Research

Distinct global shifts in genomic binding profiles of limb malformation-associated *HOXD13* mutations

Daniel M. Ibrahim,[1,2,3] Peter Hansen,[1,4] Christian Rödelsperger,[1,6] Asita C. Stiege,[2] Sandra C. Doelken,[4] Denise Horn,[4] Marten Jäger,[1,4] Catrin Janetzki,[1] Peter Krawitz,[4] Gundula Leschik,[1] Florian Wagner,[5] Till Scheuer,[1,2] Mareen Schmidt-von Kegler,[1] Petra Seemann,[1,3] Bernd Timmermann,[2] Peter N. Robinson,[1,2,4] Stefan Mundlos,[1,2,3,4] and Jochen Hecht[1,2,7]

[1]Berlin-Brandenburg Center for Regenerative Therapies (BCRT), Charité Universitätsmedizin Berlin, 13353 Berlin, Germany; [2]Max Planck Institute for Molecular Genetics, 14195 Berlin, Germany; [3]Berlin-Brandenburg School for Regenerative Therapies (BSRT), Charité Universitätsmedizin Berlin, 13353 Berlin, Germany; [4]Institute for Medical and Human Genetics, Charité Universitätsmedizin Berlin, 13353 Berlin, Germany; [5]Atlas Biolabs GmbH, 10117 Berlin, Germany

Gene regulation by transcription factors (TFs) determines developmental programs and cell identity. Consequently, mutations in TFs can lead to dramatic phenotypes in humans by disrupting gene regulation. To date, the molecular mechanisms that actually cause these phenotypes have been difficult to address experimentally. ChIP-seq, which couples chromatin immunoprecipitation with high-throughput sequencing, allows TF function to be investigated on a genome-wide scale, enabling new approaches for the investigation of gene regulation. Here, we present the application of ChIP-seq to explore the effect of missense mutations in TFs on their genome-wide binding profile. Using a retroviral expression system in chicken mesenchymal stem cells, we elucidated the mechanism underlying a novel missense mutation in *HOXD13* (Q317K) associated with a complex hand and foot malformation phenotype. The mutated glutamine (Q) is conserved in most homeodomains, a notable exception being bicoid-type homeodomains that have lysine (K) at this position. Our results show that the mutation results in a shift in the binding profile of the mutant toward a bicoid/PITX1 motif. Gene expression analysis and functional assays using in vivo overexpression studies confirm that the mutation results in a partial conversion of HOXD13 into a TF with bicoid/PITX1 properties. A similar shift was not observed with another mutation, Q317R, which is associated with brachysyndactyly, suggesting that the bicoid/PITX1-shift observed for Q317K might be related to the severe clinical phenotype. The methodology described can be used to investigate a wide spectrum of TFs and mutations that have not previously been amenable to ChIP-seq experiments.

[Supplemental material is available for this article.]

Transcription factors (TFs) determine cell type-specific gene expression and thereby govern developmental programs at all stages of development. TF binding to *cis*-regulatory elements regulates the tissue- and developmental stage-appropriate expression of target genes, and correspondingly, mutations in genes encoding TFs can lead to dramatic phenotypes in humans by disrupting developmental gene regulation programs. For instance, mutations in the TFs *HOXD13*, *RUNX2*, and *PITX2* are associated with brachydactyly, cleidocranial dysplasia, and Axenfeld-Rieger syndrome (Kozlowski and Walter 2000; Otto et al. 2002; Johnson et al. 2003). However, the molecular mechanisms that cause these developmental phenotypes have been difficult to address experimentally, particularly for mutations that do not result in a simple loss of function. In these cases the phenotypes can be mutation specific, as it has been reported for, e.g., *MSX2*, *FOXC1*, *PITX2*, or *HOXD13* (Kozlowski and Walter 2000; Wilkie et al. 2000; Johnson et al. 2003; Saleem et al. 2003). How these mutations exert their effect remains unclear, but it is likely that the mutant protein has

altered binding specificity or affinity, resulting in abnormal targets or activity (Caronia et al. 2003; Saleem et al. 2003; Zhao et al. 2007).

The development of ChIP-seq technology, which couples chromatin immunoprecipitation (ChIP) with high-throughput sequencing (seq), has enabled researchers to investigate TF binding on a genome-wide scale, opening new approaches to exploring transcriptional control mechanisms (Valouev et al. 2008; Barski and Zhao 2009; Park 2009). To date, however, ChIP-seq has largely been used to investigate the binding of wild-type TFs in cell lines, and also in tissue samples where relatively large amounts of cells are readily available (Robertson et al. 2008; Ouyang et al. 2009; Welboren et al. 2009; Wang et al. 2011; Junion et al. 2012). In order to harness ChIP-seq methodologies to investigate mutant TFs involved in hereditary diseases, a number of technical hurdles must be overcome.

Current ChIP-seq methodologies are not well suited to investigating the effects of mutations in TFs, because antibodies

typically do not distinguish between wild-type and mutant protein. Additionally, TFs often belong to gene families that consist of several close paralogs that in many cases are coexpressed. A further hindrance for analysis of TFs in developmental processes is that they are active during specific developmental time points that are difficult to recapitulate in cell culture systems and for which only very small amounts of tissue would be available in mouse models. This makes investigation of such factors essentially impossible with current ChIP-seq protocols that are based on antibodies specific to the protein of interest and that use cell culture or primary tissue samples. These difficulties apply especially to *Hox* genes, a family of highly conserved TFs. Arguably the most studied TFs in limb development are the 5′ *Hoxa* and *Hoxd* (*Hoxa/d9-13*) genes that pattern the proximo-distal axis and regulate growth and shape of the digits. Despite this, little is known about the targets of HOX proteins in limbs (Svingen and Tonissen 2006; Salsi et al. 2008; Villavicencio-Lorini et al. 2010). Genetic analyses have demonstrated that they are involved in limb patterning in the early phase of limb bud growth and later directly influence the size and shape of the skeletal elements (Muragaki et al. 1996; Zakany et al. 2004; Kuss et al. 2009; Delpretti et al. 2012).

Mutations in *HOXD13*, the most 5′ gene of the *HOXD* cluster, include polyalanine expansions and frame-shift or missense mutations, and are associated with severe digit malformations in humans (Debeer et al. 2002; Caronia et al. 2003; Johnson et al. 2003). Intriguingly, polyalanine expansions and frame-shift mutations are predominantly associated with the formation of extra digits and webbing between digits, collectively called synpolydactyly, while missense substitutions have been reported to cause distinctive shortening of the hands/feet (brachydactyly) (for a review, see Goodman 2002). The fact that these two classes of *HOXD13* mutations are associated with two distinct phenotypes suggests that the underlying patho-mechanism is also distinct. Polyalanine expansions, which lead to aggregation of mutant HOXD13 in the cytoplasm (Albrecht et al. 2004), and *HOXD13* frame-shift mutations are likely to represent loss-of-function mutations. Because of the different phenotype observed with *HOXD13* missense mutations, we hypothesized that they represent gain-of-function mutations, whereby mutant proteins acquire altered genomic binding profiles. In this case, the phenotype would result not only from a lack of activation of target genes, but also from the activation of an abnormal set of target genes.

In this work, we apply a ChIP-seq methodology designed to investigate mutations in TFs to characterize a *HOXD13* missense mutation identified in a patient with a severe form of brachydactyly/oligodactyly. The mutation, Q317K, alters the glutamine residue at position 50 of the homeodomain to a lysine residue, which is characteristic of bicoid-type homeodomain proteins including PITX1. We show that the Q317K mutation shifts the HOXD13 binding profile toward that of PITX1. The induced gene expression patterns and the phenotypic effects following injection of the constructs in chicken embryos provide further evidence for a global shift in the regulatory properties of the mutant HOXD13 toward that of PITX1. As a control, we investigated a second *HOXD13* missense mutation that is associated with a distinct clinical phenotype, and found a strikingly different shift in the binding profile, suggesting that distinct molecular effects of mutations in TFs may underlie some of the clinical variability that can be seen with different mutations in the same TF. The methodology described here can be used to investigate a wide spectrum of TFs and mutations that have not previously been amenable to ChIP-seq experiments.

Results

The *HOXD13*^{Q317K} mutation causes a complex brachydactyly/oligodactyly phenotype

We identified an individual with a complex brachydactyly/oligodactyly phenotype. She was born with normal measurements, and had normal psychomotor development and no family history of congenital malformations. The girl's hands had only four fingers each as well as severe shortening of the digits and camptodactyly. The terminal phalanges and the nails of the thumbs were absent (Fig. 1A). The left foot had only three, shortened toes, whereby the nail of the left great toe was absent. Partial cutaneous webbing (syndactyly) was seen between the other toes, and the fibular toe was deviated to the tibial side. The right foot had only four, shortened toes with partial syndactyly of two toes, and the nail of the great toe was absent (Fig. 1A).

The clinical and radiological manifestations of this girl did not follow any known malformation patterns. Because of the absence of terminal phalanges, *ROR2* mutations were initially sought and excluded. Mutation analysis of *HOXD13* revealed a heterozygous mutation that was not found in either parent and had thus occurred de novo. The mutation is not present in the dbSNP database and has not been reported elsewhere. The severe brachydatyly/oligodactly phenotype described here has not been reported as part of the polyalanine expansion mutation spectrum, nor has it been described for other *HOXD13* mutations. This c.949C>A substitution, in exon 2 of the *HOXD13* gene (NM_000523.2), converts the glutamine residue at position 317 to a lysine residue (p.Q317K). At this position, a different point mutation (Q317R) has been reported previously in a family with a very different phenotype consisting of a form of brachysyndactyly (type V) (Zhao et al. 2007). The glutamine residue affected by the Q317 mutations corresponds to position 50 (Q50) of the homeodomain, which has been shown to be crucial for the binding site specificity of homeodomain proteins (Percival-Smith et al. 1990; Fraenkel et al. 1998; Chaney et al. 2005).

The Q317 mutations in *Hoxd13* lead to distinct DNA binding properties of wild-type and mutant HOXD13 proteins

To overcome the hurdles for ChIP-seq analysis of TF-variants, we developed a cell culture-based system using mesenchymal limb bud cells and retroviral overexpression and applied it to compare *Hoxd13*^{wt} to the *Hoxd13*^{Q317K} and *Hoxd13*^{Q317R} mutants (Supplemental Fig. S1). We used the chicken limb bud micromass system (chMM) for these experiments. Cells for chMM are taken from embryonic limb bud mesenchyme where *HOXD13* is expressed physiologically. In culture, the cells undergo a chondrogenic differentiation process that can be inhibited by virally expressed *Hoxd13* (Kuss et al. 2009).

To identify binding sites, we applied a targeting system in which a protein of interest is fused to a 3xFLAG-tag epitope in an RCASBP virus. This avian-specific retrovirus (RCAS) achieves high transfection levels while inserting only one copy of the virus per cell. An important motivation for the choice of the RCASBP system is the relatively mild degree of overexpression which is achieved. Using absolute quantification of transcripts, we found virally expressed *Hoxd13* transcript levels from infected chMM to be approximately three to five times higher than endogenous *HOXD13* expression in embryonic limb buds (Supplemental Methods; Supplemental Fig. S1). Cross-linked chromatin from chMM cultures infected with tagged murine *Hoxd13*^{wt}, *Hoxd13*^{Q317K}, or
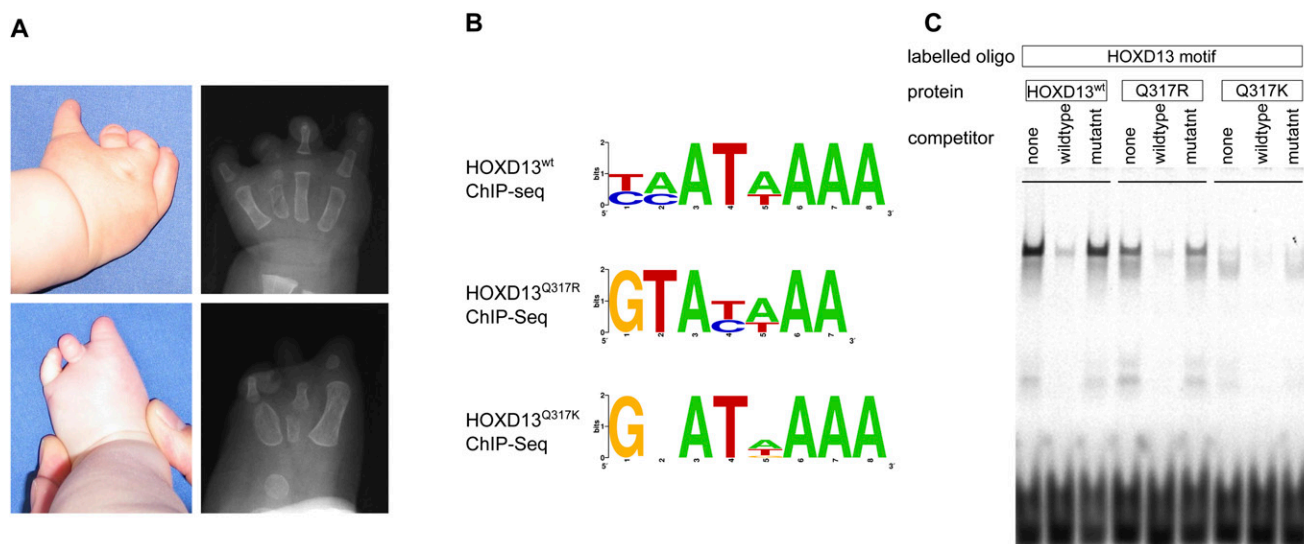
Figure 1. A HOXD13 p.Q317K mutation causes a severe limb phenotype and alters the DNA binding specificity. (*A*) Photographs and radiographs of patient with *HOXD13*^{Q317K} mutation. Missing digits (oligodactyly) and severe shortening of the digits (brachydactyly), as well as absent distal phalanges and nails are apparent. Radiographs show missing and deformed metacarpals, metatarsals, and rudimentary phalangeal bones. (*B*) The HOXD13^{Q317} mutations alter the binding motif at the 5′ end. The primary motifs found by de novo motif analysis for HOXD13^{wt} (*top*), HOXD3^{Q317R} (*middle*), and HOXD13^{Q317K} (*bottom*). Both Q317 mutant proteins show an altered 5′ end of the sequence as compared with HOXD13^{wt}. (*C*) In vitro binding of wild-type and mutant HOXD13 homeodomains to the HOXD13 motif. EMSA of purified wild-type and mutant HOXD13 homeodomains was performed using fluorescent dye-labeled double-stranded oligonucleotides containing a binding sequence for HOXD13 (CCAATAAAA). Oligonucleotides were incubated with purified wild-type and mutant HOXD13 homeodomains. Unlabeled competitor oligonucleotides reduced binding to the labeled oligonucleotide, whereas a mutant competitor did not, indicating sequence-specific binding. The HOXD13^{wt} motif is bound by HOXD13^{wt} and (weaker) HOXD13^{Q317R} but not by HOXD13^{Q317K}.

Hoxd13^{Q317R} were used for ChIP-seq. For each experiment, we performed two replicates that were subjected to quality control standards as defined by The ENCODE Project Consortium (The ENCODE Project Consortium 2011; Landt et al. 2012).

We identified 34,267 peaks for HOXD13^{wt}. To test the specificity of the identified binding sites, we conducted motif analysis using DREME (Bailey 2011). The primary motif inferred for HOXD13^{wt} showed a characteristic HOX binding motif with an AT[A/T]AAA core sequence preceded 5′ by a somewhat more loosely defined [C/T][A/C], which is largely identical to the HOXD13 motif previously identified by protein-binding microarray analysis (Fig. 1B; Berger et al. 2008).

We then applied our chMM ChIP-seq system to compare HOXD13^{Q317K} and HOXD13^{Q317R} mutant proteins with HOXD13^{wt}. After quality control and reproducibility testing, we identified 25,036 binding sites for HOXD13^{Q317K} and 21,396 for HOXD13^{Q317R}. Using these peaks, we first performed motif analysis for the mutant proteins. Strikingly, both the HOXD13^{Q317K} and the HOXD13^{Q317R} motif revealed a changed di-nucleotide at the 5′ terminus. The primary motif inferred for HOXD13^{Q317K} has a guanine residue at the 5′ end that is followed by a variable base. The HOXD13^{Q317R} motif also starts with a guanine, which is then followed by a thymine residue (Fig. 1B). Intriguingly, the AT-rich core remained largely unchanged in both motifs. Therefore, we infer that both Q317 mutations alter only the 5′ region of the DNA recognition sequence of HOXD13, while the 3′ part remains unchanged.

The HOXD13^{Q317K} mutant, but not HOXD13^{Q317R}, recognizes the PITX1-motif

We compared the HOXD13 homeodomain sequence to all other homeodomains carrying a lysine (K), like the mutant, at position 50. Sequence alignment between these homeodomains and the mutant HOXD13 homeodomain revealed the greatest similarity to the PITX-family (see Supplemental Fig. S2). Furthermore, the only homeodomain protein carrying a K at position 50 known to be involved in limb patterning is PITX1, one of the few TFs which is expressed in the hindlimb only and has been shown to confer hindlimb identity (Logan and Tabin 1999; Szeto et al. 1999). PITX1, like other K50-homeodomain TFs, binds a GGATTA motif (Tremblay et al. 1998). We hypothesized that the pathogenesis of *HOXD13*^{Q317K} might be at least partially due to aberrant transactivation of genes that are normally regulated by PITX1. We therefore performed a series of experiments to test whether HOXD13^{Q317K}, due to its altered sequence specificity, shares characteristics with PITX1.

As the motif analysis for HOXD13^{Q317K} revealed a very similar but not identical motif to PITX1, we used electrophoretic mobility shift assays (EMSA) to test the binding of wild-type and mutant homeodomains with respect to characteristic HOXD13 and PITX1 target sequences. Wild-type HOXD13 specifically bound the wild-type HOXD13 motif, but HOXD13^{Q317K} did not show appreciable binding activity. In contrast, the HOXD13^{Q317R} mutant showed considerable residual binding to the wild-type HOXD13 motif (Fig. 1C). This binding affinity was also reflected in the de novo motif analysis, which identified a HOXD13^{wt}-like motif as a secondary motif for HOXD13^{Q317R} but not for HOXD13^{Q317K} (data not shown). When we tested wild-type and mutant homeodomains for binding to the PITX1 target sequence, the Q317K mutant homeodomain efficiently bound the PITX1 motif, whereas the HOXD13^{wt} and HOXD13^{Q317R} homeodomains did not (Fig. 2B).
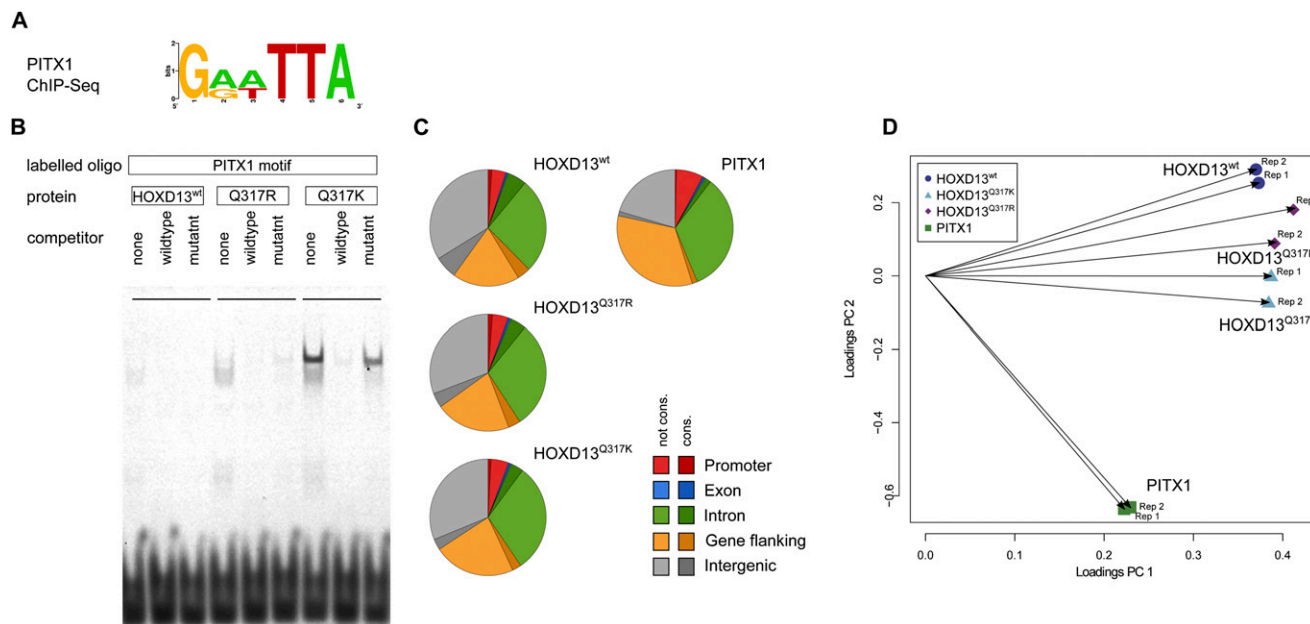
Figure 2. The HOXD13^{Q317K} mutant binds the PITX1 binding site in vitro, and the ChIP-seq profile shifts toward PITX1. (*A*) The primary PITX1 motif identified from ChIP-seq. (*B*) In vitro binding of wild-type and mutant HOXD13 homeodomains to the PITX1 motif. EMSAs were performed as in Fig. 1C using oligonucleotides containing a binding sequence for PITX1 (GGATTA). HOXD13^{Q317K} specifically binds the PITX1 motif, whereas it is not bound by HOXD13^{wt} or HOXD13^{Q317R}. (*C*) Genomic distribution of ChIP-seq peaks. The genome was divided into promoter (5 kb upstream of, 2 kb downstream from TSS), exon, intron, gene flanking (20 kb upstream, 20 kb downstream), and intergenic regions, each separated in conserved and nonconserved regions. (*D*) Principle component analysis for the coverage profiles of the uniquely mapped reads for all eight data sets (two biological replicates each for HOXD13^{wt}, HOXD13^{Q317K}, HOXD13^{Q317R}, and PITX1).

The genome-wide binding pattern of the HOXD13^{Q317K} mutant shifts toward a PITX1-like binding profile

We next wanted to see whether HOXD13^{Q317K} or HOXD13^{Q317R} were also able to recognize PITX1 binding sites in the genome. To identify PITX1 binding sites in the chicken genome, we used our established protocol to perform ChIP-seq from chMM-cultures infected with *PITX1*-expressing RCASBP-virus. We identified 40,881 peaks for PITX1 in chMM cultures. De novo motif analysis of ChIP-seq peaks readily identified the known PITX1 motif (Fig. 2A), indicating functionality of the protein.

The overall peak distribution of HOXD13^{wt}, HOXD13^{Q317K}, HOXD13^{Q317R}, and PITX1 with respect to annotated genes was similar for all factors, but HOXD13^{wt} peaks were more commonly found in conserved regions than PITX1 peaks (Fig. 2C). To compare the genome-wide distribution of protein binding we performed principle component analysis (PCA) on the read-count distribution in 500-bp windows for the eight data sets, representing two experiments each for HOXD13^{wt}, HOXD13^{Q317K}, HOXD13^{Q317R}, and PITX1. The biological replicates for each condition clustered together and the coordinates for HOXD13^{Q317K} were located at an intermediate position between those of HOXD13^{wt} and PITX1 (Fig. 2D). In contrast, the coordinates for HOXD13^{Q317R} remained closer to those of HOXD13^{wt}.

As mentioned above, we identified 34,267 HOXD13^{wt}-bound regions, 25,036 for HOXD13^{Q317K}, 21,396 for HOXD13^{Q317R}, and 40,881 for PITX1. Based on our definition of shared binding events, most peaks are TF specific, and only 45 sites are bound by all four factors. The overlap with HOXD13^{wt} or PITX1 peaks was clearly different for HOXD13^{Q317K} and HOXD13^{Q317R}. HOXD13^{Q317R} shared 27.5% (5887) of its binding sites with HOXD13^{wt}, whereas this

figure was only 6.0% (1492) for HOXD13^{Q317K} (Fig. 3C). In contrast, HOXD13^{Q317K} shared 6.6% (1671) of its binding sites with PITX1, but HOXD13^{Q317R} shared only 2.9% (633). HOXD13^{wt} shared only 1.1% (383) of its binding sites with PITX1 (Fig. 3C)

This shows that the fraction of PITX1 sites bound by HOXD13^{Q317K} is considerably larger than the overlap observed for either HOXD13^{Q317R} or HOXD13^{wt}. Additionally, HOXD13^{Q317K} and HOXD13^{Q317R} share a large set of target sites (3094) that are not bound by HOXD13^{wt} or PITX1.

Since enrichments at genomic regions are not always detected as peaks, we analyzed the distribution of reads from HOXD13^{wt}, HOXD13^{Q317K}, HOXD13^{Q317R}, and PITX1 ChIP-seq around the identified binding sites using seqMINER (Fig. 3D; Ye et al. 2011). There was no enrichment of reads at the position of PITX1 peaks in HOXD13^{wt} or HOXD13^{Q317R} ChIP-seq experiments, and the same was true vice versa. In contrast, the HOXD13^{Q317K} ChIP-seq displayed a moderate enrichment of reads at the site of PITX1 peaks, indicating HOXD13^{Q317K}-binding at PITX1 peaks. We found a similar moderate enrichment of HOXD13^{Q317K} around HOXD13^{wt} peaks. Consistent with the peak overlap, HOXD13^{Q317K} reads were enriched around HOXD13^{Q317R} peaks and vice versa. Again, the genome-wide binding of the HOXD13^{Q317K} mutant shifted toward PITX1 sites to an extent that was not found for HOXD13^{Q317R} or HOXD13^{wt}.

As an example for the changed binding behavior, Figure 3A shows the binding profiles of the four TFs at the genomic region of the *matrix-Gla protein* (*MGP*) gene. *MGP* encodes the matrix-Gla protein, which inhibits calcification in proliferating chondrocytes and blood vessels (Luo et al. 1997; Yao et al. 2011). HOXD13^{wt}, HOXD13^{Q317K}, and PITX1 bind to various sequences in the region upstream of *MGP*, while HOXD13^{Q317R} shows only weak binding
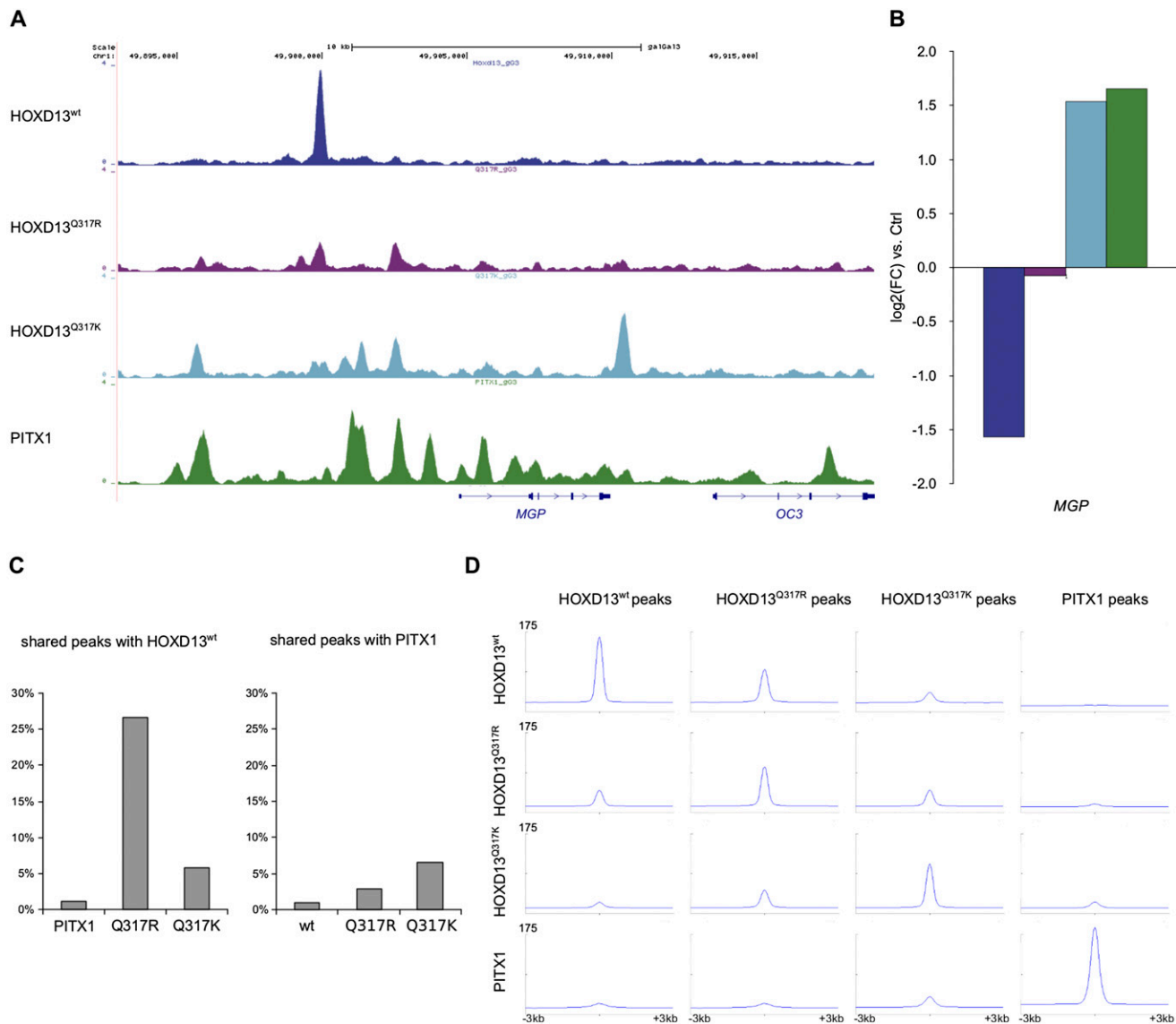
Figure 3. HOXD13^{Q317K} binding, but not HOXD13^{Q317R}, acquires PITX1-like properties. (*A*) Coverage profiles for HOXD13^{wt}, HOXD13^{Q317R}, HOXD13^{Q317K}, and PITX1 in a 20-kb region covering the *MGP* gene show the similarity in binding between HOXD13^{Q317K} and PITX1. (*B*) *MGP* is threefold up-regulated in *PITX1*- and *Hoxd13^{Q317K}*-infected chMM cultures and threefold down-regulated in *Hoxd13^{wt}* cultures (log$_2$ fold changes compared to mock-infected cultures). (*C*) Peak overlap with respect to HOXD13^{wt} and PITX1 peaks. (*Left*) PITX1 shares 1.0% (383/40,881) of its peaks with HOXD13^{wt}; 27.5% (5887/21,396) with HOXD13^{Q317R}; and 6.0% (1492/25,036) with HOXD13^{Q317K}. (*Right*) HOXD13^{wt} shares 1.1% (383/34,267) of its binding sites with PITX1; 3.0% (633/21,396) with HOXD13^{Q317R}; and 6.7% (1671/25,036) with HOXD13^{Q317K}. (*D*) HOXD13^{Q317K} shares binding properties with both HOXD13^{wt} and PITX1. Mean read densities centered (±3 kb) around the peak summits for HOXD13^{wt}, HOXD13^{Q317R}, HOXD13^{Q317K}, and PITX1 (*left* to *right*). HOXD13^{wt} (164.1 tags/50 bp; 93.5 tags/50 bp; 43.2 tags/50 bp; 23.6 tags/50 bp), HOXD13^{Q317R} (49.7 tags/50 bp; 101.2 tags/50 bp; 49.8 tags/50 bp; 19.5 tags/50 bp), HOXD13^{Q317K} (24.9 tags/50 bp; 51.9 tags/50 bp; 109.1 tags/50 bp; 25.6 tags/50 bp), and PITX1 (15.1 tags/50 bp; 24.2 tags/50 bp; 38.9 tags/50 bp; 191.2 tags/50 bp).

throughout this region. HOXD13^{wt} binds a unique site. HOXD13^{Q317K} and PITX1 share three binding sites with additional unique binding sites for each factor.

RNA-seq shows similar expression profiles induced by HOXD13^{Q317K} and PITX1

To investigate whether similarities in genomic binding profiles are reflected in gene expression patterns, we performed RNA-seq expression analysis for chMM cultures infected with *Hoxd13^{wt}*, *Hoxd13^{Q317K}*, *Hoxd13^{Q317R}*, or *PITX1*. In order to identify differentially expressed genes, expression levels were filtered according to a minimal expression of 2 RPKM in at least one experiment and at least twofold or higher differential expression in any of the cultures as compared to mock-infected cultures.

This left a total of 3118 transcripts belonging to 2506 genes that were differentially expressed in at least one of the chMM cultures (Supplemental Table 2). PCA analysis of the expression values for all differentially expressed transcripts demonstrated a similar profile as seen for ChIP-seq, with *Hoxd13^{Q317K}* being close

to *PITX1* and *Hoxd13*^{Q317R} close to *Hoxd13*^{wt} (Fig. 4A). Similarly, hierarchical clustering of transcript levels clustered *Hoxd13*^{Q317K} with *PITX1*-infected cultures, and *Hoxd13*^{Q317R} with *Hoxd13*^{wt} (Fig. 4B). For each factor, a comparable number of genes (between 845 and 1102) were differentially expressed. When we compared the genes regulated by mutant or wild-type HOXD13 with PITX1-regulated genes, we identified 436 genes coregulated by HOXD13^{Q317K} and PITX1, whereas only 262 genes were co-regulated by HOXD13^{wt} and PITX1 and 237 genes by HOXD13^{Q317R} and PITX1 (Supplemental Table 2).

To identify the genes that might contribute to the similarity of HOXD13^{Q317K} to PITX1 found in PCA and hierarchical clustering analysis, we focused on the 436 HOXD13^{Q317K}/PITX1 coregulated genes. To characterize this group of genes, we performed GO analysis on the 305 co-down-regulated and 131 co-up-regulated genes. For the down-regulated genes, GO terms related to cartilage differentiation and limb development were overrepresented, which were also overrepresented in the sets of down-regulated genes for each individual factor (Supplemental Fig. S3). Among the HOXD13^{Q317K}/PITX1 co-up-regulated genes, genes annotated to several GO categories related to muscle development as well as blood vessel development were overrepresented. These and related GO terms were specific for HOXD13^{Q317K} and PITX1 up-regulated genes and were not similarly enriched in the HOXD13^{wt} or HOXD13^{Q317R} up- or down-regulated genes (Supplemental Fig. S3).

We then wanted to see whether the genes coregulated by HOXD13^{Q317K} and PITX1 were also cobound by the factors. Therefore, we classified a gene as cobound when a HOXD13^{Q317K}/PITX1 shared peak was found in the gene body or within 25 kb up- or downstream from the gene. Examples for cobinding of HOXD13^{Q317K} and PITX1 at coregulated genes are shown in Supplemental Fig. S4. Using this classification, 13.1% (57) of the 436 coregulated genes were also cobound by HOXD13^{Q317K}/PITX1, significantly more than expected by chance (Fisher's exact $P = 1.295 \times 10^{-3}$). An analogous analysis performed for the HOXD13^{Q317R} mutant and PITX1 found 5.06% (12 out of 237) coregulated genes cobound ($P = 0.1335$) (Supplemental Fig. S5).

Overexpression of *Hoxd13*^{Q317K} and *PITXI* induces similar phenotypes

Chicken micromass cultures are derived from limb bud mesenchymal cells. Cultured at high density, they differentiate into various cell lineages, including chondrocytes which form cartilage nodules that undergo differentiation, mineralization, and finally bone formation (Ahrens et al. 1979). Viral expression of wild-type *Hoxd13* strongly inhibits chondrogenic differentiation. Instead, cells in every part of the culture except the very center do not assemble into organized structures (Fig. 5). Overexpression of *Hoxd13*^{Q317R} does not inhibit chondrocyte formation like the wild
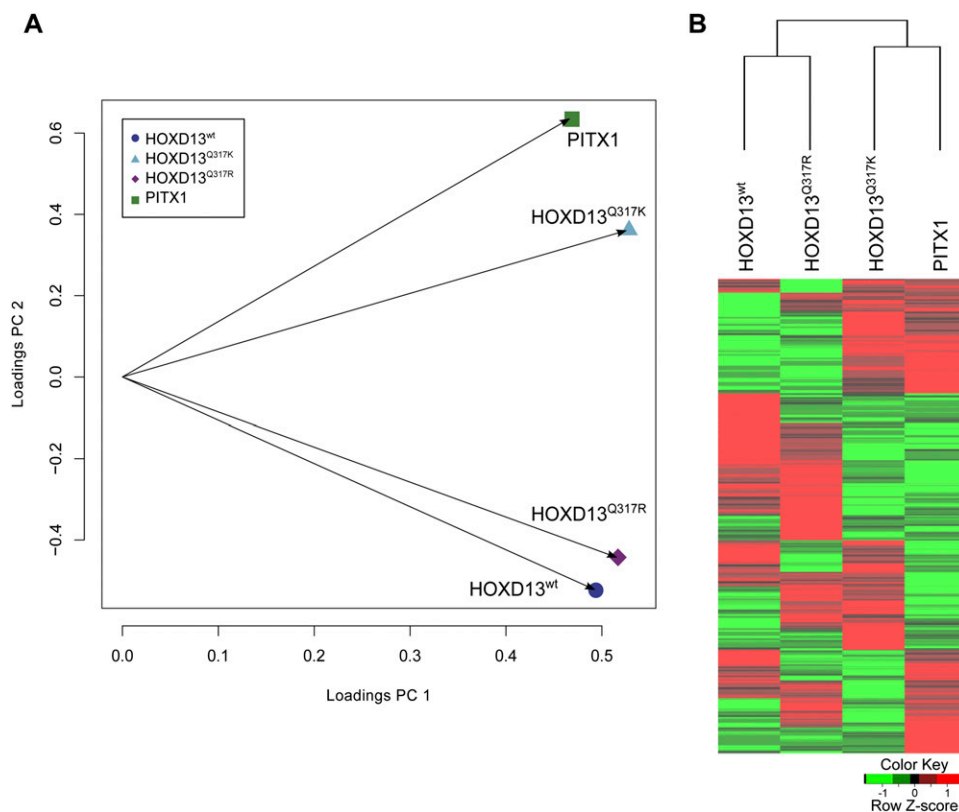
Figure 4. Gene regulation by HOXD13^{Q317K} is more similar to PITX1 than to HOXD13^{wt} or HOXD13^{Q317R}. Global expression analysis of genes regulated by HOXD13^{wt}, HOXD13^{Q317R}, HOXD13^{Q317K}, or PITX1 in chMM cultures. RNA-seq expression analysis found 3118 transcripts (2506 genes) to be at least twofold differentially regulated in *Hoxd13*^{wt}, *Hoxd13*^{Q317R}, *Hoxd13*^{Q317K}, or *PITX1* chMM compared to mock-infected cultures (RPKM cutoff 2). (*A*) Principle component analysis (PCA) of the expression values (normalized to expression values in mock-infected cultures) of these genes shows a similar clustering as seen in the ChIP-seq PCA. (*B*) Hierarchical clustering of the expression values (RPKM) clusters *Hoxd13*^{Q317K} with *PITX1* and *Hoxd13*^{Q317R} with *Hoxd13*^{wt}. Red indicates up-regulation and green indicates down-regulation, as compared to mock-infected cultures.
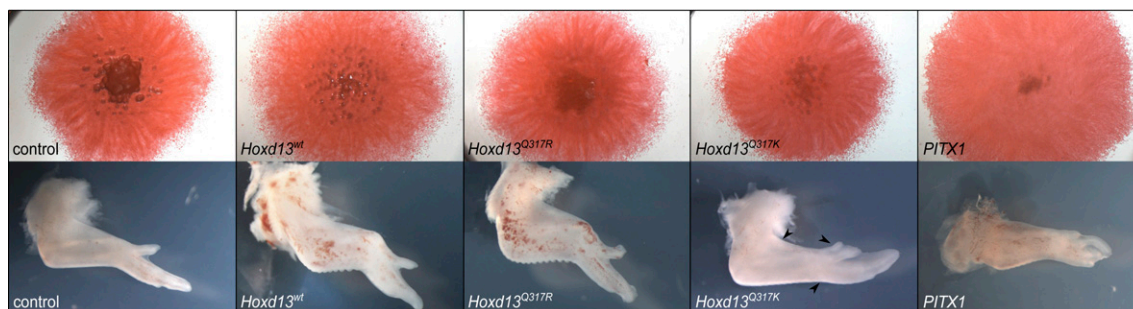
Figure 5. Some effects on cell differentiation and development are shared between $Hoxd13^{Q317K}$ and *PITX1*. (*Upper* row) Eosin-stained chMM cultures for $Hoxd13^{wt}$, $Hoxd13^{Q317R}$, $Hoxd13^{Q317K}$, and *PITX1*. Both $Hoxd13^{Q317K}$ and *PITX1* induced differentiation into thick, fibrotic tissue. (*Lower* row) Viral overexpression of $Hoxd13^{wt}$, $Hoxd13^{Q317R}$, $Hoxd13^{Q317K}$, and *PITX1* in chicken wing buds (HH36). $Hoxd13^{wt}$ and $Hoxd13^{Q317R}$ overexpression leads to an overall shortening of skeletal elements. $Hoxd13^{Q317K}$-infected wing buds lack or have a reduced forelimb-specific patagium, have a straightened wrist, and develop a fourth digit (arrowheads). *PITX1* overexpression leads to a partial hind- to forelimb transformation.

type does. The cultures, although a bit smaller in size than mock-infected chMM, appear to undergo the same chondrogenic differentiation process (Fig. 5). chMM infection with $Hoxd13^{Q317K}$ results in a completely different phenotype. Inhibition of chondrogenic differentiation is even stronger, but in contrast to $Hoxd13^{wt}$, the cells form a multilayered fibrotic tissue which radially extends from the center of the culture (Fig. 5). This radially extending fibrotic tissue was also observed in *PITX1*-expressing chMM cultures, although inhibition of chondrogenesis was not as strong as in cultures infected with $Hoxd13^{Q317K}$.

So far, our data showed that binding specificity, gene expression, and morphology of $Hoxd13^{Q317K}$-infected chMM cultures shift from *Hoxd13* to *PITX1* characteristics. To test whether this shift in specificity can also be observed in vivo, we expressed $Hoxd13^{wt}$, $Hoxd13^{Q317R}$, $Hoxd13^{Q317K}$, and *PITX1* using the same FLAG-tagged constructs as used for ChIP-seq experiments in chicken limb buds. The effects of $Hoxd13^{wt}$ and *PITX1* in this experiment have been described and are clearly distinguishable. As previously shown, expression of *PITX1* in the developing chick wing bud leads to a partial fore-to-hindlimb transformation of the infected wing (Logan and Tabin 1999; Szeto et al. 1999). This manifests itself in three characteristic features. The infected wings show an absence of the patagium (the skin that will form the surface of the wing), demonstrate straightening of the posterior wrist flexure, and develop an additional digit (Fig. 5). In contrast, ectopic expression of $Hoxd13^{wt}$ leads to a reduction in bone size but has no effect on limb patterning (Fig. 5; Goff and Tabin 1997).

Like $Hoxd13^{wt}$, viral infection of wing buds with $Hoxd13^{Q317R}$ mainly leads to an overall reduction of the limb structures (Fig. 5). We note that in a minority of cases, we observed reduced interdigital cell death or mispositioning of the wrist (Supplemental Fig. S6).

In $Hoxd13^{Q317K}$-infected wing buds, we observed the lack or reduction of the patagium as well as a straightening of the posterior flexure of the wrist. Also, a fourth digit developed. However, the position of this digit was different from that observed in *PITX1*-infected wings. This extra digit extended from the first phalanx of digit 1, while with *PITX1* overexpression, the additional digit was located completely apart from digit 1 (Fig. 5, arrowheads; Supplemental Fig. S6).

Taken together, the effects of $Hoxd13^{Q317K}$ expression in chMM cultures and limb bud injections are distinct from those induced by $Hoxd13^{wt}$ or $Hoxd13^{Q317R}$ and display a partial overlap with those of wild-type *PITX1*.

Discussion

TFs are frequently involved in the pathogenesis of disease, in particular, congenital malformation syndromes. In many instances, this is due to haploinsufficiency, i.e., the functional loss of one allele caused by deletions, nonsense mutations, frame-shifts, or other inactivating variants. Point mutations that result in amino acid alterations are, however, more difficult to interpret. Provided that a missense mutation does not simply destabilize the protein, it may induce aberrant function by mechanisms such as altering protein-protein or protein-DNA binding. Mutations in *HOXD13* are associated with a broad spectrum of limb phenotypes including webbing of fingers (syndactyly), additional fingers (polydactyly), and shortening of fingers or of the entire hand/foot (brachydactyly). Polyalanine expansions in *HOXD13* are associated with synpolydactyly, consisting of syndactyly between fingers III and IV and an additional finger located in between digits III and IV (Muragaki et al. 1996). Amino acid substitutions in the homeodomain have been described in cases with brachydactyly mainly affecting the distal thumb and/or the metacarpals (Johnson et al. 2003). Interestingly, a mutation at the identical position as the one described here but to a different amino acid, p.Q317R, causes yet another phenotype—a combination of synpolydactyly and brachydactyly (Zhao et al. 2007). The complex limb malformation that is associated with the Q317K mutation is different from the previously described phenotypes and thus extends the phenotypic spectrum associated with *HOXD13* mutations even further. The presence of several distinctive phenotypes suggests mutation-specific disease mechanisms for each type of mutation. In the case of HOXD13 polyalanine expansions, the pathomechanism appears to be a combination of loss of function and gain of function caused by protein aggregation and the interaction with other alanine repeat containing TFs (Villavicencio-Lorini et al. 2010). For missense mutations, considering the drastic differences in phenotypic expression, a common pathomechanism for all mutations is a rather unlikely scenario, and specific effects, such as alterations in binding capacity, are more likely to explain this phenomenon.

The identical Q to K substitution, as in the patient described here, has been introduced in the *Drosophila* fushi-tarazu (ftz) homeodomain. This changed the recognition motif in vitro from <u>CC</u>ATTA to <u>GG</u>ATTA, the motif of the homeodomain-protein bicoid that endogenously carries a K (lysine) at position 50 of the homeodomain (Percival-Smith et al. 1990; Schier and Gehring 1992). The effects of the homeodomain Q50K substitution have been in-

vestigated in *Drosophila*, and its altered regulatory capacity has been studied using reporter constructs (Schier and Gehring 1992; Zhao et al. 2000). Flies transgenic for an *ftz*^{Q50K} allele develop normally and do not show any malformations, although reporter constructs respond to the altered binding of ftz^{Q50K} (Schier and Gehring 1992). The ftz^{Q50K} protein binds with high affinity to constructs carrying consensus bcd binding sites and activates them; however, it fails to activate natural enhancers harboring nonconsensus binding sites, demonstrating the importance of other homeodomain sequences and sequences outside the homeodomain that are needed to confer the full regulatory capacity of the TF (Zhao et al. 2000).

In the mouse genome, there are 25 K50 (bicoid-type) homeodomain TFs (20 in humans), among them the genes of the *Otx*, *Obox*, *Rhox, Six,* and *Pitx* families. The only bicoid-type TF in vertebrates that is known to play a major role in limb development is PITX1. Studies in mouse and chicken have shown that *Pitx1*, expressed in the developing hindlimb but not in the forelimb, is required for the formation of hindlimb characteristics and produces hindlimb-like morphologies when misexpressed in forelimbs (Logan and Tabin 1999; Szeto et al. 1999).

Our results show that, in comparison with HOXD13^{wt}, HOXD13^{Q317K} leads to a clearly different regulation of target genes, many of which are also recognized and regulated by PITX1. This is likely due to the fact that HOXD13^{Q317K} is capable of binding to and regulating a substantially greater number of PITX1 targets than is observed for ftz^{Q50K} and bcd. This is supported by the change in sequence specificity induced by HOXD13^{Q317R}, which is similar but distinct, and the fact that only HOXD13^{Q317K} can recognize the PITX1 binding site in vitro, indicating that the recognition of PITX1 sites is specific to HOXD13^{Q317K}. Furthermore, the number of coregulated genes for PITX1 and HOXD13^{Q317K} is much higher than for the other factors, and shared binding sites for PITX1 and HOXD13^{Q317K} are enriched in the vicinity of the coregulated genes, also indicating shared properties of HOXD13^{Q317K} and PITX1. However, it remains difficult to assign individual ChIP-seq peaks to target genes (for review, see MacQuarrie et al. 2011; Spitz and Furlong 2012). As in our results, developmental TFs have been found to bind throughout the genome, and many peaks are found in intergenic regions that cannot be easily attributed to the regulation of a specific target (Cao et al. 2010). Therefore, the absence of a nearby peak does not exclude the regulation of a certain gene. Neither does the presence of ChIP-seq peaks in promoters of differentially expressed genes prove a regulatory function of an individual peak. Thus, while the enrichment of shared peaks near coregulated genes links cobinding to coregulation, the list of candidate genes will neither be complete nor will all the genes be direct targets.

The adoption of a novel set of regulatory targets is only possible if the new binding sites in the genome are meaningful in the regulatory context of the cell. The cells in the embryonic limb bud are responsive to *Pitx1* expression, as indicated by *Pitx1* overexpression and knockout experiments in transgenic models. In the patient, the altered binding of HOXD13^{Q317K} is therefore likely to disturb the regulatory balance during limb bud development by regulating PITX1 targets in the *HOXD13* expression domain. *PITX1* is expressed in the entire hindlimb at early stages and becomes more restricted to the middle part (zeugopod) in later stages, whereas *HOXD13* is expressed in fore- and hindlimb, first in a small domain in the posterior limb bud and then in the entire distal limb (autopod). Thus, a PITX1-like activity in the *HOXD13* expression domain can be expected to result in a major degree of misexpression. The hypothesis that normal HOXD13 function is

partially replaced by PITX1 activity is supported by the effects of overexpressing *Hoxd13*^{Q317K} and *PITX1* in chMM cultures and in infected chicken wing buds, which leads to overlapping developmental consequences in both cases, suggesting a similar set of downstream effects. Our in vitro results demonstrate the mutant is able to activate PITX1 target genes. Furthermore, the in vivo experiments show that this effect is biologically relevant, as the mutant is able to induce PITX1-like morphological changes. The specificity of these effects is further corroborated by our investigations of the HOXD13^{Q317R} mutant.

The phenotype in the *HOXD13*^{Q317K} patient is different and more severe than that described for the *HOXD13*^{Q317R} patients. The less severe phenotype of *HOXD13*^{Q317R} is possibly due to two reasons. On the one hand, HOXD13^{Q317R} still demonstrates considerable residual binding to the wild-type HOXD13 motif and may thus have remaining wild-type activity. On the other hand, the new binding sites do not seem to correspond to known TF target sites and may therefore have only minimal effects. Thus, our results show that a missense mutation in a TF can act as a gain-of-function mutation that shifts the binding profile of the relevant TF on a genome-wide scale. The individual consequences of such a shift, though, are dependent on the presence of compatible *cis*-regulatory elements/targets as well as the properties of protein domains involved in transactivation and protein-protein-interaction. This applies not only to pathogenic mutations but is also likely to be relevant in respect to TF evolution.

ChIP-seq has been used to characterize the genomic binding profiles of numerous TFs in many different experimental settings. Nevertheless, it has not yet been systematically used to investigate the effects of mutations in TFs associated with hereditary diseases. The main roadblocks consist of the difficulty in distinguishing between wild-type and mutant TFs with a primary antibody and in obtaining sufficient amounts of tissue of a relevant anatomical site and developmental stage for the analysis. We have concentrated our efforts on the chMM system in this work, but similar methodologies are likely to be useful for other organ systems that are amenable to infection with the RCASBP virus, including models for muscle, neuronal, and neural crest cell development (McCobb et al. 1990; Dorman and Johnson 1999; Etchevers 2011). While we chose the murine *Hoxd13* for our viral constructs, the very high conservation between mouse and human HOXD13 (94% protein identity, 100% identity in the homeodomain) and strong conservation of TF binding specificities (Hecht et al. 2008; Schmidt et al. 2010) makes the findings directly translatable between mouse and human. The advantage of the chMM system over other cell culture systems is that the cultures undergo a physiological differentiation process, and thus are well suited to investigating the factors that influence this process. The RCASBP system leads to overexpression of the transduced TF which might lead to false-positive binding sites. However, for the comparison of mutated TFs vs. wild type, this is not likely to be relevant. Additionally, the overexpression levels in the case of *Hoxd13* and *PITX1* in chMM cultures are still in the range of endogenous expression levels in the developing limb.

The chMM-ChIP-seq system enables a relatively quick functional characterization by comparing wild-type and mutant proteins in the genomic context and reveals not only where the mutant TF fails to bind but also if the protein changes its binding specificities. To date, the analytical tools to characterize missense mutations have been limited. Most frequently, nuclear localization of the protein is tested in transient transfection assays supplemented with gel shift assays and/or luciferase assays on reporter constructs. Failure to bind the consensus motif or to activate the

reporter result in the characterization of the mutation as a loss of function (Kozlowski and Walter 2000; Saleem et al. 2003; Zhao et al. 2007). According to these criteria, the *HOXD13*^{Q317K} mutation described here would qualify as a loss-of-function mutation. However, our data show that this is not the case. Our results suggest that missense mutations in TFs may have other effects such as the acquisition of novel functions via expansion and/or alterations of binding motifs.

Methods

Clinical analysis and Sanger sequencing

One family with brachydactyly was investigated. Mutation analysis in the *ROR2* and *HOXD13* genes was performed using standard Sanger sequencing to test both parents and the affected child. *HOXD13* mutation nomenclature is according to RefSeq Accession NM_000523.2. Primer sequences are given in the Supplemental Methods. All patients provided signed, informed consent to this study. The study was approved by the local ethics committee.

Chicken micromass cultures

Chicken micromass cultures were prepared as described elsewhere (Solursh et al. 1978). In brief, cells were obtained from 4.5-d chicken embryos (HH24) and mixed with virus before seeding. chMM cultures were harvested after nine days incubation in chicken micromass medium consisting of DMEM/Ham's F12 (Biochrom) with 10% FBS (PAA), 0.2% chicken serum (Sigma), 1% L-glutamine (Lonza), and penicillin/streptomycin (Lonza). For harvesting, chMM cultures were digested with 0.1% collagenase (Sigma) in chMM-medium to obtain a roughly single-cell suspension and fixed for 10 min on ice with 1% formaldehyde in chMM-medium. Before chromatin extraction and ChIP, fixed cells were frozen in liquid N$_2$ and stored at −80°C. For histological staining, cultures were fixed in the cell culture dish with 4% PFA in PBS overnight and stained with eosin solution (Sigma).

3xFLAG-tag vectors as a universal ChIP-seq tool

In order to generate fusion proteins for ChIP-seq that would be recognized by a single antibody under standardized conditions, we modified the pSlax-13 vector (Morgan and Fekete 1996) to allow fusion of an insert to an N-terminal triple FLAG-tag and directional cloning via ClaI and SpeI. We cloned the coding sequences for wild-type (wt) and mutant (Q317K, Q317R) murine *Hoxd13* (NM_008275.2) as well as wild-type chicken *PITX1* (NM_001167684) into a modified RCASBP(A)-vector using the ClaI and SpeI sites.

Retroviral infection of chicken limb buds

Infection of the chicken limb buds with the viruses described above was performed as described previously (Logan and Tabin 1998). Chicken embryos (VALO BioMedia) were infected by injecting virus into the presumptive wing field at HH 9-11 and incubated for the appropriate time. Embryos were fixed in 4% PFA overnight and analyzed using a dissecting microscope.

Electromobility shift assays

Cloning and expression of wild-type and mutant homeodomains in *Escherichia coli* were performed as described in the Supplemental Methods. EMSA was performed using 20 μl binding buffer (100 mM NaCl, 2 mM MgCl$_2$, 0.1 mg/mL BSA, 4 mM spermidine, 25 mM

HEPES, pH 7.5, 1× Roche complete protease inhibitor), and 1.5 μL poly(dIdC) (100 ng/μL) as well as 1 μL purified homeodomain (400 ng/μL) were added. Reactions were incubated on ice for 5 min, and then 2 μL of the Cy3-labeled oligonucleotide (1.5 pmol/μL) and—if appropriate—1 μL (15 pmol/μL) of the unlabeled competitor were added and incubated for 15 min at room temperature. Oligonucleotide sequences are given in the Supplemental Methods. Before loading of the sample onto the gel, 2 μL loading buffer (40% glycerol + 0.01% bromphenol blue) were added. Electrophoresis was performed on an 8% native polyacrylamide gel in 0.5× TBE and scanned on a FLA-5000 Scanner (Fuji).

Chromatin immunoprecipitation

ChIP was performed as described in Lee et al. (2006) with the following alterations. Cells were suspended in 1.5 mL L3 buffer, sonicated with a Diagenode Bioruptor for 45 cycles (30-sec pulse, 30-sec pause, HI power), and stored at −80°C before use. Chromatin concentration was determined by measuring the DNA content in a 150-μL sample of the sonicated chromatin. For each ChIP, 40 μL of magnetic Protein G beads (Invitrogen, #100.04D) were preincubated with 10 μg anti-FLAG M2 antibody (Sigma, F1804), and incubated with ~35 μg chromatin overnight.

ChIP-seq

For ChIP-seq library preparation, ethanol-precipitated DNA from ChIP was redissolved in 46 μL of water and subjected to library preparation using the NEBNext ChIP-seq Library Prep Master Mix for Illumina (New England Biolabs), selecting for a fragment size range of 300–450 bp according to the manufacturer's instructions. Sequencing was performed on an Illumina GAIIx sequencer with 36-bp single reads using TruSeq v5 sequencing chemistry and TruSeq v5 cluster generation kits.

RNA-seq

Total RNA from chMM cultures was isolated using peqGOLD Tri-Fast Reagent (PEQLAB) and subsequent purification with RNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. mRNA for RNA-seq libraries was purified using the Oligotex mRNA Mini Kit (Qiagen) from 6 μg of total RNA. RNA-seq libraries were constructed using the NEBNext mRNA Library Prep Master Mix Set for Illumina according to the manufacturer's instructions, selecting for a fragment size range of 300–500 bp. Sequencing was performed on an Illumina GAIIx sequencer with 115-bp single reads using TruSeq v5 sequencing chemistry and TruSeq v5 cluster generation kits.

Computational analysis

ChIP–seq
Data
In this study, sequence data for a total of eight ChIP-seq experiments were analyzed (see Supplemental Table 1). Two libraries each were generated for ChIP and input samples of HOXD13^{wt}, HOXD13^{Q317R}, HOXD13^{Q317K}, and PITX1 with an average of 27,193,329 nonredundant uniquely mappable reads.

Quality filtering and read mapping

Reads with an average Phred score lower than 28 were discarded. Quality filtered reads were mapped against the reference genome *galGal3* using the aligner BWA, allowing up to two mismatches

(Version 0.5.9-r16) (Li and Durbin 2009). Non-uniquely mapped reads were discarded and redundant reads were removed using SAMtools rmdup (Li et al. 2009). The quality of ChIP-enrichment was validated using spp (Kharchenko et al. 2008; Landt et al. 2012).

Principle component analysis (ChIP-seq)

For principle component analysis (PCA) of the coverage profiles of the uniquely mapped reads, the *galGal3* genome was divided into nonoverlapping 500-bp windows. For each sequence data set, the number of uniquely mapped reads for each window was determined, yielding count vectors of length 2,198,570. Q-mode PCA was then performed on the normalized counts of each of the eight experiments with the R function prcomp.

ChIP quality control and peak calling

Quality control and tests for reproducibility of ChIP-seq experiments were performed according to the guidelines suggested by The ENCODE Project Consortium (Li et al. 2011; Landt et al. 2012). Peak calling was performed according to Kundaje (2012) using the MACS2 peak caller (Zhang et al. 2008). Briefly, peaks for each replicate were called against pooled input controls of both biological replicates with low stringency (parameters: -g 1.0×10^9, –p 1×10^{-1},–too-large,–bw, as determined by spp in the quality control step). The final peak set was called from the pooled replicates, sorted according to *P*-value, and the threshold was determined by using the number of peaks above IDR 0.0025 from the pool-consistency IDR-analysis (see Supplemental Table 1).

Motif analysis

Motif analysis was conducted using *DREME* (MEME package, version 4.7.0) (Bailey 2011). Sequences of 150 bp surrounding the summits of the 1000 strongest peaks (sorted by *P*-value) were used as input (parameters: -e 0.05, -m 10, -g 100, -s 1, -v 2, -mink 3, -maxk 10).

Peaks in genomic regions

The *galGal3* genome was subdivided into the following five classes of regions: promoter, exon, intron, gene flanking, and intergenic on the basis of annotation data from Ensembl BioMart (database: Ensemble Genes 66; data set: WASHUC2).

Regions between 5 kb upstream of and 2 kb downstream from the gene start position were defined as promoter regions. For genes shorter than 2 kb, the downstream part of the promoter was omitted. All exon and intron regions that do not overlap with promoter regions were defined as intron and exon regions. Regions between 25 kb and 5 kb upstream of the gene start position, as well as regions between the gene end position and 20 kb downstream, were defined as gene flanking regions. Regions not covered by any of the previously defined regions were defined as intergenic regions. All classes of regions were further subdivided into the subclasses, conserved and not conserved, using the most conserved track for vertebrates of UCSC (phastConsElements7way).

To count peaks in different genomic regions, summit files were intersected with the regions defined above using *intersectBed* of BEDTools (Quinlan and Hall 2010), yielding for all peaks which are not in intergenic regions at least one association with a gene. One summit may be assigned multiple genes because the regions are overlapping to some extent.

seqMINER analysis

We used seqMINER version 1.3.3 (Ye et al. 2011). The peak summits for HOXD13[wt], HOXD13[Q317K], HOXD13[Q317R], and PITX1 were taken as reference coordinates. Subsequently, mean read densities relative to these locations were calculated for merged reads from the two biological replicates, and were calculated for each factor. The graph displays the distribution of mean read densities (tags/50 bp) centered on the peak summits (± 3000 bp).

RNA-seq

The reference transcriptome for *Gallus gallus* was downloaded from the Ensembl release 64 database using BioMart 0.7 on October 7, 2011. The database is based on the WASHUC2/GALGAL3 genome assembly and contains 23,392 transcripts (exon models) for 17,934 genes. The reads for each data set/condition were separately mapped to the reference transcriptome using bowtie (release 0.12.7). Except for the seed length ($l = 40$) and the number of allowed mismatches in the seed region ($m = 3$), the default parameters were used for the alignment. A preceding quality filtering of the reads was skipped due to Bowtie's ability to use Phred quality scores of reads during the alignment step. Reads mapping to multiple transcripts were randomly assigned to one of them. To quantify the transcript expression and normalize between the data sets, the reads per kilobase of exon model per million mapped reads (RPKM) values for all exon models were calculated (Mortazavi et al. 2008) and used for further analysis.

Hierarchical clustering and principle component analysis of expression data

The analysis of expression data was performed for the five RNA-seq data sets with RPKM values for 3118 transcripts showing a minimum RPKM value of 2 in at least one of the experiments with the additional criterion of a fold change of at least two compared to the RPKM values obtained for the control RCASBP vector. Q-mode PCA and hierarchical clustering were applied to the base-2 logarithms of the RPKM fold-change values. The PCA was performed with the R function prcomp (retx = TRUE, center = TRUE, scale = TRUE). The hierarchical clustering and dendrogram calculation was done using the Matlab 2011 function clustergram with the Euclidean metric for pairwise comparison and unweighted average distance for the linkage. GO-analysis was performed using DAVID (Huang et al. 2008, 2009).

Cobinding of TF to differentially coregulated genes

For the 3118 differentially regulated transcripts, the RPKM values were summed up for each gene and condition. RPKM fold changes relative to control cultures were calculated as for transcripts. Genes with a twofold up-regulation for HOXD13[Q317K] and PITX1 or a twofold down-regulation for HOXD13[Q317K] and PITX1 were defined as HOXD13[Q317K]/PITX1 coregulated genes. HOXD13[Q317R]/PITX1 coregulated genes were defined accordingly.

Shared peaks for HOXD13[Q317K] and PITX1 and, accordingly, HOXD13[Q317R] and PITX1 were assigned to genes if they fall in the promoter, the gene body, or gene flanking regions of annotated ENSEMBL genes (see Peaks in Genomic Regions section above). Genes with at least one shared peak for HOXD13[Q317K] and PITX1 were defined as HOXD13[Q317K]-PITX1 cobound genes. HOXD13[Q317R]-PITX1 cobound genes were defined accordingly.

Data access

Data from this study have been submitted to the NCBI Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE44799.

Acknowledgments

References

Ahrens PB, Solursh M, Reiter RS, Singley CT. 1979. Position-related capacity for differentiation of limb mesenchyme in cell culture. *Dev Biol* **69:** 436–450.

Albrecht AN, Kornak U, Böddrich A, Süring K, Robinson PN, Stiege AC, Lurz R, Stricker S, Wanker EE, Mundlos S. 2004. A molecular pathogenesis for transcription factor associated poly-alanine tract expansions. *Hum Mol Genet* **13:** 2351–2359.

Bailey TL. 2011. DREME: Motif discovery in transcription factor ChIP-seq data. *Bioinformatics* **27:** 1653–1659.

Barski A, Zhao K. 2009. Genomic location analysis by ChIP-Seq. *J Cell Biochem* **107:** 11–18.

Berger MF, Badis G, Gehrke AR, Talukder S, Philippakis AA, Peña-Castillo L, Alleyne TM, Mnaimneh S, Botvinnik OB, Chan ET, et al. 2008. Variation in homeodomain DNA binding revealed by high-resolution analysis of sequence preferences. *Cell* **133:** 1266–1276.

Cao Y, Yao Z, Sarkar D, Lawrence M, Sanchez GJ, Parker MH, MacQuarrie KL, Davison J, Morgan MT, Ruzzo WL, et al. 2010. Genome-wide MyoD binding in skeletal muscle cells: A potential for broad cellular reprogramming. *Dev Cell* **18:** 662–674.

Caronia G, Goodman FR, McKeown CME, Scambler PJ, Zappavigna V. 2003. An I47L substitution in the HOXD13 homeodomain causes a novel human limb malformation by producing a selective loss of function. *Development* **130:** 1701–1712.

Chaney BA, Clark-Baldwin K, Dave V, Ma J, Rance M. 2005. Solution structure of the K50 class homeodomain PITX2 bound to DNA and implications for mutations that cause Rieger syndrome. *Biochemistry* **44:** 7497–7511.

Debeer P, Bacchelli C, Scambler PJ, De Smet L, Fryns JP, Goodman FR. 2002. Severe digital abnormalities in a patient heterozygous for both a novel missense mutation in HOXD13 and a polyalanine tract expansion in HOXA13. *J Med Genet* **39:** 852–856.

Delpretti S, Zakany J, Duboule D. 2012. A function for all posterior Hoxd genes during digit development? *Dev Dyn* **241:** 792–802.

Dorman CM, Johnson SE. 1999. Activated Raf inhibits avian myogenesis through a MAPK-dependent mechanism. *Oncogene* **18:** 5167–5176.

The ENCODE Project Consortium. 2011. A user's guide to the encyclopedia of DNA elements (ENCODE). *PLoS Biol* **9:** e1001046.

Etchevers H. 2011. Primary culture of chick, mouse or human neural crest cells. *Nat Protoc* **6:** 1568–1577.

Fraenkel E, Rould MA, Chambers KA, Pabo CO. 1998. Engrailed homeodomain-DNA complex at 2.2 Å resolution: A detailed view of the interface and comparison with other engrailed structures. *J Mol Biol* **284:** 351–361.

Goff DJ, Tabin CJ. 1997. Analysis of Hoxd-13 and Hoxd-11 misexpression in chick limb buds reveals that Hox genes affect both bone condensation and growth. *Development* **124:** 627–636.

Goodman FR. 2002. Limb malformations and the human HOX genes. *Am J Med Genet* **112:** 256–265.

Hecht J, Stricker S, Wiecha U, Stiege A, Panopoulou G, Podsiadlowski L, Poustka AJ, Dieterich C, Ehrich S, Suvorova J, et al. 2008. Evolution of a core gene network for skeletogenesis in chordates. *PLoS Genet* **4:** e1000025.

Huang DW, Sherman BT, Lempicki RA. 2008. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4:** 44–57.

Huang DW, Sherman BT, Lempicki RA. 2009. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* **37:** 1–13.

Johnson D, Kan S-H, Oldridge M, Trembath RC, Roche P, Esnouf RM, Giele H, Wilkie AOM. 2003. Missense mutations in the homeodomain of HOXD13 are associated with brachydactyly types D and E. *Am J Hum Genet* **72:** 984–997.

Junion G, Spivakov M, Girardot C, Braun M, Gustafson EH, Birney E, Furlong EEM. 2012. A transcription factor collective defines cardiac cell fate and reflects lineage history. *Cell* **148:** 473–486.

Kharchenko PV, Tolstorukov MY, Park PJ. 2008. Design and analysis of ChIP-seq experiments for DNA-binding proteins. *Nat Biotechnol* **26:** 1351–1359.

Kozlowski K, Walter MA. 2000. Variation in residual PITX2 activity underlies the phenotypic spectrum of anterior segment developmental disorders. *Hum Mol Genet* **9:** 2131–2139.

Kundaje A. 2012. Irreproducibility Discovery Rate (IDR): A robust method for automated adaptive thresholding and reproducibility analysis of ChIP-seq data. https://sites.google.com/site/anshulkundaje/projects/idr.

Kuss P, Villavicencio-Lorini P, Witte F, Klose J, Albrecht AN, Seemann P, Hecht J, Mundlos S. 2009. Mutant Hoxd13 induces extra digits in a mouse model of synpolydactyly directly and by decreasing retinoic acid synthesis. *J Clin Invest* **119:** 146–156.

Landt SG, Marinov GK, Kundaje A, Kheradpour P, Pauli F, Batzoglou S, Bernstein BE, Bickel P, Brown JB, Cayting P, et al. 2012. ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res* **22:** 1813–1831.

Lee TI, Johnstone SE, Young RA. 2006. Chromatin immunoprecipitation and microarray-based analysis of protein location. *Nat Protoc* **1:** 729–748.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25:** 1754–1760.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup . 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25:** 2078–2079.

Li Q, Brown JB, Huang H, Bickel PJ. 2011. Measuring reproducibility of high-throughput experiments. *Annals of Applied Statistics* **5:** 1752–1779.

Logan M, Tabin C. 1998. Targeted gene misexpression in chick limb buds using avian replication-competent retroviruses. *Methods* **14:** 407–420.

Logan M, Tabin CJ. 1999. Role of Pitx1 upstream of Tbx4 in specification of hindlimb identity. *Science* **283:** 1736–1739.

Luo G, Ducy P, McKee MD, Pinero GJ, Loyer E, Behringer RR, Karsenty G. 1997. Spontaneous calcification of arteries and cartilage in mice lacking matrix GLA protein. *Nature* **386:** 78–81.

MacQuarrie KL, Fong AP, Morse RH, Tapscott SJ. 2011. Genome-wide transcription factor binding: Beyond direct target regulation. *Trends Genet* **27:** 141–148.

McCobb D, Best P, Beam K. 1990. The differentiation of excitability in embryonic chick limb motoneurons. *J Neurosci* **10:** 2974–2984.

Morgan BA, Fekete DM. 1996. Manipulating gene expression with replication-competent retroviruses. *Methods Cell Biol* **51:** 185–218.

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* **5:** 621–628.

Muragaki Y, Mundlos S, Upton J, Olsen BR. 1996. Altered growth and branching patterns in synpolydactyly caused by mutations in HOXD13. *Science* **272:** 548–551.

Otto F, Kanegane H, Mundlos S. 2002. Mutations in the RUNX2 gene in patients with cleidocranial dysplasia. *Hum Mutat* **19:** 209–216.

Ouyang Z, Zhou Q, Wong WH. 2009. ChIP-Seq of transcription factors predicts absolute and differential gene expression in embryonic stem cells. *Proc Natl Acad Sci* **106:** 21521–21526.

Park PJ. 2009. ChIP-seq: Advantages and challenges of a maturing technology. *Nat Rev Genet* **10:** 669–680.

Percival-Smith A, Muller M, Affolter M, Gehring WJ. 1990. The interaction with DNA of wild-type and mutant fushi tarazu homeodomains. *EMBO J* **9:** 3967–3974.

Quinlan AR, Hall IM. 2010. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26:** 841–842.

Robertson AG, Bilenky M, Tam A, Zhao Y, Zeng T, Thiessen N, Cezard T, Fejes AP, Wederell ED, Cullum R, et al. 2008. Genome-wide relationship between histone H3 lysine 4 mono- and tri-methylation and transcription factor binding. *Genome Res* **18:** 1906–1917.

Saleem RA, Banerjee-Basu S, Berry FB, Baxevanis AD, Walter MA. 2003. Structural and functional analyses of disease-causing missense mutations in the forkhead domain of FOXC1. *Hum Mol Genet* **12:** 2993–3005.

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

Schier AF, Gehring WJ. 1992. Direct homeodomain-DNA interaction in the autoregulation of the fushi tarazu gene. *Nature* **356:** 804–807.

Schmidt D, Wilson MD, Ballester B, Schwalie PC, Brown GD, Marshall A, Kutter C, Watt S, Martinez-Jimenez CP, Mackay S, et al. 2010. Five-vertebrate ChIP-seq reveals the evolutionary dynamics of transcription factor binding. *Science* **328:** 1036–1040.

Solursh M, Ahrens PB, Reiter RS. 1978. A tissue culture analysis of the steps in limb chondrogenesis. *In Vitro* **14:** 51–61.

Spitz F, Furlong EEM. 2012. Transcription factors: From enhancer binding to developmental control. *Nat Rev Genet* **13:** 613–626.

Svingen T, Tonissen KF. 2006. Hox transcription factors and their elusive mammalian gene targets. *Heredity* **97:** 88–96.

Szeto DP, Rodriguez-Esteban C, Ryan AK, O'Connell SM, Liu F, Kioussi C, Gleiberman AS, Izpisúa-Belmonte JC, Rosenfeld MG. 1999. Role of the Bicoid-related homeodomain factor Pitx1 in specifying hindlimb morphogenesis and pituitary development. *Genes Dev* **13:** 484–494.

Tremblay JJ, Lanctôt C, Drouin J. 1998. The pan-pituitary activator of transcription, Ptx1 (pituitary homeobox 1), acts in synergy with SF-1 and Pit1 and is an upstream regulator of the Lim-homeodomain gene Lim3/Lhx3. *Methods Enzymol* **12:** 428–441.

Valouev A, Johnson DS, Sundquist A, Medina C, Anton E, Batzoglou S, Myers RM, Sidow A. 2008. Genome-wide analysis of transcription factor binding sites based on ChIP-Seq data. *Nat Methods* **5:** 829–834.

Villavicencio-Lorini P, Kuss P, Friedrich J, Haupt J, Farooq M, Türkmen S, Duboule D, Hecht J, Mundlos S. 2010. Homeobox genes d11-d13 and a13 control mouse autopod cortical bone and joint formation. *J Clin Invest* **120:** 1994–2004.

Wang J, Kumar RM, Biggs VJ, Lee H, Chen Y, Kagey MH, Young RA, Abate-Shen C. 2011. The Msx1 homeoprotein recruits Polycomb to the nuclear periphery during development. *Dev Cell* **21:** 575–588.

Welboren W-J, van Driel MA, Janssen-Megens EM, van Heeringen SJ, Sweep FCGJ, Span PN, Stunnenberg HG. 2009. ChIP-Seq of ERα and RNA polymerase II defines genes differentially responding to ligands. *EMBO J* **28:** 1418–1428.

Wilkie AOM, Tang Z, Elanko N, Walsh S, Twigg SRF, Hurst JA, Wall SA, Chrzanowska KH, Maxson RE. 2000. Functional haploinsufficiency of the human homeobox gene MSX2 causes defects in skull ossification. *Nat Genet* **24:** 387–390.

Yao Y, Jumabay M, Wang A, Boström KI. 2011. Matrix Gla protein deficiency causes arteriovenous malformations in mice. *J Clin Invest* **121:** 2993–3004.

Ye T, Krebs AR, Choukrallah M-A, Keime C, Plewniak F, Davidson I, Tora L. 2011. seqMINER: An integrated ChIP-seq data interpretation platform. *Nucleic Acids Res* **39:** e35.

Zakany J, Kmita M, Duboule D. 2004. A dual role for Hox genes in limb anterior-posterior asymmetry. *Science* **304:** 1669–1672.

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9:** R137.

Zhao C, Dave V, Yang F. 2000. Target selectivity of bicoid is dependent on nonconsensus site recognition and protein-protein interaction. *Mol Cell Biol* **20:** 8112–8123.

Zhao X, Sun M, Zhao J, Leyva JA, Zhu H, Yang W, Zeng X, Ao Y, Liu Q, Liu G, et al. 2007. Mutations in HOXD13 underlie syndactyly type V and a novel brachydactyly-syndactyly syndrome. *Am J Hum Genet* **80:** 361–371.

Distinct global shifts in genomic binding profiles of limb malformation-associated *HOXD13* mutations

Daniel M. Ibrahim, Peter Hansen, Christian Rödelsperger, et al.

Supplemental Material	http://genome.cshlp.org/content/suppl/2013/10/04/gr.157610.113.DC1.html
References	This article cites 63 articles, 23 of which can be accessed free at: http://genome.cshlp.org/content/23/12/2091.full.html#ref-list-1
Creative Commons License	This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see http://genome.cshlp.org/site/misc/terms.xhtml). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 3.0 Unported), as described at http://creativecommons.org/licenses/by-nc/3.0/.
Email Alerting Service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**

To subscribe to *Genome Research* go to:
http://genome.cshlp.org/subscriptions