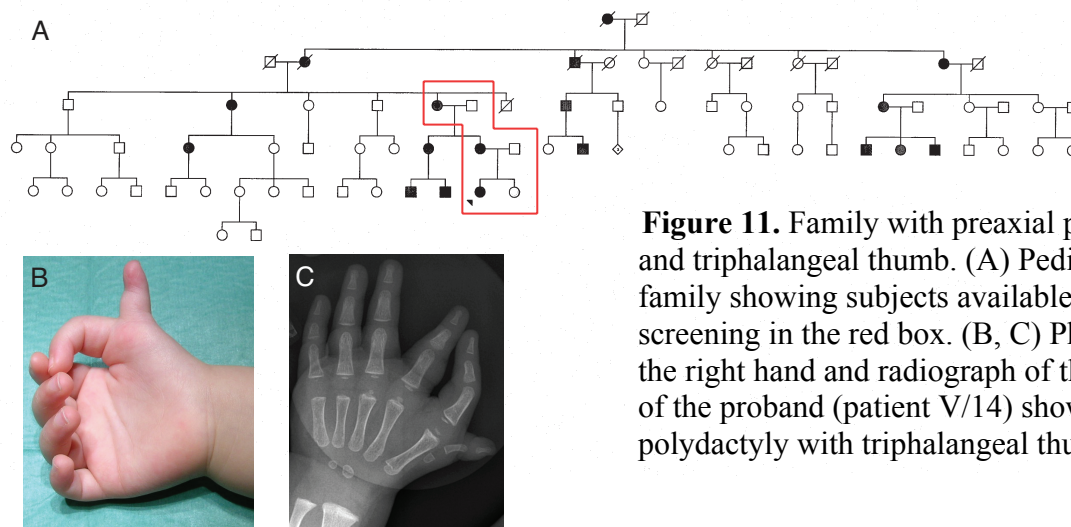


Figure 11. Family with preaxial polydactyly and triphalangeal thumb. (A) Pedigree of the family showing subjects available for DNA screening in the red box. (B, C) Photograph of the right hand and radiograph of the left hand of the proband (patient V/14) showing preaxial polydactyly with triphalangeal thumb.

Sequence Analysis

We screened a 2.1 kilobase (kb) region that encompasses the ZRS and a nearby region where mutations were shown to cause polydactyly in dogs (pZRS) (Park et al., 2008)

(chr7:156,583,564–156,585,727; UCSC Genome Browser; <http://genome.ucsc.edu>; hg19). All

affected individuals were found to be heterozygous for a 13 bp insertion (TAAGGAAGTGATT,

Figure 12A) starting at position 603 of the ZRS sequence let(Lettice et al., 2003), here named

ZRS603ins13. None of the three unaffected members tested (Figure 11A) carried this insertion.

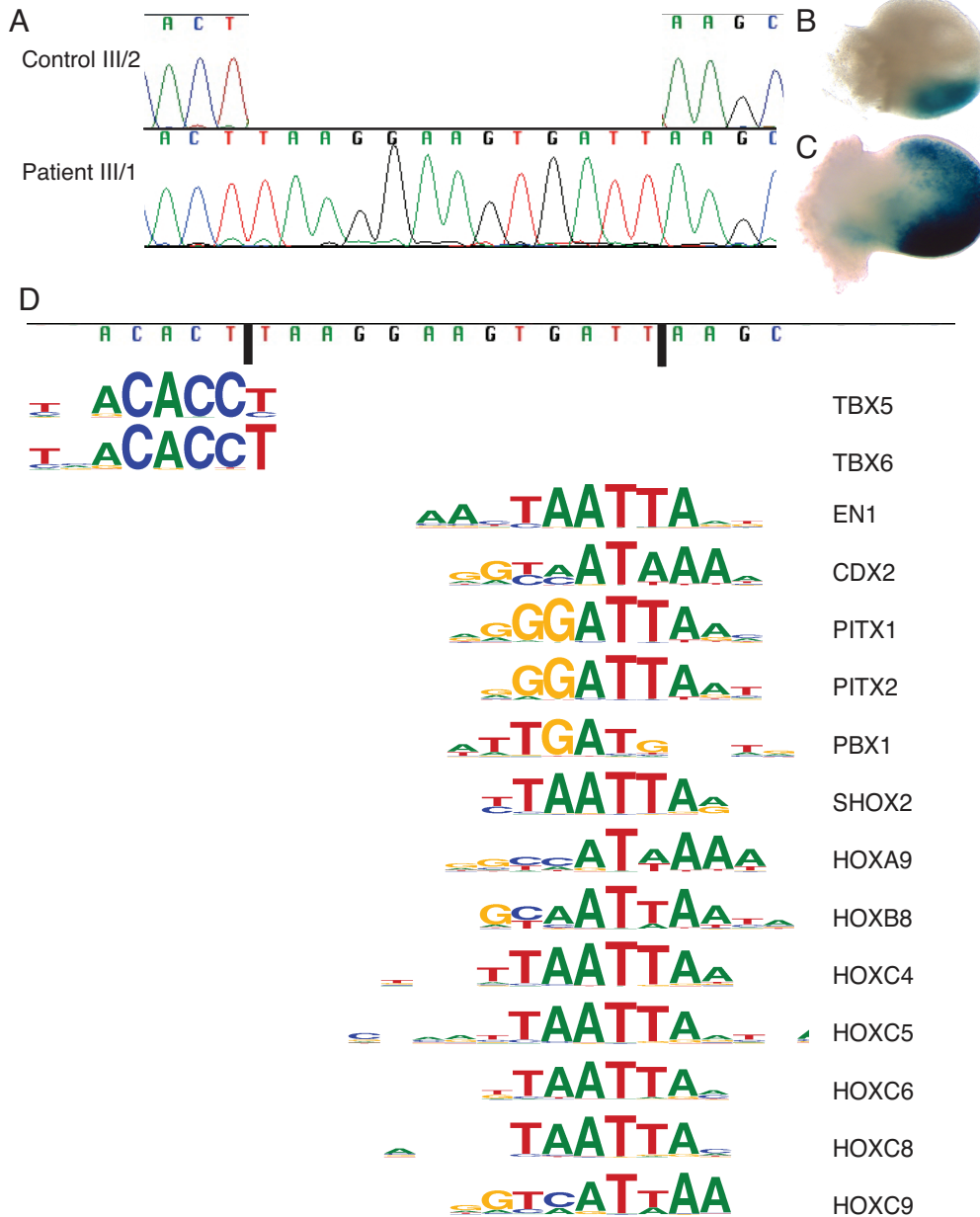


Figure 12. The insertion changes expression of a reporter gene and creates predicted transcription factor binding sites. (A) Chromatogram of the ZRS603ins13 insertion mutation. Sequencing a cloned allele of the ZRS from the proband shows an insertion of 13 bp compared to the wildtype sequence represented by the sequence from an unaffected sibling. (B) ZRS E11.5 mouse transgenic assay of human wtZRS showing posterior *LacZ* expression. (C) ZRS E11.5 mouse transgenic assay of ZRS603ins13 showing strong posterior and ectopic anterior *LacZ* expression. (D) The 13bp insertion creates binding sites for multiple limb-associated transcription factors. LOGO plots show consensus binding sequences for the indicated transcription factors.

Mouse Transgenic Enhancer Assays

The effect of the ZRS603ins13 insertion on the ZPA-specific enhancer function of the ZRS was examined using a transgenic mouse enhancer assay. Wildtype and mutant alleles of the 2.1 kb human ZRS and pZRS region were cloned into the HSP68-LacZ enhancer assay vector. This vector has an *Hsp68* minimal promoter (a promoter that is not sufficient to drive reporter expression without the presence of a functional enhancer) followed by the *LacZ* reporter gene. Transgenic mice carrying these vectors were screened at mouse embryonic (E) day 11.5 – the stage at which previous studies have seen ectopic reporter gene expression from ZRS mutation assays. The wildtype ZRS sequence recapitulates the normal *Shh* expression pattern in E11.5 mouse embryos with expression of the reporter gene in the posterior region of the limb bud in seven out of eight independent transgenic mice (Figure 12B, Appendix II). With the mutation, four out of six LacZ-positive embryos had limb expression and all four of these embryos showed ectopic anterior expression (Figure 12C) instead of the normal posterior-restricted ZPA expression pattern (Figure 12B). The expression is limited to the mesoderm of the limb bud and absent from the apical ectodermal ridge (AER) with staining at the anterior and posterior sides. While the four embryos did not show the same intensity of staining, each embryo had anterior expression that was nearly as strong as the posterior expression. It is worth noting that two of the seven limb-expressing wildtype ZRS embryos also showed a small degree of anterior limb expression that was much weaker than the posterior expression, in contrast to the ZRS603ins13 embryos where the anterior and posterior expression were similar (Appendix II). The variation between embryos is likely due to integration site or copy number differences between embryos and these differences highlight the qualitative nature of such transgenic assays. The reproducible anterior expression in ZRS603ins13 embryos is consistent with the enhancer expression pattern

observed in similar assays with ZRS mutations that have been shown to cause polydactyly (Furniss et al., 2008; Lettice et al., 2003). No other consistent expression patterns were observed in these embryos.

Transcription Factor Binding Site Analysis

To understand how this insertion could affect the expression of *SHH* in the developing limb, we screened the mutant ZRS sequence for TFBS differences from the wildtype human ZRS sequence. TFBS motifs from the UniPROBE database (Newburger and Bulyk, 2009) as well as TFBS compiled from the literature for important limb developmental transcription factors not represented in the database were used for this screen. The insertion creates two major sites that match the binding preferences for several transcription factors, including some that are important in limb development (Figure 12D). Among the detected TFBS known to be involved in limb development, are Engrailed1 (*EN1*), *TBX5*, *TBX6*, *SOX8*, and multiple *HOX* genes including *HOXA9*, *HOXB8*, and *HOXC* genes expressed in the developing limb (Nelson et al., 1996) (Figure 12D). No limb-related TFBS were predicted to be disrupted by the insertion.

Discussion

These findings indicate that ZRS603ins13, the small insertion mutation within the ZRS, is likely the cause of preaxial polydactyly and triphalangeal thumb in this family. The mutation is shown to be present in three generations of affected individuals in a fully penetrant inheritance pattern with an invariable phenotype.

The known human point mutations in the ZRS are dispersed over approximately 600 bp and cause a range of preaxial polydactyly phenotypes (VanderMeer and Ahituv, 2011). There is no clear relationship between the location of the mutation and the severity of the phenotype, but it is thought that the mutations may disrupt TFBS. The ZRS603ins13 mutation results in the creation of two motifs that are predicted to bind multiple transcription factors that are known to be expressed in early limb development including *TBX5*, *TBX6*, *Engrailed1 (EN1)*, *SOX8*, and multiple *HOX* genes. *TBX5* mutations have been shown to cause Holt Oram syndrome (MIM# 142900), a congenital defect syndrome with limb malformations that include triphalangeal thumbs. *EN1* (MIM# 131290) is a homeodomain-containing transcription factor required for ventral development of the limb and its deletion in mouse models leads to AP limb patterning defects. A consensus binding site for many *HOX* genes that are all expressed at various stages in the developing limb bud (Nelson et al., 1996) was also found in this insertion. While these transcription factors are not known to directly regulate *SHH* expression in the limb, they are present at the stages in development where abnormal binding to an important regulatory element could stimulate improper expression or interfere with the normal binding of transcription factors that should be regulating *SHH*. Additional limb-associated TFBS created by this mutation that were not detected by our computational analysis could exist and may be related to ZRS activity in the developing limb.

3.5 – ZRS402 C>T Homozygous Mutation

Clinical Reports

Family A

The proband (Figure 13A, individual V/29) has mesomelic shortening of upper and lower extremities due to hypoplastic tibias and hypoplastic radio-ulnar bones, bilateral preaxial polydactyly of toes with 2 extra toes in both sides and bilateral proximally placed hypoplastic thumbs (Figure 13D-F, Appendix II). Her hands have normal second to fifth fingers. She has rudimentary thumbs, which originate from the base of the second finger; they are shortened and laterally deviated (Figure 13D). She has mesomelic shortening of lower extremities with bowed tibias and bilateral dimples. Her feet are shortened and have five remaining toes; two extra toes were removed on each side and she has a gap with a surgical scar between the first and second toe. Her x-rays show short triphalangeal thumbs, mild shortening and fusion of radio-ulnar bones, fusion of metacarpal bones and hypoplastic tibias with bowing of the fibula bilaterally (Figure 13D-F). She was the only family member with lower limb malformations.

Both parents have abnormal thumbs. The father (Figure 13A, IV/30) has bilateral triphalangeal thumbs and a preaxial remnant with a small bone that he is able to manipulate (Figure S1).

Radiography showed a small proximal extra digit and triphalangeal thumbs in both sides (Figure 13B). The mother (Figure 13C, IV/31) has digitalization of thumbs with an extra crease in the distal interphalangeal space (see Appendix II). Radiography reveals bilateral elongated distal phalanx of the first finger; she does not have three phalanges (Figure 13C). The proband has a 5 year old sibling reported to have digitalization of thumbs who was not available for examination.

Extensive family history reveals multiple family members in both sides of the family with digitalization of thumbs and triphalangeal thumbs with and without polydactyly, but no other member with shortening of the limbs. A pedigree of 6 generations and 128 members was constructed from the reported family history (Figure 13A). The proband's parents deny consanguinity, however, both paternal and maternal grandparents were from the same small village in Mexico.

Family B

A second Mexican family of five generations with at least 12 members affected including the 4 from the nuclear family studied and more than 8 from the paternal side was ascertained (Figure 13G). The most common phenotype in this family was digitalization of thumbs and triphalangeal thumbs (Figure 13H). Two subjects had triphalangeal thumb on one hand and preaxial polydactyly on the other hand (Appendix II). Two members had unilateral preaxial polydactyly, and two had only radioulnar synostosis. No involvement of the feet was observed in any subject. DNA samples were collected from the nuclear family in generations IV and V, consisting of the father (IV/10) with radioulnar synostosis, unaffected mother, three affected siblings with bilateral triphalangeal thumbs and one unaffected son.

The phenotypes of affected individuals in both families are summarized in Table 4.

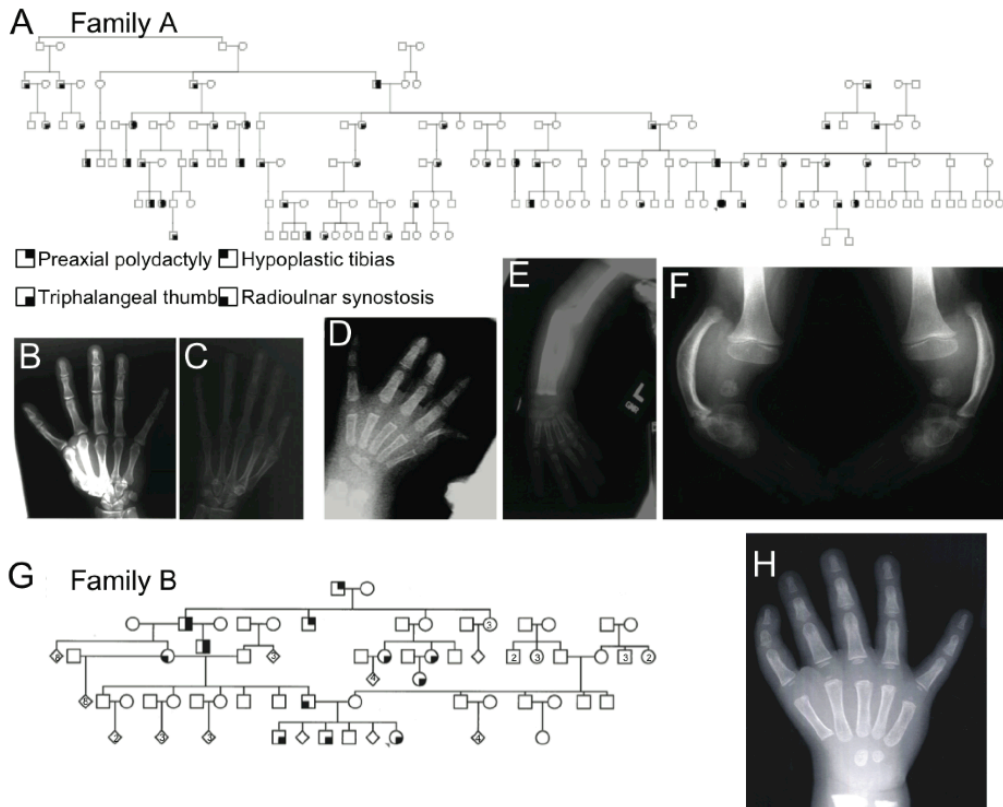


Figure 13. (A) The pedigree of family A contains 154 individuals; 31 have isolated triphalangeal thumb, 14 have preaxial polydactyly with triphalangeal thumb and only one (V/29) has Werner mesomelic syndrome. (B) The proband's father (IV/30) has bilateral triphalangeal thumbs and a preaxial remnant with a small bone that he is able to manipulate. (C) The proband's mother (IV/31) has digitalization of thumbs with an extra crease in the distal interphalangeal space and bilateral elongated distal phalanx of first fingers. (D-F) The proband (V/29) was born with preaxial polydactyly and after surgery has rudimentary thumbs. She has mild shortening and fusion of radio-ulnar bones, fusion of metacarpal bones and bilateral hypoplastic tibias with bowing of the fibula. (G) The 5-generation pedigree of family B contains 2 individuals with preaxial polydactyly, 6 individuals with isolated triphalangeal thumb, 2 individuals with PPD and TPT, and 2 individuals with mild radioulnar synostosis. (H) The proband and siblings have bilateral triphalangeal thumbs, V/12 is shown here.

Table 4. Phenotypic spectrum in the affected subjects in the ZRS 402 C>T families

Affected Subject	Preaxial polydactyly	Triphalangeal thumb	Radioulnar synostosis	Hypoplastic tibias	Preaxial polydactyly (feet)
Family A - II/1		X			
Family A - II/3		X			
Family A - II/6		X			
Family A - II/8	X	X			
Family A - II/12		X			
Family A - III/2		X			
Family A - III/4		X			
Family A - III/7	X	X			
Family A - III/11		X			
Family A - III/13		X			
Family A - III/16	X	X			
Family A - III/18		X			
Family A - III/24		X			
Family A - III/27		X			
Family A - III/29		X			
Family A - IV/1	X	X			
Family A - IV/4	X	X			
Family A - IV/5		X			
Family A - IV/8		X			
Family A - IV/11	X	X			
Family A - IV/12		X			
Family A - IV/15		X			
Family A - IV/17		X			
Family A - IV/19		X			
Family A - IV/21	X	X			
Family A - IV/22		X			
Family A - IV/30	X	X			
Family A - IV/31		X			
Family A - IV/33		X			
Family A - IV/35	X	X			
Family A - V/36	X	X			
Family A - V/2		X			
Family A - V/3		X			
Family A - V/7	X	X			
Family A - V/13		X			
Family A - V/17	X	X			
Family A - V/24		X			

Family A - V/29	X	X	X	X	X
Family A - V/30		X			
Family A - V/32		X			
Family A - V/34		X			
Family A - V/35	X	X			
Family A - VI/1		X			
Family A - VI/5	X	X			
Family A - VI/6		X			
Family A - VI/11		X			
Family B - I/1	X				
Family B - II/2	X	X			
Family B - II/6	X				
Family B - III/10			X		
Family B - III/11	X	X			
Family B - III/15		X			
Family B - III/17		X			
Family B - IV/17			X		
Family B - IV/23		X			
Family B - V/9		X			
Family B - V/11		X			
Family B - V/14		X			

Sequence Analysis

Because the phenotype was similar to that of other ZRS mutations, we screened a 2.1 kilobase (kb) region (hg19, chr7:156,583,564-156,585,727) encompassing the ZRS and the preZRS (pZRS), an adjacent region where mutations are thought to cause polydactyly in dogs (Park et al., 2008). In family A, the parents of the proband (IV/30 and IV/31) were both found to be heterozygous for a novel single base C>T mutation at position 402 of the ZRS sequence (Lettice et al., 2003), (ZRS402C>T) while the proband was found to be homozygous for this mutation (Figure 14A) with a normal copy number as confirmed by qPCR (data not shown). In Family B, the father and three affected subjects were found to be heterozygous for the same mutation. The unaffected sibling did not have the mutation. This mutation is not found in dbSNP

(<http://www.ncbi.nlm.nih.gov/SNP/>) or in the 1000 Genomes Project database

(<http://www.1000genomes.org/>) and was not detected in 59 unrelated Mexican individuals (see Appendix II).

Haplotype analysis determined that the mutations have a common origin. No common relatives were identified through these pedigrees, but they are likely to be distantly related and could be considered to be one large family.

TFBS Analysis

Transcription factor binding site analysis did not identify any limb development-related binding sites disrupted or introduced by the ZRS 402C>T mutation.

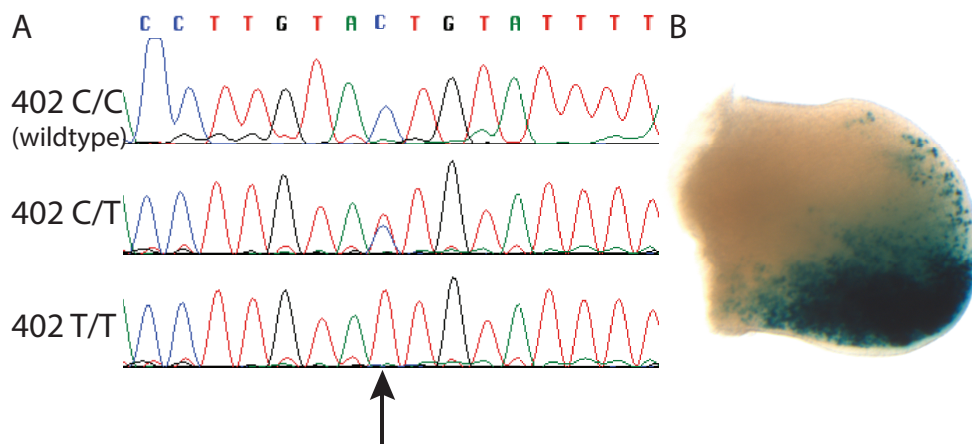


Figure 14. (A) Chromatogram traces show the heterozygous mutation in the father (IV/30) and the homozygous mutation in the proband of Family A compared to a wildtype control sample. (B) The ZRS402C>T mutant transgenic enhancer assay shows the expanded anterior expression of LacZ beyond the posterior of the limb buds of an E11.5 mouse embryo (right hindlimb shown).

Mouse Transgenic Enhancer Assays

A transgenic mouse assay was used to determine the impact of ZRS402C>T on the normal ZPA-specific expression pattern of the wildtype ZRS (Lettice et al., 2003). Transgenic mouse embryos were screened at mouse embryonic day 11.5. These embryos were compared to wildtype ZRS-HSP68-LacZ mice from our ZRS603ins13 study (Laurell et al., 2012). The wildtype ZRS sequence recapitulates normal *Shh* expression in the mouse E11.5 limb; LacZ is restricted to the posterior of the limb bud. The ZRS402C>T transgenic mice show ectopic anterior expression in six of six LacZ-positive transgenic embryos (Figure 14B, Appendix II). The anterior expression of LacZ in the ZRS402C>T embryos is consistent with other ZRS mutations in this assay (Furniss et al., 2008; Laurell et al., 2012; Lettice et al., 2003, 2008). It is worth noting that these *in vivo* assays are qualitative because of uncontrollable variables including variation in expression level between embryos, likely the result of copy number or site of integration differences. No other consistent expression patterns were observed in these embryos.

Discussion

Our findings suggest that the ZRS402 C>T mutation is the likely cause of the limb phenotypes in these families. It segregates with all assayed affected individuals, is not found in multiple human variation databases or 59 unrelated Mexican individuals and leads to ectopic expression of the ZRS in mouse E11.5 embryos. The phenotype is transmitted in an autosomal dominant manner with variable expression among the heterozygotes. While many ZRS mutations have invariable phenotypic presentation (Albuissou et al., 2011; Farooq et al., 2010; Laurell et al., 2012; Semerci et al., 2009), there are others where point mutations have variable phenotypes among affected individuals, similar to what is seen here (Gurnett et al., 2007; Lettice et al., 2003). There is at

least one family whose mutation has incomplete penetrance because phenotypically normal carriers have been identified (Gurnett et al., 2007).

Human point mutations in the ZRS are dispersed over approximately 600 bp and cause a range of limb phenotypes (VanderMeer and Ahituv, 2011). There has been only one homozygous individual reported and this individual had no change in phenotype compared to heterozygotes in the same family (Furniss et al., 2008). Here, we identified an additional homozygous individual with a more severe phenotype, WMS, compared to heterozygous family members. This is the only individual with a phenotype in the lower limbs in these two families. Interestingly, the ZRS mutations that have been reported to cause WMS are seen as heterozygous mutations at ZRS404, two bases away from this ZRS402 mutation. There is no clear relationship between the location of ZRS mutations and the severity of the phenotypes they cause, but the mutations may disrupt different TFBS. For this mutation, no candidate TFBS changes were predicted based on the TRANSFAC database. It is possible that alternative binding sites for known transcription factors were not detected or that there is differential binding of TFs whose binding motifs are not known. It is also possible that these mutations interfere with looping of the ZRS to the *Shh* promoter or the chromatin domain changes that are related to *Shh* expression, but the mechanisms of these interactions are still poorly understood (Amano et al., 2009).

The 402 C>T mutation adds to our understanding of the ZRS. The homozygous phenotype shows that the presence of two mutated copies of the ZRS can cause a more severe congenital limb abnormality. The addition of a second point mutation that can cause WMS further supports the hypothesis that mutations in the portion of the ZRS near 402 and 404 can lead to a more

severe limb phenotype. Expanding the known cohort of mutations and their resulting phenotypes will help researchers investigate the molecular methods that allow distal enhancers to regulate gene expression.

3.6 – Conclusions

The ZRS is a striking region for a number of reasons. It is one of the farthest known enhancers from its target gene at nearly 1 megabase from Sonic Hedgehog. Despite the increasing catalog of known distal regulatory elements, very few have been identified close to this range. It is also an exceptional case because despite extensive characterization of regulatory regions near *Shh*, the ZRS remains the sole enhancer for the limb bud. This is surprising because other genes appear to have multiple enhancers regulating the same tissue (Kikuta et al., 2007). *TBX4* has two hindlimb-specific enhancer elements that are separate from one another despite regulating one gene in the same tissue and time point (Menke et al., 2008). The *HOXD* cluster, another region that has been extensively characterized for regulatory activity has multiple enhancers active at different time points in limb development. Some of the many distal elements regulating *HOXD* genes are even referred to as an “archipelago” of enhancer elements (Montavon et al., 2011). Even *SHH* itself, in tissues other than the limb, uses multiple enhancers with overlapping expression domains (Jeong et al., 2008; Sagai et al., 2009). This is a common pattern for other genes and enhancers where there are multiple partially redundant “shadow enhancers” that are active in the same tissue. It is unknown why the ZRS appears to be an exception to the pattern and *SHH* has yet to have any “shadow enhancers” identified that are also active in the limb.

Another interesting feature of the ZRS is that there appear to be many sites where mutations cause human phenotypes. With other “enhancer bashing” studies, there are relatively few sites where mutations affect the function of the enhancer (Smith et al., 2013a). The ZRS, by contrast, has over 15 different single base pair mutations spread over nearly 650 bases and multiple mutations have been shown affect the expression pattern driven by the enhancer. To date, only a few of these mutations are computationally predicted to affect transcription factor binding and those predictions do not point to common pathways that explain their phenotypes.

It is possible that some of the features of the ZRS are explained by the genomic context. In the limb bud, the ZRS and the promoter of *Shh* physically interact in cells of the posterior limb, as expected, and in the anterior limb where *Shh* is not normally expressed (Amano et al., 2009).

This “poised” interaction may make the ZRS particularly sensitive to mutations that affect how transcription factors recognize the region. It is also possible that the location of the ZRS within an actively transcribed gene, *LMBRI*, forces an open chromatin structure that makes the ZRS more sensitive.

It is unclear why most mutations in the ZRS lead to relatively mild phenotypes of preaxial polydactyly or triphalangeal thumb, but a few affect the long bones of the arms and legs.

Because developmental studies have shown that intervention in limb development at different stages affects different portions of the limb – the autopod is affected by later interventions compared to the zeugopod – it is possible that there is a temporal aspect to the way that these mutations function. With the mouse *in vivo* enhancer studies that have been done on the ZRS, only a single time point can be examined in each embryo and all experiments to date have

focused on the E11.5 time point. It would be interesting to compare expression from mutated ZRS enhancers at earlier stages to see if they affect the timing of expression in addition to the location.

Ultimately, whether the ZRS is a good model for enhancers in general is hard to say at this stage. There are some features that make the ZRS appear unique, but many of these features are what have made it one of the best-characterized enhancers in mammalian development. It is essential for limb development and there are many available mutations to study. While the lessons learned from the ZRS may not apply universally to distal enhancers, they can shed light on mechanisms of transcriptional control and tissue-specific gene regulation.

3.7 - Methods

DNA Sequence Analysis

Primers were designed to cover the ZRS region (Lettice et al., 2003) and the pZRS (Park et al., 2008). Primer sequences can be found in table 5 below. Sequencing was performed by Quintara Biosciences (Quintara Biosciences, Albany CA USA) according to standard procedure and sequences were analyzed using Sequencher (Gene Codes Corporation, Ann Arbor, MI USA). Common SNPs in the region were excluded from our mutation analysis (a full list of excluded variants is in Appendix II). Any mutation that was detected was confirmed with a second independent PCR and sequencing reaction.

Table 5. Primers for human ZRS sequencing

	Forward Primer	Reverse Primer
ZRS1	TTTCAAATGCTCACTTTACATGG	TTTTATGACCAGATGACTTTTTCC
ZRS2	AGGCTGGACTTCCTACTCACTCT	GAATAAAAATGTCAGGAGGAAAAA
pZRS1	AAATTTTACATAACAATCATATG	AAGCAGCTAACTTTTATCTTGGAA
pZRS2	AGGTGACAGCAAATAATCTAAA	TGCTGAAGTGATACTGAAGAGAGG

Transcription Factor Binding Site Predictions

For the two point mutations, the MATCH algorithm function of TRANSFAC was used to compare predicted transcription factor binding for the wildtype and mutant sequences in a 15 base pair region centered on the variation site (Kel et al., 2003). TFBS predictions were limited to mammalian non-repetitive binding matrix data with a score threshold of 0.8.

For the 13 base pair insertion, a more comprehensive analysis was used. We acquired 345 transcription factor motifs from the UniPROBE database (Newburger and Bulyk, 2009) and compiled a small number of motifs for additional known limb factors from the literature including TBX5 (MIM# 601602) (Ghosh et al., 2001), TBX6 (MIM# 602427) (White and Chapman, 2005), the PBX-MEIS1 complex (MIM#s 602100, 601739) (Chang et al., 1997) and the MEIS1-HOXA9 (MIM# 142956) complex (Shen et al., 1997). Binding sites were predicted by scanning DNA sequences for motifs using our implementation of the MATCH algorithm (Kel et al., 2003) with a score threshold of 0.8.

Mouse Transgenic Analysis

The wildtype and two mutant ZRS enhancer assays were used to look at enhancer expression. The complete ZRS region with ZRS603ins13 was PCR amplified from the proband of that family with primers carrying XhoI and ApaI restriction sites (FwPrimer:

GGCCctcgagTTTCAAATGCTCACTTTACATGG, RevPrimer:

ATgggcccTGCTGAAGTGATACTGAAGAGAGG) and cloned into the Hsp68-LacZ enhancer assay vector (Kothary, et al., 1989). It was sequence-verified (Quintara Biosciences) to make sure that it had the ZRS603ins13 mutation and that no other mutations existed in the sequence.

The wildtype enhancer assay vector was generated by removing the insertion from the ZRS603ins13-Hsp68-LacZ vector (Mutagenix, Hillsborough, NJ, USA) and sequence verified.

The ZRS402C>T enhancer vector was generated by site directed mutagenesis (QuikChange Lightning, Mutagenix) and sequenced. Generation of transgenic mice and Beta galactosidase staining at embryonic (E) day 11.5 were done by Cyagen Biosciences, Inc. (Guangzhou, China).

All animal work was approved by the UCSF Institutional Animal Care and Use Committee.

Chapter 4 – ChIP-seq identifies limb signaling center enhancers

4.1 – ChIP-seq

Chromatin immunoprecipitation followed by deep sequencing (ChIP-seq) is used to study molecular interactions between DNA and proteins. ChIP-seq uses antibodies to pull down DNA-protein complexes, allowing the identification of genomic regions where specific proteins bind (Figure 15).

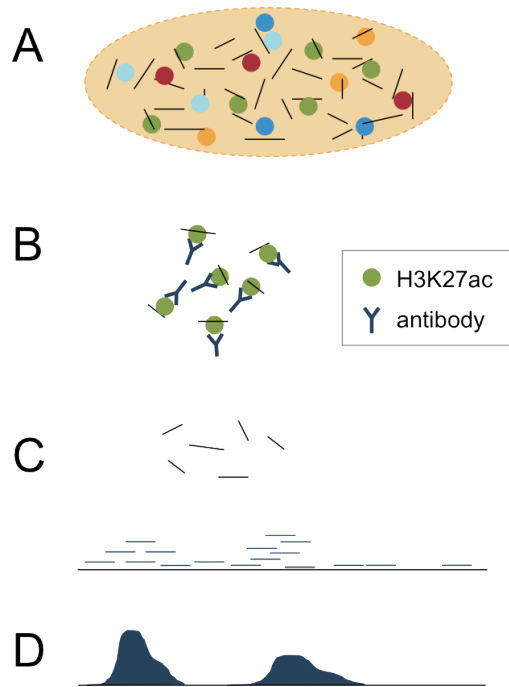


Figure 15. ChIP-seq. (A) Chromatin from a tissue of interest is crosslinked to stabilize DNA-protein interactions and sheared into small fragments. All proteins that were interacting with DNA are fixed in place (B) An antibody that recognizes a transcription factor or a modified histone is used to pull out complexes of target protein and DNA. Here an antibody recognizes the histone modification H3K27ac. (C) DNA fragments are released and sequenced so they can be mapped back to the genome. (D) Regions where many fragments map are called as “peaks” by algorithms that compare the number of reads in the ChIP sample to the number in a control.

Genome-wide binding of transcription factors and chromatin-modifying proteins is being used to locate limb enhancers with up to 80% accuracy (Visel et al., 2009a; Vokes et al., 2008), however, these enhancers show a wide range of expression patterns. This is a critical limitation of previous studies because individual hits cannot easily be mapped to limb signaling centers or pathways. To compound this limitation, valid signals from small tissue subsets can be missed because the most prevalent tissue in a sample gives much stronger signals. In one whole-limb ChIP-seq assay, 19 of the 20 validated hits showed expression in the mesenchyme, but none had expression specific to the ZRS or AER; the most important signaling regions of the developing limb bud (Visel et al., 2009a) (<http://enhancer.lbl.gov>).

The limb is patterned by specialized populations of cells. These cells are a small minority of the total cell population of the limb bud, only about 3-5%, but they are critical for axis specification and patterning. Applying ChIP-seq to the entire limb is not an efficient way to identify enhancers in these small populations because signal from such a small fraction of the limb has the potential to be lost as noise compared to the total signal. Testing signaling centers alone can aid in the detection of enhancers that are only active in a small portion of the limb and will also allow each enhancer to be assigned to a predicted expression domain.

Two signaling centers are particularly important in the development of the limb. The first is the zone of polarizing activity (ZPA), a well-studied signaling center that is important for patterning the anterior-posterior (A-P) axis (see Chapter 2) The ZPA has an important role in limb morphology, particularly in the number and identity of digits with many animal models of polydactyly displaying a second ectopic anterior ZPA in development (Blanc et al., 2002; Chan

et al., 1995; Masuya et al., 1995). The most common phenotype caused by anterior “ZPA activity” is the development of additional digits at the anterior side of the limb or the adoption of more posterior identity by anterior digits. Enhancers active in the ZPA give insight into how the A-P axis is defined and provide candidates for enhancers involved in the evolution of supernumerary digits or in cases of human polydactyly. The second is the apical ectodermal ridge (AER), the region responsible for growth of the limb along the proximal-distal (P-D) axis (see Chapter 2). The mechanisms that specify P-D cell fate are not known, but one model suggests that signals from the AER may be involved in an early patterning of limb bud mesenchyme cells and results from this assay will provide candidate genes for the “distal” signal. Without a functioning AER, the growth of the limb can be severely stunted. Mutations in enhancers that are active in the AER could cause defects in limb growth and patterning in humans.

4.2 – Genome-wide identification of enhancers in signaling centers of the developing limb

The identification of enhancers in the ZPA and AER is challenging. While previous reports have used chromatin immunoprecipitation followed by sequencing (ChIP-seq) on histone marks (Cotney et al., 2013) or E1A binding protein p300 (EP300/p300) (Visel et al., 2009b) on whole limb buds, these address the limb bud as a single tissue. The ZPA and AER are also likely to have different epigenetic signatures than the rest of the cells of the limb, but the signal would be diluted by the larger limb mesenchyme cell population. Tellingly, Visel et al. used mouse transgenic enhancer assays to show that 20 out of the 25 mouse embryonic day (E) 11.5 p300 ChIP-seq peaks are enhancers. However, none of these enhancers were expressed in the ZPA or AER.

In order to identify active enhancers that are specific to these signaling centers, we isolated E11.5 mouse embryonic limbs that were fluorescently labeled in the ZPA or AER and sorted them using fluorescence-activated cell sorting (FACS). We then carried out ChIP-seq on the sorted cells for an epigenetic mark, H3K27ac, associated with active enhancers (Creyghton et al., 2010; Rada-Iglesias et al., 2012). Subsequent mouse transgenic enhancer assays on selected ZPA and AER ChIP-seq peaks found them to be active in these signaling centers. Through this work we have identified two novel sets of signaling-center specific enhancers and potential genes that they regulate which can play important roles in limb development and morphology.

4.3 – Results and Discussion

In order to identify active enhancers within the ZPA or AER, we used transgenic mouse lines that express GFP in these regions. For the ZPA, we used *Shh-GFP-cre* mice (Harfe et al., 2004) that have a cassette containing an in frame fusion between GFP and *cre* inserted at the ATG of the mouse *Shh* gene, leading to specific GFP expression in the ZPA (See Appendix III). For the AER, we crossed *Msx2-cre* mice that expresses *cre* specifically in the AER under the control of the *Msx2* promoter (Sun et al., 2000) to GNZ (ROSA26-nGFP) mice that contain a loxP-flanked STOP cassette followed by a GFP/beta-galactosidase fusion protein sequence with an SV40 nuclear localization signal (Stoller et al., 2008), thus allowing for AER-specific GFP expression (Appendix III). Embryos were harvested at E11.5 and ZPA or AER cells were isolated from dissected limbs by FACS. Similar to previous measurements, the ZPA and AER cells each made up 2-4% of the total cells of the limb (Rock et al., 2007). These cells were subjected to ChIP-seq with an antibody for H3K27ac (Figure 16A).

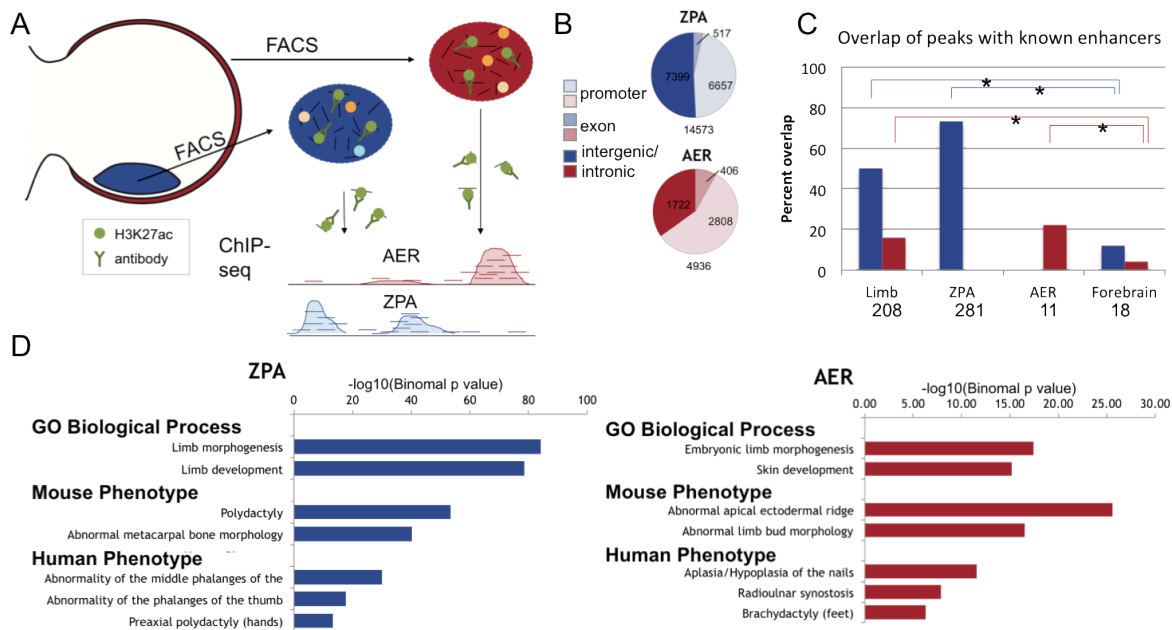


Figure 16. (A) Schematic of ChIP-seq methodology. Cells from the ZPA and AER were isolated by FACS followed by ChIP-seq. (B) The different functional categories associated with H3K27ac ZPA and AER ChIP-seq peaks. (C) ZPA and AER ChIP-seq peaks overlap functionally characterized enhancers from the VISTA Enhancer Browser. Both the ZPA and AER peaks show significant overlap with enhancers in the whole limb, but have little overlap with forebrain enhancers ($*p < 0.0001$; Fisher's exact test, one-tailed). ZPA and AER enhancers also overlap more with their respective tissue. (D) GREAT (McLean et al., 2010) shows significantly enriched terms related to ZPA and AER biological function and phenotypes.

In the ZPA, we identified 14,573 peaks. Of these peaks, 45% were at gene promoters (defined as -2500 and +500 bp from the transcription start site (TSS) based on RefSeq gene annotations), 4% in exons (excluding those also in the promoter region) and 51% in intergenic or intronic regions (Figure 16B). In the AER, we identified 4,936 peaks, 57% of which were located at gene promoters, 8% in exons and 35% in intergenic or intronic regions (Figure 16B). We next analyzed whether these peaks overlap 208 functionally validated limb enhancers from the VISTA Enhancer Browser (Visel et al., 2007). Of the 208 limb enhancers, we annotated 11 with

ZPA/posterior limb bud expression and 18 with AER/ectoderm expression. As a negative control, we used from the same database a set of 281 forebrain enhancers that did not have any limb expression. We observed a significant overlap with the whole limb set versus the forebrain set ($p < 0.0001$; Fisher's exact test, one-tailed) for both the ZPA and AER (Figure 16C).

Comparison to the ZPA and AER functional enhancers found that our ZPA and AER ChIP-seq peaks overlap more with their respective tissue, but the number of functionally characterized ZPA and AER enhancers was too low for statistical significance.

To associate biological functions to these peaks we used GREAT (McLean et al., 2010). In both ZPA and AER datasets, we observed a significant enrichment for terms related to limb development and limb malformation phenotypes (Figure 16D). The ZPA includes mouse and human phenotypic terms related to polydactyly and abnormal digit patterning which are common effects of ZPA signaling disruptions. The AER terms include abnormal limb bud morphology, aplasia of the nails, and brachydactyly, all of which are phenotypes associated with AER defects. Limb development and malformation phenotype terms are also enriched when using only non-TSS peaks (Appendix III). We also examined gene loci known to be highly expressed in the ZPA or AER and observed a correlation with H3K27ac peak presence (Appendix III). As examples, we observed a specific H3K27ac peak for *Shh* and its enhancer, ZRS, in the ZPA and for engrailed homeobox 1 (*En1*) in the AER, while *Alx4*, a gene expressed specifically in the anterior mesenchyme, did not have an H3K27ac signal in either tissue (Figure 17).

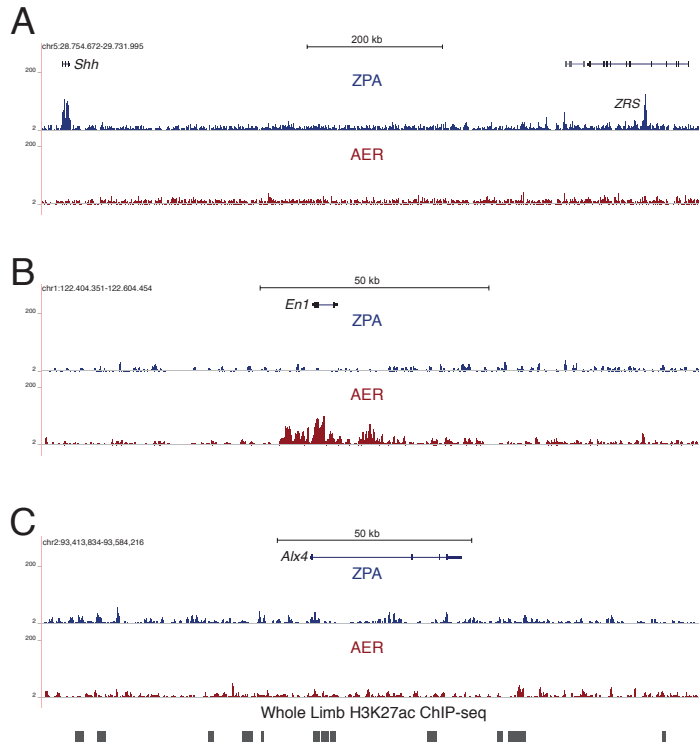


Figure 17. (A) The *Shh* gene and its enhancer, ZRS, are active in the ZPA and overlap H3K27ac. In the AER, there is no H3K27ac peak at this locus. (B) *En1*, an AER and ventral ectoderm expressed gene, has an H3K27ac peak in the AER sample, but not in the ZPA. (C) The locus around *Alx4*, a gene expressed in the anterior limb mesenchyme, does not have an H3K27ac peak in either signaling center.

We next compared the ZPA and AER ChIP-seq peaks to a published data set of H3K27ac ChIP-seq from whole limb tissue and found that, based on peak overlap of >1 base pair, there are 1,233 peaks unique to the ZPA and 715 unique to the AER (Figure 18A). We then set out to create a list of potential ZPA and AER enhancers. We removed ZPA and AER peaks that overlap with the promoter region and/or overlap (>1 base pair) with previously published E11.5 whole-limb H3K27ac and p300 ChIP-seq datasets [REFs]. The resulting lists consist of 1,233 ZPA-specific and 715 AER-specific peaks.

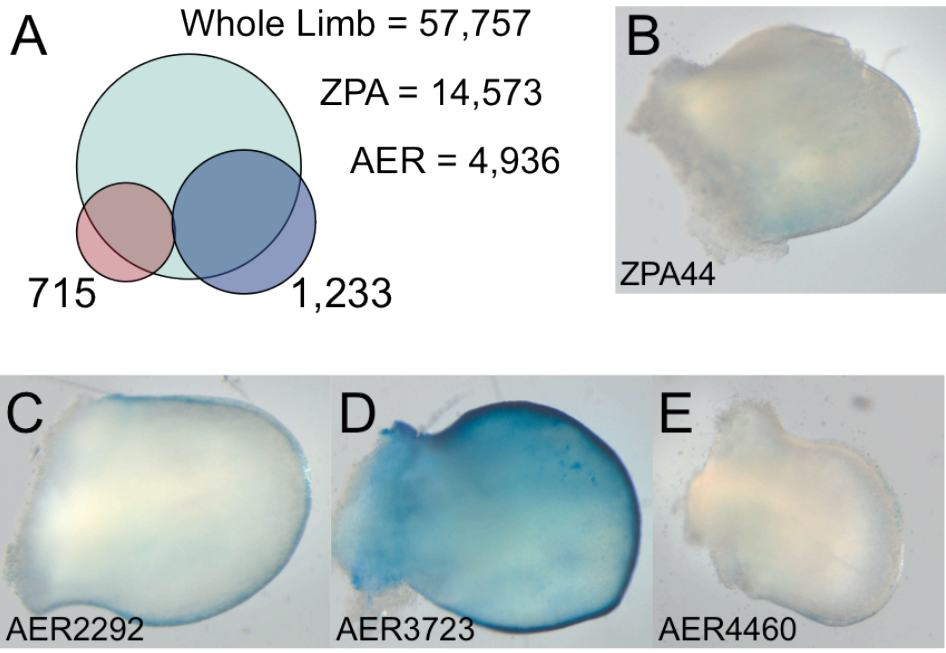


Figure 18. (A) By overlapping ZPA and AER H3K27ac peaks to whole-limb H3K27ac peaks, we obtained 1,233 ZPA and 715 AER specific ChIP-seq peaks. (B) ZPA peak 44 shows ZPA-specific limb enhancer expression in a mouse transgenic assay. (C-E) AER peaks 2292, 3723 and 4460 show AER-specific limb expression in a mouse transgenic assay.

To functionally validate that these ChIP-seq peaks are signaling center-specific, we tested five candidate enhancers from each signaling center for enhancer activity in an *in vivo* mouse assay. For these assays, we prioritized enhancers that were near genes involved in limb development (see Appendix III). These genes include transcription factors, signal receptors and genes involved in cilia transport that are required for *Shh* signaling. Of the five ZPA candidates, one showed enhancer expression in the posterior limb bud (Figure 18B). ZPA44 is located 30kb upstream of *Tcfap2b*, a gene expressed in the ZPA at levels nearly 4 times higher than in the rest of the limb (Rock et al., 2007). Out of the five AER candidates, three were positive for enhancer activity in the AER and ectoderm layer of the limb bud (Figure 18C-E). AER2292 is located 600kb upstream of *Fgfr2*, which is necessary for limb bud development (Xu et al., 1998).

AER3723 is 142kb downstream of *Sp8*, a transcription factor expressed in the AER and required for limb bud outgrowth (Bell et al., 2003). The third AER enhancer, AER4460 was one of the strongest ChIP-seq signals. AER4460 is located in the first intron of the Protein S gene (*Pros1*), 15kb upstream of ADP-ribosylation factor-like 13b (*Arl13b*), which is involved in the localization of *Shh* pathway components in the cilium (Larkins et al., 2011). The rate of *in vivo* enhancer validation for this data set was low, particularly in the ZPA. This could be due to the stringency of our peak selection. By eliminating peaks in our dataset that had any overlap to a region called as a peak in the whole limb set, we likely eliminated enhancers that are biologically specific to the ZPA, but produced a signal that was identifiable in the whole limb set. These are probably some of the strongest enhancers because we find that the distribution of peak heights in the signaling center specific sets is lower than the distribution for the entire set of either the ZPA or AER (Appendix III). For example, the ZRS, which is known to control ZPA-specific expression of *Shh*, is called as a peak in the whole limb H3K27ac dataset under some parameters (Cotney et al., 2012).

Using FACS followed by low-cell ChIP-seq on ZPA and AER fluorescently labeled E11.5 embryos, we identified thousands of potential limb signaling center peaks. Several of these sequences would have been missed by whole-limb ChIP-seq, due to these signaling centers being diluted in this tissue. We observed a strong correlation between known ZPA and AER specific gene loci (Appendix III) and H3K27ac peak prevalence. We also observed a strong correlation with previously characterized limb enhancers (Visel et al., 2007) and various biological terms using GREAT (McLean et al., 2010). Our functional mouse enhancer validation had a relatively low success rate, particularly in the ZPA, possibly due to our enhancer selection using only

peaks that were exclusively called in either ZPA or AER, thus removing potentially strong peaks whose signal was also observed in whole-limb ChIP-seq datasets.

Our study provides a novel list of putative ZPA and AER enhancers, highlighting regions in the genome which can serve as ideal candidates for nucleotide variation that can lead to evolutionary differences in limb morphology and limb malformations. Using GREAT to associate peaks with biological function, we find that these peaks are in proximity to genes with roles in limb development and patterning processes that align with the roles of the ZPA and AER. Further assays will be needed in order to functionally characterize these regions.

4.4 – Methods

Tissue collection

To isolate cells from the ZPA and AER, transgenic mouse lines were established that express GFP in one of these regions. For the ZPA, a line with an EGFP/Cre cassette in one allele of *Shh* *Shh*^{tm1(EGFP/cre)Cjt} (Harfe et al., 2004) was used. Heterozygous males were mated to CD-1 females and GFP-positive embryos were used for tissue collection. For the AER, homozygous male mice carrying an AER-active Cre transgene, tg(*Msx2-cre*)5Rem (Sun et al., 2000), were mated to females homozygous for a floxed reporter B6.129-GT(ROSA)26Sor^{tm1Joe} (Stoller et al., 2008). All mice were backcrossed onto a CD-1 strain to address any variation of embryo size between lines. All mouse work was approved by the UCSF institutional animal care and use committee.

Embryos were collected at E11.5 from timed matings. For the ZPA samples, embryos were examined briefly with a dissecting fluorescent microscope to select those carrying the GFP

allele. Both forelimbs and hindlimbs were collected and crosslinked with 1% formaldehyde for 10 minutes at room temperature. Cross-linking was then quenched with glycine and tissue was rinsed with cold PBS. Tissue was homogenized to a single cell suspension by douncing for FACS. Cells were kept on ice in a solution of 0.5 μ M EDTA, 0.05% BSA in PBS for sorting as previously described (Rock et al., 2007) and sorted on a FACS Aria II (BD Biosciences). GFP-positive cells represented approximately 2-4% of the limb tissue for both ZPA and AER. Cells were collected in the sorting buffer, pelleted by centrifugation, and flash frozen to -80°C .

ChIP-seq

ChIP was done using the LowCell#ChIP kit (Diagenode) according to the manufacturer's protocol. Cells were pooled to give approximately 70,000 cells per IP and sonicated with a Covaris sonicator (S220 Focused-ultrasonicator, Covaris) and 30ul of sheared chromatin was used for each ChIP. The antibody anti-acetyl histone H3 (Lys17) clone CMA309 (Milipore 05-1334) was used. Libraries were constructed using a pre-release beta version of the Rubicon ThruPLEX (now ThruPLEX-FD, Rubicon) library construction kit according to manufacturer's directions. For each library, 10 ul of ChIP material was used with a total of 14 cycles of amplification.

Sequencing on an Illumina HiSeq was carried out at the UC Davis Genome Center and FASTQ files were aligned to the Mus Musculus genome (UCSC build mm9) using Bowtie 0.12.8 (PMID: 19261174). A single bp mismatch was permitted and reads with multiple alignments were discarded. From 43,109,366 and 40,996,716 reads for the AER and ZPA samples respectively, 31,258,376 and 30,757,898 aligned uniquely. The input sample was sequenced more deeply,

with 111,754,456 reads total and 80,925,525 aligning uniquely. In each case, approximately 11% of sequences failed to align, and 15% of were suppressed due to multiple alignments.

The alignments were sorted and indexed using SAMtools 0.1.18 (PMID: 19505943) and converted to BED files using the utility bam2bed, a part of bedtools 2.17.0 (PMID: 20110278). The peak-finding tool SICER 1.1 (PMID: 19505939) was used to identify enriched H3K27ac islands in the AER and ZPA samples, with the input sample as the control library. The settings used were as follows: redundancy threshold=1, window size=200bp, fragment size=300bp, effective genome fraction=0.74, gap size=200bp, FDR=0.01.

Mouse transgenic enhancer assays

Ten selected peaks were PCR-amplified from mouse genomic DNA. Coordinates of peaks and primer sequences are available in Appendix III. The enhancer candidates were cloned into an HSP68-LacZ enhancer assay vector (Kothary et al., 1989) using the Gateway system (Invitrogen). Transgenic mouse assays were carried out through Cyagen Biosciences using standard procedures (Nagy et al., 2002) and embryos were examined at mouse E11.5. All mouse work was approved by the UCSF institutional animal care and use committee.

Chapter 5 – Conclusions and Future Directions

5.1 – Summary

This work has focused on the role of enhancers that are active in development and may play important roles in tissue patterning. Using a model of enhancer activity where isolated phenotypes that frequently occur as part of syndromes, we identified multiple novel mutations in an enhancer that cause preaxial polydactyly and triphalangeal thumb. To expand our knowledge of enhancers that are active in signaling centers of the tetrapod limb, I used ChIP-seq for active histone signatures and identified enhancers that are active specifically in the ZPA and AER of the mouse limb bud. These enhancers are likely candidates for the location of mutations that cause additional human limb phenotypes that are related to defects in patterning and development. The number of isolated limb malformations with no identified gene coding mutation suggests that many of them could be caused by *cis*-regulatory mutations. Based on our understanding of model organism development and these candidate enhancers, it will be possible to continue to identify enhancers where changes in sequence cause changes in human phenotypes.

5.2 – Evidence for additional *cis*-regulatory mutations

There are likely to be many more *cis*-regulatory mutations that cause limb malformations that have not yet been observed in human patients. For example, several *cis*-regulatory mutations have been shown to cause limb malformations in model organisms but have not yet been found in humans. One example is a duplication of a noncoding region 5' of *SRY* (sex determining Y)-

box 9 (*SOX9*), that is associated with brachydactyly-anonychia (Kurth et al., 2009). *SOX9* is a gene that is involved in chondrocyte differentiation and skeletal development. Inactivation of *Sox9* specifically in the mouse limbs leads to their complete absence, but the primary axes of the limb appear to be patterned correctly (Akiyama et al., 2002). A transgenic mouse designed to overexpress *Sox9* specifically in the limb mesenchyme showed polydactyly as well as short, broad digits (Akiyama et al., 2007). Enhancers in the duplicated region have not been identified, but the phenotype of the human duplication is consistent with the mouse overexpression phenotype and suggests that changing the expression of *SOX9* may be able to cause multiple limb malformation phenotypes.

It has been shown that removing genes from their normal chromosomal context by inversions or translocations can lead to limb malformations. This appears to be caused by the removal of the genes from their regulatory environment, perhaps separating a gene from its enhancers. One such example is a chromosomal rearrangement near the homeobox A (*HOXA*) gene cluster that was found to lead to postaxial polysyndactyly (Lodder et al., 2009). In this patient, a balanced inversion on chromosome 7 that does not disrupt developmentally important genes removes the *HOXA* cluster from a region of several putative *cis*-regulatory elements more than 1 megabase away from the cluster. It is possible that other cases of postaxial polydactyly could be caused by mutations in *HOXA*-associated *cis*-regulatory elements. Chromosomal breakpoints near the homeobox D (*HOXD*) cluster also cause various limb malformations without disrupting genes (Dlugaszewska et al., 2006). Similarly, a translocation with a breakpoint near the parathyroid hormone-like hormone (*PTHLH*) gene, an important chondrogenic regulator, was shown to lead to brachydactyly type E by downregulating gene expression (Maass et al., 2010).

Misexpression of genes due to genetic defects other than noncoding regulatory mutations suggests that regulatory mutations could affect those genes and lead to limb malformations. One such example is a duplication that includes the T-box 4 (*TBX4*) gene and causes isolated clubfoot (Alvarado et al., 2010). In other studies, isolated clubfoot was shown to be caused by mutations in paired-like homeodomain 1 (*PITX1*), which is thought to directly regulate the expression of *TBX4* (Gurnett et al., 2008; Logan and Tabin, 1999). Together, these provide a compelling argument that expression levels of *TBX4* may be related to this limb malformation. Two hindlimb-specific *TBX4* enhancers have been identified in the mouse (Menke et al., 2008) providing excellent candidates for *cis*-regulatory regions that may harbor mutations leading to human isolated clubfoot (Lu et al., 2012). Other cases of small duplications or deletions causing limb malformations, presumably by altering gene expression levels, have been reported (Schluth-Bolard et al., 2008; Tsai et al., 2009; van der Zwaag et al., 2010).

The depth of our current understanding of the genes and gene networks involved in tetrapod limb development provides a long list of candidate genes for which changes in *cis*-regulation could cause human limb malformations. Enhancers proposed to regulate some of these genes have already been identified (Abbasi et al., 2010; Cretokos et al., 2008; Durand et al., 2009; Feng et al., 2008; Sasaki et al., 2002). The study of human limb malformations has also led to the identification of specific genomic loci associated with various limb phenotypes. In cases where no coding mutations can be detected within the locus, the causal mutation may be in a *cis*-regulatory element. One such example is the split hand-foot malformation (SHFM), a limb malformation that is linked to 6 different genomic loci, only two of which have a coding

mutation shown to cause SHFM. Recently, a study using ChIP-seq for the transcription factor tumor protein p63 (*TP63*) identified an enhancer within the SHFM1 locus that is thought to control the expression of distal-less homeobox 5 & 6 (*DLX5/6*) genes specifically in the limb AER (Kouwenhoven et al., 2010). Fine mapping of a patient with an 880 kb deletion was shown to encompass this enhancer, suggesting that its removal might lead to this phenotype (Kouwenhoven et al., 2010). In addition, studies in model organisms can also lead to the identification of *cis*-regulatory elements where mutations cause limb malformations in the model, but where mutations have not yet been identified in human patients (Feng et al., 2008; Liska et al., 2009).

5.3 – Interpreting regulatory mutations

Deciphering the effect of *cis*-regulatory mutations is difficult in part because the mechanisms by which enhancers affect gene expression are poorly understood. Because transcription factors are known to bind to enhancers, a mutation that destroys a TFBS sequence could affect the recruitment of that transcription factor. However, many TFBS sequences are degenerate and a mutation might not have a strong impact on TF binding *in vivo*. Additionally, these binding sites often occur in clusters so mutations in a single site may not be critical (Gotea et al., 2010). It would also be possible for a mutation to create a new TFBS that could recruit the wrong TF to that region and conceivably cause ectopic expression or prevent normal expression. A better understanding of the flexibility of TFBS sequence and the importance of multiple sites would assist in the understanding of *cis*-regulatory mutation consequences.

Enhancers have been shown to physically “loop” to the promoters of the genes they regulate. This kind of long-range genomic interaction is correlated to gene expression, but remains poorly understood (reviewed in (Sexton et al., 2009)). This interaction is likely related to the mechanisms enhancers use to activate transcription. For example, it has been shown that the ZRS enhancer interacts with the *SHH* promoter only in regions of the limb where *SHH* is “poised” for transcription: the posterior ZPA and part of the anterior mesenchyme (Amano et al., 2009). In addition to this looping interaction, the same study showed that the looped *SHH*–ZRS locus moves out of its chromosome territory only in the ZPA region where it is actively transcribed. Other studies show a role for nuclear matrix proteins in the interaction of SHH and the ZRS (Zhao et al., 2009). While the exact mechanisms of looping remain unclear, it appears to be one way that cell-type specific DNA interactions are formed in a pattern that correlates with gene expression (Kagey et al., 2010).

5.4 – Outlook

It is reasonable to speculate that *cis*-regulatory elements other than enhancers play a causative role in limb malformations. The best-studied enhancer, the ZRS, clearly fulfills the canonical role of an enhancer – upregulating a specific gene in *cis* – and yet mutations within it can increase *SHH* expression beyond the normal level. Understanding the function of *cis*-regulatory elements like silencers and insulators is compounded by the lack of robust and reproducible *in vivo* assays for these elements. Even the commonly used enhancer assays are not quantitative because of copy number and integration site variability. The creation of functional assays to quantitatively validate *cis*-regulatory elements in an *in vivo* system will allow more rigorous study of regulatory mutations than is currently possible. Further advancements in tools to

identify interactions between regions of DNA and nuclear organization will also allow better identification and understanding of how regulatory elements control gene expression (van Steensel and Dekker, 2010). Additionally, better identification of enhancers that are directly involved in axis specification and pattern formation in the limb will help inform the search for causal mutations in cases of isolated malformations.

References

Abbasi, A.A., Pappas, Z., Malik, S., Bangs, F., Schmidt, A., Koch, S., Lopez-rios, J., and Grzeschik, K.H. (2010). Human intronic enhancers control distinct sub-domains of Gli3 expression during mouse CNS and limb development. *BMC Dev. Biol.* *10*.

Akiyama, H., Chaboissier, M.C., Martin, J.F., Schedl, A., and de Crombrughe, 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. *J Bone Min. Res* *17*, 1071.

Akiyama, H., Stadler, H.S., Martin, J.F., Ishii, T.M., Beachy, P.A., Nakamura, T., and de Crombrughe, B. (2007). Misexpression of Sox9 in mouse limb bud mesenchyme induces polydactyly and rescues hypodactyly mice. *Matrix Biol.* *26*, 224–233.

Al-Qattan, M.M., Al Abdulkareem, I., Al Haidan, Y., and Al Balwi, M. (2012). A novel mutation in the SHH long-range regulator (ZRS) is associated with preaxial polydactyly, triphalangeal thumb, and severe radial ray deficiency. *Am. J. Med. Genet. Part A* *158A*, 2610–2615.

Al-Qattan, M.M., Shamseldin, H.E., Al Mazyad, M., Al Deghaither, S., and Alkuraya, F.S. (2013). Genetic heterogeneity in type III familial cutaneous syndactyly and linkage to chromosome 7q36. *Am. J. Med. Genet. Part A* *161*, 1579–1584.

Albuisson, J., Isidor, B., Giraud, M., Pichon, O., Marsaud, T., David, A., Le Caignec, C., Bezieau, S., Giraud, Marsaud, et al. (2011). Identification of two novel mutations in Shh long-range regulator associated with familial pre-axial polydactyly. *Clin. Genet.* *79*, 371–377.

Alvarado, D.M., Aferol, H., McCall, K., Huang, J.B., Techy, M., Buchan, J., Cady, J., Gonzales, P.R., Dobbs, M.B., and Gurnett, C.A. (2010). Familial isolated clubfoot is associated with recurrent chromosome 17q23.1q23.2 microduplications containing TBX4. *Am J Hum Genet* *87*, 154–160.

Amano, T., Sagai, T., Tanabe, H., Mizushima, Y., Nakazawa, H., and Shiroishi, T. (2009). Chromosomal Dynamics at the Shh Locus: Limb Bud-Specific Differential Regulation of Competence and Active Transcription. *16*, 47–57.

Ariza-Cosano, A., Visel, A., Pennacchio, L.A., Fraser, H.B., Gómez-Skarmeta, J.L., Irimia, M., and Bessa, J. (2012). Differences in enhancer activity in mouse and zebrafish reporter assays are often associated with changes in gene expression. *BMC Genomics* *13*, 713.

Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell* *129*, 823–837.

- Bell, S.M., Schreiner, C.M., Waclaw, R.R., Campbell, K., Potter, S.S., and Scott, W.J. (2003). Sp8 is crucial for limb outgrowth and neuropore closure. *Proc. Natl. Acad. Sci. U. S. A.* *100*, 12195–12200.
- Bellen, H.J. (1999). Ten years of enhancer detection: Lessons from the fly. *Plant Cell* *11*, 2271–2281.
- Bessa, J., Tena, J.J., de la Calle-Mustienes, E., Fernandez-Minan, A., Naranjo, S., Fernandez, A., Montoliu, L., Akalin, A., Lenhard, B., Casares, F., et al. (2009). Zebrafish Enhancer Detection (ZED) Vector: A New Tool to Facilitate Transgenesis and the Functional Analysis of cis-Regulatory Regions in Zebrafish. *Dev Dyn* *238*, 2409–2417.
- Birnbaum, R.Y., Everman, D.B., Murphy, K.K., Gurrieri, F., Schwartz, C.E., and Ahituv, N. (2012a). Functional characterization of tissue-specific enhancers in the DLX5/6 locus. *Hum. Mol. Genet.* *21*, 4930–4938.
- Birnbaum, R.Y., Clowney, E.J., Agamy, O., Kim, M.J., Zhao, J., Yamanaka, T., Pappalardo, Z., Clarke, S.L., Wenger, A.M., Nguyen, L., et al. (2012b). Coding exons function as tissue-specific enhancers of nearby genes. *Genome Res.* *22*, 1059–1068.
- Blanc, I., Bach, A., and Robert, B. (2002). Unusual pattern of Sonic hedgehog expression in the polydactylous mouse mutant Hemimelic extra-toes. *Int. J. Dev. Biol.* *46*, 969–974.
- Blow, M.J., McCulley, D.J., Li, Z.R., Zhang, T., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F., et al. (2010). ChIP-Seq identification of weakly conserved heart enhancers. *Nat. Genet.* *42*, 806–U107.
- Booker, B.M., Murphy, K.K., and Ahituv, N. (2013). Functional analysis of limb enhancers in the developing fin. *Dev. Genes Evol.* *223*, 395–399.
- Butler, J.E.F., and Kadonaga, J.T. (2002). The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes Dev.* *16*, 2583–2592.
- Cecchini, K.R., Banerjee, A.R., and Kim, T.H. (2009). Towards a genome-wide reconstruction of cis-regulatory networks in the human genome. *Semin. Cell Dev. Biol.* *20*, 842–848.
- Chan, D.C., Laufer, E., Tabin, C., and Leder, P. (1995). Polydactylous limbs in Strong's luxoid mice result from ectopic polarizing activity. *Development* *121*, 1971–1978.
- Chang, C.P., Jacobs, Y., Nakamura, T., Jenkins, N.A., Copeland, N.G., and Cleary, M.L. (1997). Meis proteins are major in vivo DNA binding partners for wild-type but not chimeric Pbx proteins. *Mol Cell Biol* *17*, 5679–5687.

Cho, T., Baek, G., Lee, H., Moon, H., Yoo, W., and Choi, I. (2013). Tibial hemimelia-polydactyly-five-fingered hand syndrome associated with a 404 G>A mutation in a distant sonic hedgehog cis-regulator (ZRS): a case report. *J Pediatr Orthop B* 22, 219–221.

Clamp, M., Fry, B., Kamal, M., Xie, X.H., Cuff, J., Lin, M.F., Kellis, M., Lindblad-Toh, K., and Lander, E.S. (2007). Distinguishing protein-coding and noncoding genes in the human genome. *Proc Natl Acad Sci USA* 104, 19428–19433.

Clarke, S.L., VanderMeer, J.E., Wenger, A.M., Schaar, B.T., Ahituv, N., and Bejerano, G. (2012). Human Developmental Enhancers Conserved between Deuterostomes and Protostomes. *PLoS Genet.* 8, e1002852.

Cotney, J., Leng, J., Oh, S., DeMare, L.E., Reilly, S.K., Gerstein, M.B., and Noonan, J.P. (2012). Chromatin state signatures associated with tissue-specific gene expression and enhancer activity in the embryonic limb. *Genome Res.* 22, 1069–1080.

Cotney, J., Leng, J., Yin, J., Reilly, S.K., DeMare, L.E., Emera, D., Ayoub, A.E., Rakic, P., and Noonan, J.P. (2013). The evolution of lineage-specific regulatory activities in the human embryonic limb. *Cell* 154, 185–196.

Cretekos, C.J., Wang, Y., Green, E.D., Martin, J.F., Rasweiler, J.J., and Behringer, R.R. (2008). Regulatory divergence modifies limb length between mammals. *Genes Dev.* 22, 141–151.

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. U. S. A.* 107, 21931–21936.

Cuddapah, S., Jothi, R., Schones, D.E., Roh, T.Y., Cui, K.R., and Zhao, K.J. (2009). Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. *Genome Res* 19, 24–32.

Dahn, R.D., and Fallon, J.F. (2000). Interdigital regulation of digit identity and homeotic transformation by modulated BMP signaling. *Science* (80-.). 289, 438–441.

Dai, L., Guo, H., Meng, H., Zhang, K., Hu, H., Yao, H., and Bai, Y. (2013). Confirmation of genetic homogeneity of syndactyly type IV and triphalangeal thumb-polysyndactyly syndrome in a Chinese family and review of the literature. *Eur. J. Pediatr.* 172, 1467–1473.

Dermitzakis, E.T., Reymond, A., and Antonarakis, S.E. (2005). Conserved non-genic sequences - an unexpected feature of mammalian genomes. *Nat. Rev. Genet.* 6, 151–157.

- Długaszewska, B., Silahtaroglu, A., Menzel, C., Kubart, S., Cohen, M., Mundlos, S., Tümer, Z., Kjaer, K., Friedrich, U., Ropers, H.H., et al. (2006). Breakpoints around the HOXD cluster result in various limb malformations. *J Med Genet* 43, 111–118.
- Durand, C., Bangs, F., Signolet, J., Decker, E., Tickle, C., and Rappold, G. (2009). Enhancer elements upstream of the SHOX gene are active in the developing limb. *Eur. J. Hum. Genet.* 1–6.
- Farabee, W.C. (1903). Hereditary and sexual influence in meristic variation: a study of digital malformations in man. Harvard University PhD Thesis.
- Farooq, M., Troelsen, J.T., Boyd, M., Eiberg, H., Hansen, L., Hussain, M.S., Rehman, S. ur, Azhar, A., Ali, A., Bakhtiar, S.M., et al. (2010). Preaxial polydactyly/triphalangeal thumb is associated with changed transcription factor-binding affinity in a family with a novel point mutation in the long-range cis-regulatory element ZRS. *Eur. J. Hum. Genet.* 18, 1–4.
- Feng, W., Huang, J., Zhang, J., and Williams, T. (2008). Identification and analysis of a conserved Tcfap2a intronic enhancer element required for expression in facial and limb bud mesenchyme. *Mol. Cell. Biol.* 28, 315–325.
- Fernandez-Teran, M., and Ros, M.A. (2008). The Apical Ectodermal Ridge: morphological aspects and signaling pathways . *Int J Dev Biol* 52, 857–871.
- Fisher, S., Grice, E.A., Vinton, R.M., Bessling, S.L., Urasaki, A., Kawakami, K., and McCallion, A.S. (2006). Evaluating the biological relevance of putative enhancers using Tol2 transposon-mediated transgenesis in zebrafish. *Nat. Protoc.* 1, 1297–1305.
- Furniss, D., Lettice, L.A., Taylor, I.B., Critchley, P.S., Giele, H., Hill, R.E., and Wilkie, A.O.M. (2008). A variant in the sonic hedgehog regulatory sequence (ZRS) is associated with triphalangeal thumb and deregulates expression in the developing limb. *Hum. Mol. Genet.* 17, 2417–2423.
- Géraudie, J., and Landis, W.J. (1982). The fine structure of the developing pelvic fin dermal skeleton in the trout *Salmo gairdneri*. *Am. J. Anat.* 163, 141–156.
- Gerstein, M.B., Kundaje, A., Hariharan, M., Landt, S.G., Yan, K.-K., Cheng, C., Mu, X.J., Khurana, E., Rozowsky, J., Alexander, R., et al. (2012). Architecture of the human regulatory network derived from ENCODE data. *Nature* 489, 91–100.
- Ghosh, T.K., Packham, E.A., Bonser, A.J., Robinson, T.E., Cross, S., and Brook, J.D. (2001). Characterization of the TBX5 binding site and analysis of mutations that cause Holt-Oram syndrome. *Hum Mol Genet* 10, 1983–1994.

Gotea, V., Visel, A., Westlund, J.M., Res, G., Nobrega, M.A., Pennacchio, L.A., and Ovcharenko, I. (2010). Homotypic clusters of transcription factor binding sites are a key component of human promoters and enhancers. *Genome Res* 20, 565–577.

Grandel, H., and Brand, M. (2011). Zebrafish limb development is triggered by a retinoic acid signal during gastrulation. *Dev Dyn* 240, 1116–1126.

Grandel, H., and Schulte-Merker, S. (1998). The development of the paired fins in the zebrafish (*Danio rerio*). *Mech. Dev.* 79, 99–120.

Gurnett, C.A., Bowcock, A.M., Dietz, F.R., Morcuende, J.A., Murray, J.C., and Dobbs, M.B. (2007). Two novel point mutations in the long-range SHH enhancer in three families with triphalangeal thumb and preaxial polydactyly. *Am. J. Med. Genet. Part A* 143A, 27–32.

Gurnett, C.A., Alaei, F., Kruse, L.M., Desruisseau, D.M., Hecht, J.T., Wise, C.A., Bowcock, A.M., and Dobbs, M.B. (2008). Asymmetric lower-limb malformations in individuals with homeobox PITX1 gene mutation. *Am J Hum Genet* 83, 616–622.

Harfe, B.D., Scherz, P.J., Nissim, S., Tian, H., McMahon, A.P., and 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.

Heintzman, N.D., Hon, G.C., Hawkins, R.D., Kheradpour, P., Stark, A., Harp, L.F., Ye, Z., Lee, L.K., Stuart, R.K., Ching, C.W., et al. (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* 459, 108–112.

Heus, H.C., Hing, A., van Baren, M.J., Joosse, M., Breedveld, G.J., Wang, J.C., Burgess, A., Donnis-Keller, H., Berglund, C., Zguricas, J., et al. (1999). A physical and transcriptional map of the preaxial polydactyly locus on chromosome 7q36. *Genomics* 57, 342–351.

Hui, C.C., and Joyner, A.L. (1998). A mouse model of Greig cephalopolysyndactyly syndrome: the extra-toes' mutation contains an intragenic deletion of the Gli3 gene. *Nat. Genet.* 19, 404.

Ianakev, P., Baren, M.J. Van, Daly, M.J., Toledo, S.P.A., Cavalcanti, M.G., Neto, J.C., Silveira, E.L., Heutink, P., and Kilpatrick, M.W. (2001). Acheiropodia Is Caused by a Genomic Deletion in C7orf2, the Human Orthologue of the Lmbr1 Gene. *Analysis* 38–45.

Jeong, Y., Leskow, F.C., El-Jaick, K., Roessler, E., Muenke, M., Yocum, A., Dubourg, C., Li, X., Geng, X., Oliver, G., et al. (2008). Regulation of a remote Shh forebrain enhancer by the Six3 homeoprotein. *Nat. Genet* 40, 1348–1353.

- Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S., et al. (2010). Mediator and cohesin connect gene expression and chromatin architecture. *Nature* *467*, 430–435.
- Kel, A.E., Gobling, E., Reuter, I., Cheremushkin, E., Kel-Margoulis, O. V, and Wingender, E. (2003). MATCH: a tool for searching transcription factor binding sites in DNA sequences. *Nucleic Acids Res* *31*, 3576–3579.
- Kengaku, M., Capdevila, J., Rodriguez-Esteban, C., De La Peña, J., Johnson, R.L., Izpisua Belmonte, J.C., and Tabin, C.J. (1998). Distinct WNT pathways regulating AER formation and dorsoventral polarity in the chick limb bud. *Science* *280*, 1274–1277.
- Kikuta, H., Laplante, M., Navratilova, P., Komisarczuk, A.Z., Engström, P.G., Fredman, D., Akalin, A., Caccamo, M., Sealy, I., Howe, K., et al. (2007). Genomic regulatory blocks encompass multiple neighboring genes and maintain conserved synteny in vertebrates. *Genome Res.* *17*, 545–555.
- Kim, T.H., Abdullaev, Z.K., Smith, A.D., Ching, K.A., Loukinov, D.I., Green, R.D., Zhang, M.Q., Lobanenko, V. V, and Ren, B. (2007). Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell* *128*, 1231–1245.
- Klopocki, E., Ott, C.E., Benatar, N., Ullmann, R., Mundlos, S., and Lehmann, K. (2008). A microduplication of the long range SHH limb regulator (ZRS) is associated with triphalangeal thumb-polysyndactyly syndrome. *J. Med. Genet.* *45*, 370–375.
- Knudsen, T.B., and Kochhar, D.M. (1982). The Hemimelic Extra Toes mouse mutant: Historical perspective on unraveling mechanisms of dysmorphogenesis. *Birth Defects Res. Part C-Embryo Today-Reviews* *90*, 155–162.
- Korn, R., Schoor, M., Neuhaus, H., Henseling, U., Soininen, R., Zachgo, J., and Gossler, A. (1992). Enhancer trap integrations in mouse embryonic stem-cells give rise to staining patterns in chimeric embryos with a high frequency and detect endogenous genes. *Mech Dev* *39*, 95–109.
- Kothary, R., Clapoff, S., Darling, S., Perry, M.D., Moran, L.A., and Rossant, J. (1989). Inducible expression of an hsp68-lacZ hybrid gene in transgenic mice. *Development* *105*, 707–714.
- Kouwenhoven, E.N., van Heeringen, S.J., Tena, J.J., Oti, M., Dutilh, B.E., Alonso, M.E., de la Calle-Mustienes, E., Smeenk, L., Rinne, T., Parsaulian, L., et al. (2010). Genome-wide profiling of p63 DNA-binding sites identifies an element that regulates gene expression during limb development in the 7q21 SHFM1 locus. *PLoS Genet.* *6*.

Kurth, I., Klopocki, E., Stricker, S., van Oosterwijk, J., Vanek, S., Altmann, J., Santos, H.G., van Harssel, J.J.T., de Ravel, T., Wilkie, A.O.M., et al. (2009). Duplications of noncoding elements 5' of SOX9 are associated with brachydactyly-anonychia. *Nat. Genet.* *41*, 862–863.

Landis, W.J., and Géraudie, J. (1990). Organization and development of the mineral phase during early ontogenesis of the bony fin rays of the trout *Oncorhynchus mykiss*. *Anat. Rec.* *228*, 383–391.

Larkins, C.E., Aviles, G.D.G., East, M.P., Kahn, R.A., and Caspary, T. (2011). *Arl13b* regulates ciliogenesis and the dynamic localization of Shh signaling proteins. *Mol. Biol. Cell* *22*, 4694–4703.

Laurell, T., VanderMeer, J.E., Wenger, A.M., Grigelioniene, G., Nordenskjöld, A., Arner, M., Ekblom, A.G., Bejerano, G., Ahituv, N., and Nordgren, A. (2012). A novel 13 base pair insertion in the sonic hedgehog ZRS limb enhancer (ZRS/LMBR1) causes preaxial polydactyly with triphalangeal thumb. *Hum. Mutat.* *33*, 1063–1066.

Lettice, L.A., Horikoshi, T., Heaney, S.J.H., van Baren, M.J., van der Linde, H.C., Breedveld, G.J., Jooisse, M., Akarsu, N., Oostra, B.A., Endo, N., et al. (2002). Disruption of a long-range cis-acting regulator for Shh causes preaxial polydactyly. *Proc. Natl. Acad. Sci. U.S.A.* *99*, 7548–7553.

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

Lettice, L.A., Hill, A.E., Devenney, P.S., and Hill, R.E. (2008). Point mutations in a distant sonic hedgehog cis-regulator generate a variable regulatory output responsible for preaxial polydactyly. *Hum Mol Genet* *17*, 978–985.

Levine, M. (2010). Transcriptional Enhancers in Animal Development and Evolution. *Curr Biol* *20*, R754–R763.

Li, H., Wang, C.Y., Wang, J.X., Wu, G.S., Yu, P., Yan, X.Y., Chen, Y.G., Zhao, L.H., and Zhang, Y.P. (2009). Mutation analysis of a large Chinese pedigree with congenital preaxial polydactyly. *Eur. J. Hum. Genet.* *17*, 604–610.

Li, Q., Ritter, D., Yang, N., Dong, Z., Li, H., Chuang, J.H., and Guo, S. (2010). A systematic approach to identify functional motifs within vertebrate developmental enhancers. *Dev. Biol.* *337*, 484–495.

Liska, F., Snajdr, P., Sedová, L., Seda, O., Chylíková, B., Slámová, P., Krejčí, E., Sedmera, D., Grim, M., Krenová, D., et al. (2009). Deletion of a conserved noncoding sequence in *Plzf*

intron leads to Plzf down-regulation in limb bud and polydactyly in the rat. *Dev. Dyn.* 238, 673–684.

Lodder, E.M., Eussen, B.H., van Hassel, D. a C.M., Hoogeboom, a J.M., Poddighe, P.J., Coert, J.H., Oostra, B. a, de Klein, A., and de Graaff, E. (2009). Implication of long-distance regulation of the HOXA cluster in a patient with postaxial polydactyly. *Chromosome Res.* 17, 737–744.

Logan, M., and Tabin, C.J. (1999). Role of Pitx1 upstream of Tbx4 in specification of hindlimb identity. *Science* (80-.). 283, 1736–1739.

Lu, P., Yu, Y., Perdue, Y., and Werb, Z. (2008). The apical ectodermal ridge is a timer for generating distal limb progenitors. *Development* 135, 1395–1405.

Lu, W., Bacino, C.A., Richards, B.S., Alvarez, C., VanderMeer, J.E., Vella, M., Ahituv, N., Sikka, N., Dietz, F.R., Blanton, S.H., et al. (2012). Studies of TBX4 and chromosome 17q23.1q23.2: An uncommon cause of nonsyndromic clubfoot. *Am. J. Med. Genet. Part A* 158A, 1620–1627.

Maas, S.A., and Fallon, J.F. (2004). Isolation of the chicken Lmbr1 coding sequence and characterization of its role during chick limb development. *Dev. Dyn.* 229, 520–528.

Maas, S.A., Suzuki, T., and Fallon, J.F. (2011). Identification of spontaneous mutations within the long-range limb-specific Sonic hedgehog enhancer (ZRS) that alter Sonic hedgehog expression in the chicken limb mutants oligozeugodactyly and silkie breed. *Dev Dyn* 240, 1212–1222.

Maass, P.G., Wirth, J., Aydin, A., Rump, A., Stricker, S., Tinschert, S., Otero, M., Tsuchimochi, K., Goldring, M.B., Luft, F.C., et al. (2010). A cis-regulatory site downregulates PTHLH in translocation t(8;12)(q13;p11.2) and leads to Brachydactyly Type E. *Hum Mol Genet* 19, 848–860.

Maston, G.A., Evans, S.K., and Green, M.R. (2006). Transcriptional Regulatory Elements in the Human Genome. *Annu. Rev. Genomics Hum. Genet.* 7, 29–59.

Masuya, H., Sagai, T., Wakana, S., Moriwaki, K., and Shiroishi, T. (1995). A duplicated zone of polarizing activity in polydactylous mouse mutants . *Genes Dev.* 9, 1645–1653.

Masuya, H., Sezutsu, H., Sakuraba, Y., Sagai, T., Hosoya, M., Kaneda, H., Miura, I., Kobayashi, K., Sumiyama, K., Shimizu, A., et al. (2007). A series of ENU-induced single-base substitutions in a long-range cis-element altering Sonic hedgehog expression in the developing mouse limb bud. *Genomics* 89, 207–214.

Materna-Kirylyuk, A., Jamsheer, A., Wisniewska, K., Wieckowska, B., Limon, J., Borszewska-Kornacka, M., Sawulicka-Oleszczuk, H., Szwalkiewicz-Warowicka, E., and Latos-Bielenska,

- A. (2013). Epidemiology of isolated preaxial polydactyly type I: data from the Polish Registry of Congenital Malformations (PRCM). *BMC Pediatr.* *13*, 26.
- McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol* *28*, 495–U155.
- Menke, D.B., Guenther, C., and 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.
- Mercader, N. (2007). Early steps of paired fin development in zebrafish compared with tetrapod limb development. *Dev. Growth Differ.* *49*, 421–437.
- Mercader, N., Selleri, L., Criado, L., Pallares, P., Parras, C., Cleary, M.L., and Torres, M. (2009). Ectopic *Meis1* expression in the mouse limb bud alters P-D pattern in a *Pbx*-independent manner. *Int J Dev Biol* *53*, 1483–1494.
- Montavon, T., Soshnikova, N., Mascrez, B., Joye, E., Thevenet, L., Splinter, E., de Laat, W., Spitz, F., and Duboule, D. (2011). A Regulatory Archipelago Controls Hox Genes Transcription in Digits. *Cell* *147*, 1132–1145.
- Moore, K.L., and Persaud, T.V.N. (1998). *The Developing Human: Clinically oriented embryology* (Philadelphia: Saunders).
- Nagy, A., Gertsenstein, M., Vintersten, K., and Behringer, R. (2002). *Manipulating the mouse embryo: A laboratory manual* (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).
- Nelson, C.E., Morgan, B.A., Burke, A.C., Laufer, E., DiMambro, E., Murtaugh, L.C., Gonzales, E., Tessarollo, L., Parada, L.F., and Tabin, C. (1996). Analysis of Hox gene expression in the chick limb bud. *Development* *122*, 1449–1466.
- Neumann, C.J., Grandel, H., Gaffield, W., Schulte-Merker, S., and Nüsslein-Volhard, C. (1999). Transient establishment of anteroposterior polarity in the zebrafish pectoral fin bud in the absence of sonic hedgehog activity. *Development* *126*, 4817–4826.
- Newburger, D.E., and Bulyk, M.L. (2009). UniPROBE: an online database of protein binding microarray data on protein-DNA interactions. *Nucleic Acids Res* *37*, D77–D82.
- Parinov, S., Kondrichin, I., Korzh, V., and Emelyanov, A. (2004). Tol2 transposon-mediated enhancer trap to identify developmentally regulated zebrafish genes in vivo. *Dev Dyn* *231*, 449–459.

- Park, K., Kang, J., Subedi, K.P., Ha, J.H., and Park, C. (2008). Canine polydactyl mutations with heterogeneous origin in the conserved intronic sequence of LMBR1. *Genetics* 179, 2163–2172.
- Pennacchio, L.A., Ahituv, N., Moses, A.M., Prabhakar, S., Nobrega, M.A., Shoukry, M., Minovitsky, S., Dubchak, I., Holt, A., Lewis, K.D., et al. (2006). In vivo enhancer analysis of human conserved non-coding sequences. *Nature* 444, 499–502.
- Perissi, V., Jepsen, K., Glass, C.K., and Rosenfeld, M.G. (2010). Deconstructing repression: evolving models of co-repressor action. *Nat. Rev. Genet.* 11, 109–123.
- Petrykowska, H.M., Vockley, C.M., and Elnitski, L. (2008). Detection and characterization of silencers and enhancer-blockers in the greater CFTR locus. *Genome Res* 18, 1238–1246.
- Qu, S.M., Tucker, S.C., Ehrlich, J.S., Levorse, J.M., Flaherty, L.A., Wisdom, R., and Vogt, T.F. (1998). Mutations in mouse *Aristaless-like4* cause Strong's luxoid polydactyly. *Development* 125, 2711–2721.
- Rada-Iglesias, A., Bajpai, R., Prescott, S., Brugmann, S.A., Swigut, T., and Wysocka, J. (2012). Epigenomic Annotation of Enhancers Predicts Transcriptional Regulators of Human Neural Crest. *Cell Stem Cell* 11, 633–648.
- Rando, O.J. (2012). Combinatorial complexity in chromatin structure and function: revisiting the histone code. *Curr. Opin. Genet. Dev.* 22, 148–155.
- Riddle, R.D., Johnson, R.L., Laufer, E., and Tabin, C. (1993). Sonic-hedgehog mediates the polarizing activity of the ZPA. *Cell* 75, 1401–1416.
- Ritter, D.I., Li, Q., Kostka, D., Pollard, K.S., Guo, S., and Chuang, J.H. (2010). The importance of being cis: evolution of orthologous fish and mammalian enhancer activity. *Mol. Biol. Evol.* 27, 2322–2332.
- Rock, J.R., Lopez, M.C., Baker, H. V, and Harfe, B.D. (2007). Identification of genes expressed in the mouse limb using a novel ZPA microarray approach. *Gene Expr. Patterns* 8, 19–26.
- Sagai, T., Masuya, H., Tamura, M., Shimizu, K., Yada, Y., Wakana, S., Gondo, Y., Noda, T., and Shiroishi, T. (2004). Phylogenetic conservation of a limb-specific, cis-acting regulator of Sonic hedgehog (*Shh*). *Mamm. Genome* 15, 23–34.
- Sagai, T., Hosoya, M., Mizushina, Y., Tamura, M., and Shiroishi, T. (2005). Elimination of a long-range cis-regulatory module causes complete loss of limb-specific *Shh* expression and truncation of the mouse limb. *Development* 132, 797–803.

- Sagai, T., Amano, T., Tamura, M., Mizushina, Y., Sumiyama, K., and Shiroishi, T. (2009). A cluster of three long-range enhancers directs regional Shh expression in the epithelial linings. *Development* 136, 1665–1674.
- Sasaki, H., Yamaoka, T., Ohuchi, H., Yasue, A., Nohno, T., Kawano, H., Kato, S., Itakura, M., Nagayama, M., and Noji, S. (2002). Identification of cis-elements regulating expression of Fgf10 during limb development. *Int J Dev Biol* 46, 963–967.
- Saunders, J.W., and Gasseling, M.T. (1968). Ectodermal-mesenchymal interactions in the origin of limb symmetry. In *Epithelial-Mesenchymal Interactions*, R.F. and R.E. Billingham, ed. (Baltimore: Williams and Wilkins), pp. 78–97.
- Schluth-Bolard, C., Till, M., Labalme, A., Rey, C., Banquart, E., Fautrelle, A., Martin-Denavit, T., Le Lorc'h, M., Romana, S.P., Lazar, V., et al. (2008). TWIST microdeletion identified by array CGH in a patient presenting Saethre-Chotzen phenotype and a complex rearrangement involving chromosomes 2 and 7. *Eur J Hum Genet* 51, 156–164.
- Schwabe, G., and Mundlos, S. (2004). Genetics of congenital hand anomalies. *Handchir. Mikrochir. Plast. Chir.* 36, 85–97.
- Semerci, C.N., Demirkan, F., Ozdemir, M., Biskin, E., Akin, B., Bagci, H., and Akarsu, N.A. (2009). Homozygous feature of isolated triphalangeal thumb-preaxial polydactyly linked to 7q36: no phenotypic difference between homozygotes and heterozygotes. *Clin Genet* 76, 85–90.
- Sexton, T., Bantignies, F., and Cavalli, G. (2009). Genomic interactions: Chromatin loops and gene meeting points in transcriptional regulation. *Semin. Cell Dev. Biol.* 20, 849–855.
- Sharpe, J., Lettice, L., Hecksher-Sorensen, J., Fox, M., Hill, R., and Krumlauf, R. (1999). Identification of Sonic hedgehog as a candidate gene responsible for the polydactylous mouse mutant Sasquatch. *Curr. Biol.* 9, 97–101.
- Shen, W.F., Montgomery, J.C., Rozenfeld, S., Moskow, J.J., Lawrence, H.J., Buchberg, A.M., and Largman, C. (1997). AbdB-like Hox proteins stabilize DNA binding by the Meis1 homeodomain proteins. *Mol Cell Biol* 17, 6448–6458.
- Smale, S.T., and Kadonaga, J.T. (2003). The RNA polymerase II core promoter. *Ann Rev Biochem* 72, 449–479.
- Smith, R.P., Riesenfeld, S.J., Holloway, A.K., Li, Q., Murphy, K.K., Feliciano, N.M., Orecchia, L., Oksenberg, N., Pollard, K.S., and Ahituv, N. (2013a). A compact, in vivo screen of all 6-mers reveals drivers of tissue-specific expression and guides synthetic regulatory element design. *Genome Biol.* 14, R72.

Smith, R.P., Taher, L., Patwardhan, R.P., Kim, M.J., Inoue, F., Shendure, J., Ovcharenko, I., and Ahituv, N. (2013b). Massively parallel decoding of mammalian regulatory sequences supports a flexible organizational model. *Nat. Genet.* 45, 1021–1028.

Van Steensel, B., and Dekker, J. (2010). Genomics tools for unraveling chromosome architecture. *Nat Biotechnol* 28, 1089–1095.

Stoller, J.Z., Degenhardt, K.R., Huang, L., Zhou, D.D., Lu, M.M., and Epstein, J.A. (2008). Cre reporter mouse expressing a nuclear localized fusion of GFP and β -galactosidase reveals new derivatives of Pax3-expressing precursors. *Genesis* 46, 200–204.

Summerbell, D. (1974). Quantitative analysis of effect of excision of AER from chick limb-bud. *J. Embryol. Exp. Morphol.* 32, 651–660.

Sun, M., Ma, F., Zeng, X., Liu, Q., Zhao, X.L., Wu, F.X., Wu, G.P., Zhang, Z.F., Gu, B., Zhao, Y.F., et al. (2008). Triphalangeal thumb-polysyndactyly syndrome and syndactyly type IV are caused by genomic duplications involving the long range, limb-specific SHH enhancer. *J Med Genet* 45, 589–595.

Sun, X., Lewandoski, M., Meyers, E.N., Liu, Y.H., Maxson, R.E., and Martin, G.R. (2000). Conditional inactivation of Fgf4 reveals complexity of signalling during limb bud development. *Nat. Genet.* 25, 83–86.

Thurman, R.E., Rynes, E., Humbert, R., Vierstra, J., Maurano, M.T., Haugen, E., Sheffield, N.C., Stergachis, A.B., Wang, H., Vernot, B., et al. (2012). The accessible chromatin landscape of the human genome. *Nature* 489, 75–82.

Tsai, L.P., Liao, H.M., Chen, Y.J., Fang, J.S., and Chen, C.H. (2009). A novel microdeletion at chromosome 2q31.1-31.2 in a three-generation family presenting duplication of great toes with clinodactyly. *Clin Genet* 75, 449–456.

VanderMeer, J.E., and Ahituv, N. (2011). cis-Regulatory Mutations are a Genetic Cause of Human Limb Malformations. *Dev. Dyn.* 240, 920–930.

VanderMeer, J.E., and Ahituv, N. (2012). cis-Regulatory Enhancer Mutations are a Cause of Human Limb Malformations. In *Gene Regulatory Sequences and Human Disease*, pp. 73–93.

VanderMeer, J.E., Afzal, M., Alyas, S., Haque, S., Ahituv, N., and Malik, S. (2012). A novel ZRS mutation in a Balochi tribal family with triphalangeal thumb, pre-axial polydactyly, post-axial polydactyly, and syndactyly. *Am. J. Med. Genet. Part A* 158A, 2031–2035.

Vandermeer, J.E., Lozano, R., Sun, M., Daentl, D., Jabs, E.W., Wilcox, W.R., and Ahituv, N. (2014). A novel ZRS mutation leads to preaxial polydactyly type 2 in a heterozygous form and Werner Mesomelic syndrome in a homozygous form. Submitted.

- Visel, A., Minovitsky, S., Dubchak, I., and Pennacchio, L.A. (2007). VISTA Enhancer Browser—a database of tissue-specific human enhancers. *Nucleic Acids Res.* *35*, D88–D92.
- Visel, A., Akiyama, J.A., Shoukry, M., Afzal, V., Rubin, E.M., and Pennacchio, L.A. (2009a). Functional autonomy of distant-acting human enhancers. *Genomics* *93*, 509–513.
- Visel, A., Blow, M.J., Li, Z.R., Zhang, T., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F., et al. (2009b). ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* *457*, 854–U112.
- Vokes, S.A., Ji, H.K., Wong, W.H., and McMahon, A.P. (2008). A genome-scale analysis of the cis-regulatory circuitry underlying sonic hedgehog-mediated patterning of the mammalian limb. *Genes Dev.* *22*, 2651–2663.
- Wang, Z.B., Zang, C.Z., Rosenfeld, J.A., Schones, D.E., Barski, A., Cuddapah, S., Cui, K.R., Roh, T.Y., Peng, W.Q., Zhang, M.Q., et al. (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat. Genet.* *40*, 897–903.
- West, A.G., Gaszner, M., and Felsenfeld, G. (2002). Insulators: many functions, many mechanisms. *Genes Dev* *16*, 271–288.
- White, P.H., and Chapman, D.L. (2005). Dll1 is a downstream target of Tbx6 in the paraxial mesoderm. *Genesis* *42*, 193–202.
- Wieczorek, D., Pawlik, B., Li, Y., Akarsu, N.A., Caliebe, A., May, K.J.W., Schweiger, B., Vargas, F.R., Balci, S., Gillessen-Kaesbach, G., et al. (2010). A specific mutation in the distant Sonic Hedgehog (SHH) cis-regulator (ZRS) causes Werner mesomelic syndrome (WMS) while complete ZRS duplications underlie Haas type polysyndactyly and preaxial polydactyly (PPD) with or without triphalangeal thumb. *Hum Mutat* *31*, 81–89.
- Woolfe, A., Goodson, M., Goode, D.K., Snell, P., McEwen, G.K., Vavouri, T., Smith, S.F., North, P., Callaway, H., Kelly, K., et al. (2005). Highly Conserved Non-Coding Sequences Are Associated with Vertebrate Development. *PLoS Biol.* *3*, e7.
- Wu, L.Q., Liang, D.S., Niikawa, N., Ma, F., Sun, M., Pan, Q., Long, Z.G., Zhou, Z.M., Yoshiura, K., Wang, H., et al. (2009). A ZRS Duplication Causes Syndactyly Type IV With Tibial Hypoplasia. *Am J Med Genet* *149A*, 816–818.
- Xu, X., Weinstein, M., Li, C., Naski, M., Cohen, R.I., Ornitz, D.M., Leder, P., and Deng, C. (1998). Fibroblast growth factor receptor 2 (FGFR2)-mediated reciprocal regulation loop between FGF8 and FGF10 is essential for limb induction. *Development* *125*, 753–765.

Yano, T., and Tamura, K. (2013). The making of differences between fins and limbs. *J. Anat.* 222, 100–113.

Yonei-Tamura, S., Abe, G., Tanaka, Y., Anno, H., Noro, M., Ide, H., Aono, H., Kuraishi, R., Osumi, N., Kuratani, S., et al. (2008). Competent stripes for diverse positions of limbs/fins in gnathostome embryos. *Evol. Dev.* 10, 737–745.

Zguricas, J., Heus, H., Morales-Peralta, E., Breedveld, G., Kuyt, B., Mumcu, E.F., Bakker, W., Akarsu, N., Kay, S.P.J., Hovius, S.E.R., et al. (1999). Clinical and genetic studies on 12 preaxial polydactyly families and refinement of the localisation of the gene responsible to a 1.9 cM region on chromosome 7q36. *J Med Genet* 36, 33–40.

Zhao, J., Ding, J., Li, Y.Q., Ren, K.Q., Sha, J.H., Zhu, M.S., and Gao, X. (2009). HnRNP U mediates the long-range regulation of Shh expression during limb development. *Hum Mol Genet* 18, 3090–3097.

Van der Zwaag, P. a, Dijkhuizen, T., Gerssen-Schoorl, K.B.J., Colijn, A.W., Broens, P.M. a, Flapper, B.C.T., and van Ravenswaaij-Arts, C.M. a (2010). An interstitial duplication of chromosome 13q31.3q32.1 further delineates the critical region for postaxial polydactyly type A2. *Eur. J. Med. Genet.* 53, 45–49.

Appendix I

Zebrafish 'ancient enhancers' results

I tested multiple homologs of two different enhancers that had deep homology in protostomes and deuterostomes. These enhancers were given names to reflect the nearest gene. The Id gene is expressed in many embryonic tissues and is regulated by BMP and TGF-beta. There are multiple Id genes in humans and in zebrafish.

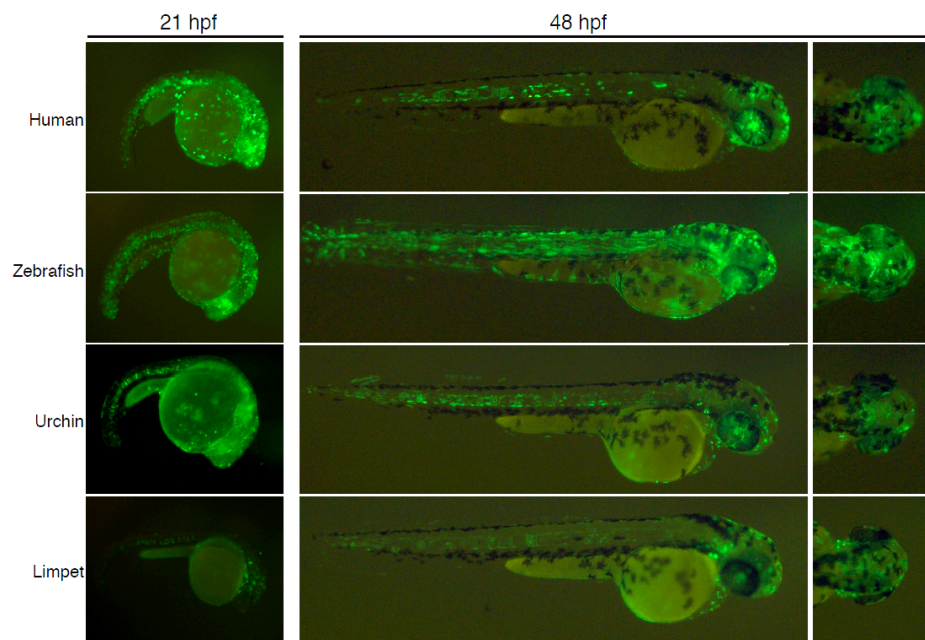


Figure S1. Expression of homologous versions of the Id enhancer in zebrafish at three developmental stages. Expression is strongest in the muscles of the trunk and the brain as seen in representative embryos.

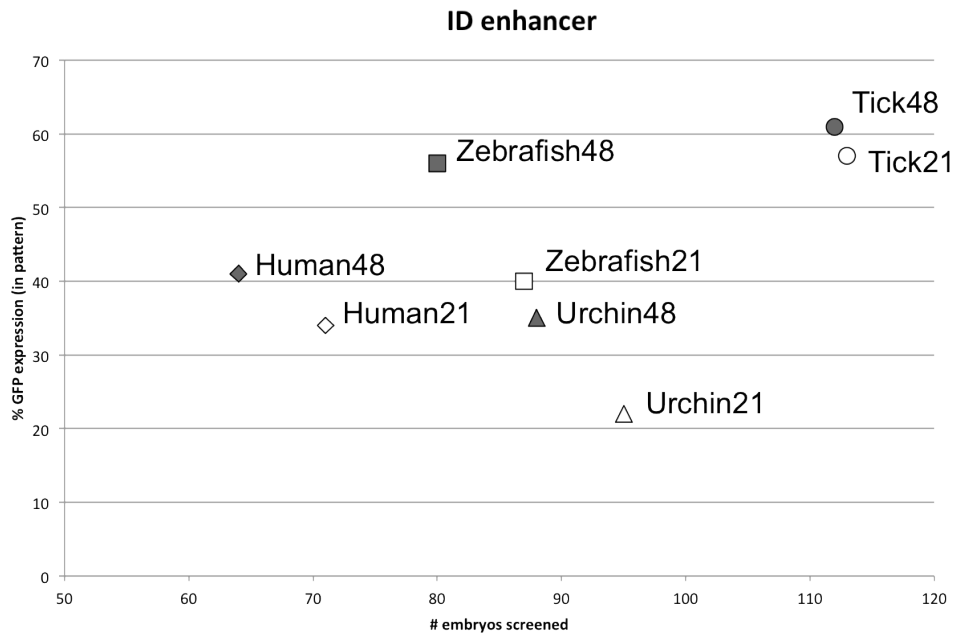


Figure S2. The percent of fish expressing GFP from the Id enhancer homologs increased over time. Expression is in the pattern seen in Figure S1 with expression in the brain and muscle.

The other enhancer is near the gene *Znf503*, a transcription factor expressed in multiple tissues of mouse embryo including brain and nervous system tissues.

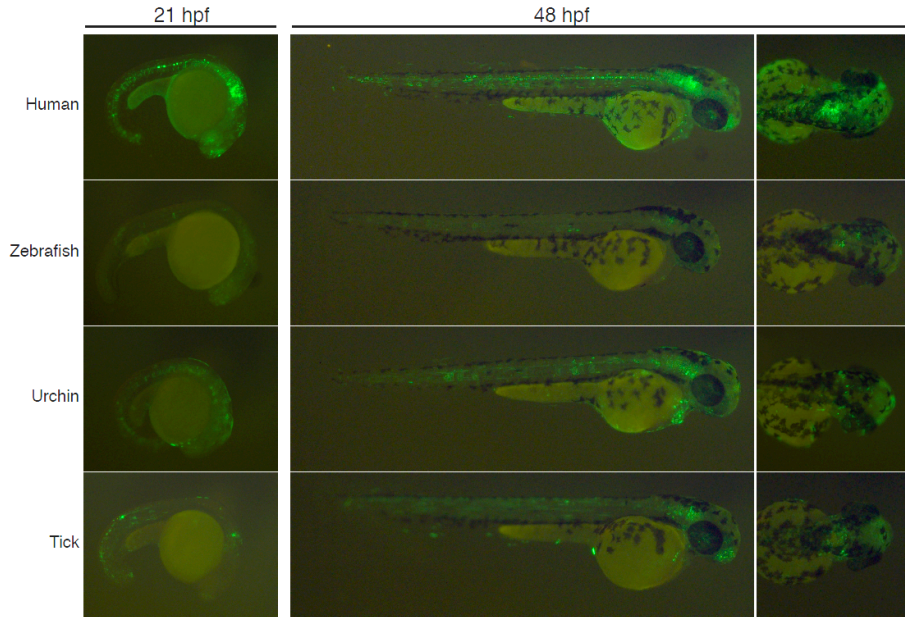


Figure S3. Expression of homologous versions of the Znf503 enhancer in zebrafish at three developmental stages. Expression is strongest in the hindbrain and spine at all time points as seen in representative embryos.

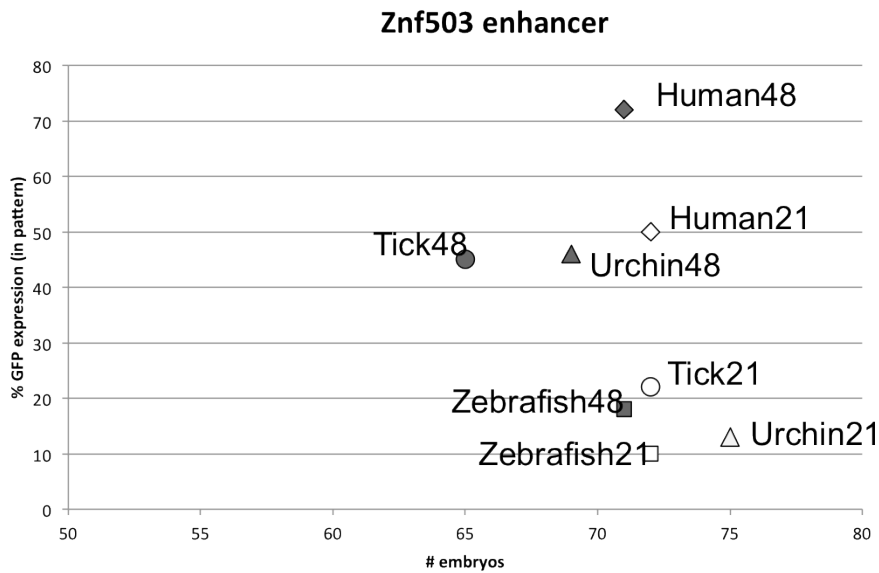


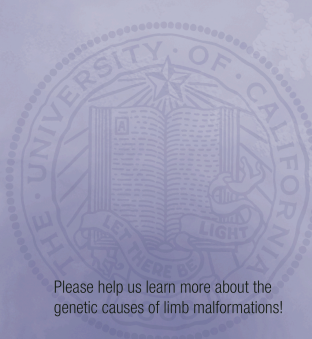
Figure S4. The percent of fish expressing GFP from the Id enhancer homologs increased over time. Expression is in the pattern seen in Figure S3 with expression in the brain and spine.

The reproducibility of the enhancer pattern from these distantly related species highlights the robustness of the zebrafish enhancer assay. This study also benefited from the ability to screen these enhancers over time in the zebrafish because the patterns of expression changed subtly as different tissues developed.


Patient recruitment brochure and poster

The study will:

- Help us learn more about the genetic causes of limb malformations
- Lead to improved patient counseling
- Aid in the development of genetic tests
- Increase our understanding of limb development and limb malformations
- Help inform individuals about how limb malformations are passed on to future generations




Please help us learn more about the genetic causes of limb malformations!




a study to identify

UCSF Limb Study

the genetic causes of limb malformations




If you are interested in participating in our study or have any questions, please contact:



Dr. Nadav Ahituv, PhD
 Institute for Human Genetics
 UCSF
 513 Parnassus Ave., Box 0794
 Health Sciences East, 901H
 San Francisco, CA 94143
 Tel.: 415 476 1838
 Fax: 415 502 0720
 Email: nadav.ahituv@ucsf.edu
 Web: <http://bts.ucsf.edu/ahituv/limb.html>


University of California
 San Francisco




Institute for Human Genetics

We are looking for individuals who have limb malformations such as:

- Syndactyly (fused or webbed fingers/toes)
- Polydactyly (> 5 fingers/toes)
- Split hand and foot
- Reduction anomalies (< 5 fingers/toes)
- Brachydactyly (short finger/toes)
- Clinodactyly (bent pinky fingers)
- Arachnodactyly (extra long fingers)
- Camptodactyly (bent little finger)



New advances in genetics have greatly increased our understanding of how mutations can cause isolated limb malformations – limb malformations that occur separately from any other clinical problems.



Many studies suggest that mutations responsible for isolated limb malformations can occur in regions that regulate genes that are important for limb development. These regions are called “enhancers” and act like light switches to activate genes at the right time and place. If there is a mutation in the enhancer, the gene is not turned “on” properly.

What is required of study participants?

- 1 Brief informed consent interview
- 2 Short questionnaire
- 3 DNA sample collected from blood or saliva

– We will provide a kit for saliva collection that can be returned to us in a prepaid envelope


Patient privacy is very important to us and all information and samples are kept confidential

Collection of saliva for a DNA sample is easy, painless and can be done in the participant’s home. Participants do not need to travel to UCSF.

- Club feet
- Congenital radioulnar synostosis
- Congenital radial head dislocation


Our research is focused on identifying enhancers that are important for limb development and identifying mutations in these enhancers that may be responsible for isolated limb malformations.

For adults



Collect saliva in the cup provided

For infants & young children



Use the sponges provided to collect saliva

Genes sequenced for different limb phenotypes

Before sequencing enhancers in patients with particular phenotypes, we established a list of candidate genes that should be sequenced to rule out coding mutations.

Polydactyly – SHH, ZRS (SHH enhancer), GLI3

Syndactyly – ZRS, HOXD13, GJA1, SHH, FBLN1

Clubfoot – TBX4, PITX1

Split hand/foot malformation – p63, BTRC, WNT10B, FBXW4

Subject phenotype questionnaires

Because the limb malformation phenotypes that we study often occur in combination with other phenotypes as a syndrome, we want to be sure that we get detailed phenotype information from subjects. Below is an example questionnaire for subjects with Split hand/foot malformation with notes on how specific phenotypes guide our search for enhancers that may harbor mutations.

Similar information is requested for all limb malformation phenotypes.

Split Hand/Food Malformation Phenotype information for USCF Limb Study

Understanding the detailed phenotypes of patients greatly helps us in identification of candidate enhancers, which may be mutated to cause their limb malformations. We appreciate your help and time in providing us with detailed phenotypes for both the limb malformation as well as other testing and non-limb phenotypes that can help us direct our research regarding specific patients. This page lists some of the specific information that will help us. There is another Microsoft Excel sheet that provides spaces to fill in the information for each patient. If there is other clinical information that you think is relevant, please include that as well.

Thank you.

Detail on each limb

- Oligodactyly, bidactyly, monodactyly?
 - o If monodactyly, is anterior or posterior element retained?
- Central cleft with syndactyly?
- Evidence of triphalangeal thumb?
- Polydactyly?
- Fingernail dysplasia?
- Long bone phenotype?

Other phenotypes

- Cleft lip or cleft palate?
- Dental abnormalities (tooth shape, oligodontia)?
- Dysplastic (misshapen) or low set ears?
- Deafness?
- Congenital heart defect?
- Seizures?
- Developmental delay/mental retardation?

Previous testing (molecular and other)

- Radiographic evaluation to determine TPT, syndactyly, identity of retained digits?
- Any family history/pedigree?
 - o Any tests for SHFM linkage?
- Sequencing of P63 gene?
 - o Other sequencing?
- Karyotype testing?
- Other tests for chromosome rearrangements or duplications?

Split Hand/Tooth Malformation Phenotype information for USCF Limb Study

How we can use this information

Detail on each limb

- Oligodactyly, bidactyly, monodactyly?
 - o If monodactyly, is anterior or posterior element retained?

Monodactyly retaining the posterior element is usually SHFM3 or SHFM4 while monodactyly retaining the anterior element is SHFM5 [1, 2].

- Central cleft with syndactyly, aka typical ectrodactyly?

Phenotype severity of hands compared to feet: feet>>hands suggests SHFM5, 1.

Feet>=hands suggests SHFM4, 3 [3].

- Evidence of triphalangeal thumb?
- Polydactyly?
- Fingernail dysplasia?

Anterior involvement (TPT, proximally placed thumb, [maybe fingernail dysplasia]) is associated with SHFM3 [1, 4]

- Long bone phenotype?

Other phenotypes

- Cleft lip or cleft palate?

Cleft lip suggests SHFM4, 1 or 5; cleft palate suggests SHFM5, maybe 1 or 4 [3].

- Dental abnormalities (tooth shape, oligodontia)?

Dental involvement would suggest SHFM4, 1, 5 [3].

- Dysplastic (misshapen) or low set ears?

Rules out SHFM4, suggests SHFM5, 1 [3].

- Deafness?

Suggests SHFM1. Rules out SHFM5.

- Congenital heart defect?

Suggests SHFM5, maybe 1. Not SHFM2, 4 [5].

- Seizures?

Suggests SHFM5, not at all in SHFM4[3].

- Developmental delay/mental retardation?

Related to karyotypic abnormalities [3]. Found with many SHFM1, SHFM5 patients.

*****SHFM2 patients have only limb-specific phenotypes. They show no craniofacial (clefting), MR, or ectodermal involvement [3].

Previous testing (molecular and other)

- Radiographic evaluation of limbs to determine TPT, syndactyly, identity of retained digits?

Cannot screen for TPT without radiographic exam [1].

- Any family history/pedigree?
 - o Any tests for SHFM linkage?
- Sequencing of P63 gene?
 - o Other sequencing?
- Karyotype testing?
- Other tests for chromosome rearrangements or duplications?

Should be obvious [2, 6-9].

1. Elliott, A.M., M.H. Reed, and J.A. Evans, *Triphalangeal thumb in association with split hand/foot: A phenotypic marker for SHFM3?* Birth Defects Research Part a-Clinical and Molecular Teratology, 2007. **79**(1): p. 58-61.
2. Basel, D., M.W. Kilpatrick, and P. Tsipouras, *The expanding panorama of split hand foot malformation*. American Journal of Medical Genetics Part A, 2006. **140A**(13): p. 1359-1365.
3. Elliott, A.M. and J.A. Evans, *Genotype-phenotype correlations in mapped split hand foot malformation (SHFM) patients*. American Journal of Medical Genetics Part A, 2006. **140A**(13): p. 1419-1427.
4. Elliott, A.M., et al., *Discrepancies in upper and lower limb patterning in split hand foot malformation*. Clinical Genetics, 2005. **68**(5): p. 408-423.
5. Elliott, A.M. and J.A. Evans, *The association of split hand foot malformation (SHFM) and congenital heart defects*. Birth Defects Research Part a-Clinical and Molecular Teratology, 2008. **82**(6): p. 425-434.
6. Faiyaz-Ul-Haque, M., et al., *Fine mapping of the X-linked split-hand/split-foot malformation (SHFM2) locus to a 5.1-Mb region on Xq26.3 and analysis of candidate genes*. Clinical Genetics, 2005. **67**(1): p. 93-97.
7. Gurnett, C.A., et al., *Evidence for an additional locus for split Hand/Foot malformation in chromosome region 8q21.11-q22-3*. American Journal of Medical Genetics Part A, 2006. **140A**(16): p. 1744-1748.
8. Lezirovitz, K., et al., *A novel locus for split-hand/foot malformation associated with tibial hemimelia (SHFLD syndrome) maps to chromosome region 17p13.1-17p13.3*. Human Genetics, 2008. **123**(6): p. 625-631.
9. Everman, D.B., et al., *Frequency of genomic rearrangements involving the SHFM3 locus at chromosome 10q24 in syndromic and non-syndromic split-hand/foot malformation*. American Journal of Medical Genetics Part A, 2006. **140A**(13): p. 1375-1383.

Appendix II

ZRS common variants and known mutations

Known variants in the ZRS and pZRS region (hg19 coordinates:) were excluded from candidate mutation analysis.

Variant	Coordinates (hg19)	Phenotype	Frequency
rs182395968	chr7:156,583,772		A: 99.908% (2176 / 2178); G: 0.092% (2 / 2179)
739 A>G	chr7:156,583,831	PPD, TPT	rare
rs140304968	chr7:156,583,907		C: 0.092% (2 / 2179); T: 99.908% (2176 / 2178)
621 C>G	chr7:156,583,949	PPD, TPT	rare
619 C>T	chr7:156,583,951	PPD, TPT	rare
603 ins13	chr7:156,583,967	PPD, TPT	rare
rs185289890	chr7:156,584,038		C: 0.138% (3 / 2179); T: 99.862% (2175 / 2178)
rs79564846	chr7:156,584,063		C: 99.954% (2177 / 2178); G: 0.046% (1 / 2179)
463 T>G	chr7:156,584,107	PPD, TPT	rare
404, multiple	chr7:156,584,166	PPD, TPT, WMS	rare
402 C>T	chr7:156,584,168	PPD, TPT	rare
396 C>T	chr7:156,584,174	PPD	rare
334 T>G	chr7:156,584,236	PPD	rare
329 T>C	chr7:156,584,241	PPD	rare
305 A>T	chr7:156,584,266	PPD	rare
297 G>A	chr7:156,584,273	PPD	rare
295 T>C	chr7:156,584,275	PPD, TPT	rare
284 C>A	chr7:156,584,283	PPD	rare
rs181332317	chr7:156,584,427		C: 0.092% (2 / 2179); T: 99.908% (2176 / 2178)
105 C>G	chr7:156,584,465	PPD, TPT	rare
rs10254391	chr7:156,584,567		C: 40.883% (1296 / 3170); G: 59.117% (1874 / 3170)
rs76511713	chr7:156,584,578		C: 0.305% (7 / 2296); T: 99.695% (2289 / 2296)
rs78890432	chr7:156,584,615		A: 99.173% (2279 / 2298); C: 0.827% (19 / 2298)
rs77295788	chr7:156,584,625		C: 0.918% (20 / 2178); T: 99.082% (2158 / 2178)
rs118161223	chr7:156,584,865		A: 99.782% (2293 / 2298); G: 0.218% (5 / 2298)
rs10229091	chr7:156,584,904		C: 48.698% (2113 / 4339); T: 51.302% (2226 / 4339)
rs143004276	chr7:156,584,934		C: 99.816% (2174 / 2178); T: 0.184% (4 / 2177)
rs148379729	chr7:156,585,040		C: 99.679% (2171 / 2178); T: 0.321% (7 / 2178)
rs78347935	chr7:156,585,137		A: 96.482% (2331 / 2416); G: 3.518% (85 / 2416)
rs6949624	chr7:156,585,184		C: 42.318% (1154 / 2727); G: 57.682% (1573 / 2727)

rs147094304	chr7:156,585,242		A: 99.908% (2176 / 2178); T: 0.092% (2 / 2179)
rs183965251	chr7:156,585,306		A: 0.138% (3 / 2179); G: 99.862% (2175 / 2178)
rs191686472	chr7:156,585,378		G: 0.138% (3 / 2179); T: 99.862% (2175 / 2178)

Human subjects sequenced for ZRS mutations

Many human DNA samples were screened for mutations in the ZRS region. The table below lists the samples with polydactyly, phenotype details and results of sequencing.

Sample	Phenotype	Family History	Category	Sequenced	Mutation
14	Extra finger on left hand attach to last finger.	Y	PAPA	n, phenotype	
15	I was born with 2 extra fingers, one on each hand. They both had fingernails.	Y	polydactyly	n, phenotype	
24	"Super digit" on both hands		polydactyly	n, phenotype	
66	Extra finger attached to outside of little finger on left hand		PAPA	n, phenotype	
86	Central polydactyly both hands. Curvature of both little fingers. Flexed left index finger. Left little toe webbed, right little toe laterally displaced.	Y	polydactyly	n, phenotype	
88	6 digits L and R. Metacarpals with "y" duplication. Web toe L stn digit.	Y	PPD	n, phenotype	
191	bilateral symmetrical triphalangeal thumbs and an extra hypoplastic radial thumb with normal number of phalanges		PPD, TPT	y	ZRS603ins13
192	bilateral symmetrical triphalangeal thumbs and an extra hypoplastic radial thumb with normal number of phalanges		PPD, TPT	y	ZRS603ins13
193	bilateral symmetrical triphalangeal thumbs and an extra hypoplastic radial thumb with normal number of phalanges		PPD, TPT	y	ZRS603ins13
292	Mirror polydactyly of the foot		PPD		
336	Rare Wassel 6 type thumb duplication with thumb hypoplasia, unilateral	N	PPD	y	
340	bilateral triphalangeal thumbs + duplication	Y	PPD, TPT	y	ZRS323 A>G
341	bilateral triphalangeal thumbs + duplication, one is associated to a Wassel Type 7 thumb duplication	Y	PPD, TPT	y	ZRS323 A>G
344	thumb duplication Wassel type 2 unilateral isolated	N	PPD	y	
345	triphalangeal thumb bilateral		TPT	y	
358	triphalangeal thumb bilateral		TPT	y	ZRS295 T>C
359	unilateral isolated (difficult) thumb duplication and triphalangism	N	PPD	y	
362	unilateral isolated thumb duplication, Wassel Type 2	N	PPD	y	
363	Polydactyly, bilateral in foot	Y	PPD	y	
365	Preaxial Polydactyly, bilateral	Y	PPD	y	

366	Polydactyly, bilateral	N	PPD	y	
370	Triphalangeal thumb family (affected subject), bilateral	Y	TPT	y	ZRS278 C>A
372	Triphalangeal thumb family (affected subject), bilateral	Y	TPT	y	ZRS278 C>A
387	Born with small extra finger off of right thumb. It was tied off when baby and now look like a small wart. Instead of 2 tendons running the length from the thumb to wrist, has only 1	Y (mother and maternal grandfather with RA)	PPD	n, phenotype	
413	bilateral postaxial polydactyly	Y; grandfather	PAPA	n, phenotype	
414	bilateral postaxial polydactyly	N	PAPA	n, phenotype	
415	bilateral postaxial polydactyly	N	PAPA	n, phenotype	
416	bilateral postaxial polydactyly	Y; polydactyly in mother's family	PAPA	n, phenotype	
417	Right hand preaxial polydactyly	N	PPD	y	
418	Right hand preaxial polydactyly	N	PPD	y	
419	Left hand preaxial polydactyly	N	PPD	y	
420	Right hand preaxial polydactyly	N	PPD	y	
421	Right hand preaxial polydactyly	Y; paternal grandmother	PPD	y	
422	bilateral hand preaxial polydactyly with triphalangeal thumb on left side	N	PPD, TPT	y	
423	Right hand preaxial polydactyly	N	PPD	y	
424	Right hand preaxial polydactyly	N	PPD	y	
425	Right hand preaxial polydactyly	N	PPD	y	
426	Right hand preaxial polydactyly	Y; maternal grandfather	PPD	y	
427	Right hand preaxial polydactyly	N	PPD	y	
428	Left hand preaxial polydactyly	N	PPD	y	
429	Right hand preaxial polydactyly	N	PPD	y	
430	Left hand preaxial polydactyly	N	PPD	y	
431	Right hand preaxial polydactyly	N	PPD	y	
466	Polydactyly	Y	polydactyly	y	ZRS278 C>A
469	Polydactyly	Y	polydactyly	y	ZRS278 C>A
471	Polydactyly	Y	polydactyly	y	ZRS278 C>A
481	Polydactyly	N	polydactyly		
505	Polydactyly, mirror images of pinkie digit on both hands and feet	Y, Mother and Uncle	PAPA	n, phenotype	
508	Pre-axial Polydactyly RH	Y, brothers	PPD		
530	Postaxial Polydactyly (LH)		PAPA	n, phenotype	
533	Postaxial Polydactyly, bilateral	Y, cousin w poly	PAPA	n, phenotype	
534	Preaxial Polydactyly, RH		PPD		
535	Postaxial Polydactyly, bilateral		PAPA	n, phenotype	
536	Postaxial Polydactyly, bilateral	N	PAPA	n, phenotype	

547	Bilateral Polydactyly		polydactyly	n, phenotype	
565	polydactyly, extra digit on each foot, as well as syndactyly, 3 toes on each	Y; maternal aunt	polydactyly	n, phenotype	
572	Polydactyly, extra thumb LH	Y; great uncle, paternal	PPD	n, phenotype	
576	Polydactyly, extra toe on top of L pinky toe		PPD	n, phenotype	
579	Polydactyly, Extra right digit	Y; paternal great aunt, second cousin, third cousin	PAPA	n, phenotype	
592	Polydactyly, extra toe on LF, syndactyly as well		polydactyly	n, phenotype	
595	Polydactyly, Extra toe on RF		polydactyly	n, phenotype	
603	Polydactyly, 2 thumbs, RH	Y; uncle, cousin	PPD	n, phenotype	
623	Bilateral Extra bone causing bend thumbs	Y	TPT	n, phenotype	
625	LH: extra bone causing bent thumb RH: Polydactyly, extra finger between 1/2 digit	Y	TPT/PPD		
628	LH: webbed extra thumb RH: extra thumb, LF: webbed extra toe RF: extra toe	Y	PPD	n, phenotype	
640	Polydactyly	N	polydactyly	poor DNA	
642	Bilateral Polydactyly hands, Synpoly, feet	N	polydactyly	poor DNA	
678	Bilateral Polydactyly, feet	Y; great uncle, etc.	polydactyly	poor DNA	
679	Bilateral Polydactyly, hands and feet	Y; brother, father, uncle	polydactyly	n, part of pedigree	
683	Crossed Type Polydactyly	Y	polydactyly	y	
685	Preaxial Polydactyly	N	PPD	y	
686	Postaxial Polydactyly	N	PAPA	y	
687	Preaxial Polydactyly	N	PPD	poor DNA	
688	Postaxial Polydactyly	N	PAPA	poor DNA	
689	Preaxial Polydactyly	N	PPD	poor DNA	
690	Preaxial Polydactyly	N	PPD	y	
691	Postaxial Polydactyly	N	PAPA	y	
692	Preaxial Polydactyly	N	PPD	y	
693	Preaxial Polydactyly	N	PPD	y	
694	Postaxial Polydactyly	N	PAPA	y	
695	Preaxial Polydactyly	N	PPD	poor DNA	
696	Postaxial Polydactyly	N	PAPA	y	
697	Preaxial Polydactyly	N	PPD	no, other pheno	
698	Postaxial Polydactyly	N	PAPA	poor DNA	
699	Postaxial Polydactyly	Y	PAPA		
701	Preaxial Polydactyly	Y	PPD	y	
703	Postaxial Polydactyly	N	PAPA	y	
704	Preaxial Polydactyly	Y	PPD		

705	Postaxial Polydactyly	N	PAPA	y	
706	Postaxial Polydactyly	Y	PAPA	y	
707	Postaxial Polydactyly	Y	PAPA	poor DNA	
708	Preaxial Polydactyly	N	PPD	y	
709	Preaxial Polydactyly	N	PPD	y	
711	Postaxial Polydactyly	Y	polydactyly	y	
712	Postaxial Polydactyly		polydactyly		
713	Preaxial Polydactyly	N	PPD	y	
715	Preaxial Polydactyly	N	PPD	y	
720	Postaxial Polydactyly	N	polydactyly		
722	Preaxial Polydactyly - hands	N	PPD	y	
724	Postaxial Polydactyly	Y	polydactyly	y	
725	Postaxial Polydactyly	N	polydactyly	no, other pheno	
726	Preaxial Polydactyly	N	PPD	y	
728	Postaxial Polydactyly		polydactyly	y	
730	Preaxial Polydactyly	N	PPD	y	
731	Postaxial Polydactyly	Y	polydactyly	y	
733	Preaxial Polydactyly	N	PPD	y	
735	Postaxial Polydactyly	Y	polydactyly	y	
737	Bilateral Extra Thumb Bone, LH: webbing digit 3/4, Bilateral Polydactyly (feet)	Y	TPT	n, phenotype	
741	Bilateral Polydactyly, hands		polydactyly		
790	All 4 limbs- Postaxial Polydactyly	N	PAPA	n, phenotype	
792	Duplicate left thumb	N	PPD	n, phenotype	
796	Accessory Right Thumb	N	PPD		
797	Duplicated Left little toe	N	PAPA	n, phenotype	
857	Double hypoplastic left thumb	N	PPD		
859	LF: extra toe	N	polydactyly	n, phenotype	
861	LF: extra toe	N	polydactyly	n, phenotype	
866	Bilateral Polydactyly: extra digits 1/5	N	polydactyly	n, phenotype	
870	RF: extra digit	N	polydactyly		
872	RF: extra toe	N	polydactyly		
876	Bilateral Extra Toe	N	polydactyly		
878	Bilateral Extra digit, Hands and Feet	N	polydactyly		
880	Extra little toe	N	polydactyly		
884	Duplicate Right thumb	N	PPD		
886	Right Hand Postaxial Polydactyly	Y; feet syndactyly on maternal side, maternal uncle	PAPA	n, phenotype	
889	Bilateral Polydactyly	Y; maternal side	polydactyly		

Supplementary data for ZRS mutation articles

ZRS603ins13

Supplementary Methods

Human DNA sequencing

Six family members were available for testing. Written informed consent was obtained from all participants or their legal guardians. The proband, her mother, father and sister were available for clinical examination. Genomic DNA was extracted from peripheral blood using standard methods (Qiagen kit EZ1). Primer sequences used for PCR and sequencing of the ZRS were obtained from Lettice et al. (Lettice, et al., 2003). Sequencing was performed using Big Dye terminator v3.1 cycle sequencing kit (Applied Biosystems) according to standard procedure and analyzed with the ABI-3730 DNA Analyzer (Applied Biosystems). The sequences were analyzed using Seqscape v2.5 (Applied Biosystems) and Sequencher (Gene Codes Corporation). The study was approved by the Regional Ethical Review Board in Stockholm, Sweden (protocol number 2010/193032) and the UCSF Committee on Human Research (protocol 10-03111).

Transcription Factor Binding Site Predictions

We acquired 345 transcription factor motifs from the UniPROBE database (Newburger and Bulyk, 2009) and compiled a small number of motifs for additional known limb factors from the literature including *TBX5* (MIM *601602) (Ghosh, et al., 2001), *TBX6* (MIM *602427) (White and Chapman, 2005), the *PBX-MEIS1* complex (MIM *602100 and MIM *601739) (Chang, et al., 1997) and the *MEIS1-HOXA9* (MIM *142956) complex (Shen, et al., 1997). Binding sites were predicted by scanning DNA sequences for motifs using our implementation of the MATCH algorithm (Kel, et al., 2003) with a score threshold of 0.8.

Mouse Transgenic Analysis

The complete ZRS region was PCR amplified from patient II/1 with primers carrying XhoI and ApaI restriction sites (FwPrimer: GGCCctcgagTTTCAAATGCTCACTTTACATGG, RevPrimer: ATgggcccTGCTGAAGTGATACTGAAGAGAGG) and cloned into the *Hsp68-LacZ* enhancer assay vector (Kothary, et al., 1989). It was sequence-verified (Quintara Biosciences) to make sure that it had the ZRS603_604ins13 mutation and that no other mutations existed in the sequence. The wildtype enhancer assay vector was generated by removing the insertion from the ZRS603_604ins13 - Hsp68-LacZ vector (Mutagenix, Hillsborough, NJ, USA) and sequence verified. Generation of transgenic mice and Beta galactosidase staining at embryonic (E) day 11.5 were done by Cyagen Biosciences, Inc. (Guangzhou, China). All animal work was approved by the UCSF Institutional Animal Care and Use Committee.

Supplementary Methods References

Chang CP, Jacobs Y, Nakamura T, Jenkins NA, Copeland NG, Cleary ML. 1997. Meis proteins are major in vivo DNA binding partners for wild-type but not chimeric Pbx proteins. *Mol Cell Biol* 17:5679-5687.

Ghosh TK, Packham EA, Bonser AJ, Robinson TE, Cross SJ, Brook JD. 2001. Characterization of the TBX5 binding site and analysis of mutations that cause Holt-Oram syndrome. *Hum Mol Genet* 10:1983-1994.

Kel AE, Gobling E, Reuter I, Cheremushkin E, Kel-Margoulis OV, Wingender E. 2003. MATCH: a tool for searching transcription factor binding sites in DNA sequences. *Nucleic Acids Res* 31:3576-3579.

Kothary R, Clapoff S, Darling S, Perry MD, Moran LA, Rossant J. 1989. Inducible expression of an hsp68-lacZ hybrid gene in transgenic mice. *Development* 105:707-714.

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

Newburger DE, Bulyk ML. 2009. UniPROBE: an online database of protein binding microarray data on protein-DNA interactions. *Nucleic Acids Res* 37:D77-D82.

Shen WF, Montgomery JC, Rozenfeld S, Moskow JJ, Lawrence HJ, Buchberg AM, Largman C. 1997. AbdB-like Hox

proteins stabilize DNA binding by the Meis1 homeodomain proteins. *Mol Cell Biol* 17:6448-6458.

White PH, Chapman DL. 2005. Dll1 is a downstream target of Tbx6 in the paraxial mesoderm. *genesis* 42:193-202.

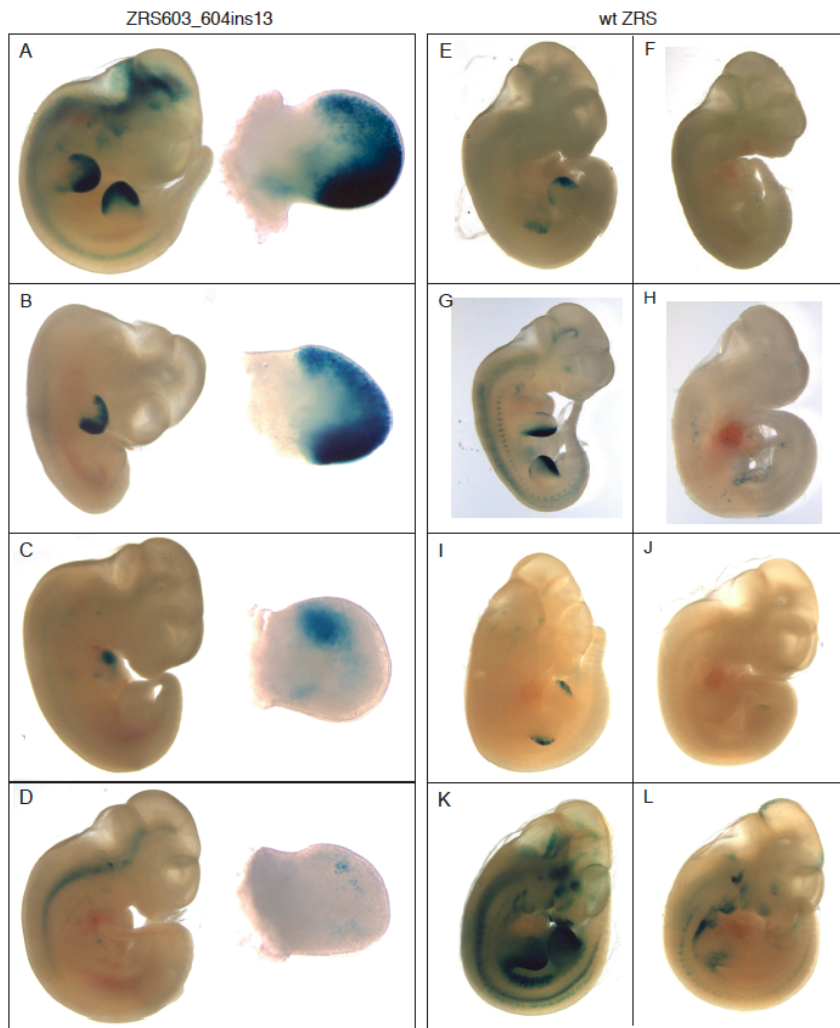


Figure S5. Whole-mount and dissected forelimbs show the anterior limb LacZ expression driven by ZRS603_604ins13 in multiple transgenic E11.5 mice compared to predominantly posterior expression driven by wtZRS. (A) This embryo has strong expression in the anterior as well as posterior of both forelimbs and hindlimbs. (B) This embryo shows strong posterior and anterior limb expression in both forelimbs and weak expression in hindlimbs, which is difficult to see with the twisted orientation of the embryo. (C) This embryo has strong staining only in the forelimbs with most expression in the anterior part of the limb. (D) The fourth embryo has weak expression only in the forelimbs, but does show anterior limb expression. (E-L) Seven of the eight (all except L) wtZRS transgenic embryos show some posterior limb expression of LacZ. E and G also show a small degree of anterior limb LacZ expression, but this is much weaker than

the posterior expression, in contrast to all four of the mutant ZRS603_604ins13 transgenic embryos (A-D).

ZRS 402 C>T

Supplementary Methods

Unaffected Population Genotyping

A restriction digest assay was used to genotype 59 unrelated Mexican individuals with no congenital malformations. The C to T change at position 402 prevents the enzyme RsaI from cutting the DNA. A 472 base region encompassing the mutation and a second, invariable, RsaI restriction site was amplified by PCR with primers Hi5f2 and Hi5R2. After digestion with RsaI, the wildtype sequence was digested into three bands (84bp, 170bp and 218bp) while the mutant sequence digests to two bands (84bp and 388bp). See supplementary figure S2.

Supplementary Results

Full Clinical Report: Family A

We performed clinical examinations and collected DNA samples from a family of 3 affected members. The proband is an 8 year old female who at birth was noted to have mesomelic shortening of upper and lower extremities due to hypoplastic tibias and hypoplastic radio-ulnar bones, bilateral preaxial polydactyly of toes with 2 extra toes in both sides and bilateral proximally placed hypoplastic thumbs. She is able to ambulate, achieved milestones at appropriate ages and is otherwise healthy with normal cognition. Her extra toes were removed and her thumbs were repositioned.

At age 8, the proband is 103.5cm in height (50th percentile for a 5 year old female). She is significantly short with mesomelic shortening of the limbs, especially in the lower extremities

and has limited extension of the elbows as well as limited wrist supination and pronation. She is microcephalic with normal facial appearance and no obvious dysmorphic features (Head circumference is 48.3 cm, <3-SD). Her chest is normal in size and symmetry. She has normal curvature of spine. Her hands have normal second to the fifth finger. She has rudimentary thumbs, which originate from the base of the second finger; they are shortened and laterally deviated. Her thighs are of normal length, but the medial segment of lower extremities is short and bowed with bilateral dimples. There is a mild limitation of extension of her knees. Her feet are shortened and have five remaining toes. She has a gap with a surgical scar between the first and second toe. Deep tendon reflexes were normal. Her x-rays show short triphalangeal thumbs, mild shortening and fusion of radio-ulnar bones, fusion of metacarpal bones and hypoplastic tibias with bowing of the fibula bilaterally.

Both parents have abnormal thumbs. The mother has digitalization of thumbs with an extra crease in the distal interphalangeal space. Her thumbs are medially deviated. Radiography reveals bilateral elongated distal phalanx of first finger; she does not have three phalanges. Her feet are normal and lower extremity x-rays show no abnormalities. The father has bilateral triphalangeal thumbs and a preaxial remnant with a small bone that he is able to manipulate. On his right palm he has a transverse palmar crease and on the left bridged transverse palmar crease. Radiography showed a small proximal extra digit and triphalangeal thumbs in both sides. His feet are normal. The proband has a 5 year old sibling reported to have digitalization of thumbs who was not available for examination.

Extensive family history reveals multiple family members in both sides of the family with digitalization of thumbs and triphalangeal thumb with and without polydactyly, but no other member with shortening of the limbs. A pedigree of 6 generations and 128 members was constructed from the reported family history. This family has 31 members with isolated triphalangeal thumbs, 14 members with preaxial polydactyly and triphalangeal thumbs and the single proband with preaxial polydactyly and long bone involvement. The proband's parents deny consanguinity, however, both paternal and maternal grandparents were from the same small village in Mexico.

Full Clinical Report: Family B

The proband (V/14) has clinical abnormalities that are limited to the thumbs which are triphalangeal with a delta-shaped middle phalanx of the thumb bilaterally. The result is an ulnar deviation of the terminal two phalanges of the thumb and the first interphalangeal joint. The positioning of the thumb is such that opposition of the long finger appears to be possible.

Two older siblings (V/9 and V/11) were previously seen and have congenital triphalangeal thumbs associated with delta phalanx. Another sibling (V/10) was examined because of laxity of the thumb at the metacarpophalangeal joint. This is present on both sides and x-rays reveal that there are only two phalanges in the thumbs. This subject can rotate at the wrists normally.

Three of four children of these parents have triphalangeal thumbs, and yet the father, who is the parent with the positive family history does not have any thumb anomalies. He does have radioulnar synostosis on one side. This does not interfere with his functioning. In discussion, he

mentioned that both the proband's father and paternal grandmother have an extra carpal bone on the radial side of the carpus which is at the greater multangular level. Although this is described as a normal variant in the radiologic literature, the question is raised whether this could be a minor manifestation of the triphalangeal thumb gene.

Supplementary Figures

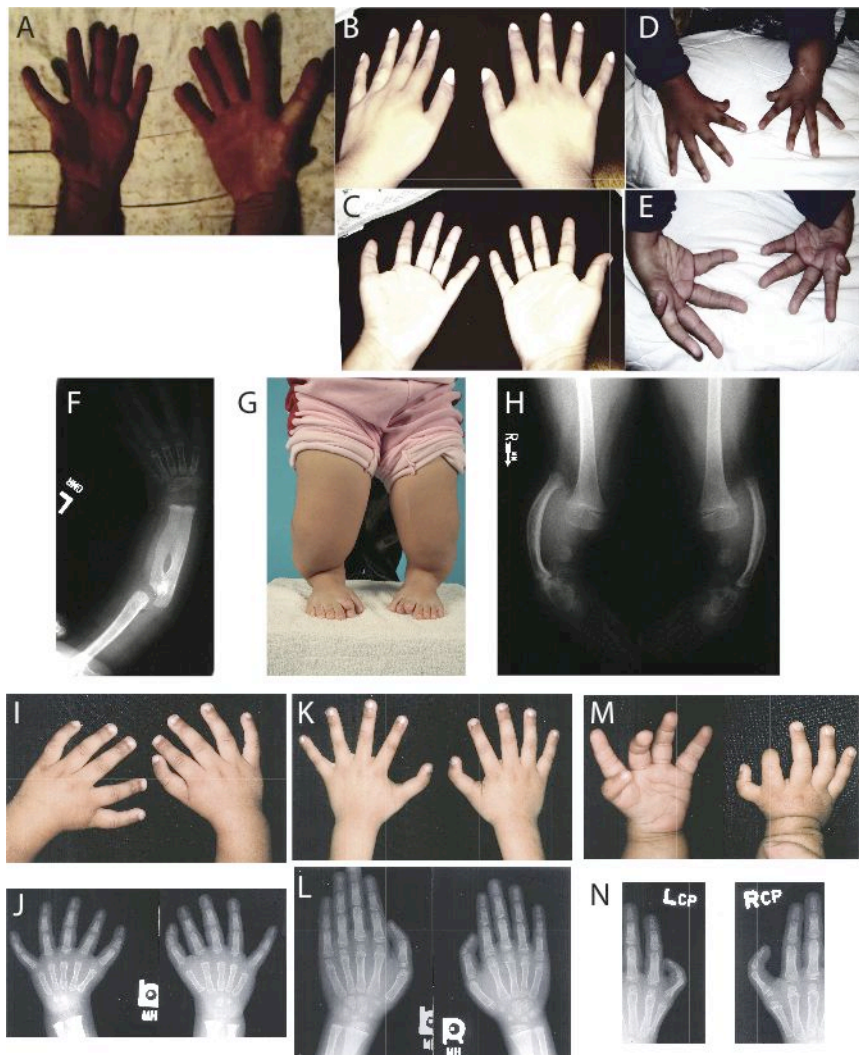


Figure S6. Additional images of phenotypes in Families A (A-H) and B (I-O). (A) The proband's father has bilateral triphalangeal thumbs and a preaxial remnant with a small bone that he is able to manipulate. (B,C) The proband's mother has digitalization of thumbs with an extra crease in the distal interphalangeal space and bilateral elongated distal phalanx of first fingers.

(D-H) After surgical removal of extra digits, the proband has rudimentary thumbs, which originate from the base of the second finger; they are shortened and laterally deviated. She has mild shortening and fusion of radio-ulnar bones with fusion of metacarpal bones and hypoplastic tibias with bowing of the fibula bilaterally (I,J) In Family B, V/9 (K,L) V/10 and (M,N) V/11 have bilateral triphalangeal thumbs.

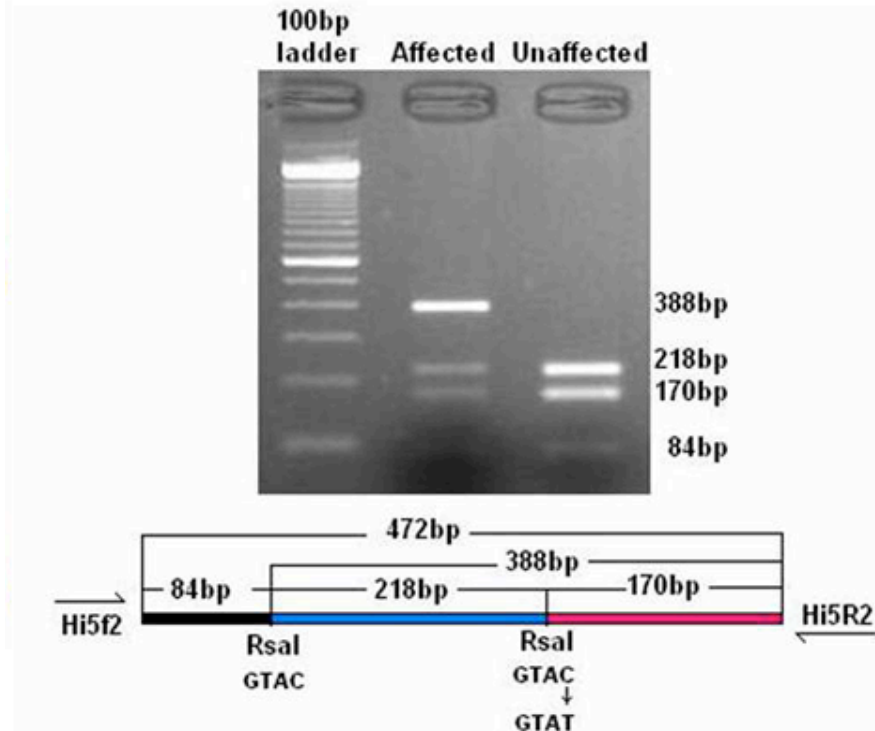


Figure S7. A restriction digest assay was used to genotype unrelated control individuals. The gel shows representative results for affected (*C/T* heterozygous) and unaffected (*C/C* homozygous wildtype) samples and a diagram indicating the restriction sites in the amplified region.

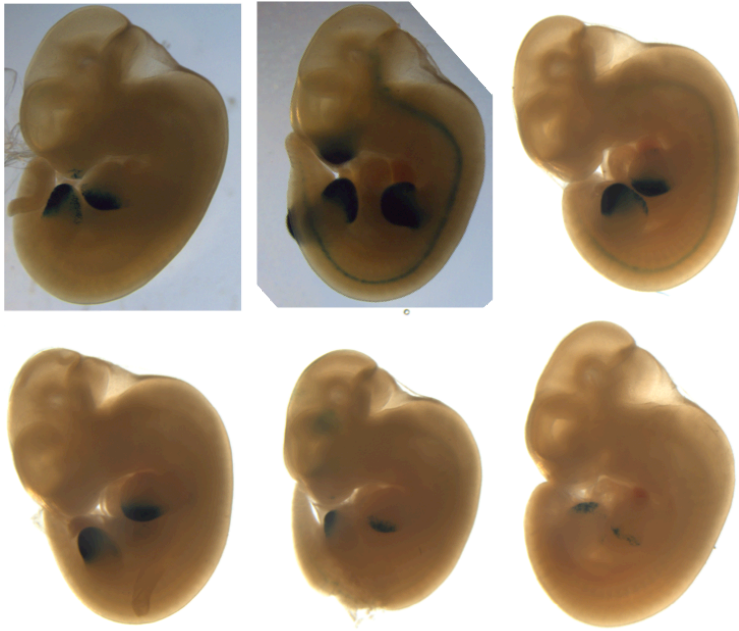


Figure S8. All six transgenic embryos show LacZ staining in the limb buds. In all embryos, there is expression of LacZ that extends out of the normal posterior ZPA into the anterior limb bud.

Appendix III

Supplemental Methods

Visualizing ChIP-seq data

To produce coverage pileups, AER and ZPA reads were extended to the predicted fragment size of 300bp using the slopBed utility in bedtools. A BED file encompassing the mm9 genome in 100bp intervals was generated, and overlapping AER, ZPA and input reads were counted in each interval using the bedtools utility coverageBed. The counts were then normalized by dividing by the total number of aligned reads for the respective sample, after which the normalized input signal was subtracted off the normalized AER and ZPA signals. The AER and ZPA signals were then scaled by a factor of 80,925,525 (the # of input reads) to facilitate viewing.

Table S1. Primers for mouse transgenic assays

Peak	Coordinates (mm9)	PrimerF	PrimerR	Nearest gene (location)
aer203	chr1:122510000-122512399	GTAGGAGGAGGCTTCATTGC	CTCCAGCTCGCCATACACTA	En1 (+12136)
aer2292	chr7:136794000-136796399	CCAAGAGCATATCCCATCT	CAGTGATAGACAGCCCTGGA	Fgfr2 (+615122)
aer3723	chr12:119942200-119943599	AGAGAGAGAGAGCACCTCATCAA	GACATCTGGGCTCACTCCT	Sp8 (-141902)
aer308	chr1:170693400-170694599	GGCCGACCTGTTACTTGITT	GGGGATAGCAGGCTACTCAG	Pbx1 (-331611)
aer4460	chr16:62861400-62862199	TGCCATTTTATTGCAAGTCAG	TCAGCAGATTATTCGTAAGTGA	Arl13b (+15345)
zpa6706	chr8:60898600-60899199	CATGGTTATTTGTCCACATTTT	GCAATGCTTTTCTTTTGTGG	Hand2
zpa12651	chr16:49332600-49333199	TTTAAAAATATTGAAAAGGTGGCTA	GCAGTAGACCTTAGGGGAGCA	Irf57 (-366507)
zpa1563	chr2:93377800-93378599	GCTCGGTACTTGCTCAGGAC	CTCTCTCCCAAATCCCTTC	Alx4 (-104391)
zpa2794	chr3:130665200-130665799	TACCAAGCACTTACTACAGGCTACA	ATAGGGGCGGTCCTTTTATT	Lef1 (-147889)
zpa44	chr1:19231200-19231999	CCTGTGGTCTTCTCCTTTG	CATCCTGTCACCAACCCTCT	Tcfap2b (+29465)

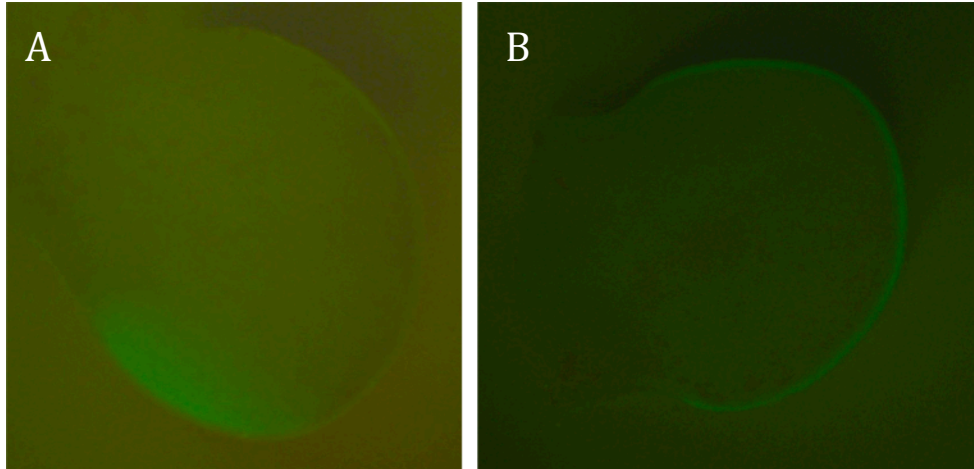


Figure S1. ZPA- and AER-specific GFP expression. (A) A limb from E11.5 $Shh^{tm1(EGFP/cre)Cjt}$ mice showing ZPA specific expression. (B) An E11.5 limb from a homozygous male mice carrying the Cre transgene ($tg(Msx2-cre)$ crossed to a floxed reporter ($B6.129-GT(ROSA)26Sor^{tm1,Joe}/J$) showing specific GFP expression in the AER.

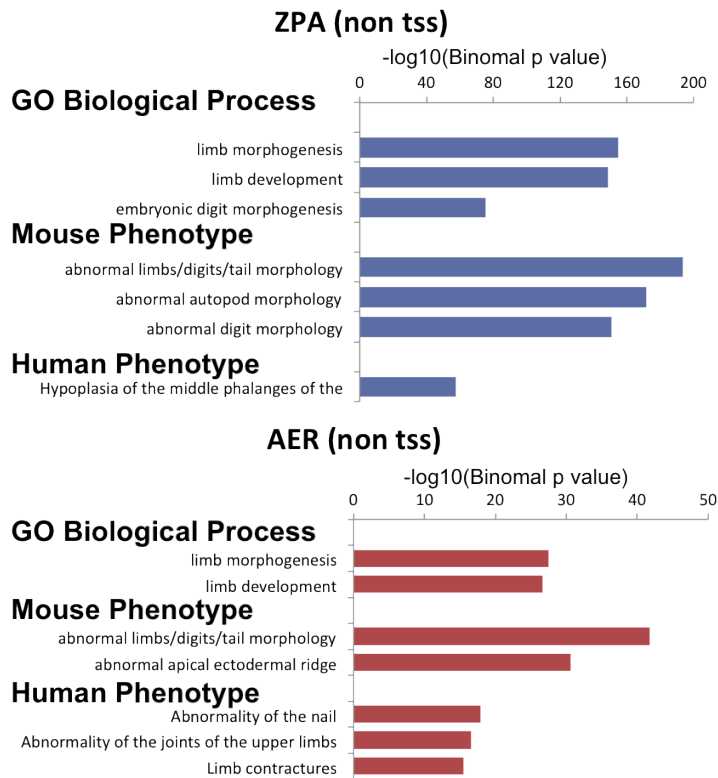


Figure S2. GREAT results for non-TSS peaks from the ZPA and AER ChIP-seq datasets show GO terms and phenotypes associated with the roles of the ZPA and AER in limb development.

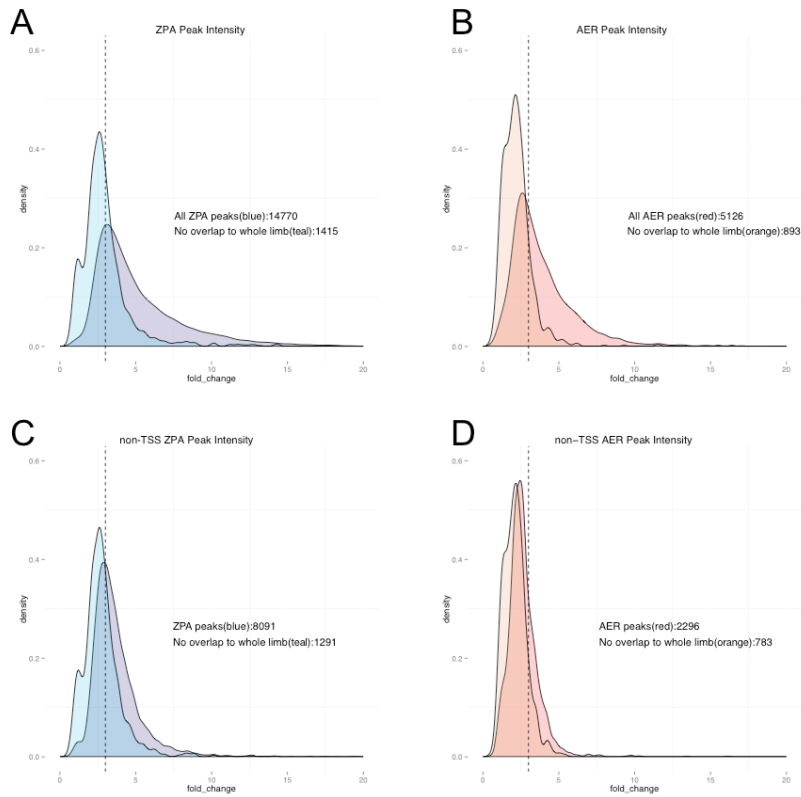


Figure S3. Peak height and signaling center specificity. The distribution of peak signal intensity shows that the ZPA and AER-specific peaks have a weaker signal in general compared to the entire ZPA and AER sets. (A) The distribution of all ZPA peaks (blue) compared to ZPA peaks with no overlap to whole limb (teal) shows that the peaks with no overlap are generally weaker. (B) The distribution of all AER peaks (red) compared to AER peaks with no overlap to whole limb (orange) shows that the peaks with no overlap are generally weaker. (C,D) The same trend is seen when looking only at non-TSS peaks from either data set. This suggests that some of the very strongest enhancer signals are the ones that overlap with whole-limb signal.

Publishing Agreement

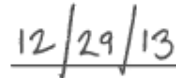
It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

Please sign the following statement:

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.



Author Signature



Date