RESEARCH ARTICLE

# *LncRNA-HIT* Functions as an Epigenetic Regulator of Chondrogenesis through Its Recruitment of p100/CBP Complexes

Hanqian L. Carlson[1], Jeffrey J. Quinn[2], Yul W. Yang[2], Chelsea K. Thornburg[3], Howard Y. Chang[2], H. Scott Stadler[1,4]*

1 Skeletal Biology Program, Shriners Hospitals for Children, Portland, Oregon, United States of America,
2 Program in Epithelial Biology, Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, California, United States of America, 3 Department of Biochemistry & Molecular Biology, Michigan State University, East Lansing, Michigan, United States of America, 4 Department of Molecular and Medical Genetics, Oregon Health & Science University, Portland, Oregon, United States of America

* hss@shcc.org

## Abstract

Gene expression profiling in E 11 mouse embryos identified high expression of the long noncoding RNA (lncRNA), *LNCRNA-HIT* in the undifferentiated limb mesenchyme, gut, and developing genital tubercle. In the limb mesenchyme, *LncRNA-HIT* was found to be retained in the nucleus, forming a complex with p100 and CBP. Analysis of the genome-wide distribution of *LncRNA-HIT*-p100/CBP complexes by ChIRP-seq revealed *LncRNA-HIT* associated peaks at multiple loci in the murine genome. Ontological analysis of the genes contacted by *LncRNA-HIT*-p100/CBP complexes indicate a primary role for these loci in chondrogenic differentiation. Functional analysis using siRNA-mediated reductions in *LncRNA-HIT* or p100 transcripts revealed a significant decrease in expression of many of the *LncRNA-HIT*-associated loci. *LncRNA-HIT* siRNA treatments also impacted the ability of the limb mesenchyme to form cartilage, reducing mesenchymal cell condensation and the formation of cartilage nodules. Mechanistically the *LncRNA-HIT* siRNA treatments impacted pro-chondrogenic gene expression by reducing H3K27ac or p100 activity, confirming that *LncRNA-HIT* is essential for chondrogenic differentiation in the limb mesenchyme. Taken together, these findings reveal a fundamental epigenetic mechanism functioning during early limb development, using *LncRNA-HIT* and its associated proteins to promote the expression of multiple genes whose products are necessary for the formation of cartilage.

## Author Summary

A fundamental problem studied by skeletal biologists is the development of regenerative therapies to replace cartilage tissues impacted by injury or disease, which for individuals affected by osteoarthritis represents nearly half of all of all adults over the age of sixty five. To date, no therapies exist to promote sustained cartilage regeneration, as we have not

been able to recapitulate the programming events necessary to instruct cells to form articular cartilage without these cells continuing to differentiate into bone. Our analysis of the early programming events occurring during cartilage formation led to the identification of *LncRNA-HIT* a long noncoding RNA that is essential for the differentiation of the embryonic limb mesenchyme into cartilage. A genome wide analysis of *LncRNA-HIT's* distribution in the mesenchyme revealed strong association between *LncRNA-HIT* and numerous genes whose products facilitate cartilage formation. In the absence of *LncRNA-HIT*, the expression of these chondrogenic genes is severely reduced, impacting the differentiation of these cells into cartilage. Mechanistically, *LncRNA-HIT* regulates these pro-chondrogenic genes by recruiting p100 and CBP to these loci, facilitating H3K27ac and transcriptional activation. *LncRNA-HIT* also appears to be present in most vertebrate species, suggesting that the epigenetic program regulated by this lncRNA may represent a fundamental mechanism used by many species to promote cartilage formation.

## Introduction

In the animal kingdom, embryogenesis proceeds through the coordinated expression of genes whose products mediate the formation of complex tissues and structures. While proteins encoded by mRNAs contribute extensively to the regulation of these developmental processes, recent studies of the human and mouse genomes suggest that long noncoding RNAs (lncRNAs) play an essential role in coordinating the expression of genes required for tissue formation and organ development [1–5].

As gene regulatory molecules, lncRNAs modulate target gene expression using a variety of mechanisms, particularly in the nucleus where they can function as decoys, scaffolds, guides, or even enhancers [6]. As decoys, lncRNAs can titrate away transcription factors (*GAS5*-Glucocorticoid receptor), the transcriptional machinery (*DHFR minor*-TFIIB) or even splicing factors (*MALAT1*- SR splicing factors) to modulate the expression of target genes [7–9]. As a component of their secondary structures, lncRNAs such as *HOTAIR*, *ANRIL*, and *Kcnq1ot1* provide protein-specific scaffolds, assembling enzyme complexes such as the polycomb repressive complexes 1 or 2 (PRC1, PRC2), LSD-1, and CoREST-HDAC to facilitate changes in histone lysine methylation or acetylation to enforce the transcriptional state of a specific locus [10–14]. Nuclear lncRNAs may also function as guides to localize protein complexes to specific chromosomal regions. Notably, the lncRNA, *XIST*, functions as a guide to recruit proteins such as YY1 or components of PRC2 to promote X chromosome inactivation [15–19]. Finally a growing body of evidence indicates that lncRNAs may also function as enhancer RNAs, using chromosomal looping to place proteins bound by the lncRNA proximal to genes to facilitate their regulation [5, 20–22].

A key group of genes regulated by lncRNAs are the Hox genes, a conserved family of developmental transcription factors that exhibit temporally- and spatially-restricted domains of expression and function [5, 23–29]. Evidence for the unique functions of the vertebrate *Hox* lncRNAs was first shown with the *trans*-acting functions of *HOTAIR*, which is expressed from the HoxC locus to recruit PRC2 to the HOXD cluster, where it mediates H3K27 methylation to repress the expression of several 5' HoxD genes [4]. During development, the 5' HoxD proteins regulate limb and axial skeleton development, suggesting that de-repression of the 5' HoxD genes, by perturbations in *HOTAIR* expression, would most likely affect these same skeletal elements [30–32]. Recent studies confirm this hypothesis, as mice lacking *HOTAIR* exhibit malformations of carpal/metacarpal elements in the limb as well as homeotic transformations of

the lumbar, sacral, and caudal vertebrae, phenotypes consistent with de-repression of *Hoxd11* and *Hoxd13* [29].

Proximal to the 5' HoxA gene cluster is the lncRNA *HOTTIP*, which functions as an enhancer lncRNA to regulate the expression of 5' HoxA genes to control the growth and elongation of zeugopod and autopod skeletal elements [5]. Mechanistically, *HOTTIP* modulates gene expression by chromosomal looping, placing its recruited Trithorax/WDR5/MLL protein complexes proximal to 5' HoxA genes to facilitate H3K4me3 and gene expression [5,33]. *HOTTIP* function has also been associated with endochondral ossification [34]. This finding in conjunction with the functional studies of *HOTAIR* indicate a role for lncRNAs associated with the vertebrate Hox genes as necessary components for the development, patterning, and maturation of skeletal tissues [5,29,34]. Interestingly, a second lncRNA, *LncRNA-HIT*, has also been identified within the HOXA locus [35]. While initially characterized as a TGFβ-induced modulator of epithelial-mesenchymal transformations in tumor cells [35] our subsequent analysis of this lncRNA indicates that it is also expressed in the early limb where we hypothesized a role for this lncRNA in mediating chondrogenic differentiation. Characterization of subcellular localization of *LncRNA-HIT* by RNA FISH indicated the transcript is predominantly localized to the nucleus where it associates with p100/CBP complexes, suggesting that *LncRNA-HIT* may regulate gene expression by recruiting modulators of H3K27 acetylation. To identify the genes most likely to be regulated by *LncRNA-HIT*, ChIRP-seq was performed using pre-chondrogenic limb bud tissue. From this analysis *LncRNA-HIT* was found to be associated within 25 kb of numerous pro-chondrogenic genes including: *Bmpr1b*, *Gli2*, *Col14a1*, *Adam17*, *Kdelr2*, *Pik3cb*, *Hoxa13*, *Hoxa11*, *Ncam1*, and *Gpc6*. Loss of function analyses confirmed *LncRNA-HIT*'s role as a modulator of pro-chondrogenic expression as, H3K27ac, and near peak gene expression were significantly reduced by *LncRNA-HIT*-specific siRNA treatments in the limb mesenchyme. Moreover, chondrogenic differentiation was also significantly reduced in limb mesenchyme treated with siRNAs targeting either *LncRNA-HIT* or the p100 transcript, confirming that the *LncRNA-HIT* and its recruited protein complex are necessary to maintain H3K27ac and chondrogenic gene expression which facilitates differentiation. Combined, these findings identify an epigenetic mechanism functioning during limb skeletogenesis, using an lncRNA to coordinate the expression genes necessary to direct undifferentiated limb mesenchyme towards a chondrogenic state.

## Results

### Developmental expression of *LncRNA-HIT* reveals both distinct and similar domains of expression with 5' *HOX* genes

The *LncRNA-HIT* transcript was first identified as the full length cDNA, *9530018H14RIK*, by the RIKEN Mouse Gene Encyclopedia Project and was mapped as a single exon to mouse chromosome 6 between *Hoxa11* and *Hoxa13* by genome sequencing (S1 Fig) [36]. Conservation analysis of the *LncRNA-HIT* cDNA sequence using BLAST (NCBI) revealed a single 253 bp region present in most vertebrate species (S1 Fig). Characterization of *LncRNA-HIT*'s protein coding potential using the Coding Potential Assessment Tool (CPAT) revealed multiple stop codons in all six reading frames and a protein coding probability of 0.084, well below the 0.44 threshold predicted for protein coding genes (S1 and S2 Figs) (CPAT version 1.2; http://rna-cpat.sourceforge.net, [37]). In comparison, CPAT analysis of neighboring genes, *Hottip* and *Hoxa13*, revealed scores of 0.049 and 0.99 respectively, confirming the poor coding potential of the lncRNA *Hottip* as well as the favorable protein coding potential of *Hoxa13* (S2 Fig). Finally analysis of the potential initiation codons present in the *LncRNA-HIT* transcript also suggested poor protein coding potential, as these sites lack a Kozak consensus sequence and exhibited
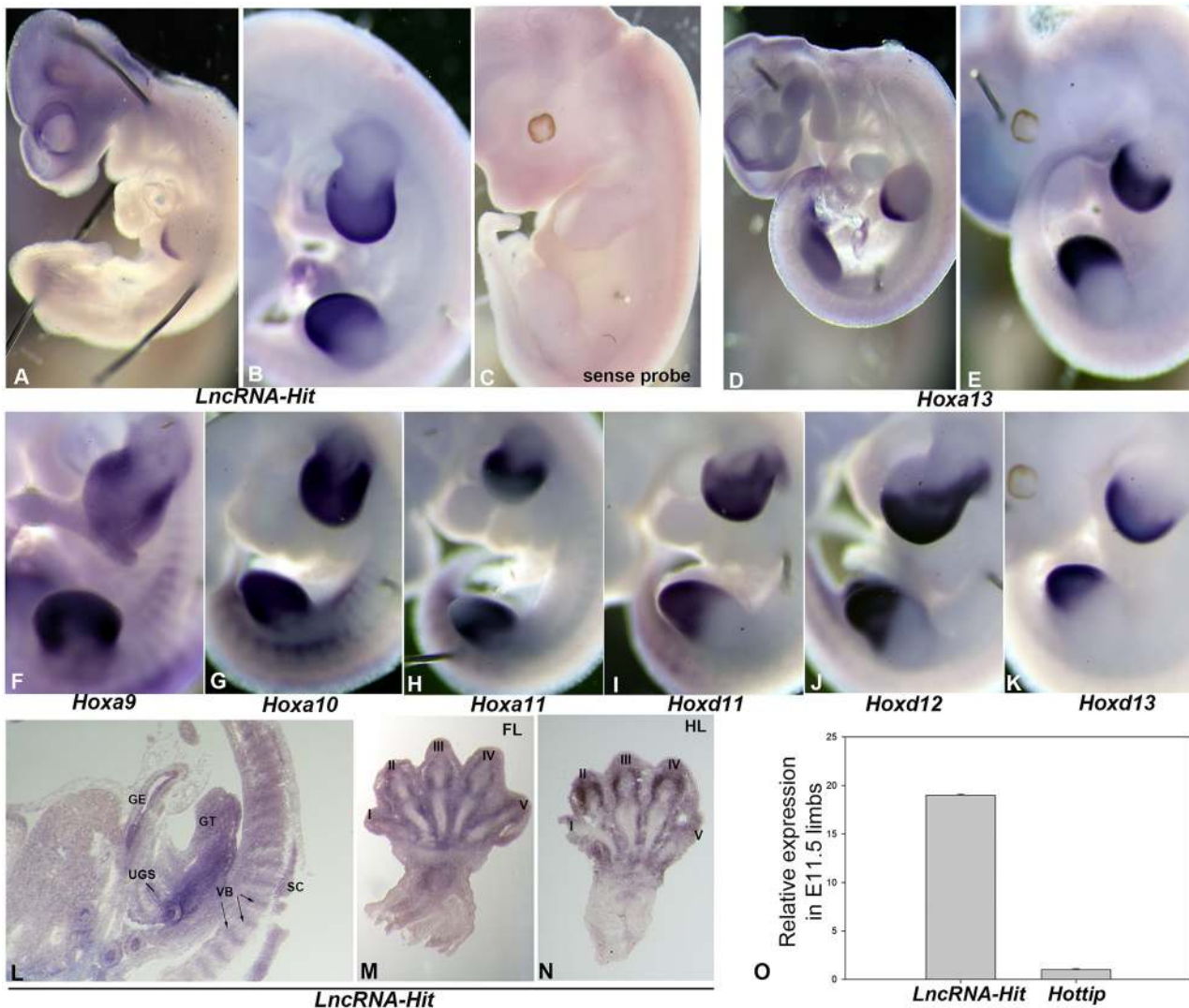
**Fig 1. Expression analysis of the lncRNA *LncRNA-HIT*. (A)** *In situ* hybridization using an antisense *LncRNA-HIT* riboprobe detects the transcript as early as E 10.5 in the distal limb which expands throughout the limb bud at E 11.5 **(B)**. **(C)** *In situ* hybridization using a sense orientation *LncRNA-HIT* riboprobe detects no *LncRNA-HIT* transcripts, confirming the unidirectional transcription of *LncRNA-HIT* in the same orientation as the 5' HoxA genes in the developing limb. **(D and E)** Analysis of *Hoxa13* expression shows a high level of overlap with *LncRNA-HIT* in the distal limb. **(F-H)** Analysis of *Hoxa9-Hoxa11* expression in the E 11.5 distal limb reveals some overlap with *LncRNA-HIT* in the limb bud. **(I-K)** Analysis of *Hoxd11-Hoxd13* expression in the E 11.5 distal limb reveals some overlap with *LncRNA-HIT*. **(L)** *LncRNA-HIT* is also expressed in the developing genital tubercle, gut epithelium, urogenital sinus, spinal cord, and vertebral bodies at E 13.5. **(M-N)** *LncRNA-HIT* expression is detected in the digit perichondrial tissues, digit joint fields, and in the developing carpal/tarsal skeletal elements in E 13.5 forelimbs and hindlimbs. **(O)** Relative fold expression of *LncRNA-HIT* and *Hottip* expression in the E 11.5 limbs as determined by qRTPCR. Values represent the *Gapdh* normalized average expression of *LncRNA-HIT* and Hottip in the limb calculated from three independent analyses. Bars represent the standard deviation of the mean from the three independent assays. UGS = urogenital sinus, VB = vertebral body, GT = genital tubercle, GE = gut epithelium, SC = spinal cord.

doi:10.1371/journal.pgen.1005680.g001

poor translational potential as determined by the NetStart software package (S2 Fig), http://www.cbs.dtu.dk/services/NetStart/, [38–39].

Localization of the *LncRNA-HIT* transcript by *in situ* hybridization revealed expression in the pre-chondrogenic limb mesenchyme as early as embryonic day (E) 10.5 (Fig 1A). By E 11.5 *LncRNA-HIT* expression is expanded to a greater portion of the limb bud, encompassing the majority of the pre-chondrogenic tissues that normally express members of the 5' HoxA and D gene clusters including: *Hoxa9-13* and *Hoxd11-13* (Fig 1B). By E 13.5, *LncRNA-HIT* expression

continued to follow the expression pattern of *Hoxa13*; particularly in the limb perichondrial tissues, gut, genital tubercle, and urogenital sinus (Fig 1L–1N) [27, 40–42]. *In situ* hybridization using the same *LncRNA-HIT* sequence transcribed in the opposite orientation revealed no expression in the limb or other embryonic tissues, confirming reports [43] of its unidirectional transcription in the same orientation as the HoxA genes (Fig 1C, S1 Fig). Finally, quantitation of *LncRNA-HIT* expression in the E 11.0 limb confirmed that it is highly expressed (Ct = 21.4 ± 0.8) at levels nearly nineteen-fold greater than *Hottip*, which resides approximately 5 kb from *LncRNA-HIT* on mouse chromosome 6 (Fig 1O, S1 Fig).

## The *LncRNA-HIT* transcript is localized to the nucleus where it binds p100/CBP complexes

A recent analysis of the subcellular localization of 61 lncRNAs revealed a majority localized to the nucleus, recruiting proteins to mediate histone modifications, chromatin accessibility, and gene expression [44]. To gain insight into the potential function of the *LncRNA-HIT*, we first examined its sequence composition, which identified the same nuclear retention motif present in the *BORG* lncRNA [45] (Fig 2, S1 Fig). Analysis of *LncRNA-HIT's* subcellular localization by single molecule RNA FISH detected the lncRNA in the nucleus where it was distributed diffusely and in larger foci in the limb bud mesenchyme (Fig 2A–2D). In contrast, *Gapdh* transcripts localized primarily to the cytoplasm in these cells, a finding consistent with the protein-coding function of the *Gapdh* transcript (Fig 2E and 2F). Finally, pre-treatment of the fixed limb mesenchyme with RNase A resulted in the complete loss of detection of the *LncRNA-HIT* and *Gapdh* signals, indicating that the signals detected by each FISH probe set corresponds to RNA hybridization rather than hybridization to the corresponding sequence present in the chromosomal DNA. (Fig 2G and 2H).

To determine which proteins interact with *LncRNA-HIT* in the nucleus, we first tested whether it functions similarly to its chromosomal neighbor, *Hottip*, which recruits WDR5 to facilitate H3K4me3 [5, 46]. Immunoprecipitation of WDR5 from Flag-WDR5 293T cells transfected with an *LncRNA-HIT* expression vector revealed no enrichment of the *LncRNA-HIT* transcript, suggesting a separate mechanism for *LncRNA-HIT* function in the nucleus (S3 Fig). To identify the proteins binding to *LncRNA-HIT*, an RNA affinity assay was used to isolate limb proteins preferentially binding to the *LncRNA-HIT* transcript. Mass spectroscopy analysis of the U1 control and *LncRNA-HIT* elution fractions identified multiple proteins in the *LncRNA-HIT* and U1 control elution fractions (S4 Fig). As an initial filter, proteins common to both the U1 and *LncRNA-HIT* elution fractions were excluded from subsequent analysis. Several proteins exclusive to the *LncRNA-HIT* elution fractions were also excluded from subsequent analysis as they could not be reproducibly detected in replicate elution fractions. After these initial exclusions, a single protein was identified to be consistently enriched (> 23-fold average enrichment) in only the *LncRNA-HIT* elution fractions. The single protein was identified as p100, a 100 Kd transcriptional co-factor that partners with creb binding protein (CBP) to recruit histone acetyltransferase activity to the STAT6 locus (Fig 3, S4 Fig, and Methods) [47–50]. Analysis of *Snd1* expression which encodes the p100 protein confirmed that p100 is co-expressed with *LncRNA-HIT* in many of the same embryonic regions including the undifferentiated fore- and hindlimb mesenchyme as well as in the developing genital tubercle (Fig 3A–3D). We next evaluated whether p100 and CBP form a complex with endogenous *LncRNA-HIT* in the limb mesenchyme. Western blot analysis of the limb bud proteins co-precipitating with biotinylated-*LncRNA-HIT* revealed enrichment of p100 and CBP, suggesting that both proteins may form a complex with *LncRNA-HIT* (Fig 3E and 3F). Testing this hypothesis, we examined whether endogenous *LncRNA-HIT* present in the limb mesenchyme
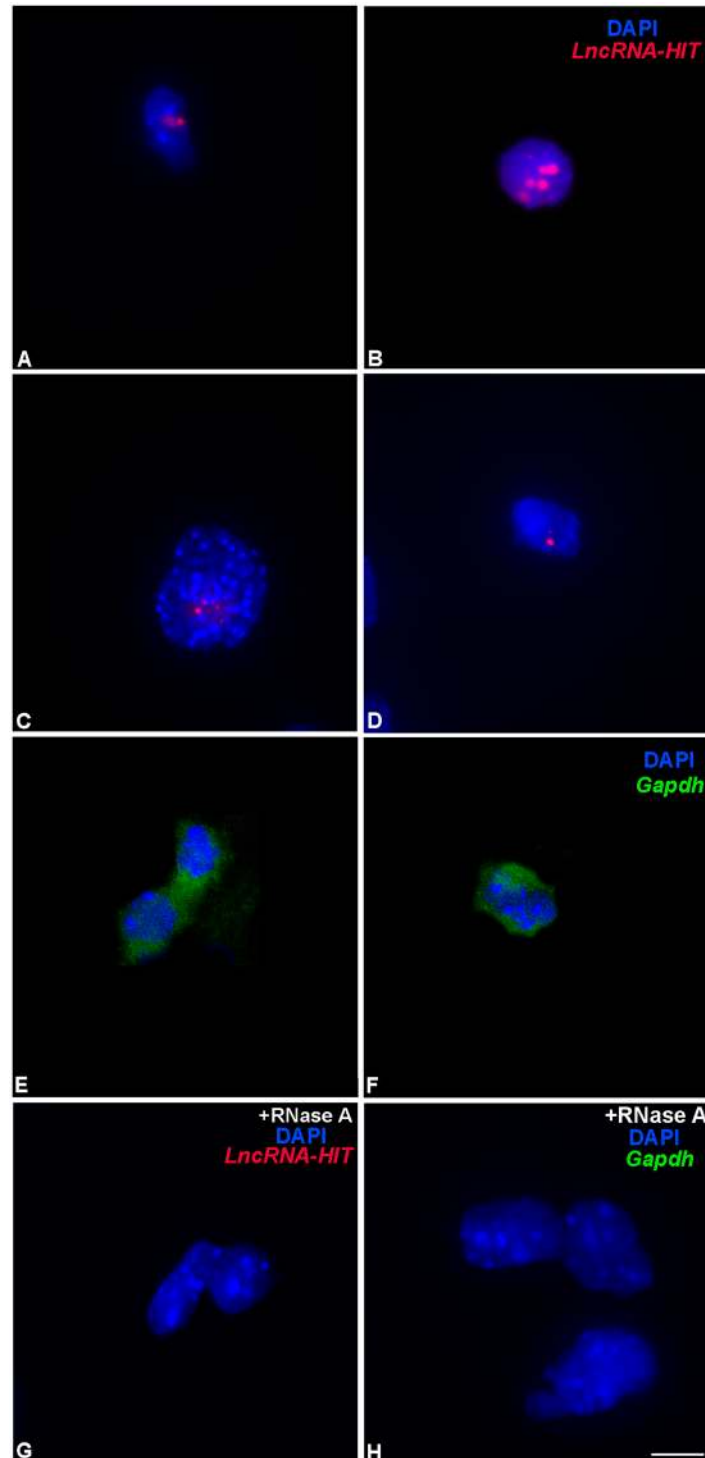
**Fig 2. Subcellular localization of *LncRNA-HIT* and *Gapdh* transcripts in limb mesenchyme using single molecule RNA fluorescent in situ hybridization (FISH). (A-D)** *LncRNA-HIT* transcripts are detected in the nucleus of undifferentiated limb mesenchyme at E 11.0. Red signal = detection of *LncRNA-HIT* probe sets labeled with CalFluor610. Blue signal = DAPI staining of the nuclear DNA. **(E** and **F)** *Gapdh* transcripts are detected in the cytoplasm of undifferentiated limb mesenchyme. Green signal = detection of Gaph probe sets labeled with CalFluor610 and pseudo-colored green. Blue signal = DAPI staining of the nuclear DNA. **(G)** Negative control using RNase A prior to hybridization with the *LncRNA-HIT* probe sets reveals no detected *LncRNA-HIT* (red signal) in the nucleus confirming the detected signal in panels A-D represent hybridization

with the *LncRNA-HIT* transcript. The nuclear DNA was unaffected by the RNase A treatment and stained positively with DAPI (blue signal). **(H)** Negative control using RNase A prior to hybridization with the *Gapdh* probe sets reveals no detected *Gapdh* transcript (green signal) in the cytoplasm confirming the detected signal in panels E and F represent hybridization with the *Gapdh transcript*. The nuclear DNA was unaffected by the RNase A treatment and stains positively with DAPI (blue signal). Bar = 10 μm.

would co-precipitate with p100 and CBP. Immunoprecipitation of p100 and CBP from limb bud mesenchyme revealed a consistent enrichment of *LncRNA-HIT* (> 2-fold) compared to parallel precipitations using IgG, confirming that the endogenous lncRNA forms a complex with p100 and CBP in the pre-chondrogenic limb mesenchyme (Fig 3G).
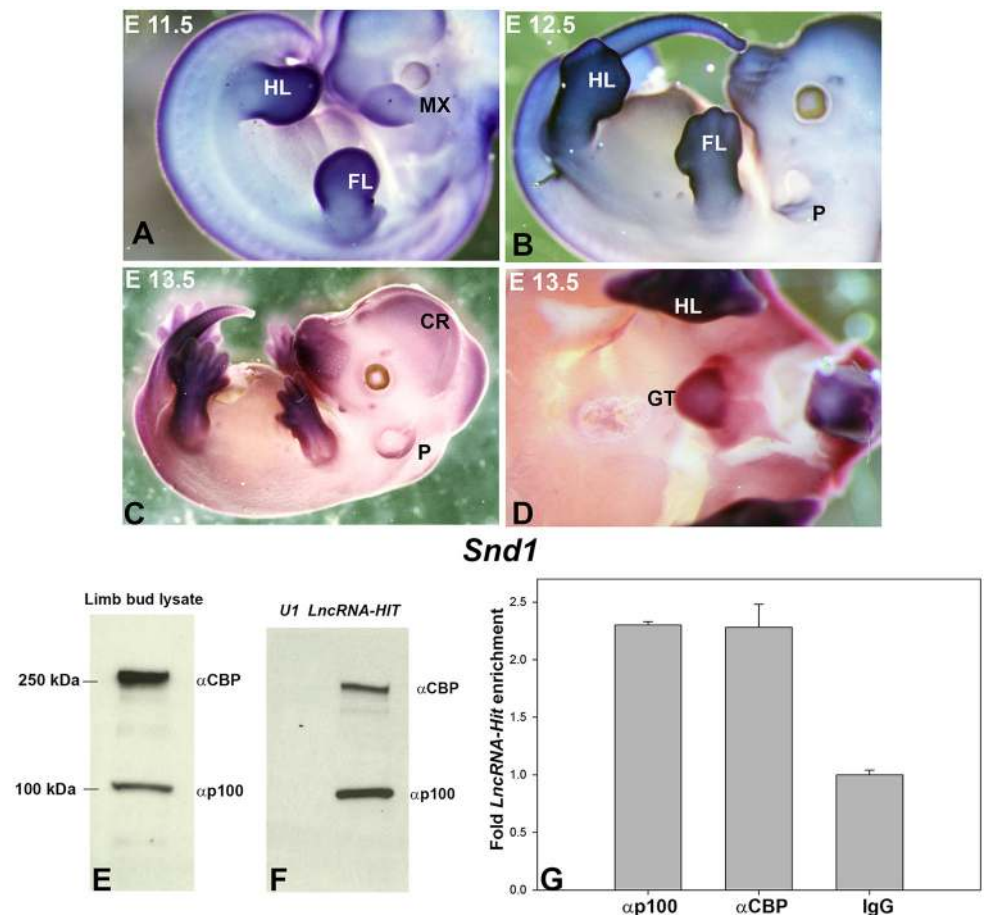


**Fig 3. *LncRNA-HIT* recruits a complex of p100 and CBP in the limb. (A-D)** *Snd1* which encodes p100 is expressed in many of the same regions as *LncRNA-HIT* including the limbs at E11-13.5, spinal cord, and genital tubercle. FL = forelimb, HL = hindlimb, MX = maxillary component of the first branchial arch, SC = spinal cord, P = pinnae. **(E)** Western blot analysis using antibodies specific for CBP and p100 confirm the presence of both proteins in cell lysates from E 11.5 limb buds. **(F)** Streptavidin-mediated precipitation of biotinylated *LncRNA-HIT* and *U1* RNA transcripts reveals co-precipitation of CBP and p100 from limb mesenchyme, confirming recruitment of both proteins by the lncRNA in the limb mesenchyme. **(G)** RIP analysis of p100 and CBP from the limb mesenchyme reveals co-precipitation of endogenous *LncRNA-HIT* transcript, confirming that CBP, p100, and *LncRNA-HIT* are present in a complex in the limb bud mesenchyme. Values represent the mean fold enrichment of *LncRNA-HIT* after precipitation with antibodies specific for CBP, p100 from three independent assays. Control precipitations were performed in parallel using murine IgG. Bars represent the standard deviation of the mean from the three independent assays.

## LncRNA-HIT contacts multiple chondrogenic loci in the developing limb mesenchyme

The nuclear retention of *LncRNA-HIT* as well as its ability to bind p100/CBP complexes suggested the lncRNA may function at the chromosomal level to modulate gene expression in the developing limb. To identify the chromosomal regions contacted by *LncRNA-HIT*, ChIRP-seq was used to precipitate chromatin fractions associated with the lncRNA from E 11.0 limb bud mesenchyme.

The analysis pipeline previously described [51] was used to process the *LncRNA-HIT*-associated peaks as determined by ChIRP-seq. DNA isolated from the precipitated chromatin was submitted to Elim Biopharm (Hayward, CA) for library preparation and next generation sequencing. The sequenced fragments were assembled against murine genome (NCBI37/mm9) using Bowtie and peaks were ranked by MACS using their assigned p-value [52–53]. Visual inspection of the ranked peaks using the UCSC genome browser revealed a marked drop-off in peak signal strength beyond peak 775 (p-value = 5.9 x $10^{-57}$). Based on this result, the top 775 peaks were selected to identify candidate genes potentially regulated by *LncRNA-HIT*. From the initial cohort of 775 peaks, 173 peaks were excluded from subsequent studies as they were also present in parallel assays using the same *LncRNA-HIT* probe sets and a glial cell line control that does not express *LncRNA-HIT*. The remaining 602 peaks were evaluated for their association with *cis*-regulatory elements in the murine genome using the Genomic Regions of Enrichment of Annotations software (GREAT) (S5 Fig) [54]. Interestingly, while 588 peaks were identified as associating with one or more *cis*-regulatory elements, a correlative function for these associated *cis*-regulatory regions was not identified by the GREAT software package (S5 Fig). Next, because GREAT provides only a computational prediction of function, a second analysis was performed focusing on high confidence peak-to-gene relationships using peaks mapping within 25 kb of a known gene. After this analysis, 42 peak-associated candidate genes were identified as potential targets for regulation by *LncRNA-HIT*-p100/CBP complexes (Fig 4). To validate the regulation of these genes by *LncRNA-HIT*-p100/CBP complexes, E 11.0 limb mesenchymal cells were transfected with two siRNAs specific for the *LncRNA-HIT* transcript (-6.15-fold knockdown, Fig 4) and evaluated for changes in gene expression in three independent assays using qRTPCR. From this analysis, 28 near-peak genes exhibited decreased expression in response to the *LncRNA-HIT* siRNA treatments with fold-change differences ranging from -1.5- to -9.75 compared to transfections using scrambled siRNA controls (Fig 4). The remaining near-peak genes exhibited no change in expression in response to the *LncRNA-HIT* siRNA treatments (6 genes) or could not be detected in the E 11.0 limb mesenchyme by qRTPCR (8 genes) (Fig 4). The 28 genes exhibiting decreased expression in response to the *LncRNA-HIT* siRNA treatments were queried as a group for ontological function using the AmiGO 2 term enrichment software which identified proximal/distal pattern formation and regulation of chondrocyte differentiation (p≤ 0.05) as the most significant ontological categories, suggesting a role for *LncRNA-HIT* in the regulation of genes required for limb chondrogenesis (S6 Fig, and Methods).

To test whether *LncRNA-HIT* regulates chondrogenesis in the limb mesenchyme, an *in vitro* micromass chondrogenesis assay was used, as it recapitulates many of the cellular events occurring during chondrogenesis including mesenchymal cell condensation, cartilage extracellular matrix expression, and cartilage nodule formation and is amenable to siRNA-mediated knockdown of pro-chondrogenic genes including *Runx1, Ctgf, Notch*, and *Angptl4* [27, 55–62]. Transfection of limb mesenchyme with siRNAs specific for *LncRNA-HIT* significantly reduced *LncRNA-HIT* RNA levels by nearly eighty percent, confirming the effectiveness of the siRNAs to target *LncRNA-HIT* for degradation in the micromass assay (n = 6 independent replicates, Fig 5). Analysis of the micromass assays transfected with the *LncRNA-HIT* siRNAs revealed a

| Gene | Fold-change after LncRNA-HIT siRNA | Peak location relative to gene | p-value |
|---|---|---|---|
| Pik3cb | -9.75 (.15) | Intron | 1076.46 |
| LncRNA-HIT | -6.15 (.15) | Nascent | 3100.00 |
| Col14a1 | -5.10 (.30) | Intron | 1414.53 |
| Bmpr1b | -4.75 (.49) | Exon | 562.61 |
| Pbx1 | -3.70 (.10) | Intron | 893.22 |
| D15ertd621e | -3.65 (1.06) | 14 kb upstream | 838.75 |
| Adam17 | -3.15 (.10) | Exon | 899.13 |
| Dtl | -3.00 (.30) | Intro | 1095.14 |
| Fcho2 | -2.93 (.10) | Intron | 1715.00 |
| Hoxa13† | -2.90 (.05) | 1 kb downstream | 3100.00 |
| Arl15 | -2.46 (.25) | Intron | 710.66 |
| Pcsk6 | -2.45 (.49) | Intron | 1234.53 |
| Ell2 | -2.23 (.10) | Exon | 1467.11 |
| Kdelr2 | -2.20 (.14) | 8 kb upstream | 1056.53 |
| Hoxa11† | -2.10 (.35) | 11 kb upstream† | 3100.00 |
| Ctnnb1 | -1.88 (.49) | Exon | 599.22 |
| Zbtb40 | -1.73 (.30) | Intron | 912.21 |
| Cdh11 | -1.76 (.10) | 18 kb downstream | 931.33 |
| Ctgf | -1.68 (.51) | Exon | 584.17 |
| Adamts5 | -1.67 (.05) | Exon | 593.39 |
| Gli2 | -1.65 (.19) | Intron | 748.85 |
| Stard7 | -1.63 (.10) | Exon | 1246.12 |
| Rbms3 | -1.61 (.18) | Intron | 733.40 |
| Gpc6 | -1.60 (.08) | Intron | 1353.69 |
| Trio | -1.57 (.14) | 7 kb upstream | 1004.04 |
| Mospd4 | -1.56 (.26) | 8 kb upstream | 847.09 |
| Glt25d1 | -1.56 (.10) | 15 kb downstream | 1186.92 |
| Ncam1 | -1.50 (.22) | Intron | 722.89 |
| Alox5ap | -1.50 (.30) | Intron | 1142.99 |
| Ralgps2 | no change | 16 kb upstream | 750.78 |
| Col5a3 | no change | Intron | 611.46 |
| Acss3 | no change | Intron | 2062.97 |
| Akap6 | no change | Intron | 1095.52 |
| Lynx1 | no change | Intron | 941.61 |
| Sap130 | no change | Exon | 746.90 |
| Pde11a | not detected | Intron | 799.97 |
| Vwa5b1 | not detected | Intron | 597.72 |
| Tmem132d | not detected | Intron | 1404.44 |
| Unc13a | not detected | Intron | 1186.92 |
| Tpo | not detected | Intron | 781.19 |
| Mctp1 | not detected | Intron | 990.63 |
| Nuggc | not detected | 9 kb upstream | 773.81 |
| Ppp1r32 | not detected | Exon | 830.44 |

**Fig 4. *LncRNA-HIT* peaks are significantly associated with multiple loci which respond to siRNA-mediated reductions in the lncRNA.** Values represent mean fold-change decrease in limb bud expression after siRNA treatments compared to scrambled siRNA controls from three independent assays. Standard deviation for these analyses is depicted in parentheses. p-values for the associated peaks is depicted by the transformation -10 $\log_{10}$ (p-value) as determined by MACS analysis. †Loci located in *cis* to the nascent *LncRNA-HIT* peak.
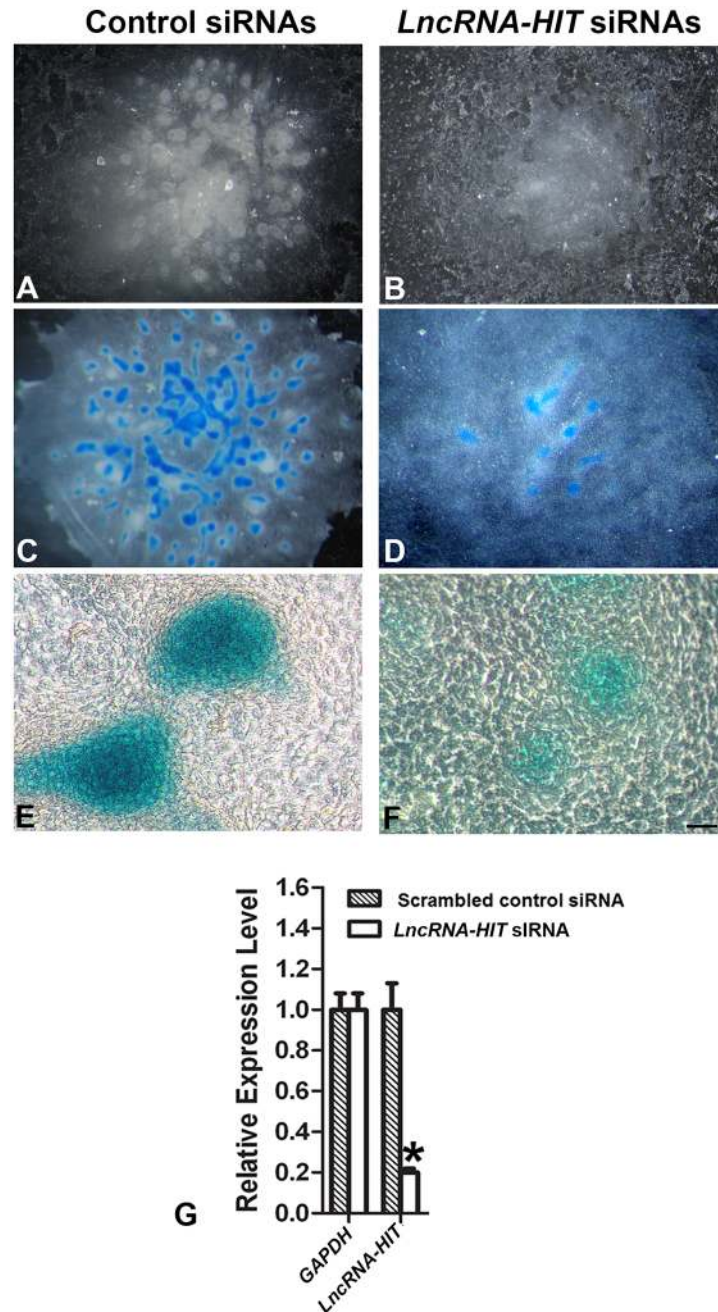
**Fig 5. Chondrogenic differentiation is impacted by siRNA-mediated reductions in *LncRNA-HIT* in micromass assays. (A)** Mesenchymal cell condensation is unaffected after transfection of the dissociated limb mesenchyme with a scrambled siRNA control. **(C and E)** After condensation, the scrambled control transfected limb mesenchyme are competent to differentiate into cartilage nodules that stain positively for the cartilage marker alcian blue. **(B,D,F)** Parallel assays transfecting the same limb mesenchyme with two siRNAs specific for *LncRNA-HIT* result in a loss in cell condensation (panel B) as well as severe reductions in the formation of cartilage nodules that stain with alcian blue (panels D and F). **(G)** Quantitation of the *LncRNA-HIT* levels 24H after transfection with the *LncRNA-HIT* siRNAs or scrambled siRNAs reveal an eighty percent reduction in endogenous *LncRNA-HIT* levels compared to parallel transfections using a scrambled siRNA control. Values represent the average relative expression calculated from six independent replicates. Bars represent the standard deviation of the mean from the six independent assays. A Student's t-test was used to determine significance of the averaged *LncRNA-HIT* expression values compared to the scrambled *LncRNA-HIT* siRNA control. A significant difference is indicated by an asterisk.

doi:10.1371/journal.pgen.1005680.g005

substantial reduction in chondrogenesis, resulting in the formation of fewer cartilage nodules that stained with alcian blue (n = 6 independent assays, Fig 4). In contrast, parallel micromass assays using the same populations of limb mesenchyme cells transfected with a scrambled control siRNA confirmed that the transfected limb mesenchyme was competent to undergo chondrogenic differentiation, exhibiting robust cell condensation and the formation of multiple alcian blue positive cartilage nodules (n = 6 independent assays, Fig 5).

Next, if chondrogenesis is facilitated by *LncRNA-HIT's* recruitment of p100/CBP complexes, then siRNA-mediated reduction in *Snd1*, which encodes p100, should also affect the chondrogenic capacity of the undifferentiated limb mesenchyme. Testing this hypothesis, micromass assays using limb mesenchyme transfected with *Snd1* siRNAs produced a similar loss in *Snd1* transcript levels (75% knockdown) and reduced chondrogenesis, resulting in the formation of few cartilage nodules staining with alcian blue (n = 6 independent assays, Fig 6). Control transfections using a scrambled *Snd1* siRNA consistently resulted in robust cartilage formation producing numerous cartilage nodules staining positively for alcian blue (n = 6 independent assays, Fig 5), supporting the hypothesis that *LncRNA-HIT* and its recruited proteins are essential for chondrogenic differentiation of the early limb mesenchyme.

## *LncRNA-HIT* exhibits enhancer RNA function to promote H3K27ac to facilitate gene expression

The decrease in *Hoxa13* and *Hoxa11* expression in response to the *LncRNA-HIT* siRNA treatments (Fig 4) suggested that *LncRNA-HIT* may be functioning as an enhancer RNA, using its recruitment of p100/CBP to its site of transcription to promote neighboring 5' HoxA gene expression through its maintenance of H3K27ac. Testing this hypothesis, we first examined whether siRNA-mediated reductions in *LncRNA-HIT* or *Snd1* impact expression of additional genes proximal to the nascent site of *LncRNA-HIT* transcription. qRTPCR analysis of limb mesenchyme treated with *LncRNA-HIT*- or *Snd1* (p100)-specific siRNAs revealed significant reductions in expression for all 5'HoxA gene members (Fig 7), suggesting that *LncRNA-HIT* is functioning as an enhancer lncRNA to mediate expression of the 5' HoxA genes. Interestingly, no changes in expression were detected for *Creb5* and *Skap2*, which flank the HoxA cluster, and for *Hoxd13*, suggesting a specific effect on the 5' HoxA genes which are proximal to the site of *LncRNA-HIT* transcription (Fig 7). Changes in 3' HoxA gene expression (*Hoxa1-Hoxa7*) could not be determined, as their expression in the limb mesenchyme is below the level of detection by qRTPCR. Finally *Col2a1*, a strong indicator of chondrogenic differentiation, also exhibited reduced expression in response to the *LncRNA-HIT* and *Snd1* siRNA treatments confirming the disruption in chondrogenesis exhibited by the micromass assays (Figs 5–7).

Next we hypothesized that H3K27ac should also be affected by siRNA-mediated reductions in *LncRNA-HIT* as CBP recruitment would be concomitantly affected in the limb mesenchyme. Mapping of the H3K27ac-tagged chromosomal regions in the E 10.5 limb was previously reported (GEO Dataset: GSE30641, [63]). Using this dataset, the H3K27ac sites proximal to the *LncRNA-HIT* -associated loci were examined for changes using an acetylation-specific H3K27 antibody and quantitative chromatin immunoprecipitation (qChIP) (Fig 8). Starting with the site of nascent *LncRNA-HIT* transcription, qChIP analysis of fourteen H3K27ac-tagged chromosomal regions between *Hoxa11* and *Hoxa13* revealed consistent reductions in precipitated fragment enrichment in response to the *LncRNA-HIT* siRNA treatments (n = 3 independent assays) (Fig 7). Most notable was a greater than five-fold reduction in H3K27ac fragment enrichment for the region associated with exon 2 of *Hoxa13* (Fig 8, fragment 7). A greater than 3-fold decrease in H3K27ac chromatin fragment enrichment was also detected in the *Hoxa13* intronic region (Fig 8, fragment 11) and in chromatin fragments more proximal to *Hoxa11*
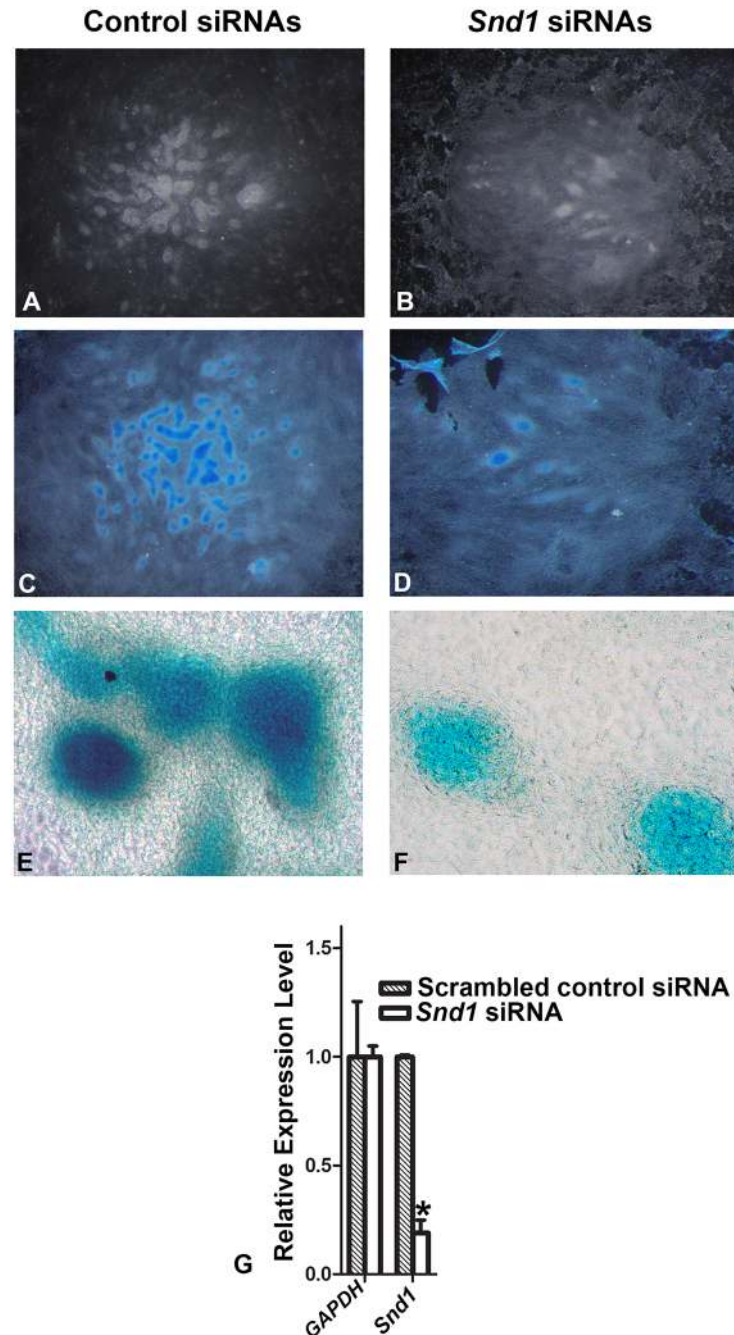
**Fig 6. Chondrogenic differentiation is impacted by siRNA-mediated reductions in Snd1 (p100) in micromass assays. (A)** Mesenchymal cell condensation is unaffected after transfection of the dissociated limb mesenchyme with a scrambled siRNA control. **(C and E)** After condensation, the scrambled control transfected limb mesenchyme are competent to differentiate into cartilage nodules that stain positively for the cartilage marker alcian blue. **(B,D,F)** Parallel assays transfecting the same limb mesenchyme with a *Snd1* siRNA cocktail severely impacts mesenchymal cell condensation producing few pre-cartilaginous condensations that can differentiate into mature cartilage nodules that stain positively with alcian blue. **(G)** Quantitation of the *Snd1* levels 24H after transfection with the *Snd1 siRNA cocktail* or scrambled siRNA controls reveal a seventy percent reduction in endogenous *Snd1* levels compared to parallel transfections using the scrambled siRNA controls. Values represent the average relative expression calculated from six independent replicates. Bars represent the standard deviation of the mean from the six independent assays. A Student's t-test was used to determine significance of the averaged *Snd1* expression values compared to the scrambled *siRNA* controls. A significant difference is indicated by an asterisk.

doi:10.1371/journal.pgen.1005680.g006
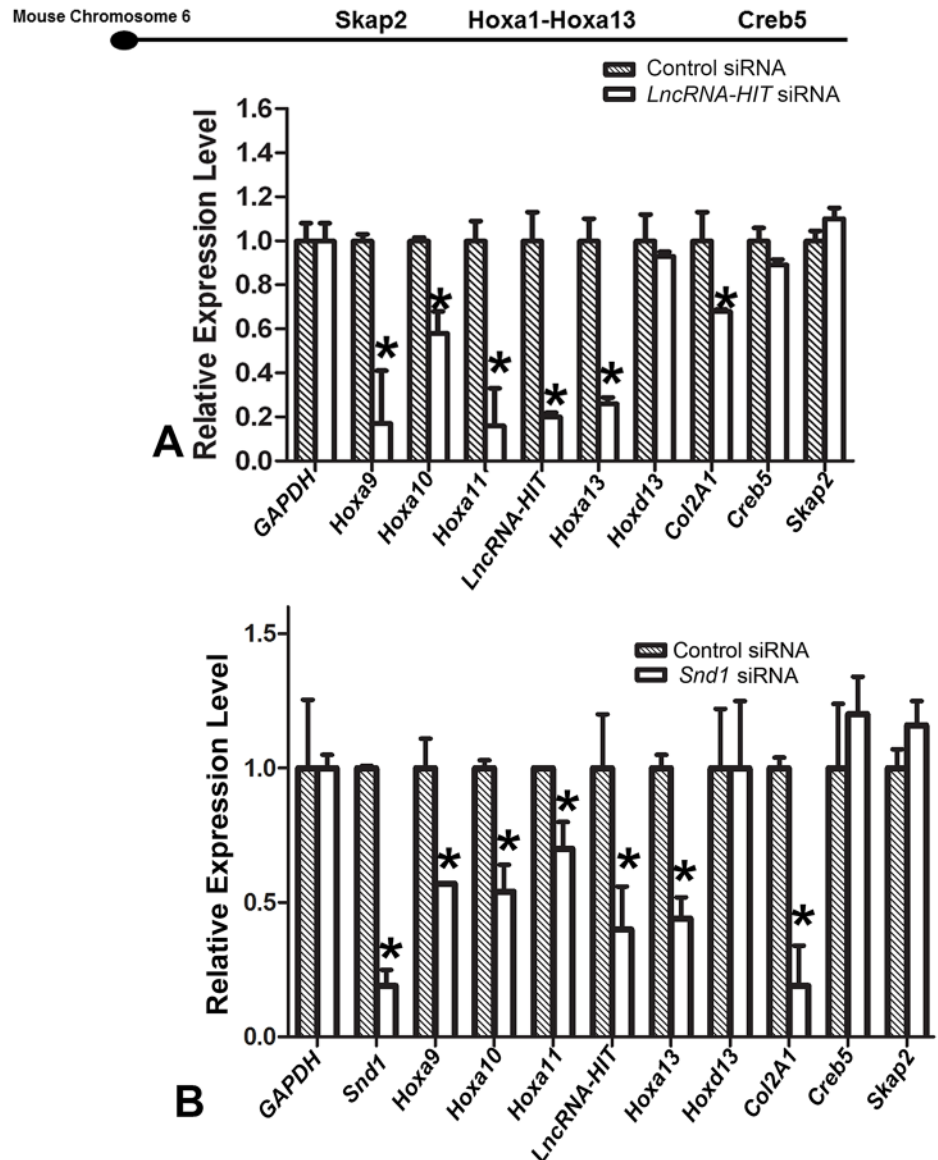
**Fig 7. siRNA-mediated reduction in *LncRNA-HIT* or *Snd1* results in reduced levels of 5' HoxA gene expression. (A)** Relative gene expression after transfection with *LncRNA-HIT* siRNAs or scrambled control siRNAs. Values represent average expression levels calculated from six independent replicates. Bars represent the standard deviation of the mean from the six independent replicates. Asterisks denote a significant changes in gene expression as determined a Student's t test. **(B)** Relative gene expression after transfection with *Snd1* siRNAs or scrambled control siRNAs. Values represent average expression levels calculated from six independent replicates. Bars represent the standard deviation of the mean from the six independent replicates. Asterisks denote a significant changes in gene expression as determined a Student's t test.

doi:10.1371/journal.pgen.1005680.g007

(Fig 8, fragment 1). Several H3K27ac regions were also identified within the *LncRNA-HIT* locus (Fig 8B). qChIP analysis of these regions revealed reduced enrichment for the entire region, with fold change decreases ranging from 1.5 to 2.8 (Fig 8, fragments 2–6).

Expanding this analysis, we examined additional chromosomal loci bound by *LncRNA-HIT* for changes in H3K27ac fragment enrichment in response to the siRNA treatments. For this analysis, we selected the H3K27ac sites most proximal to five loci exhibiting the greatest
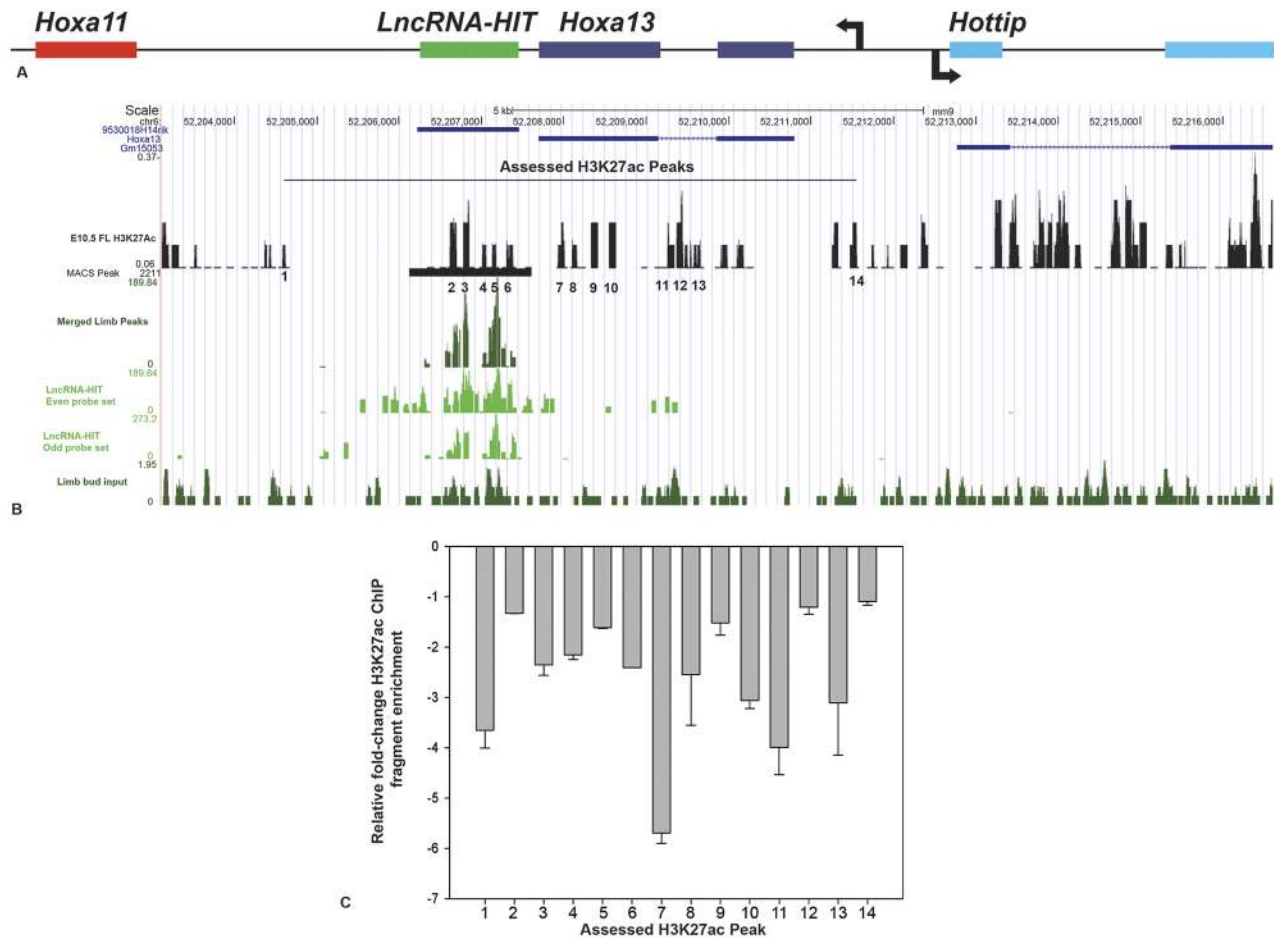
**Fig 8. siRNA-mediated loss of *LncRNA-HIT* affects H3K27ac at its site of nascent transcription. (A)** Relative location of *LncRNA-HIT* and its nearest neighbors, *Hoxa11* and *Hoxa13*. **(B)** Visualization of the nascent *LncRNA-HIT* (green peaks) detected by ChIRP-seq using the USCS Genome Browser. Peaks identified by the even and odd *LncRNA-HIT* probe sets are indicated on the left axis using even and odd probe sets specific for the *LncRNA-HIT* lncRNA and displayed by the UCSC genome browser. The localization of the fourteen H3K27ac peaks (black peaks) proximal to the *LncRNA-HIT* associated peaks as determined by overlaying the H3K27ac limb bud ChIP-seq dataset, GSE3064 [62]. **(C)** qChIP Analysis of the fourteen H3K27ac regions proximal to the nascent *LncRNA-HIT* peaks revealed a loss in fragment enrichment in response to the *LncRNA-HIT* siRNA treatments in limb bud mesenchyme. Values represent the mean fold change H3K27ac fragment enrichment compared to parallel treatments using scrambled *LncRNA-HIT* siRNA controls for three independent assays. Error bars represent the standard deviation of the mean for the three independent assays.

doi:10.1371/journal.pgen.1005680.g008

decrease in expression in response to the *LncRNA-HIT* siRNA treatments: *Pik3cb*, *Col14a1*, *Bmpr1b*, *Pbx1*, and *D15ertd621e* (Fig 4). qChIP analysis revealed consistent reductions in immunoprecipitated H3K27ac fragments for four of the five candidate loci (n = 3 independent assays) including *Pik3cb* (> 4-fold decrease), *Col14a1* (> 3-fold decrease), *Pbx1* (>1.4-fold decrease) and *D15ertd621e* (>2-fold decrease) (Fig 9). Attempts to amplify the immunoprecipitated H3K27ac fragments proximal to the *LncRNA-HIT* associated region at the *Bmpr1b* locus were unsuccessful using several primer pair combinations and sheared chromatin from the *LncRNA--HIT*—or control-siRNA treated samples, suggesting that H3K27ac in this region is insufficient to facilitate immunoprecipitation by the acetylation-specific H3K27 antibody (Fig 9).

Finally to test whether changes in H3K27ac peak enrichment was specifically affected by the *LncRNA-HIT* siRNA treatments we examined loci lacking associated *LncRNA-HIT* peaks for changes in H3K27ac fragment enrichment. Analysis of H3K27ac fragments associated with the promoter regions of *Creb5* and *Skap2* revealed no changes in fragment enrichment in response
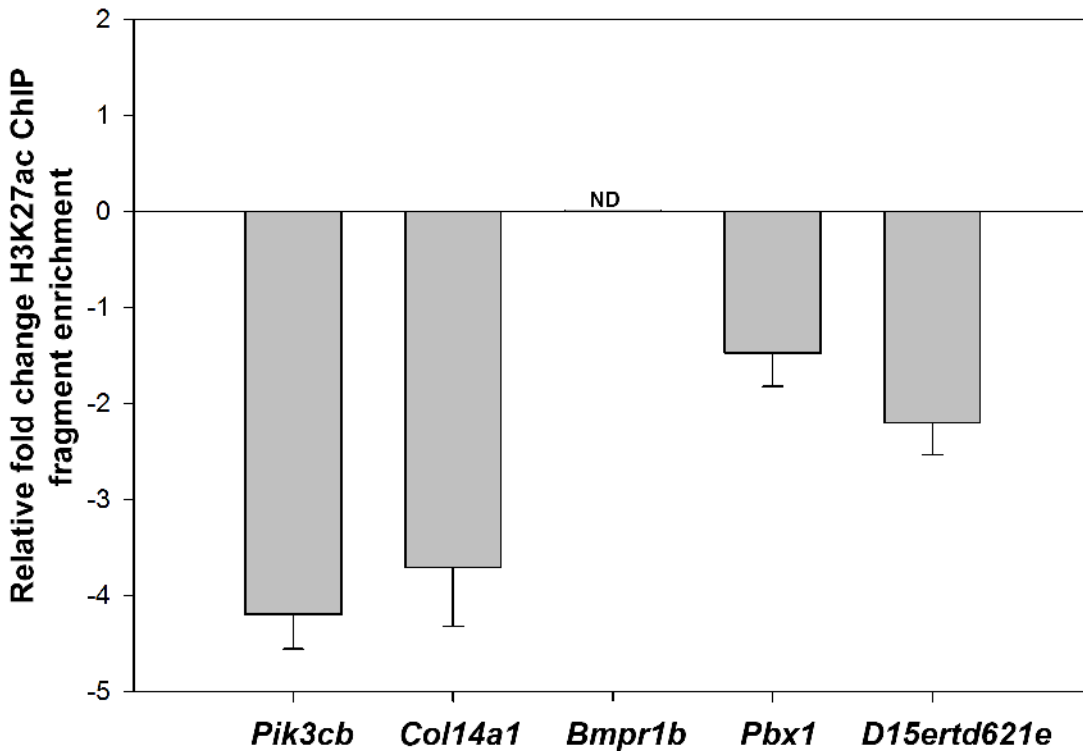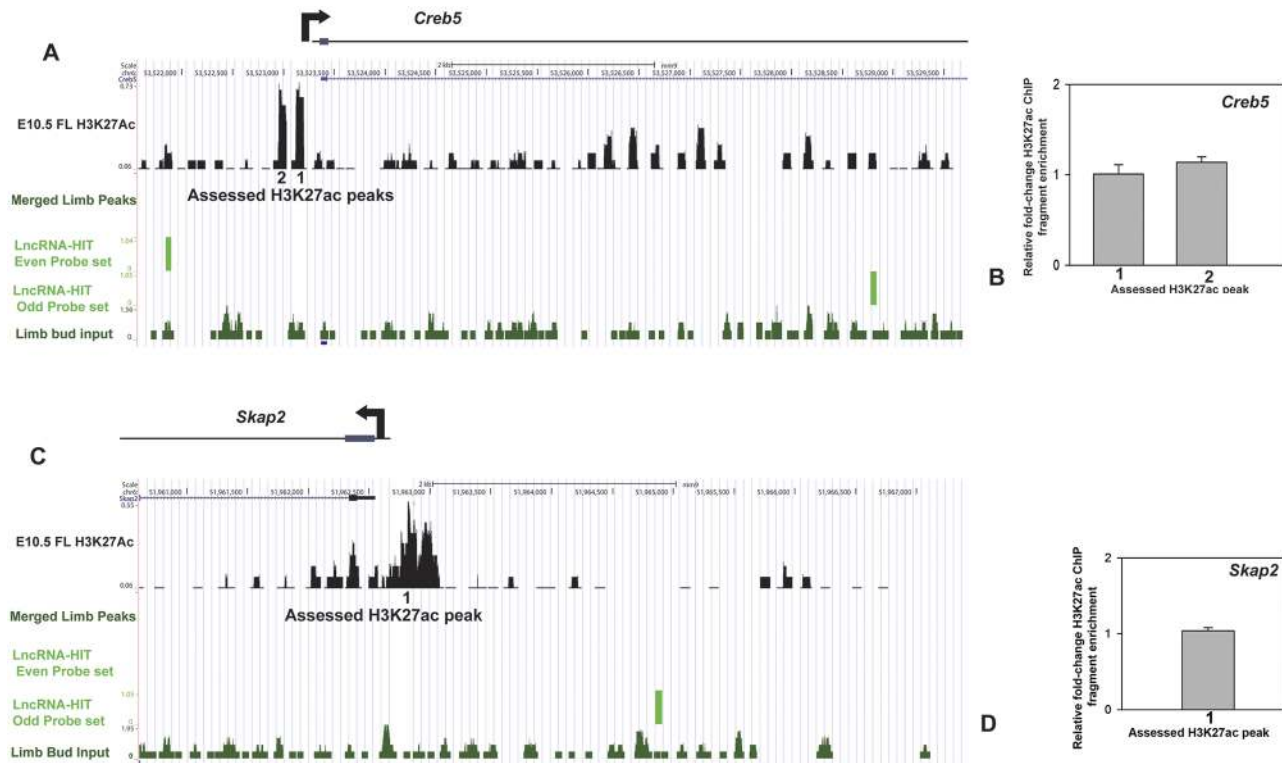
**Fig 9. qChIP analysis of H3K27ac at the top five loci impacted by the *LncRNA-HIT* siRNA treatments.** Values represent the mean fold change H3K27ac fragment enrichment compared to parallel treatments using scrambled *LncRNA-HIT* siRNA controls for three independent assays. Error bars represent the standard deviation of the mean for the three independent assays. ND = no qPCR product detected for the H3K27ac peak proximal to *Bmpr1b*.

doi:10.1371/journal.pgen.1005680.g009

to *LncRNA-HIT* siRNA treatments, supporting the conclusion that *LncRNA-HIT*-p100/CBP is functioning to maintain H3K27ac status in regions bound by the lncRNA-protein complex (Fig 10).

## *LncRNA-HIT* and p100 cooperate to promote gene expression

We next examined whether the recruitment of p100 by *LncRNA-HIT* contributes to the regulation of near-peak gene expression. To address this question, an RNA tethering assay was used to dissect the gene-regulatory contributions of *LncRNA-HIT* and p100 towards the activation of a synthetic UAS-luciferase reporter (Fig 11). Recruitment of *LncRNA-HIT* to the reporter locus was facilitated by adding five copies of the *BoxB* transcript to the *LncRNA-HIT* RNA. The *BoxB* transcript is strongly bound by the RNA binding protein λN, which when fused to the GAL4 DNA binding domain facilitates the recruitment of the *BoxB-LncRNA-HIT* transcript to the UAS luciferase locus (Fig 11) [5,64]. Co-transfection of the UAS-luciferase reporter, λN-GAL4, and p100-GAL4 expression vectors resulted in no activation of the UAS luciferase reporter, indicating that the recruitment of these factors to the synthetic UAS locus is not sufficient to activate gene expression (Fig 11). Similarly, co-transfection of the UAS-luciferase reporter and expression vectors encoding λN-GAL4, p100-GAL4, and the *BoxB* transcript lacking *LncRNA-HIT* also resulted in no activation of the UAS luciferase reporter, indicating that the recruitment of *BoxB*, p100, and λN to the synthetic locus is not sufficient to activate gene expression (Fig 11A and 11C). Additionally, luciferase expression was not detected in cells transfected with UAS-luciferase reporter and expression vectors encoding *BoxB-LncRNA-HIT* and λN-GAL4, suggesting that the recruitment of *LncRNA-HIT* to the UAS locus is

**Fig 10. qChIP analysis of H3K27ac at *Creb5* and *Skap2* which lack LncRNA-HIT associated peaks. (A)** *Creb5*-assocaited H3K27ac peaks assessed for changes in fragment enrichment in response to LncRNA-HIT siRNA treatments. **(B)** qChIP analysis of H3K27ac peak enrichment in response the LncRNA-HIT siRNA treatments at the *Creb5* locus. Values represent the mean fold change H3K27ac fragment enrichment relative to parallel treatments using a scrambled siRNA control. Bars represent the standard deviation of the mean for three independent assays. **(C)** *Skap2*-associated H3K27ac peak assessed for changes in fragment enrichment in response to LncRNA-HIT siRNA treatments. **(D)** qChIP analysis of H3K27ac peak enrichment in response the LncRNA-HIT siRNA treatments at the *Skap2* locus. Value represents the mean fold change H3K27ac fragment enrichment relative to parallel treatments using a scrambled siRNA control. Bar represents the standard deviation of the mean for three independent assays.

doi:10.1371/journal.pgen.1005680.g010

not sufficient to activate gene expression (Fig 11B and 11C). Finally, co-transfections using the UAS-luciferase reporter, *BoxB-LncRNA-HIT*, λN-GAL4, and p100-GAL4 expression vectors resulted in strong activation of the luciferase reporter in a dosage-dependent manner, indicating that co-recruitment of p100 and *LncRNA-HIT* to a locus is required to activate gene expression (Fig 11A and 11C).

## Discussion

The development of the vertebrate limb requires a coordinated series of cellular events to facilitate the formation of bone and articular cartilage. Initiating this process is the proliferation and migration of mesenchymal progenitor cells from the lateral plate mesoderm. Once in the limb bud, the mesenchymal cells condense forming cartilage templates for all skeletal tissues. Finally the chondrocytes within these templates mature, forming bone through endochondral ossification or remain as a specialized population of chondrocytes to form the articular cartilage [65–68]. Interestingly, while SOX9 appears to be essential for many of the cellular processes required for skeletal development, emerging evidence suggests that an epigenetic component, mediated by lncRNAs, may also play an important role in the formation, maintenance, or pathology of the appendicular skeletal tissues [5,34,69–73]. Given the large number of lncRNAs expressed by the human (>15,000) and mouse genomes (>7600), it is likely that
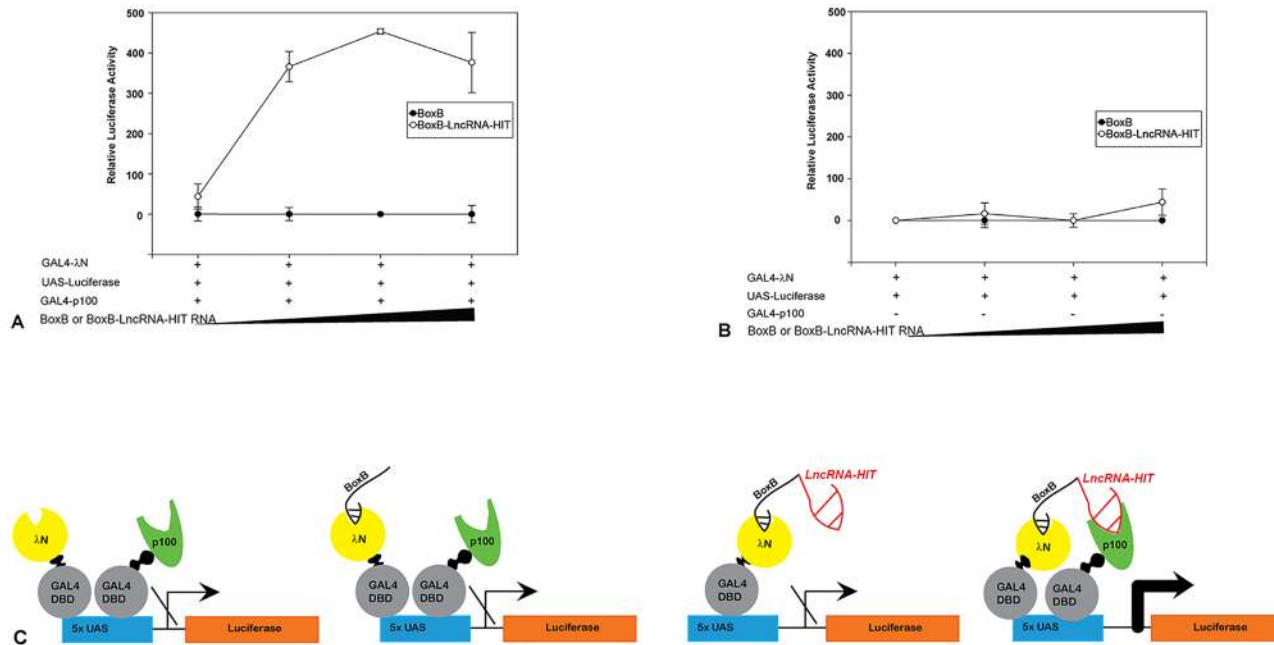
**Fig 11. Co-recruitment of *LncRNA-HIT* and p100 is required to stimulate gene expression from a synthetic locus. (A)** Analysis of UAS-luciferase reporter activation after GAL4-λN-BoxB*LncRNA-HIT* and GAL4-p100 recruitment. **(B)** Analysis of UAS-luciferase reporter activation in the absence of GAL4-p100. **(C)** Model of UAS-luciferase reporter activation in response to the recruitment of *LncRNA-HIT* and p100. For panels A and B, increasing dosages (0, 50, 100, 250, ng) of the BoxB or BoxB-*LncRNA-HIT* transcripts are represented by the black triangle.

doi:10.1371/journal.pgen.1005680.g011

additional roles for these molecules will be identified in the appendicular cartilage and bone (GENCODE ver. 22, [74]). In this report, we identify the lncRNA, *LncRNA-HIT*, as an essential component for chondrogenesis, regulating multiple genes to facilitate the formation of cartilage from undifferentiated limb mesenchyme.

## *LncRNA-HIT* regulates BMP signaling to facilitate chondrogenesis

Ontological analysis of loci contacted and regulated by *LncRNA-HIT* suggests that the lncRNA coordinates the expression of genes whose products mediate chondrogenic differentiation in the limb mesenchyme. This conclusion is consistent with the decrease in cartilage nodule formation in response to the *LncRNA-HIT* and *Snd1* siRNA treatments. Early studies of cartilage formation identified disruptions in the initial condensation of the limb mesenchyme as a primary cause of reduced cartilage formation [75–76]. More importantly the regulation of mesenchymal condensation was recently shown to require BMP-SMAD4 signaling, independent of SOX9 function [77]. This finding provides a mechanism to explain the loss of cartilage formation in the micromass assays, as *Bmpr1b*, a major receptor for GDF5 and other BMPs during limb development, is significantly down-regulated in response to the *LncRNA-HIT* siRNA treatments (Fig 4) [78–80]. In the absence of BMPR1B, the primary chondrogenic phenotype exhibited by humans and mice is the loss of prechondrogenic condensations and skeletal element hypoplasia, particularly in the distal limb, a region that also strongly expresses *LncRNA--HIT* (Fig 1) [80–83]. Reductions in *Bmpr1b* expression in response to the *LncRNA-HIT* siRNA treatments could also reflect a loss in p100 enrichment. This conclusion is supported by the RNA tethering assay which determined that interactions between *LncRNA-HIT* and p100 were sufficient to activate gene expression from a synthetic locus. By this mechanism, *LncRNA-HIT* could regulate *Bmpr1b* as well as additional loci by stabilizing p100 at the *Bmpr1b* locus;

providing a functional explanation for the 4.75-fold decrease in *Bmpr1b* expression in response to the *LncRNA-HIT* siRNA treatments. Alternatively, the down-regulation of *Bmpr1b* expression in response to the *LncRNA-HIT* siRNA treatments could also reflect indirect regulation of *Bmpr1b* by another *LncRNA-HIT*-regulated gene product.

A second component contributing to the control of chondrogenesis by *LncRNA-HIT* is its regulation of *Hoxa13 and Hoxa11* which previous studies indicate can regulate additional components of the BMP-signaling cascade including *Bmp2* and *Bmp7* by HOXA13 and *Runx2 by* HOXA11 [84–85]. Thus in the absence of *LncRNA-HIT* function, the expression of receptor, ligand, and transcriptional components of the BMP signaling pathway are reduced, which in combination, explains the reductions in mesenchymal cell condensation and cartilage nodule formation exhibited by the micromass assays treated with the *LncRNA-HIT* siRNAs.

## *LncRNA-HIT* functions as an enhancer lncRNA

The down-regulation of 5' HoxA genes in response to *LncRNA-HIT* siRNA treatments suggests that *LncRNA-HIT* may be functioning as an enhancer lncRNA. Recent studies of enhancer lncRNAs indicate that these molecules regulate the expression of nearby genes by recruiting chromatin modifying proteins to increase the accessibility of the chromosomal region to gene-regulatory factors [6,22,86]. In the limb mesenchyme, our detected decrease in H3K27ac and 5' HoxA gene expression in response to *LncRNA-HIT* siRNA treatments support a role for *LncRNA-HIT* as an enhancer lncRNA using its associated proteins (p100 and CBP) to facilitate expression of the surrounding 5' HoxA genes (Fig 12). It is interesting to speculate that the presence of *LncRNA-HIT* within the 5' HoxA cluster in many vertebrate genomes (S1 Fig) may reflect its conservation as an essential enhancer of 5' HoxA gene expression, providing a fundamental epigenetic mechanism to promote chondrogenic differentiation; presumably through its recruitment of the p100/CBP complexes. P100/CBP complexes are important components of the STAT6-dependent enhanceosome which functions as a transcriptional enhancer of genes involved in the interleukin 4 gene regulatory cascade [47–48,87]. In limb mesenchyme,
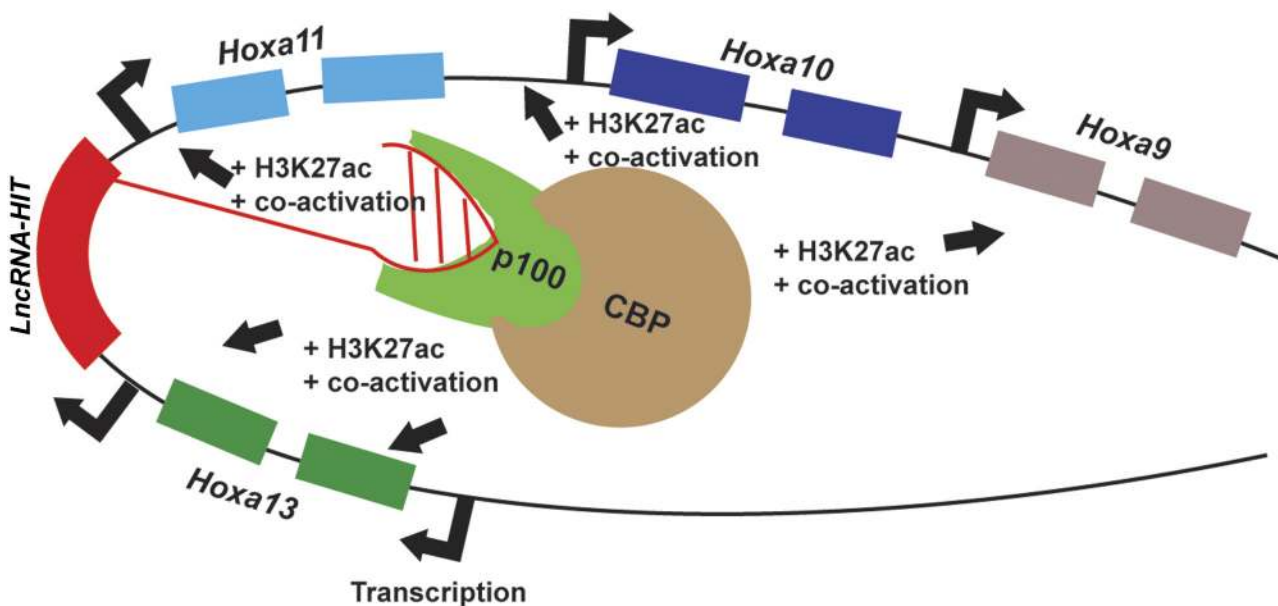


Fig 12. **Model of enhancer RNA function of *LncRNA-HIT* at the 5' HoxA locus.**

*LncRNA-HIT* may be working in a similar capacity, forming a complex with p100 and CBP to regulate pro-chondrogenic gene expression.

Interestingly, the results from the RNA tethering assay also suggest that interactions between *LncRNA-HIT* and p100 may be required to confer functional competence to the RNA protein complex, as the activation of the luciferase reporter occurred only when p100 and *LncRNA-HIT* were co-expressed. Recent studies support this conclusion, as mutations in WDR5 that prevent lncRNA binding impacts WDR5-MLL chromatin occupancy, maintenance of H3K4 trimethylation, and the regulation of genes necessary for embryonic stem cell self-renewal [46]. Taken together these findings suggest that *LncRNA-HIT* promotes gene expression by two mechanisms, first: by binding to p100 to stabilize the protein to promote its co-activator function and second: through its recruitment of CBP to promote chromatin accessibility by maintaining H3K27ac which in the 5' HoxA region facilitates the expression of multiple factors whose products instruct the limb mesenchyme to form skeletal tissues [27, 30–31,42, 88–89].

## *LncRNA-HIT* may have additional functions in onset and progression of some cancers and in the development of other tissues

Transcriptional profiling of breast tumor cell lines initially identified *LncRNA-HIT* as a TGFβ-induced transcript [35]. Interestingly, TGFβ1/SMAD signaling can also promote *Snd1* (p100) expression which can facilitate epithelial to mesenchymal transitions (EMT) in mammary tumor cell lines, a causal event in tumor cell migration and invasion [90–94]. In these tissues, Tgfβ should induce both *LncRNA-HIT* and *Snd1* (p100) which would activate *Hoxa13*, which has been previously shown to mediate the formation of vascular tissues, providing a mechanism to promote tumor vascularization and growth [95].

Interestingly several studies investigating predicative biomarkers of cancer progression and patient outcomes in hepatocellular carcinoma have implicated the expression *HOXA13* and *HOTTIP* as indicators of poor prognosis with high disease progression [96–97]. Given that the physical distance between *LNCRNA-HIT*, *HOXA13*, and *HOTTIP* is only 6.5 kb in both humans and mice, it is possible that *LNCRNA-HIT's* function may also be co-opted in these cancers. If human *LNCRNA-HIT* is determined to be expressed in hepatocellular carcinoma, a ChIRP-seq approach could be used to determine the loci contacted by the lncRNA-protein complex, providing additional therapeutic targets for this disease. Moreover, several non-chondrogenic regions also express *LncRNA-HIT* during murine development including the gut epithelium and the genital tubercle, suggesting that the lncRNA may have additional roles in the formation of these tissues. Using the approaches outlined in this study, including ChIRP-seq and RNA affinity chromatography the proteins interacting with *LncRNA-HIT* in these tissues as well as the loci contacted by the protein-lncRNA complexes can be identified providing new insights into the developmental regulation of these tissues. To discern the *in vivo* functions of *LncRNA-HIT* during development will be more challenging, as targeted disruptions of *LncRNA-HIT* should affect 5' HoxA gene expression either by disrupting activation mediated by the enhancer RNA functions of *LncRNA-HIT* -p100/CBP complexes or by disrupting *cis*-regulatory elements required for 5' HoxA gene expression. By either mechanism 5' HoxA gene expression would be reduced, creating phenotypes that cannot be solely attributed to the loss of *LncRNA-HIT* function.

*LncRNA-HIT* is also capable of functioning as a guide RNA as it can recruit the same proteins to loci independent of its site of transcription. By this process *LncRNA-HIT*-p100/CBP complexes can regulate a larger cohort of genes. As an epigenetic modulator of chondrogenesis, *LncRNA-HIT*, may be one of several non-coding RNA species required for the formation and/

or maintenance of cartilage tissues including micro RNAs MiR-145 and 337 as well as the small nucleolar RNAs U38 and U48 which are elevated in the serum of individuals affected by injury-induced osteoarthritis [98–99]. These studies in conjunction with the present work indicates a higher level of gene regulation is necessary for the initial formation of cartilage tissues as well as its loss during osteoarthritis, providing additional targets to exploit as therapies to minimize joint tissue loss as a consequence of injury or disease.

## Materials and Methods

### Ethics statement

All investigations using mice were certified as compliant with AVMA guidelines by the OHSU IACUC board prior to implementation following an approved mouse use protocol IS00001648 to HSS. Euthanasia was conducted using CO2 gas until all signs of movement and respiration ceased, followed by prompt thoracotomy. This method is consistent with the recommendations of the Panel on Euthanasia of the American Veterinary Medical Association. PHS 398/2590 (Rev. 06/09).

### Mouse tissues and gestational timing

Embryonic tissues were derived from timed matings of wild type Swiss Webster mice (Charles River Labs). Gestational development was measured in embryonic days (E) where E = 0.5 reflects the detected day of a vaginal plug.

### RNA FISH

Limb bud mesenchyme was dissected from E 11.0 embryos, dissociated, and grown in culture in two chamber cover-glass slides (Nunc Lab- TEKII, Thermo Scientific) as described [55]. After 24H the cells were fixed for 10 min using 4% formaldehyde/PBS and prepared for RNA FISH following the protocol described by the Lanctôt laboratory (http://lanctotlab.org/en/ protfish_rnafisholigo20mer.html). The *LncRNA-HIT* RNA FISH probe sets were synthesized and labeled with CalFluor Red 610 by Stellaris Biosearch Technologies (S1 Table). Cells were imaged using a Zeiss Axioplan epifluorecence microscope fitted with a 100x 1.4 NA oil objective and a cooled monochrome CCD camera using a 2 second exposure. Detected signals were pseudo-colored red (*LncRNA-HIT*), green (*Gapdh*), and blue (DAPI) using the Zeiss AxioVision digital image processing software (release 4.7.2).

### p100 and CBP RNA Immunoprecipitation (RIP)

Approximately 1mg of E 11.5 limb bud lysates were used to immunoprecipitate p100 and CBP with endogenous *LncRNA-HIT* transcripts following the previously described RIP procedure [5, 68] using antibodies specific for SND1 (p100) (Cat. Ab65078 Abcam, Cambridge, MA) and CBP (Cat. 7389 Cell Signaling, Danvers, MA). After immunoprecipitation of p100, CBP, or IgG, the co-immunoprecipitated RNA was extracted using TRIzol following the manufacturer's protocol (ThermoFisher). After extraction, the isolated RNA was treated with RQ1 Rnase-Free Dnase (Cat. M6101 Promega, Madison, WI) and evaluated for enrichment of *LncRNA-HIT* using qRTPCR as described in the RNA isolation and qRTPCR section of the Methods. Three independent immunoprecipitation assays were performed for p100, CBP, and the IgG control analysis of *LncRNA-HIT* co-immunoprecipitation.

## Micromass tissue culture

Embryos were collected at E 10.5 as described [55]. Cells were transfected with two small interfering RNAs (siRNAs) specific for *LncRNA-HIT* or with a scrambled siRNA control at a concentration of 10 nM using Trifectin as recommended (IDT, Coralville, Iowa). 24H after transfection, the cells were trypsinized and resuspended to a final concentration of $2x10^7$ cells/ml and seeded as 10 ul drops into 60 mm Falcon tissue cultures as described [56]. The *LncRNA-HIT* siRNA sequences used were:

1. Antisense 5' UUAAGGUCACAGACCACCUUGGAGGGU 3'

2. Antisense 5' CCUCUCUCUCCCUCCCUCCUUUCCUUU 3'

For *Snd1* repression, limb mesenchyme was transfected using a commercial *Snd1* siRNA cocktail at 30 nM concentrations as recommended by the manufacturer (Life Technologies: Cat: 4390771, Grand Island, NY).

## ChIRP-seq analysis

25 oligonucleotide probes corresponding to the murine *LncRNA-HIT* transcript were selected using the Stellaris Probe Designer Software set to masking level 5 (biosearchtech.com) (S2 Table). The oligonucleotide sequences were synthesized with an 18-atom spacer followed by a biotin tag located at the 3' end and purified by HPLC as described by the manufacturer (IDT, Coralville, Iowa). The autopod regions of the E 11 limbs were dissected and digested for 10 minutes in sterile PBS containing 0.1% trypsin and 0.1% collagenase as described [56]. Approximately 2000 embryonic limb buds were required to produce the 200mg cells needed for each ChIRP assay. Cell lysis and chromatin shearing and streptavidin bead precipitation were accomplished using a Bioruptor instrument (Diagenode) as previously described [50]. As a control, chromatin from murine glial cells (a kind gift from Dr. Peter Hurlin), that do not express *LncRNA-HIT*, were used in parallel assays and hybridized with the same *LncRNA-HIT* probes. The precipitated DNA was quantified using a Qubit 2.0 DNA fluorometer (Life Technologies) and submitted to Elim Biopharm (Hayward, CA) for library preparation and next generation sequencing using an Ilumina HiSeq 2500 sequencer.

## Data analysis

The *LncRNA-HIT* ChIRP-seq libraries were mapped, normalized, and analyzed using the ChIRP-seq analysis pipeline previously described [51]. Raw reads from each ChIRP-seq library were uniquely mapped to the reference mouse genome (mm9) using bowtie-0.12.9 and normalized to total read count as described [52–53]. Even and odd *LncRNA-HIT* ChIRP samples were merged by taking the minimum value of the even and odd tracks at each genomic position. The mouse glial cell line that does not express *LncRNA-HIT* was used as a control to identify false hybridization peaks detected by the *LncRNA-HIT*-specific even and odd probe pools. Peaks were called by MACS using the *LncRNA-HIT* ChIRP merge sample and the DNA input, with an initial threshold p-value $\leq 5.9 \times 10^{-57}$ corresponding to 775 peaks based on visual inspection of the peaks using the UCSC Genome Browser. Peaks were additionally filtered by visual inspection with UCSC Genome Browser excluding any peak with a corresponding peak in the glial controls. The limb-specific peaks were then submitted for analysis using the Genomic Regions of Enrichment of Annotations Tool (GREAT) [54]. Following GREAT the peaks were additionally filtered to identify high confidence peaks mapping within 25 kb of a known gene resulting in the identification of 42 *LncRNA-HIT*-associated genes by visualization using

the UCSC Genome Browser. The *LncRNA-HIT* ChIRP-seq data sets have been submitted to the NCBI GEO repository under the accession number, GSE70986.

## Peak validation

42 near-peak genes were validated for regulation by *LncRNA-HIT* -p100/CBP using siRNA-mediated reduction of the native *LncRNA-HIT* transcript followed by gene-specific qRTPCR in limb mesenchyme primary cultures. For each gene, a minimum of three independent *LncRNA-HIT* siRNA treatments followed by qRTPCR was performed. A scrambled siRNA control was used in parallel assays for each validation experiment. Genes exhibiting decreased expression $\leq$ -1.5-fold, were assessed for their collective ontological function using the GO consortium ontology tool kit, AmiGo version 2.1.4, using the criteria: experimental biological processes in *M. musculus* (http://amigo2.berkeleybop.org/amigo) [100].

## RNA Isolation and qRTPCR

RNA was isolated using TRIzol (Life Technologies, Grand Island, NY) as instructed by the manufacturer. Gene expression levels were quantitated using a real-time quantitative reverse transcriptase polymerase chain reaction method (qRTPCR). First-strand cDNA was synthesized using an ImProm-II Reverse Transcription System (Promega, Madison, WI). A minimum of three independent samples were used for qRTPCR using a SYBR Green PCR Super Mix and an IQ5 thermal cycler according to the manufacturer's instructions (BioRad Hercules, CA). Fold change expression levels were determined after normalization of the amplification products to *Gapdh* expression using the BioRad IQ5 software suite. Student's t-tests were used to determine statistical significance. Data was plotted using Sigmaplot 10.0 (Systat, San Jose, CA). Primer sequences used for the qRTPCR assays are presented in S3 Table.

## Alcian blue staining

Cartilage nodule formation in the siRNA-treated micromass cultures was visualized by staining the micromass cultures with Alcian blue 8GX (Canemeo Inc, QC, Canada), as described [101]. Alcian blue stained nodules were photographed using a Leica MZFLIII stereoscope fitted with a Canon EOS 40D digital camera.

## RNA tethering assay

RNA tethering assays were performed using expression vectors encoding GAL4-p100 andGAL4-λN as described [5, 64]. Expression plasmids containing five copies of the *BoxB* transcript or containing the same five copies of the *BoxB* transcript fused to the *LncRNA-HIT* transcript (*BoxB-LncRNA-HIT*) were subsequently transfected into NG108-15 cells with the GAL4-λN and/or the GAL4-p100 expression vectors using increasing dosages of the *BoxB* or *BoxB-LncRNA-HIT* vectors at 0, 50, 100, or 250 ng per assay and assessed for luciferase expression 48H after transfection using the Dual-Luciferase Reporter Assay System (Promega, Madison, WI). All analyses were performed in triplicate and the average and standard deviation were plotted using Sigmaplot 10.1.

## *LncRNA-HIT* WDR5 interaction analysis

Flag-WDR5 293T cells were transfected with a *LncRNA-HIT* expression vector and evaluated for WDR5-*LncRNA-HIT* interactions by RNA protein co-immunoprecipitation as described [102].

## RNA affinity assay

Mouse *LncRNA-HIT* and *U1* genes were cloned into the pCDNA3 vector (Life Technologies, Grand Island, NY) and transcribed using T7 RNA polymerase and biotin-16-UTP as described by the manufacturer (Roche, Indianapolis, IN). Cell lysates from E11.0 wild type embryos limbs which were homogenized in mRIPA buffer and centrifuged as described [103]. 1 mg of total protein was incubated with either the *LncRNA-HIT* or *U1* transcripts tagged with biotin-16-UTP 12 H at 4°C. Following incubation, the RNA-protein mixtures were incubated with 100 μl prewashed streptavidin beads for 2.5 H at 4°C. After incubation, the streptavidin beads were washed five times in ice cold PBS and combined with 50 μl of an SDS elution buffer containing 125 mM Tris/HCl pH 6.8, 20% glycerol, 4% SDS, 10% 2-mercaptoethanol and 0.02% bromphenol blue. After two rounds of elution, the recovered protein mixture was heated to 95°C for 5 minutes and fractionated by SDS-PAGE electrophoresis using a 4–12% gradient gel (Nu Page, Invitrogen). The gels were stained with Coomassie blue and the individual bands digested with trypsin and submitted for identification by mass spectroscopy by the OHSU Proteomics Core Facility as described [104].

## *In situ* hybridization

The antisense riboprobe specific for mouse *LncRNA-HIT* was generated using PCR amplification of the gene-specific sequence. The amplifying primers were: 5'-AGAGGAGGTTCCCA-GACTCC-3', and 5'-GCACACAAACACTGATATGCAA-3'. Riboprobe synthesis and *in situ* hybridization was performed on mouse embryos as described [105].

## Quantitative Chromatin immunoprecipitation (qChIP)

Limb buds from E 11.0 embryos were dissected, and the mesenchymal cells were transfected with the previously described *LncRNA-HIT* siRNAs. 24 H after transfection, qChIP assays were performed as described [106–108], using an anti-acetyl histone H3 (lys27) antibody (Cat. 07–360, Millipore, Billerica, MA) or mouse IgG (Diagenode: C15400001) as a negative control. DNA sequences specific to H3K27ac peaks proximal to *LncRNA-HIT* -associated peaks were identified using an overlay of the E 10.5 limb bud H3K27ac ChIP-seq dataset [63] and the UCSC Genome Browser. Primers were selected to amplify the DNA regions containing the H3K27ac *peaks* using Primer3 (simgene.com) (S4 Table). ChIP-enriched H3K27ac fragments were quantified by qPCR using an IQ5 quantitative PCR instrument (BioRad, Hercules, CA) as described [107–108].

## Supporting Information

**S1 Fig. *LncRNA-HIT* gene sequence, conservation and translational analysis. (A)** Location of *LncRNA-HIT*, *Hoxa11*, *Hoxa13*, and *Hottip* on mouse Chromosome 6. Numerical values represent predicted physical distances in kilobase pairs. Arrows represent the direction of gene transcription. **(B)** Sequence of *9530018H14RIK*. Conserved region is underlined in red. Sequences underlined in black represent the potential RNA nuclear retention signal. **(C)** Translation of the *LncRNA-HIT* transcript in all six reading frames. Pink Bars represent stop codons in each reading frame. Green Bars represent initiation codons present in each reading frame. **(D)**. *LncRNA-HIT* (9530018H14Rik) probe sets included on the Affymetrix MOE 430 2.0 Mouse gene chip. **(E)** Conservation of the sequence in multiple vertebrate species. (TIF)

**S2 Fig. Assessment of *LncRNA-HIT*, *Hoxa13*, and *Hottip* for protein coding potential. (A)** Coding potential assessment using CPAT which determined low coding potential for

*LncRNA-HIT* and *Hottip*, and high coding potential for *Hoxa13*. **(B)** Initiation codon translational analysis using NetStart for *LncRNA-HIT*, *Hoxa13*, and *Hottip*. Stop codons are depicted with the letter N, favorable initiation codons are depicted with the letter i.
(TIF)

**S3 Fig. RIP qRTPCR analysis of *LncRNA-HIT*, *Hotair*, or *U1* RNA binding to WDR5.** Blue bars represent gene-specific enrichment after precipitation of FLAG-tagged WDR5 compared to an IgG precipitated control (pink bars). Note that the precipitation of WDR5 did not produce an enrichment of the *LncRNA-HIT* transcript.
(TIF)

**S4 Fig. *LncRNA-HIT* and U1 RNA affinity chromatography identifies a specific interaction between *LncRNA-HIT* and p100 in E 11.0 limb bud cell lysates.** Proteins bound to the *LncRNA-HIT* RNA were eluted using increasing molar amounts of NaCl or Guanidine and fractionated using PAGE. Individual bands were sequences by mass spectroscopy. A prominent 100 Kd band (asterisk) eluting from the *LncRNA-HIT* column but not the U1 column was identified as p100 by protein mass spectrometry.
(TIF)

**S5 Fig. GREAT analysis of *LncRNA-HIT* peak distribution in the murine genome. (A)** Genomic regions with associated gene regulatory regions. Red bar represent the number of *LncRNA-HIT* peaks associating with genomic regions lacking a known *cis*-regulatory function. Open bars represent the number of *LncRNA-HIT* peaks associating with previously characterized *cis*-regulatory element capable of regulating one or more genes. **(B).** Distance between the *LncRNA-HIT* associated genomic regions known translational start sites (TSS). **(C)** Absolute physical distance between *LncRNA-HIT* associated peaks and TSS sites. **(D)** GO analyses based on the association of the lncRNA with known cis-regulatory elements. Note that no functional terms were detected for the *LncRNA-HIT* peaks, suggesting the lncRNA may associate with novel cis-regulatory elements or may function at non *cis* regulatory regions to regulate near-peak gene expression.
(TIF)

**S6 Fig. Ontological analysis of the *LncRNA-HIT*-regulated genes using AMIGO 2.0 for experimental biological processes.** Note that significant ontologies were detected for proximal distal pattern formation and the regulation of chondrocyte differentiation.
(TIF)

**S1 Table. Nucleotide sequences for the LncRNA-HIT RNA FISH probe set.**
(DOCX)

**S2 Table. LncRNA-HIT ChIRP probe set.**
(DOCX)

**S3 Table. qRTPCR primers used to validate LncRNA-HIT regulation of peak-associated genes.**
(DOCX)

**S4 Table. qPCR primers used to detect changes in H3K27ac fragment enrichment.**
(DOCX)

**S5 Table. Sequences associated with murine H3K27ac peaks proximal to the promoter regions of *Skap*2 and *Creb5*.**
(DOCX)

## Acknowledgments

## Author Contributions

Conceived and designed the experiments: HSS HYC. Performed the experiments: HLC YWY CKT. Analyzed the data: HLC HSS JJQ. Contributed reagents/materials/analysis tools: HYC YWY. Wrote the paper: HSS HYC HLC.

## References

1. Okazaki Y, Furuno M, Kasukawa T, Adachi J, Bono H, Kondo S, et al. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. Nature. 2002; 420(6915):563–73. PMID: 12466851

2. Mouse Genome Sequencing C, Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, et al. Initial sequencing and comparative analysis of the mouse genome. Nature. 2002; 420(6915):520–62. PMID: 12466850

3. Waterston RH, Lander ES, Sulston JE. On the sequencing of the human genome. Proceedings of the National Academy of Sciences of the United States of America. 2002; 99(6):3712–6. PMID: 11880605

4. Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell. 2007; 129 (7):1311–23. PMID: 17604720

5. Wang KC, Yang YW, Liu B, Sanyal A, Corces-Zimmerman R, Chen Y, et al. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. Nature. 2011; 472(7341):120–4. doi: 10.1038/nature09819 PMID: 21423168

6. Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. Annu Rev Biochem. 2012; 81:145–66. doi: 10.1146/annurev-biochem-051410-092902 PMID: 22663078

7. Martianov I, Ramadass A, Serra Barros A, Chow N, Akoulitchev A. Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. Nature. 2007; 445(7128):666–70. PMID: 17237763

8. Tripathi V, Ellis JD, Shen Z, Song DY, Pan Q, Watt AT, et al. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Molecular cell. 2010; 39(6):925–38. doi: 10.1016/j.molcel.2010.08.011 PMID: 20797886

9. Kino T, Hurt DE, Ichijo T, Nader N, Chrousos GP. Noncoding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. Sci Signal. 2010; 3(107):ra8. doi: 10.1126/scisignal.2000568 PMID: 20124551

10. Zappulla DC, Cech TR. RNA as a flexible scaffold for proteins: yeast telomerase and beyond. Cold Spring Harb Symp Quant Biol. 2006; 71:217–24. PMID: 17381300

11. Tsai MC, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, et al. Long noncoding RNA as modular scaffold of histone modification complexes. Science. 2010; 329(5992):689–93. doi: 10.1126/science.1192002 PMID: 20616235

12. Yap KL, Li S, Munoz-Cabello AM, Raguz S, Zeng L, Mujtaba S, et al. Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. Molecular cell. 2010; 38(5):662–74. doi: 10.1016/j.molcel.2010.03.021 PMID: 20541999

13. Kotake Y, Nakagawa T, Kitagawa K, Suzuki S, Liu N, Kitagawa M, et al. Long non-coding RNA ANRIL is required for the PRC2 recruitment to and silencing of p15(INK4B) tumor suppressor gene. Oncogene. 2011; 30(16):1956–62. doi: 10.1038/onc.2010.568 PMID: 21151178

14. Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature. 2010; 464(7291):1071–6. doi: 10.1038/nature08975 PMID: 20393566

15. Lyon MF. Gene action in the X-chromosome of the mouse (Mus musculus L.). Nature. 1961; 190:372–3. PMID: 13764598

16. Brockdorff N, Ashworth A, Kay GF, McCabe VM, Norris DP, Cooper PJ, et al. The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. Cell. 1992; 71(3):515–26. PMID: 1423610

17. Lee JT. Lessons from X-chromosome inactivation: long ncRNA as guides and tethers to the epigenome. Genes & development. 2009; 23(16):1831–42.

18. Jeon Y, Lee JT. YY1 tethers Xist RNA to the inactive X nucleation center. Cell. 2011; 146(1):119–33. doi: 10.1016/j.cell.2011.06.026 PMID: 21729784

19. da Rocha ST, Boeva V, Escamilla-Del-Arenal M, Ancelin K, Granier C, Matias NR, et al. Jarid2 Is Implicated in the Initial Xist-Induced Targeting of PRC2 to the Inactive X Chromosome. Molecular cell. 2014; 53(2):301–16. doi: 10.1016/j.molcel.2014.01.002 PMID: 24462204

20. Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, et al. Widespread transcription at neuronal activity-regulated enhancers. Nature. 2010; 465(7295):182–7. doi: 10.1038/nature09033 PMID: 20393465

21. De Santa F, Barozzi I, Mietton F, Ghisletti S, Polletti S, Tusi BK, et al. A large fraction of extragenic RNA pol II transcription sites overlap enhancers. PLoS Biol. 2010; 8(5):e1000384. doi: 10.1371/journal.pbio.1000384 PMID: 20485488

22. Orom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G, et al. Long noncoding RNAs with enhancer-like function in human cells. Cell. 2010; 143(1):46–58. doi: 10.1016/j.cell.2010.09.001 PMID: 20887892

23. Harding K, Wedeen C, McGinnis W, Levine M. Spatially regulated expression of homeotic genes in Drosophila. Science. 1985; 229(4719):1236–42. PMID: 3898362

24. Utset MF, Awgulewitsch A, Ruddle FH, McGinnis W. Region-specific expression of two mouse homeo box genes. Science. 1987; 235(4794):1379–82. PMID: 2881353

25. Chadwick R, McGinnis W. Temporal and spatial distribution of transcripts from the Deformed gene of Drosophila. The EMBO journal. 1987; 6(3):779–89. PMID: 16453753

26. Zakany J, Fromental-Ramain C, Warot X, Duboule D. Regulation of number and size of digits by posterior Hox genes: a dose-dependent mechanism with potential evolutionary implications. Proceedings of the National Academy of Sciences of the United States of America. 1997; 94(25):13695–700. PMID: 9391088

27. Stadler HS, Higgins KM, Capecchi MR. Loss of Eph-receptor expression correlates with loss of cell adhesion and chondrogenic capacity in Hoxa13 mutant limbs. Development. 2001; 128(21):4177–88. PMID: 11684655

28. Kmita M, Duboule D. Organizing axes in time and space; 25 years of colinear tinkering. Science. 2003; 301(5631):331–3. PMID: 12869751

29. Li L, Liu B, Wapinski OL, Tsai MC, Qu K, Zhang J, et al. Targeted disruption of Hotair leads to homeotic transformation and gene derepression. Cell Rep. 2013; 5(1):3–12. doi: 10.1016/j.celrep.2013.09.003 PMID: 24075995

30. Davis AP, Witte DP, Hsieh-Li HM, Potter SS, Capecchi MR. Absence of radius and ulna in mice lacking hoxa-11 and hoxd-11. Nature. 1995; 375(6534):791–5. PMID: 7596412

31. Fromental-Ramain C, Warot X, Messadecq N, LeMeur M, Dolle P, Chambon P. Hoxa-13 and Hoxd-13 play a crucial role in the patterning of the limb autopod. Development. 1996; 122(10):2997–3011. PMID: 8898214

32. Davis AP, Capecchi MR. A mutational analysis of the 5' HoxD genes: dissection of genetic interactions during limb development in the mouse. Development. 1996; 122(4):1175–85. PMID: 8620844

33. Schuettengruber B, Chourrout D, Vervoort M, Leblanc B, Cavalli G. Genome regulation by polycomb and trithorax proteins. Cell. 2007; 128(4):735–45. PMID: 17320510

34. Kim D, Song J, Han J, Kim Y, Chun CH, Jin EJ. Two non-coding RNAs, MicroRNA-101 and HOTTIP contribute cartilage integrity by epigenetic and homeotic regulation of integrin-alpha1. Cellular signalling. 2013; 25(12):2878–87. doi: 10.1016/j.cellsig.2013.08.034 PMID: 24018042

35. Richards EJ, Zhang G, Li ZP, Permuth-Wey J, Challa S, Li Y, et al. Long non-coding RNAs (LncRNA) regulated by transforming growth factor (TGF) beta: LncRNA-HIT-mediated TGFbeta-induced epithelial to mesenchymal transition in mammary epithelia. The Journal of biological chemistry. 2015; 290 (11):6857–67. doi: 10.1074/jbc.M114.610915 PMID: 25605728

36. Kawai J, Shinagawa A, Shibata K, Yoshino M, Itoh M, Ishii Y, et al. Functional annotation of a full-length mouse cDNA collection. Nature. 2001; 409(6821):685–90. PMID: 11217851

37. Wang L, Park HJ, Dasari S, Wang S, Kocher JP, Li W. CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. Nucleic acids research. 2013; 41(6):e74. doi: 10.1093/nar/gkt006 PMID: 23335781

38. Kozak M. At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. Journal of molecular biology. 1987; 196(4):947–50. PMID: 3681984

39. Pedersen AG, Nielsen H. Neural network prediction of translation initiation sites in eukaryotes: perspectives for EST and genome analysis. Proc Int Conf Intell Syst Mol Biol. 1997; 5:226–33. PMID: 9322041

40. Morgan EA, Nguyen SB, Scott V, Stadler HS. Loss of Bmp7 and Fgf8 signaling in Hoxa13-mutant mice causes hypospadia. Development. 2003; 130(14):3095–109. PMID: 12783783

41. Shaut CA, Saneyoshi C, Morgan EA, Knosp WM, Sexton DR, Stadler HS. HOXA13 directly regulates EphA6 and EphA7 expression in the genital tubercle vascular endothelia. Developmental dynamics: an official publication of the American Association of Anatomists. 2007; 236(4):951–60.

42. Perez WD, Weller CR, Shou S, Stadler HS. Survival of Hoxa13 homozygous mutants reveals a novel role in digit patterning and appendicular skeletal development. Developmental dynamics: an official publication of the American Association of Anatomists. 2010; 239(2):446–57.

43. Bono H, Kasukawa T, Furuno M, Hayashizaki Y, Okazaki Y. FANTOM DB: database of Functional Annotation of RIKEN Mouse cDNA Clones. Nucleic acids research. 2002; 30(1):116–8. PMID: 11752270

44. Cabili MN, Dunagin MC, McClanahan PD, Biaesch A, Padovan-Merhar O, Regev A, et al. Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution. Genome Biol. 2015; 16:20. doi: 10.1186/s13059-015-0586-4 PMID: 25630241

45. Zhang B, Gunawardane L, Niazi F, Jahanbani F, Chen X, Valadkhan S. A novel RNA motif mediates the strict nuclear localization of a long noncoding RNA. Mol Cell Biol. 2014; 34(12):2318–29. doi: 10.1128/MCB.01673-13 PMID: 24732794

46. Yang YW, Flynn RA, Chen Y, Qu K, Wan B, Wang KC, et al. Essential role of lncRNA binding for WDR5 maintenance of active chromatin and embryonic stem cell pluripotency. Elife. 2014; 3:e02046. doi: 10.7554/eLife.02046 PMID: 24521543

47. Valineva T, Yang J, Palovuori R, Silvennoinen O. The transcriptional co-activator protein p100 recruits histone acetyltransferase activity to STAT6 and mediates interaction between the CREB-binding protein and STAT6. The Journal of biological chemistry. 2005; 280(15):14989–96. PMID: 15695802

48. Valineva T, Yang J, Silvennoinen O. Characterization of RNA helicase A as component of STAT6-dependent enhanceosome. Nucleic acids research. 2006; 34(14):3938–46. PMID: 16914450

49. Goenka S, Cho SH, Boothby M. Collaborator of Stat6 (CoaSt6)-associated poly(ADP-ribose) polymerase activity modulates Stat6-dependent gene transcription. The Journal of biological chemistry. 2007; 282(26):18732–9. PMID: 17478423

50. Shaw N, Zhao M, Cheng C, Xu H, Saarikettu J, Li Y, et al. The multifunctional human p100 protein 'hooks' methylated ligands. Nature structural & molecular biology. 2007; 14(8):779–84.

51. Chu C, Qu K, Zhong FL, Artandi SE, Chang HY. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. Molecular cell. 2011; 44(4):667–78. doi: 10.1016/j.molcel.2011.08.027 PMID: 21963238

52. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009; 10(3):R25. doi: 10.1186/gb-2009-10-3-r25 PMID: 19261174

53. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008; 9(9):R137. doi: 10.1186/gb-2008-9-9-r137 PMID: 18798982

54. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, et al. GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol. 2010; 28(5):495–501. doi: 10.1038/nbt.1630 PMID: 20436461

55. Solursh M, Ahrens PB, Reiter RS. A tissue culture analysis of the steps in limb chondrogenesis. In Vitro. 1978; 14(1):51–61. PMID: 203522

56. Owens EM, Solursh M. Cell-cell interaction by mouse limb cells during in vitro chondrogenesis: analysis of the brachypod mutation. Developmental biology. 1982; 91(2):376–88. PMID: 7095270

57. Gamer LW, Cox KA, Small C, Rosen V. Gdf11 is a negative regulator of chondrogenesis and myogenesis in the developing chick limb. Developmental biology. 2001; 229(2):407–20. PMID: 11203700

58. Miura T, Shiota K. Extracellular matrix environment influences chondrogenic pattern formation in limb bud micromass culture: experimental verification of theoretical models. Anat Rec. 2000; 258(1):100–7. PMID: 10603453

59. Mathieu M, Iampietro M, Chuchana P, Guerit D, Djouad F, Noel D, et al. Involvement of angiopoietin-like 4 in matrix remodeling during chondrogenic differentiation of mesenchymal stem cells. The Journal of biological chemistry. 2014; 289(12):8402–12. doi: 10.1074/jbc.M113.539825 PMID: 24505142

60. Song JJ, Aswad R, Kanaan RA, Rico MC, Owen TA, Barbe MF, et al. Connective tissue growth factor (CTGF) acts as a downstream mediator of TGF-beta1 to induce mesenchymal cell condensation. J Cell Physiol. 2007; 210(2):398–410. PMID: 17111364

61. Wang J, Wang X, Holz JD, Rutkowski T, Wang Y, Zhu Z, et al. Runx1 is critical for PTH-induced onset of mesenchymal progenitor cell chondrogenic differentiation. PLoS One. 2013; 8(9):e74255. doi: 10.1371/journal.pone.0074255 PMID: 24058535

62. Tian Y, Xu Y, Fu Q, Chang M, Wang Y, Shang X, et al. Notch inhibits chondrogenic differentiation of mesenchymal progenitor cells by targeting Twist1. Mol Cell Endocrinol. 2015; 403:30–8. doi: 10.1016/j.mce.2015.01.015 PMID: 25596548

63. Cotney J, Leng J, Oh S, DeMare LE, Reilly SK, Gerstein MB, et al. Chromatin state signatures associated with tissue-specific gene expression and enhancer activity in the embryonic limb. Genome Res. 2012; 22(6):1069–80. doi: 10.1101/gr.129817.111 PMID: 22421546

64. Baron-Benhamou J, Gehring NH, Kulozik AE, Hentze MW. Using the lambdaN peptide to tether proteins to RNAs. Methods in molecular biology. 2004; 257:135–54. PMID: 14770003

65. Milaire J, Mulnard J. Histogenesis in 11-day mouse embryo limb buds explanted in organ culture. J Exp Zool. 1984; 232(2):359–77. PMID: 6502102

66. Ng LJ, Wheatley S, Muscat GE, Conway-Campbell J, Bowles J, Wright E, et al. SOX9 binds DNA, activates transcription, and coexpresses with type II collagen during chondrogenesis in the mouse. Developmental biology. 1997; 183(1):108–21. PMID: 9119111

67. Zhao Q, Eberspaecher H, Lefebvre V, De Crombrugghe B. Parallel expression of Sox9 and Col2a1 in cells undergoing chondrogenesis. Developmental dynamics: an official publication of the American Association of Anatomists. 1997; 209(4):377–86.

68. Kronenberg HM. Developmental regulation of the growth plate. Nature. 2003; 423(6937):332–6. PMID: 12748651

69. Akiyama H, Chaboissier MC, Martin JF, Schedl A, de Crombrugghe B. The transcription factor Sox9 has essential roles in successive steps of the chondrocyte differentiation pathway and is required for expression of Sox5 and Sox6. Genes & development. 2002; 16(21):2813–28.

70. Akiyama H, Kim JE, Nakashima K, Balmes G, Iwai N, Deng JM, et al. Osteo-chondroprogenitor cells are derived from Sox9 expressing precursors. Proceedings of the National Academy of Sciences of the United States of America. 2005; 102(41):14665–70. PMID: 16203988

71. Liu Q, Zhang X, Dai L, Hu X, Zhu J, Li L, et al. Long noncoding RNA related to cartilage injury promotes chondrocyte extracellular matrix degradation in osteoarthritis. Arthritis Rheumatol. 2014; 66 (4):969–78. doi: 10.1002/art.38309 PMID: 24757148

72. Maass PG, Rump A, Schulz H, Stricker S, Schulze L, Platzer K, et al. A misplaced lncRNA causes brachydactyly in humans. J Clin Invest. 2012; 122(11):3990–4002. doi: 10.1172/JCI65508 PMID: 23093776

73. Goldring MB, Marcu KB. Epigenomic and microRNA-mediated regulation in cartilage development, homeostasis, and osteoarthritis. Trends Mol Med. 2012; 18(2):109–18. doi: 10.1016/j.molmed.2011.11.005 PMID: 22178468

74. Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, et al. GENCODE: the reference human genome annotation for The ENCODE Project. Genome Res. 2012; 22(9):1760–74. doi: 10.1101/gr.135350.111 PMID: 22955987

75. Cottrill CP, Archer CW, Wolpert L. Cell sorting and chondrogenic aggregate formation in micromass culture. Developmental biology. 1987; 122(2):503–15. PMID: 3596020

76. Thorogood PV, Hinchliffe JR. An analysis of the condensation process during chondrogenesis in the embryonic chick hind limb. J Embryol Exp Morphol. 1975; 33(3):581–606. PMID: 1176861

77. Lim J, Tu X, Choi K, Akiyama H, Mishina Y, Long F. BMP-Smad4 signaling is required for precartilaginous mesenchymal condensation independent of Sox9 in the mouse. Developmental biology. 2015; 400(1):132–8. doi: 10.1016/j.ydbio.2015.01.022 PMID: 25641697

78. ten Dijke P, Yamashita H, Sampath TK, Reddi AH, Estevez M, Riddle DL, et al. Identification of type I receptors for osteogenic protein-1 and bone morphogenetic protein-4. The Journal of biological chemistry. 1994; 269(25):16985–8. PMID: 8006002

79. Nishitoh H, Ichijo H, Kimura M, Matsumoto T, Makishima F, Yamaguchi A, et al. Identification of type I and type II serine/threonine kinase receptors for growth/differentiation factor-5. The Journal of biological chemistry. 1996; 271(35):21345–52. PMID: 8702914

80. Baur ST, Mai JJ, Dymecki SM. Combinatorial signaling through BMP receptor IB and GDF5: shaping of the distal mouse limb and the genetics of distal limb diversity. Development. 2000; 127(3):605–19. PMID: 10631181

81. Lehmann K, Seemann P, Stricker S, Sammar M, Meyer B, Suring K, et al. Mutations in bone morpho-genetic protein receptor 1B cause brachydactyly type A2. Proceedings of the National Academy of Sciences of the United States of America. 2003; 100(21):12277–82. PMID: 14523231

82. Chimal-Monroy J, Rodriguez-Leon J, Montero JA, Ganan Y, Macias D, Merino R, et al. Analysis of the molecular cascade responsible for mesodermal limb chondrogenesis: Sox genes and BMP signaling. Developmental biology. 2003; 257(2):292–301. PMID: 12729559

83. Suzuki T, Hasso SM, Fallon JF. Unique SMAD1/5/8 activity at the phalanx-forming region determines digit identity. Proceedings of the National Academy of Sciences of the United States of America. 2008; 105(11):4185–90. doi: 10.1073/pnas.0707899105 PMID: 18334652

84. Knosp WM, Scott V, Bachinger HP, Stadler HS. HOXA13 regulates the expression of bone morpho-genetic proteins 2 and 7 to control distal limb morphogenesis. Development. 2004; 131(18):4581–92. PMID: 15342482

85. Gross S, Krause Y, Wuelling M, Vortkamp A. Hoxa11 and Hoxd11 regulate chondrocyte differentia-tion upstream of Runx2 and Shox2 in mice. PLoS One. 2012; 7(8):e43553. doi: 10.1371/journal.pone. 0043553 PMID: 22916278

86. Ilott NE, Heward JA, Roux B, Tsitsiou E, Fenwick PS, Lenzi L, et al. Long non-coding RNAs and enhancer RNAs regulate the lipopolysaccharide-induced inflammatory response in human mono-cytes. Nat Commun. 2014; 5:3979. doi: 10.1038/ncomms4979 PMID: 24909122

87. Yang J, Aittomaki S, Pesu M, Carter K, Saarinen J, Kalkkinen N, et al. Identification of p100 as a coac-tivator for STAT6 that bridges STAT6 with RNA polymerase II. The EMBO journal. 2002; 21 (18):4950–8. PMID: 12234934

88. Wellik DM, Capecchi MR. Hox10 and Hox11 genes are required to globally pattern the mammalian skeleton. Science. 2003; 301(5631):363–7. PMID: 12869760

89. Xu B, Wellik DM. Axial Hox9 activity establishes the posterior field in the developing forelimb. Pro-ceedings of the National Academy of Sciences of the United States of America. 2011; 108(12):4888–91. doi: 10.1073/pnas.1018161108 PMID: 21383175

90. Kang Y, He W, Tulley S, Gupta GP, Serganova I, Chen CR, et al. Breast cancer bone metastasis mediated by the Smad tumor suppressor pathway. Proceedings of the National Academy of Sciences of the United States of America. 2005; 102(39):13909–14. PMID: 16172383

91. Nam JS, Suchar AM, Kang MJ, Stuelten CH, Tang B, Michalowska AM, et al. Bone sialoprotein medi-ates the tumor cell-targeted prometastatic activity of transforming growth factor beta in a mouse model of breast cancer. Cancer Res. 2006; 66(12):6327–35. PMID: 16778210

92. Roberts AB, Russo A, Felici A, Flanders KC. Smad3: a key player in pathogenetic mechanisms dependent on TGF-beta. Ann N Y Acad Sci. 2003; 995:1–10.

93. Yu L, Liu X, Cui K, Di Y, Xin L, Sun X, et al. SND1 Acts Downstream of TGFbeta1 and Upstream of Smurf1 to Promote Breast Cancer Metastasis. Cancer Res. 2015; 75(7):1275–86. doi: 10.1158/0008-5472.CAN-14-2387 PMID: 25596283

94. Liu X, Dong L, Zhang X, Wang B, Wang X, Li H, et al. Identification of p100 target promoters by chro-matin immunoprecipitation-guided ligation and selection (ChIP-GLAS). Cell Mol Immunol. 2011; 8 (1):88–91. doi: 10.1038/cmi.2010.47 PMID: 20921938

95. Shaut CA, Keene DR, Sorensen LK, Li DY, Stadler HS. HOXA13 Is essential for placental vascular patterning and labyrinth endothelial specification. PLoS Genet. 2008; 4(5):e1000073. doi: 10.1371/journal.pgen.1000073 PMID: 18483557

96. Quagliata L, Matter MS, Piscuoglio S, Arabi L, Ruiz C, Procino A, et al. Long noncoding RNA HOT-TIP/HOXA13 expression is associated with disease progression and predicts outcome in hepatocellu-lar carcinoma patients. Hepatology. 2014; 59(3):911–23. doi: 10.1002/hep.26740 PMID: 24114970

97. Pan TT, Jia WD, Yao QY, Sun QK, Ren WH, Huang M, et al. Overexpression of HOXA13 as a poten-tial marker for diagnosis and poor prognosis of hepatocellular carcinoma. Tohoku J Exp Med. 2014; 234(3):209–19. PMID: 25341685

98. Yang B, Guo H, Zhang Y, Chen L, Ying D, Dong S. MicroRNA-145 regulates chondrogenic differentia-tion of mesenchymal stem cells by targeting Sox9. PLoS One. 2011; 6(7):e21679. doi: 10.1371/journal.pone.0021679 PMID: 21799743

99. Zhang L, Yang M, Marks P, White LM, Hurtig M, Mi QS, et al. Serum non-coding RNAs as biomarkers for osteoarthritis progression after ACL injury. Osteoarthritis Cartilage. 2012; 20(12):1631–7. doi: 10. 1016/j.joca.2012.08.016 PMID: 22944527

100. Carbon S, Ireland A, Mungall CJ, Shu S, Marshall B, Lewis S, et al. AmiGO: online access to ontology and annotation data. Bioinformatics. 2009; 25(2):288–9. doi: 10.1093/bioinformatics/btn615 PMID: 19033274

101.    Francis-West PH, Abdelfattah A, Chen P, Allen C, Parish J, Ladher R, et al. Mechanisms of GDF-5 action during skeletal development. Development. 1999; 126(6):1305–15. PMID: 10021348

102.    Gomez JA, Wapinski OL, Yang YW, Bureau JF, Gopinath S, Monack DM, et al. The NeST long ncRNA controls microbial susceptibility and epigenetic activation of the interferon-gamma locus. Cell. 2013; 152(4):743–54. doi: 10.1016/j.cell.2013.01.015 PMID: 23415224

103.    Izumi H, McCloskey A, Shinmyozu K, Ohno M. p54nrb/NonO and PSF promote U snRNA nuclear export by accelerating its export complex assembly. Nucleic acids research. 2014; 42(6):3998–4007. doi: 10.1093/nar/gkt1365 PMID: 24413662

104.    Wilmarth PA, Taube JR, Riviere MA, Duncan MK, David LL. Proteomic and sequence analysis of chicken lens crystallins reveals alternate splicing and translational forms of beta B2 and beta A2 crystallins. Invest Ophthalmol Vis Sci. 2004; 45(8):2705–15. PMID: 15277495

105.    Manley NR, Capecchi MR. The role of Hoxa-3 in mouse thymus and thyroid development. Development. 1995; 121(7):1989–2003. PMID: 7635047

106.    Cho OH, Rivera-Perez JA, Imbalzano AN. Chromatin immunoprecipitation assay for tissue-specific genes using early-stage mouse embryos. J Vis Exp. 2011;(50: ).

107.    Mukhopadhyay A, Deplancke B, Walhout AJ, Tissenbaum HA. Chromatin immunoprecipitation (ChIP) coupled to detection by quantitative real-time PCR to study transcription factor binding to DNA in Caenorhabditis elegans. Nat Protoc. 2008; 3(4):698–709. doi: 10.1038/nprot.2008.38 PMID: 18388953

108.    Caby S, Pierce RJ. Quantitative chromatin immunoprecipitation (Q-ChIP) applied to Schistosoma mansoni. Mol Biochem Parasitol. 2009; 166(1):77–80. doi: 10.1016/j.molbiopara.2009.02.014 PMID: 19428676