SMAD4: a Multifunctional Regulator of Limb Bud Initiation and Outgrowth

Inauguraldissertation
zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch Naturwissenschaftlichen Fakultät
der Universität Basel

Von

Julie Gamart

aus Lille, Frankreich

Basel, 2017

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von

Prof. Dr. Rolf Zeller (Dissertationsleiter und Fakultätsverantwortlicher),

Dr. Vanessa Ribes (Korreferentin)

Basel, den 12. Dezember 2017

Dekan Prof. Dr. Martin Spiess

1. Table of contents

2. Summary

During mouse embryonic development, the spatio-temporal expression of genes is controlled by both interlinked signalling pathways and interactions between transcription factors and their target *cis*-regulatory modules. To gain global insights into the roles of a *trans*-acting transcriptional regulator in a specific tissue, the genome-wide profiling of its target regulatory regions and their association with the putative target genes are essential. Therefore, I have combined several types of genome-wide analyses such as ChIP-seq using epitope-tagged transcription factors with ATAC-seq and RNA-seq to study the functions of HAND2 and SMAD4 during heart and limb bud development, respectively.

In *Hand2*-deficient embryos, we observed that cells of the atrioventricular canal do not undergo the endothelial-mesenchymal transition that underlies cardiac cushion development. By combining HAND2^{3xF} ChIP-seq and RNA-seq analysis, we have identified the HAND2 gene regulatory network involved in these processes and show that HAND2 is a key regulator of heart valve development.

Limb bud outgrowth and patterning are regulated by a self-regulatory feedback signalling system operating between the SHH and FGF signalling pathways that critically depends on the BMP antagonist GREMLIN1. However, the establishment of these signalling feedback loops requires initiation of *Gremlin1* expression by high BMP activity. For my PhD research, I have investigated the roles of the BMP signalling pathway during limb bud initiation by studying the functions of the BMP signal transducer SMAD4. By combining genome-wide SMAD4^{3xF} ChIP-seq, ATAC-seq and RNA-seq analyses, I am able to show that SMAD4 participates in activation of *Gremlin1* expression by interacting with *Grem1* coding exon 2 (a putative regulatory region). Furthermore, the identification of the SMAD4 gene regulatory network reveals multiple functions of SMAD4 during the onset of limb bud development. Especially, SMAD4

directly regulates target genes involved in limb bud outgrowth and patterning. Rather unexpected, my analysis reveals that SMAD4 directly regulates cholesterol homeostasis and controls the gradient and activity of the SHH signalling pathway during early limb bud development.

3. List of Abbreviations

7DHC	7-dehydro-cholesterol
AER	Apical Ectodermal Ridge
AP	Antero-posterior
ATAC	Assay for transposase-accessible chromatin
BMP	Bone morphogenetic protein
BMPR	BMP receptor
BRE	BMP responsive element
BSA	Bovine serum albumin
cDNA	Complementary DNA
ChIP	Chromatin Immunoprecipitation
co-Smad	common-SMAD
Cq	Quantification cycle
CRM	*Cis*-regulatory module
CTCF	CCCTC-binding Factor
DAB	3,3'-Diaminobenzidine
DBH	Dopamine beta-hydroxylase
DEG	Differentially expressed gene
DIG	Digoxigenin
DMEM	Dulbecco's phosphate buffered saline
DMSO	Dimethyl sulfoxide
DV	Dorso-ventral
E	Embryonic day
EDTA	Ethylenediaminetetraacetic acid
EGTA	ethylene glycol-bis(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid
EMFI	Embryonic mouse fibroblast
EMT	Epithelial-mesenchymal transition
ES cell	Embryonic stem cell
FBS	Fetal bovine serum
FCS	Fetal calf serum
FGF	Fibroblast growth factor
FMN1	Formin1
GCR	Global control region
GFP	Green fluorescent protein
GO	Gene ontology
GRE1	GLI responsive element 1
GREM1	Gremlin1
GRN	Gene regulatory network
GRS1	Gremlin regulatory sequence 1
HAND2	Heart, Autonomic nervous system and Neural crest Derivatives 2
hCG	Human chorionic gonadotropin
HMCO	Human Mouse Chicken Opossum
HR	Homologous Recombination
hrs	Hours
ID1,2,3	Inhibitor of Differentiation 1,2,3
IHC	Immunohistochemistry
LIF	Leukemia Inhibitory Factor
MH	MAD homology domain
MSX2	Msh homeobox 2
O.C.T.	Optimum cutting temperature
ON	Overnight
PBS	Phosphate Buffered Saline
PBT	PBS with 0.1% Tween 20
PD	Proximo-distal
PFA	Paraformaldehyde
polyA	Polyadenylation
PRRX1,2	Paired Related homeobox 1,2
pSMAD	phosphorylated SMAD
Ptch1	Patched 1
PTH1R	parathyroid hormone 1 receptor
R-Smad	Receptor SMAD
RPL19	Ribosomal Protein L19
rpm	Revolutions per minute
RT	Room Temperature
RT-qPCR	Real-time quantitative PCR
SAG	smoothened antagonist
SBE	SMAD binding element
SD	Standard deviation
SDS	Sodium Dodecyl Sulfate
SHH	Sonic Hedgehog
SMAD	Small Mothers Against Decapentaplegic
SMO	Smoothened
SNAI1	Snail family zinc finger 1
TF	Transcription Factor
TGFβ	Transforming Growth Factor Beta

TSS	Transcriptional Start Site	**WISH**	Whole-mount *in situ* Hybridization
UTR	Untranslated Region	**Wnt**	Wingless-type MMTV integration site
UCSC	University of California Santa Cruz	**WT**	Wild-Type
WB	Western Blot	**ZPA**	zone of polarizing activity
		ZRS	ZPA regulatory sequence

4. Introduction

4.1. *Cis*-regulatory modules and regulation of gene expression

A major open question is how pools of undifferentiated progenitors acquire positional information and how their differentiation generates the different cell-types that give rise to functional tissues and organs during embryonic development. Cell fates are controlled by cells responding to different extracellular stimuli such as graded signals and morphogens. These cell-cell signalling interactions trigger the intracellular signalling cascades that control the robustness of spatio-temporal gene expression, which is controlled by specific interactions of transcription factor (TF) complexes with gene regulatory regions called *cis*-regulatory modules (CRMs).

The activity of CRMs is tissue-specific and results in enhancing or repressing the transcription of target genes (Bulger and Groudine, 2011; Levine, 2010; Ong and Corces, 2011). Gene expression is controlled either by one specific or several CRMs which can have additive, synergistic or redundant functions and provide a robust regulation of gene expression (Barolo, 2012; Frankel et al., 2010). CRMs are located either in vicinity to the promoter, intragenic or intergenic regions at distances varying from several kilobases (kb) to 2-3 megabases (Mb, Krivega and Dean, 2012; Zeller and Zuniga, 2007). They define the *cis*-regulatory landscape of the target gene, whose sizes vary according to the number of CRMs and their distance to the promoter. During limb bud development, many of them have been identified and a recent genome-wide study has characterized and classified more than 400 limb-associated *cis*-regulatory landscapes (Andrey et al., 2017). One of the best studied cases of long range regulation is the ZPA regulatory sequence (ZRS), which is a CRM located around 800kb upstream of its transcriptional target gene encoding the Sonic Hedgehog (SHH) ligand (Lettice et al., 2003). The

ZRS is highly conserved and is located in the intron 5 of the *Lmbr1* gene (Zeller and Zuniga, 2007). In general, the activity of CRMs correlates with open or accessible chromatin, which due to their sensitivity to transposases, can be identified by ATAC-seq (Buenrostro et al., 2013, 2015). The interactions of CRMs with TF-complexes can be studied by chromatin immunoprecipitaton followed by deep-sequencing (ChIP-seq, Figure 1A, Visel et al., 2009).

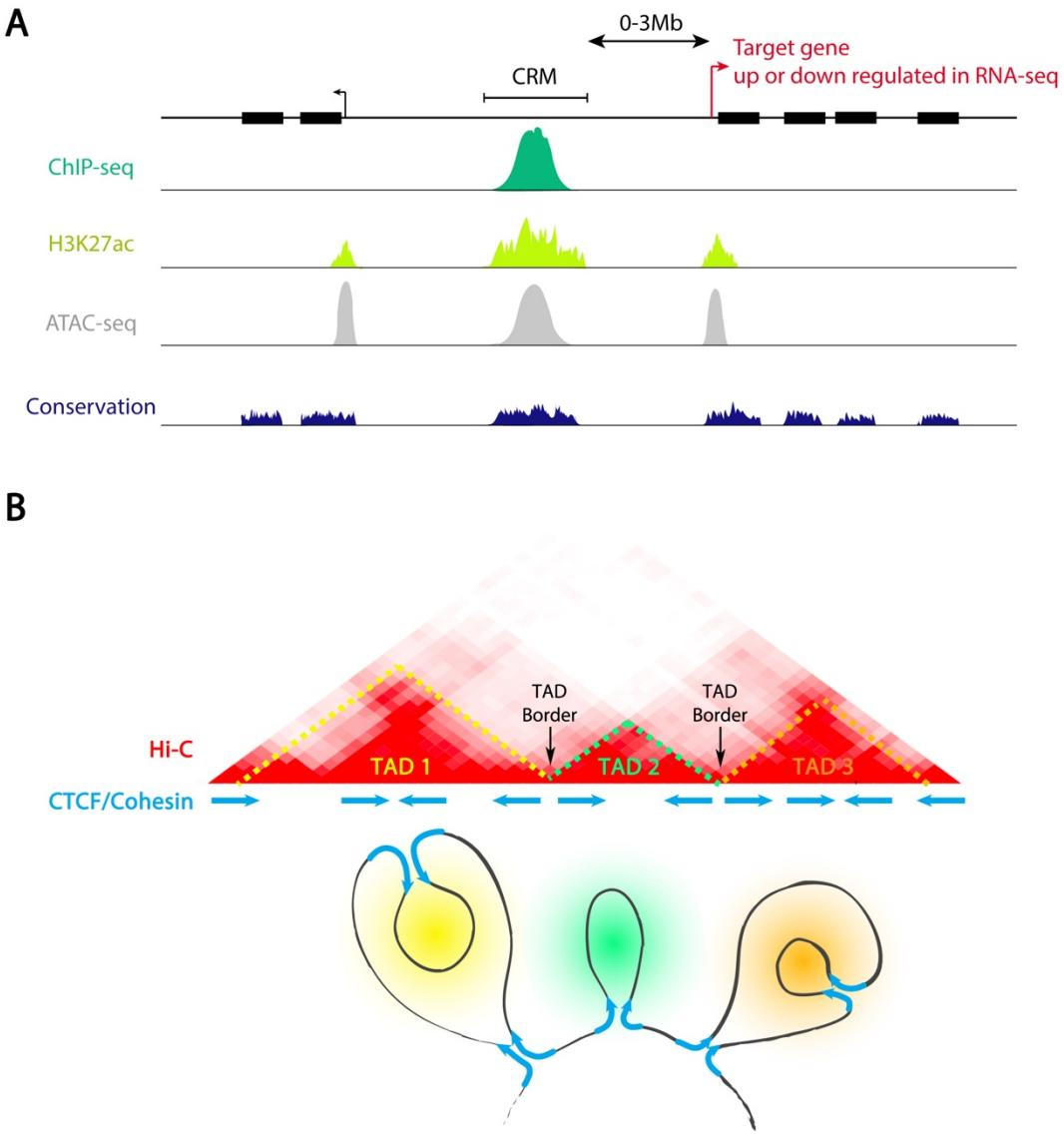


Figure 1: *Cis*-regulatory modules and 3D genome architecture
(A) Schematic representation of a *cis*-regulatory module (CRM) enriched by ChIP-seq detecting specific transcription factor complexes (dark green) and decorated by H3K27ac marks (light green). ATAC-seq shows that the CRM is overlapping a region of open chromatin (grey). Conserved sequences are indicated in dark blue. Bioinformatics analysis

associates this CRM with the closest gene, whose expression is altered in the transcriptome of a mouse mutant lacking the transcription factor of interest. The distance between the CRM and its target gene can be up to 3Mb. (B) Representation of the 3D genomic architecture at the locus of interest and the likely chromatin loops that form depending on the orientation of the CTCF/Cohesin binding sites. Two converging CTCF binding sites will initiate a chromatin looping within a particular topologically associated domain (TAD). Diverging CTCF binding sites will form TAD boundaries.

In addition, post-translational modifications of histones in nucleosomes such as acetylation of lysine 27 and/or monomethylation of lysine 4 of the histone 3 are used as a readout of the activating trans-regulatory activity (H3K27ac and H3K4me1, Figure 1A, Shlyueva et al., 2014). The evolutionary sequence conservation among different species is an excellent criteria for identifying functionally and evolutionary relevant CRMs (Lopez-Rios et al., 2014), even though mere sequence conservation does not mean that CRM functions are conserved (Nelson and Wardle, 2013).

Interactions among CRMs with promoter regions are detected using Circular Chromosome Conformation Capture (4C) and long-range interactions at a genomic scale are identified by Hi-C (Matharu and Ahituv, 2015; Mora et al., 2016; Whalen et al., 2016). Interactions between distant acting-CRMs and the basal transcriptional machinery located at the promoter of the target gene are facilitated by chromatin three-dimensional conformation changes and looping within the genomic landscape encompassing the target transcription units (Dixon et al., 2012; Mora et al., 2016; Stevens et al., 2017). The chromatin looping mediates the contacts between the relevant CRMs and the transcription start site (TSS) of the target gene (Matharu and Ahituv, 2015). The interactions of Cohesin and CCCTC-binding factor (CTCF) complexes with their genomic target regions are essential for the formation of chromatin loops: it has been shown that the orientation of the CTCF binding site provides the loops with directionality (Rao et al., 2014). CTCF-sites in convergent orientations are predominantly found in chromatin loops and are associated with domains of chromatin interactions or topologically associated domain (TADs, Figure 1B, Dixon et al., 2012; Rao et al., 2014). In contrast, enrichment of divergent

Cohesin-CTCF complexes is associated with borders between TADs (Figure 1B, Gómez-Marín et al., 2015).

However, to associate the distant-acting CRMs to their real target genes can be challenging. It is now state of the art to use genome-wide approaches to try to associate CRMs to their transcriptional target genes. For instance, a CRM in an open chromatin region that interacts specifically with complexes containing TF-X, can be associated to closest gene, whose expression is altered when TF-X is genetically inactivated (Figure 1A). In addition, the *cis*-regulatory effects of a distant CRM are in general limited to transcription units located in the same TAD as the CRM (Dixon et al., 2012). The *cis*-regulatory potential of candidate CRMs can be functionally assessed by their capacity to drive the expression of a reporter gene in transgenic mouse embryos.

The developing limb bud is an excellent model to study how signalling interactions impact on gene regulation. Research from our group and others has shown that limb outgrowth and patterning depends on interconnected gene regulators networks (GRNs) that interact in a spatio-temporally highly dynamic fashion (reviewed by Zuniga, 2015).

4.2. Signalling interactions during early limb bud development

The tetrapod limb bud originates from the lateral plate mesoderm as a bulge of mesenchymal cells enveloped in an epithelium of ectodermal cells. Mouse limb buds grow out perpendicular to the body axis and are patterned along three distinct axes - the proximo-distal (PD), antero-posterior (AP) and dorso-ventral axis (DV, Zeller et al., 2009). In the nascent limb bud mesenchyme, interactions between the transcriptional regulators HAND2 and GLI3 and between HAND2 and the ZRS are essential to activate *Shh* expression specifically in the cells of

the zone of polarizing activity (ZPA, Galli et al., 2010; Osterwalder et al., 2014). During outgrowth, limb bud development is controlled by two signalling centers: the SHH signalling in the ZPA and the apical-ectodermal ridge (AER), which produces several FGF ligands. Between these two signalling centers, the SHH/GREM1/FGF feedback loop is established as the expression of Gremlin1 (GREM1), a Bone Morphogenic Protein (BMP) antagonist, is activated (Khokha et al., 2003; Michos et al., 2004; Zuniga and Zeller, 1999; Zuniga et al., 2004). During limb bud initiation (embryonic days E9.5-E10.0, Figure 2A), high levels of BMP activity are required to establish a functional AER and to initiate *Grem1* expression in the posterior limb bud mesenchyme (Bénazet et al., 2009; Nissim et al., 2006; Zeller et al., 2009). The increase in *Grem1* expression progressively lowers BMP activity in the limb bud mesenchyme, which in turn regulates AER length (Bénazet et al., 2009). Between E10.25 to E11.0, the propagation of the SHH/GREM1/FGF signalling feedback loop (Figure 2B) induces the proliferative expansion of the limb bud mesenchymal progenitors (LMPs). After E11.0, the number of mesenchymal cells refractory to *Grem1* expression is increased in the posterior mesenchyme in concert with distal-anterior expansion of the *Grem1* expression domain (Bénazet et al., 2009; Nissim et al., 2006). The increase of AER-FGFs negatively regulates *Grem1* expression and contributes to the termination of the epithelial-mesenchymal feedback signalling system (Figure 2C, Verheyden and Sun, 2008). During this termination phase, the BMP activity increases, which promotes the exit of the digit progenitors toward chondrogenic differentiation (Bénazet et al., 2012; Lopez-Rios et al., 2012; Pizette and Niswander, 2001).

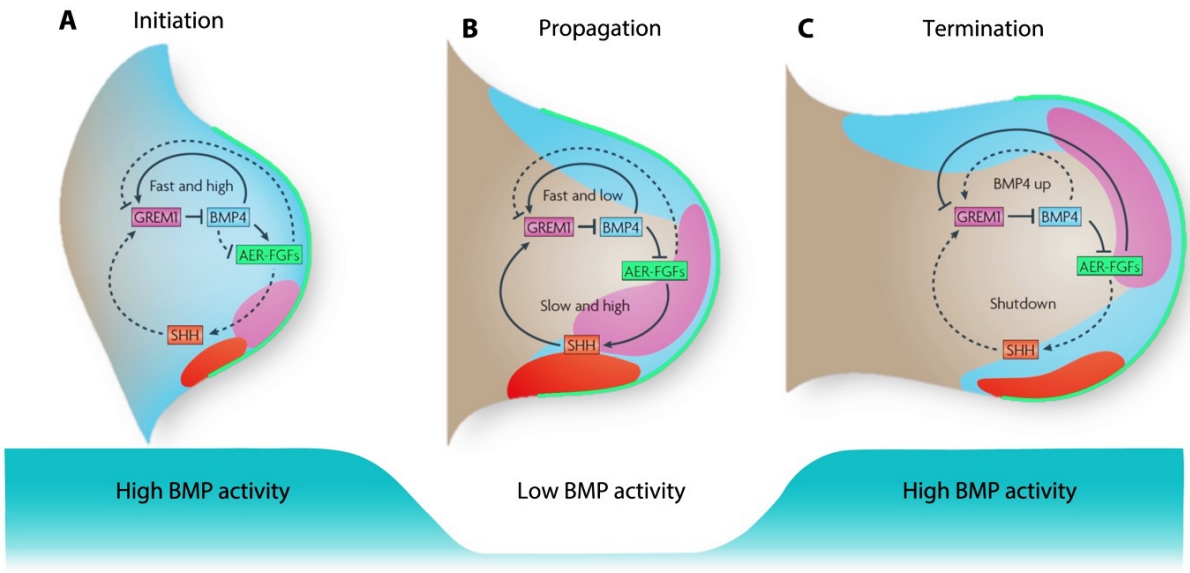


Figure 2: The feedback signalling system that controls limb bud outgrowth
(A) Initiation phase. High BMP activity (blue) in the mesenchyme maintains the AER-*Fgf* expression (green) and activates *Grem1* expression (pink). (B) Propagation phase: GREM1 antagonises BMP activity. This enables AER-FGF signalling to up-regulate *Shh* expression in the ZPA (red), which in turn enhances *Grem1* expression. These interactions define the SHH/GREM1/FGF feedback signalling system. (C) Termination phase: *Grem1* expression is down-regulated as its expression is increasingly inhibited by AER-FGFs and the descendants of *Shh* expressing cells become refractory to *Grem1* expression. This termination of feedback signalling and GREM1-mediated antagonism results in a renewed increase of BMP activity, which allows LMPs to initiate chondrogenic differentiation. Figure adapted from Zeller et al., 2009.

The study of *cis*-regulatory landscapes and the identification of CRMs involved in limb bud development leads to a better understanding of how these signalling feedback loops are interlinked. For instance, we have more knowledge now on how *Shh* expression is regulated by transcriptional regulators such as HOX proteins, HAND2 and ETS transcription factors interacting with the ZRS (Capellini et al., 2006; Galli et al., 2010; Lettice et al., 2012; Osterwalder et al., 2014). The *Grem1* *cis*-regulatory landscape has also been intensively investigated to understand how it integrates inputs from different signalling pathways (Li et al., 2014; Zeller and Zuniga, 2007; Zuniga et al., 2012). In addition, genome-wide analysis of key TF such as HAND2 (Osterwalder et al., 2014), HOXD13 (Sheth et al., 2016) or TWIST (Lee et al., 2014) allowed to decipher their specific roles during limb bud development. However, many TF

are expressed in different developing tissues, like HAND2 which is also involved in heart development. The same genome-wide approach allows us to study if the same or different HAND2 GRNs control heart and limb bud development.

4.3. The BMP signalling pathway during limb bud development

BMP ligands belong to the transforming growth factor beta (TGFβ) superfamily. To activate the BMP signalling pathway, BMP ligands bind two types of transmembrane receptors (BMPR-IA/IB and BMPR-II), which form hetero-tetrameric complexes. This triggers the phosphorylation of the receptor-associated SMAD protein (R-SMADs) SMAD1, SMAD5 and SMAD8. The phosphorylated R-SMADs form a complex with SMAD4, which then translocates to the nucleus and controls the expression of target genes via BMP responsive elements (BREs). SMAD6 and SMAD7 are inhibitory SMADs (I-SMADs), which negatively regulate BMP signalling transduction. TGFβ signalling pathway acts with the same mechanism, but the signal is transduced via the R-SMADs SMAD2 and SMAD3 that form complexes with SMAD4. Therefore, SMAD4 participates in mediating the transcriptional regulation in response to both BMP and TGFβ signal transduction (Pignatti et al., 2014, Figure 3). SMAD4 contains two evolutionary conserved MAD homology (MH) domains separated by a linker. The MH1 domain located in the N-terminal part of SMAD4 interacts with DNA, mediates the protein-protein interactions and contains the nuclear localization signal (NLS). The MH2 domain located in the C-terminal part is phosphorylated, regulates SMAD oligomerisation and contains the nuclear export signal (NES, Moustakas and Heldin, 2009).

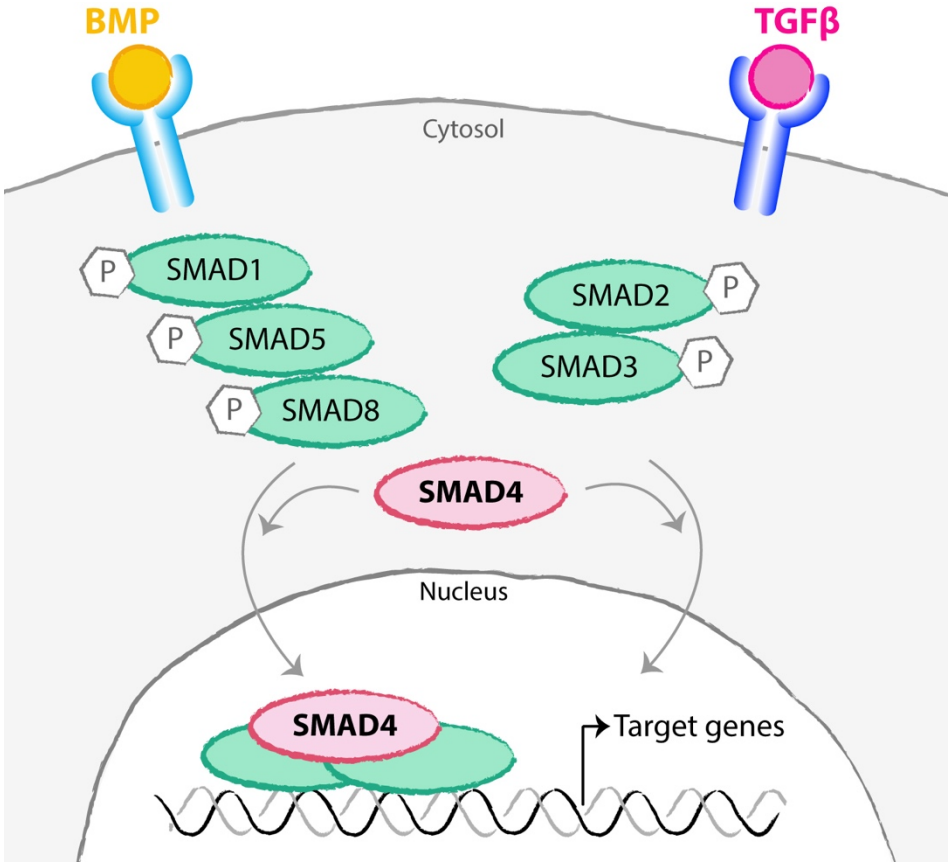


Figure 3: The BMP/TGFβ signalling pathways
BMP ligands (yellow) bind the transmembrane BMP-receptors (light blue) and induce phosphorylation of SMAD1, SMAD5 and SMAD8 (green). These phospho-R-SMADs form a complex with SMAD4 (red) that translocates into the nucleus to regulate the expression of BMP target genes. TGFβ ligands (pink) bind the TGFB-receptors (dark blue) to phosphorylate SMAD2 and SMAD3 (green). A complex with SMAD4 is also formed that translocates into the nucleus to control the expression of TGFβ target genes. SMAD4 is the common nuclear mediator of BMP and TGFβ signal transduction.

In the developing limb bud, three BMP ligands are expressed: BMP2, BMP4 and BMP7. Cooperation between the TGFβ and BMP signalling pathways induces the chondrogenic program in forelimb buds after E10.5 (Karamboulas et al., 2010). However, outgrowth and patterning require BMP signalling, while TGFβ signalling is not required for early limb bud development (Bénazet et al., 2009; Pignatti et al., 2014). Therefore, we can investigate the roles of the BMP signalling during the onset and early phase of limb bud outgrowth (from E9.75 to E10.0) by studying the functions of the SMAD4 protein. As mentioned above, two phases of high BMP activity are required for normal limb bud development.

During the late phase (E11.0-E12.0), high BMP activity is required to induce chondrogenic differentiation and SMAD4 functions in the initial compaction and the onset of chondrogenic differentiation progenitor cells (Bénazet et al., 2012). During the onset of limb bud development (E9.5-E10.0), BMP4 and SMAD4 are required to initiate *Grem1* expression and for the formation of a functional AER (Benazet and Zeller, 2013; Bénazet et al., 2009, 2012). However, little is known about the range of transcriptional target genes controlled by BMP4/SMAD4 during limb bud initiation. Mouse embryos deficient for *Smad4* die before E7.5 due to gastrulation defects (Chu et al., 2004). As the *Prx1*-CRE transgene is active in limb bud mesenchyme from E9.5 onward (Logan et al., 2002), conditional inactivation of *Smad4* in the limb bud mesenchyme (Figure 4B) is used to investigate the roles of SMAD4 and the BMP signalling pathway during early limb bud development (Bénazet et al., 2012).

GREM1 is a crucial node in the SHH/GREM/FGF feedback signalling system. *Grem1* expression is activated by BMP signalling during the onset of limb bud development and its dynamic expression from E9.5 to E12.0 is regulated by SHH, Wnt and FGF signalling (Figure 4A, B, Bénazet et al., 2009; Capdevila et al., 1999; Scherz et al., 2004; Verheyden and Sun, 2008; Zuniga and Zeller, 1999 and unpublished results). The *Grem1* transcription unit consists of two exons with a coding exon 2. Its transcriptional regulation is controlled by a large *cis*-regulatory landscape extending 3' of *Grem1* into the neighbouring *Formin1* (*Fmn1*) gene. The *Fmn1-Grem1* genomic landscape contains highly conserved non-coding regions able to drive the expression of a *LacZ* reporter in the limb bud mesenchyme in a pattern similar to *Grem1* expression (Li et al., 2014; Zuniga et al., 2012). However, how these CRMs scattered throughout the *Grem1 cis*-regulatory landscape integrate the different signalling inputs into the dynamic spatio-temporal gene expression remains unclear (Figure 4C).

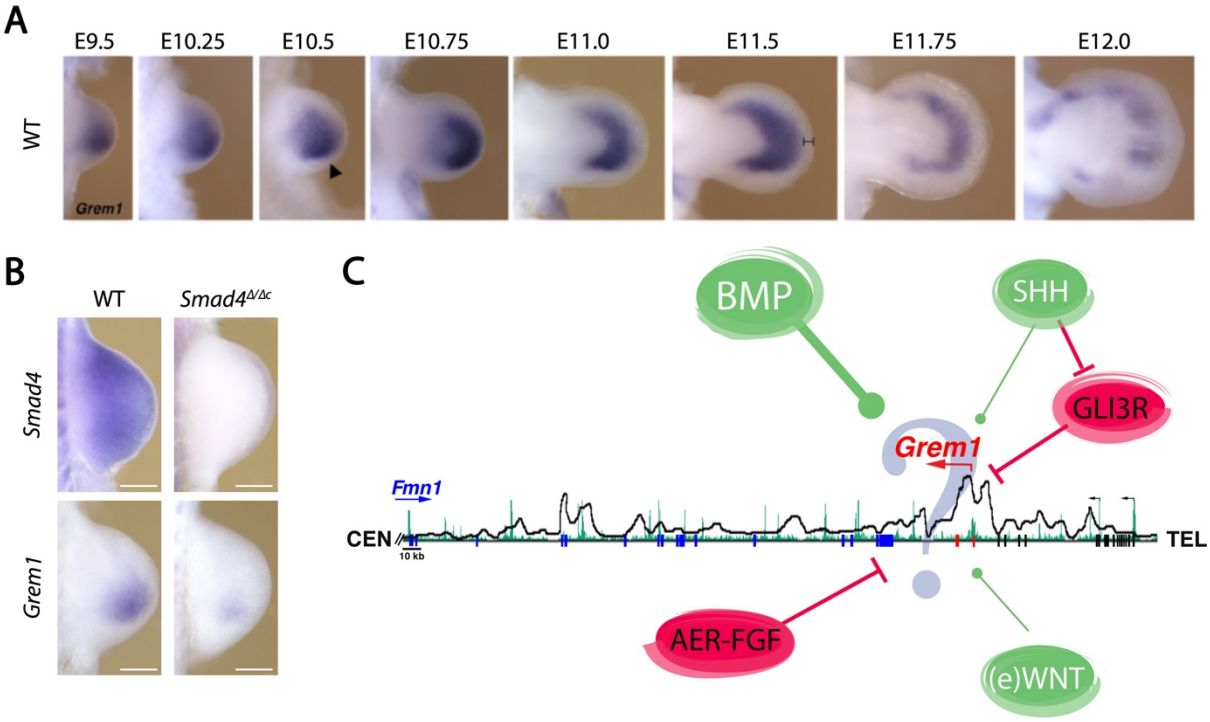


Figure 4: The dynamic *Grem1* expression is initiated by BMP/SMAD4 and regulated by integration of different signalling inputs
(A) WISH for *Grem1* in forelimbs from E9.5 to E12.0. Adapted from Zuniga et al., 2012. (B) WISH for *Smad4* and *Grem1* in WT and *Smad4*$^{\Delta/\Delta c}$ forelimbs at E10.0. Scale bars: 100μm. (C) Schematic representation of the *Grem1* genomic landscape that must integrate inputs from BMP, SHH, WNT and FGF signalling pathways. The *Grem1* transcription unit consists of 2 exons (in red) and is located upstream of Fmn1 transcription unit with 24 exons (in blue).

Little is known about the direct transcriptional targets of SMAD4 and potential additional functions of BMP signalling during the onset of limb bud outgrowth. A negative feedback between BMP-SHH has been described in chick limb buds (Bastida et al., 2009). In particular, BMP activity restricts *Shh* transcription in the ZPA by interfering with FGF and Wnt signalling. However, the mechanism of competition between the SHH and BMP signalling pathways is not understood in mouse limb buds.

4.4. SHH signalling pathway and cholesterol functions

During limb bud development, the SHH ligand is produced by the ZPA cells in the posterior mesenchyme (Figure 2, in red) and its graded distribution along the AP axis participates in establishing AP polarity in the early limb bud (Zhu and Mackem, 2017) and controls proliferation of LMPs (Zhu et al., 2008). Therefore, spatio-temporally controlled regulation of *Shh* expression and diffusion of the morphogenetic signal are essential for normal limb development. In particular, loss of *Shh* expression truncates distal limb skeletal structures while its anterior ectopic expression causes polydactylies (Chiang et al., 1996, 2001, Büscher et al., 1997). Therefore, one of the most common human congenial malformation is the preaxial polydactyly which is characterized by mutations in the ZRS affecting its activity and inducing ectopic anterior *Shh* expression (Sagai et al., 2004; Zeller et al., 2009). In contrast, absence of an active ZRS characterizes limbless reptiles and amphibians (Sagai et al., 2004). Recently, it has been shown that replacement of the mouse ZRS by the snake ZRS sequence is sufficient to “serpentize” the mouse, resulting in a limbless mouse (Kvon et al., 2016).

Autoproteolytic cleavage of the SHH protein releases the N-term signalling domain that is then covalently modified by palmitic acid (N-term) and cholesterol (C-term) moieties (Figure 5A). The palmilation and the cholesterylation modulate the long-range signalling of SHH (Chen et al., 2004; Li et al., 2006a). The cholesterol modification impacts on the SHH diffusion as it will be retained in the plasma membrane, which restrict its free mobility (Li et al., 2006a; Peters et al., 2004; Tukachinsky et al., 2012). The AP spread of SHH ligand is thus controlled to avoid too anterior expansion of the gradient (Li et al., 2006a; Zeng et al., 2001; Zhu and Scott, 2004). In addition to cholesterol modification of the SHH ligand, its receptor Patched1 (Ptch1) and HHIP1 (HH-interacting protein 1) are also involved for restricting the spread of the SHH ligand. Interestingly, these both SHH-interacting proteins are also direct transcriptional targets of

SHH, which points to the fact that cells responding to SHH up-regulate negative regulators of SHH signal transduction (Briscoe et al., 2001; Chen and Struhl, 1996; Chuang et al., 2003). Activation of signal transduction occurs as a consequence of SHH binding to its twelve-pass transmembrane receptor Ptch1, which in turn releases the inhibition of Smoothened (SMO) by Ptch1 at the primary cilia. As a consequence, SMO is cholesterol-modified at both its C- and N-terminal domains, which results in a change to its active conformation (Huang et al., 2016; Xiao et al., 2017). Intracellular signal transduction is triggered by modification of GLI3 full length (GLI-FL) to its GLI3-activator isoform (GLI-A), which translocates to the nucleus and activates/up-regulates the expression of SHH target genes (Figure 5B, Briscoe and Thérond, 2013). Cholesterol modification of the SHH ligand and SMO proteins is essential for graded SHH signalling as it regulates the diffusion of the ligand and activation of signal transduction in the limb bud mesenchymal cells responding to SHH signalling.

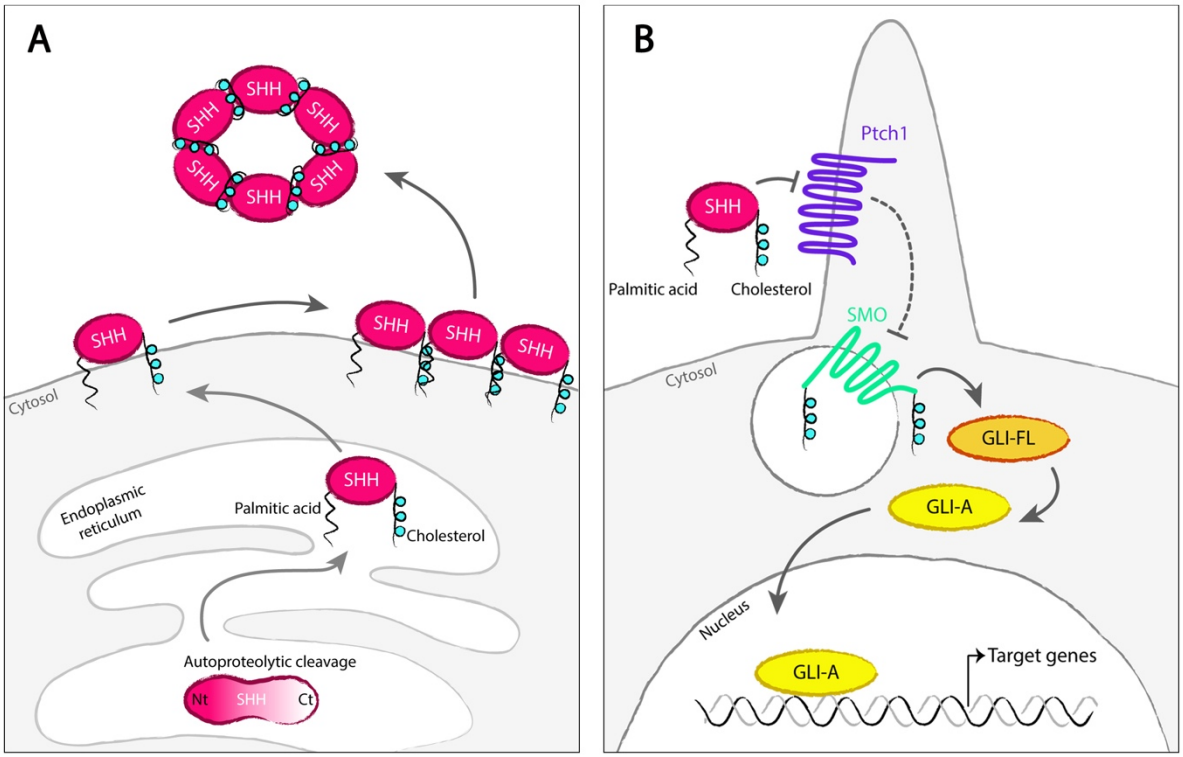


Figure 5: Cholesterol-modifications are crucial for SHH signalling
(A) SHH is processed by autoproteolytic cleavage in the endoplasmic reticulum. The amino-terminal SHH signalling peptide is covalently modified by palmitic acid at the N-terminus and by cholesterol at its C-terminus. This step is crucial for regulating the diffusion

of the SHH ligand. Secreted soluble SHH ligands form multimers, complexed by the cholesterol. (B) Activation of SHH signalling in responding cells: SHH binds its receptor Patched1 (Ptch1), which relieves the repression of Smoothened (SMO). Active SMO is covalently cholesterol modified at both the C- and N-terminal ends. GLI-FL (GLI3-full length) becomes GLI-A (Gli3-activator) and participates in regulating the expression of SHH target genes.

In addition to its roles in steroid hormones, neuroactive steroids, oxysterols and bile acids synthesis, the cholesterol homeostasis is critical for normal growth and embryonic development. Many human malformation syndromes are caused by an inborn error of cholesterol synthesis (Porter and Herman, 2011). During early embryonic development, maternal sterols are a major source of cholesterol until about E12.0, but the synthesized sterols by the embryo eventually become the primary source of cholesterol (Tint et al., 2006). The importance of maternal cholesterol for normal embryonic development was shown by treating pregnant rats with inhibitors of cholesterol biosynthesis. This gives birth defects which phenocopy the *Shh* loss-of-function phenotype (Porter et al., 1996, Roux et al., 1964). However, the importance of endogenous cholesterol synthesis by the embryo is revealed by the fact that many inborn errors of metabolism are associated with severe development defects (Porter and Herman, 2011). Cholesterol biosynthesis is complex and involves many enzymes and both positive and negative regulatory feedback loops (Figure 6). It includes the TFs sterol regulatory element-binding protein 1 or 2 (SREBF1/2), which directly enhances the expression of enzymes depending on intracellular cholesterol concentration (Ye and DeBose-Boyd, 2011). The Smith-Lemli-Optitz Syndrome (SLOS), lathosterolosis and desmosterolosis are the most common congenital metabolic diseases caused by genetic mutations affecting enzymes controlling the final steps of cholesterol biosynthesis: these mutations cause accumulation of specific precursor sterols.

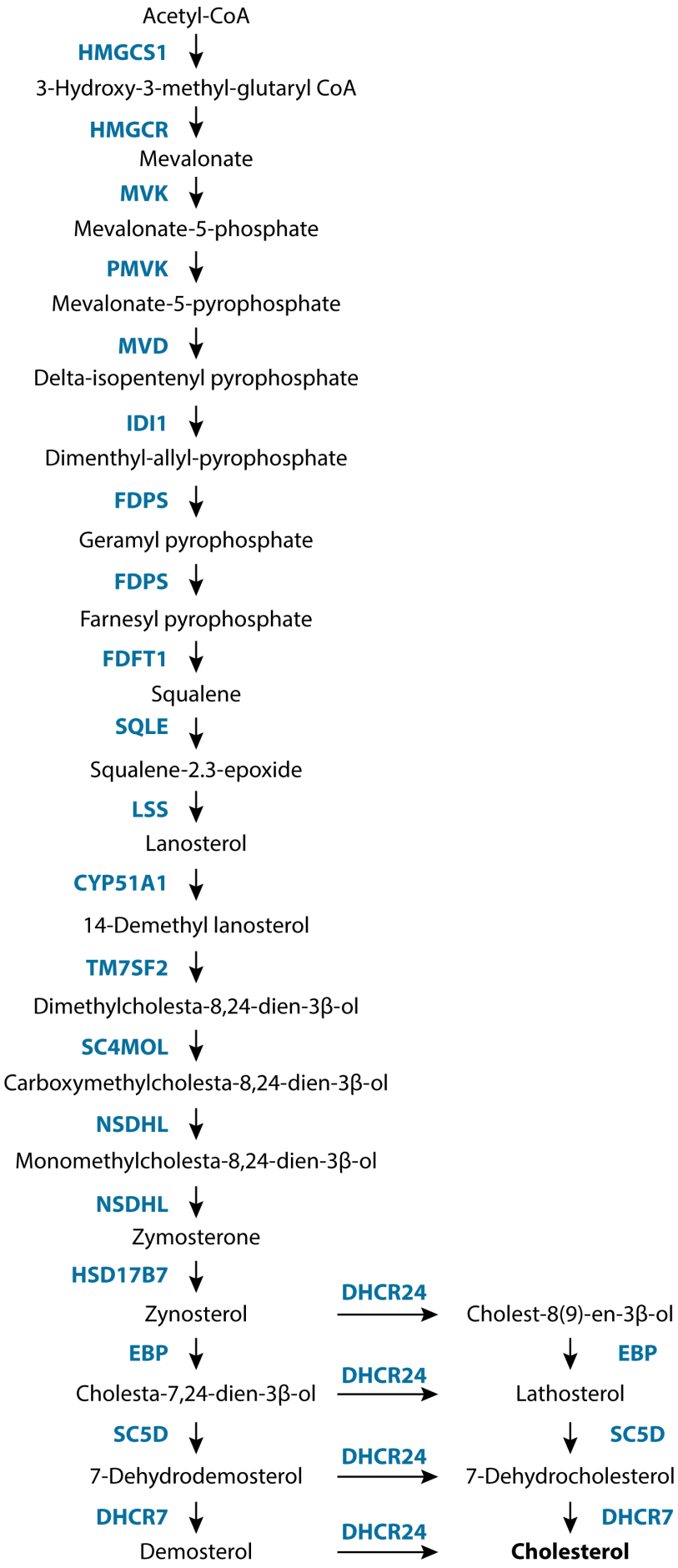

Acetyl-CoA
HMGCS1
3-Hydroxy-3-methyl-glutaryl CoA
HMGCR
Mevalonate
MVK
Mevalonate-5-phosphate
PMVK
Mevalonate-5-pyrophosphate
MVD
Delta-isopentenyl pyrophosphate
IDI1
Dimenthyl-allyl-pyrophosphate
FDPS
Geramyl pyrophosphate
FDPS
Farnesyl pyrophosphate
FDFT1
Squalene
SQLE
Squalene-2.3-epoxide
LSS
Lanosterol
CYP51A1
14-Demethyl lanosterol
TM7SF2
Dimethylcholesta-8,24-dien-3β-ol
SC4MOL
Carboxymethylcholesta-8,24-dien-3β-ol
NSDHL
Monomethylcholesta-8,24-dien-3β-ol
NSDHL
Zymosterone
HSD17B7
Zynosterol
DHCR24
Cholest-8(9)-en-3β-ol
EBP
EBP
Cholesta-7,24-dien-3β-ol
DHCR24
Lathosterol
SC5D
SC5D
7-Dehydrodemosterol
DHCR24
7-Dehydrocholesterol
DHCR7
DHCR7
Demosterol
DHCR24
Cholesterol

Figure 6: The cholesterol biosynthesis pathway
Cholesterol is synthetized from Acetyl-coA by 20 different enzymatic reactions whose each corresponding enzymes are indicated in blue.

In SLOS patients the DHCR7 enzyme is mutated, which results in accumulation of the 7-dehydro-cholesterol (7DHC, Porter, 2008). In lathosterolosis patients, the *Sc5d* gene is mutated and the lathosterol intermediate is accumulated. In desmosterolosis patients the *Dhcr24* gene mutated, which results in accumulation of desmosterol (Porter and Herman, 2011). The diagnosis involves detection of cholesterol intermediate forms by gas chromatography / mass spectrometry (GC/MS). These metabolic diseases have a large spectrum of phenotypes and are characterized by altered facial features, mental retardation and limb defects (Porter and Herman, 2011). Many of the associated malformations affect tissues and organs that critically depend on SHH signalling during embryonic development (Cooper et al., 2003). In addition, mouse models for SLOS and lathosterolosis show a defective response to SHH signalling due to impaired SMO activity (Blassberg et al., 2016; Cooper et al., 2003). This corroborates the importance of the cholesterol modification for SHH signalling and signal transduction during embryonic development.

As mentioned before mouse embryonic development depends initially (up to about E12.0) on both embryonic cholesterol biosynthesis and uptake of maternal cholesterol absorption via the placenta (Tint et al., 2006). As the uptake of maternal cholesterol decreases in parallel to the increase in endogenous cholesterol synthesis, both processes are tightly balanced to ensure normal cholesterol levels in the embryo (Cohen, 2008; Tint et al., 2006). As cholesterol is water-insoluble, its transport requires the association of low lipoproteins (LDL) with apolipoprotein B (ApoB, Figure 7). In particular, the cholesterol uptake by cells depends on the amount of LDL-receptors (LDL-R) present at cell membranes. These receptors internalize the LDL complexes by endocytosis and are then recycled to the cell membrane. The LDL endosome

fuses with a lysosome, which results in degradation of the LDL moiety and releases the cholesterol (Figure 7, left part). The cholesterol uptake is controlled by the proprotein convertase subtilisin/kexin type-9 (PCSK9), which binds the LDL-R and targets the receptor for lysosomal degradation (Figure 7, right part, Lagace, 2014; Lagace et al., 2006).

These studies reveal the complexity of cholesterol metabolism in embryos, but it is still unknown how this complex regulation is linked to and modulated by the molecular interactions and gene regulatory networks that control embryonic development.

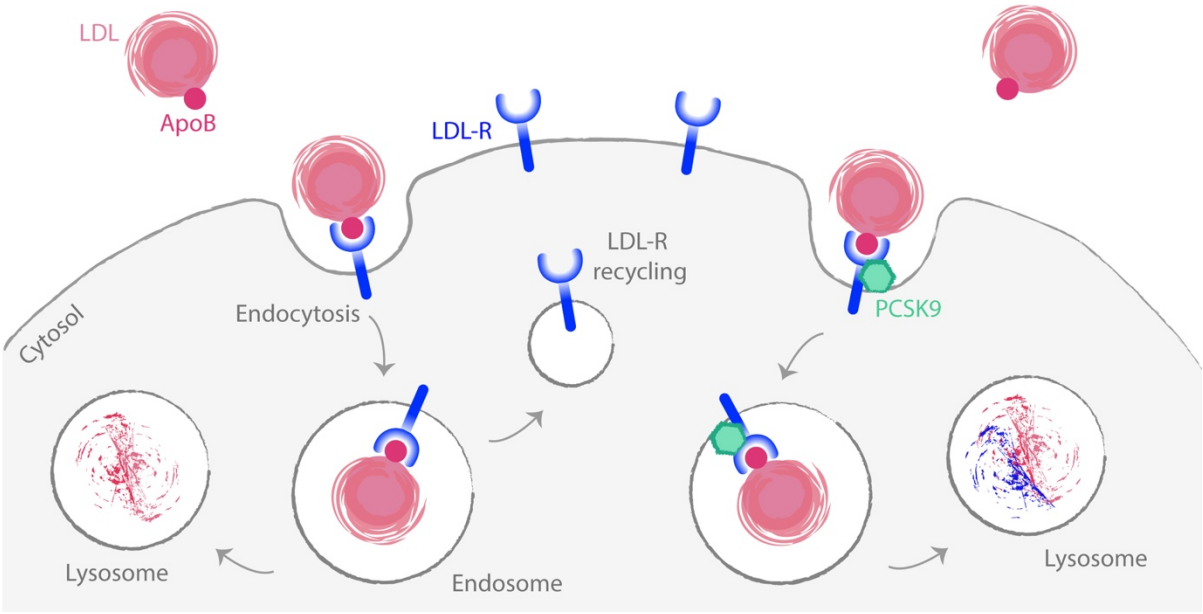


Figure 7: Cholesterol uptake and regulation by cells
The extracellular cholesterol is transported by the low density lipoproteins (LDL, pink) in a complex with the ApoB protein (dark pink). The cell membrane contains LDL-receptors (LDL-R, blue), which recognize and bind the ApoB-LDL complex. The complex is internalized by endocytosis and the LDL-R is then recycled to the cell membrane. The LDL-endosome fuses with a lysosome that degrades the LDL and releases the cholesterol into the cell. Degradation of the LDL receptor: the PCSK9 convertase (in green) binds to the LDL-R and blocks its recycling. Thereby LDL/LDL-R complex is degraded by fusion with a lysosome, which decreases the LDL-R concentration in the cell membrane.

5. Aims of the Thesis

Understanding the mechanisms that drive cell fates during the embryogenesis is a fascinating question. During my PhD in the laboratory of Prof. Dr. Rolf Zeller and Dr. Aimée Zuniga, I could continue to deepen my knowledge about developmental genetics acquired during my Master by focussing on heart and limb development.

First, I had the opportunity to participate in the study of transcriptional gene regulation during heart development. In particular, I was involved in the identification of direct transcriptional targets of HAND2, which is an essential transcription factor for heart development. We have discovered the HAND2-dependent gene regulatory network controling the endothelial to mesenchymal transition during atrioventricular canal development (the publication describing these results is presented in the first part of my thesis in its published version).

In the second part of my thesis, I analysed the roles of the BMP signalling pathway during the early limb bud development. I have focussed my study on the activity of the BMP signalling transducer SMAD4. I had the chance to perform and combine different types of genome-wide analyses to identify the gene regulatory networks controlled by SMAD4. As high BMP activity is essential for the initiation and outgrowth of the limb bud, I have identified and studied the multiple functions of SMAD4 during this decisive early phase of limb bud development. First, I studied the role of SMAD4 in the activation of the BMP antagonist *Gremlin1*, which is essential for establishment of the signalling feedback loop that drives limb bud outgrowth. Then, my research identified the SMAD4-dependent gene regulatory network that is essential for patterning the early limb bud. Finally, I have identified a completely novel and highly relevant mechanism whereby SMAD4 controls cholesterol homeostasis and indirectly modulates the spread of the SHH signal and its signal transduction activity.

6. Results

6.1. Identification of the direct targets of HAND2 during heart development

"HAND2 Target Gene Regulatory Networks Control Atrioventricular Canal and Cardiac Valve Development"

Laurent, F., Girdziusaite, A., **Gamart, J**., Barozzi, I., Osterwalder, M., Akiyama, J.A., Lincoln, J., Lopez-Rios, J., Visel, A., Zuniga, A., et al. (2017). Cell Rep. *19*, 1602–1613.

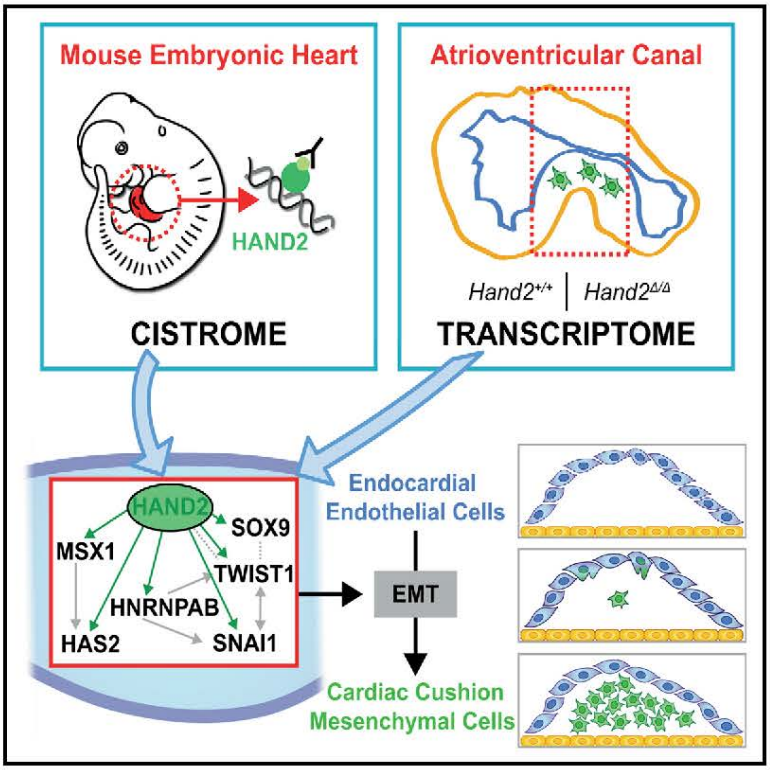
This article by Laurent et al. focuses on the functional requirement of HAND2 in endothelial to mesenchymal transition (EMT) and subsequent migration of the delaminating cells into the cardiac cushions of the atrioventricular canal (AVC). To this publication, I have contributed to various experiments such as blinded determination of cell numbers for the heart explants shown in Figures 3, 6b and S5c. Together with Ausra Girdziusaite and Rolf Zeller I have participated in the isolation of AVCs for RNA-seq analysis (Figure 4). I have performed the ChIP-qPCR analysis using *Hand2*$^{3xF/+}$ hearts to verify the HAND2 transcriptional target genes that are involved in EMT processes and AVC morphogenesis (Figure 4). In addition, I have performed the bioinformatics analysis using the HAND2 ChIP-seq datasets, to identify the HAND2-interacting regions located in the topologically associated domains (TADs) of HAND2 direct target genes of interest (data not shown). Finally, I have prepared the graphical abstract together with Frédéric Laurent.

Article

Cell Reports

HAND2 Target Gene Regulatory Networks Control Atrioventricular Canal and Cardiac Valve Development

Graphical Abstract



Authors

Frédéric Laurent, Ausra Girdziusaite, Julie Gamart, ..., Axel Visel, Aimée Zuniga, Rolf Zeller

Correspondence

rolf.zeller@unibas.ch

In Brief

Laurent et al. combine ChIP-seq with transcriptome analysis to identify the HAND2 target gene network that controls the EMT and mesenchymal cell migration during cardiac cushion formation in the atrioventricular canal (AVC). The HAND2 transcriptional targets include *Snai1*, whose re-expression in *Hand2*-deficient AVC explants partially restores mesenchymal cell migration.

Highlights

- HAND2 controls development of the AVC cardiac cushions forming mitral/tricuspid valves
- HAND2 is a key regulator of the EMT underlying cardiac cushion mesenchyme formation
- Identification of the HAND2 target gene networks that control EMT and AVC development
- HAND2 acts upstream of the EMT key regulator *Snai1* in AVC and other embryonic tissues

Accession Numbers

GSE73368
GSE94246

Laurent et al., 2017, Cell Reports *19*, 1602–1613
May 23, 2017
http://dx.doi.org/10.1016/j.celrep.2017.05.004

OPEN ACCESS
CellPress

Cell Reports
Article

HAND2 Target Gene Regulatory Networks Control Atrioventricular Canal and Cardiac Valve Development

Frédéric Laurent,[1,7] Ausra Girdziusaite,[1,7] Julie Gamart,[1] Iros Barozzi,[2] Marco Osterwalder,[1,2] Jennifer A. Akiyama,[2] Joy Lincoln,[3] Javier Lopez-Rios,[4] Axel Visel,[2,5,6] Aimée Zuniga,[1,8] and Rolf Zeller[1,8,9,*]

[1]Developmental Genetics, Department of Biomedicine, University of Basel, 4058 Basel, Switzerland
[2]Functional Genomics Department, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[3]Center for Cardiovascular and Pulmonary Research, The Research Institute at Nationwide Children's Hospital, Columbus, OH 43205, USA
[4]Development and Evolution, Department of Biomedicine, University of Basel, 4058 Basel, Switzerland
[5]U.S. Department of Energy Joint Genome Institute, Walnut Creek, CA 94598, USA
[6]School of Natural Sciences, University of California, Merced, CA 95343, USA
[7]These authors contributed equally
[8]Senior author
[9]Lead Contact
*Correspondence: rolf.zeller@unibas.ch
http://dx.doi.org/10.1016/j.celrep.2017.05.004

SUMMARY

The HAND2 transcriptional regulator controls cardiac development, and we uncover additional essential functions in the endothelial to mesenchymal transition (EMT) underlying cardiac cushion development in the atrioventricular canal (AVC). In *Hand2*-deficient mouse embryos, the EMT underlying AVC cardiac cushion formation is disrupted, and we combined ChIP-seq of embryonic hearts with transcriptome analysis of wild-type and mutants AVCs to identify the functionally relevant HAND2 target genes. The HAND2 target gene regulatory network (GRN) includes most genes with known functions in EMT processes and AVC cardiac cushion formation. One of these is *Snai1*, an EMT master regulator whose expression is lost from *Hand2*-deficient AVCs. Re-expression of *Snai1* in mutant AVC explants partially restores this EMT and mesenchymal cell migration. Furthermore, the HAND2-interacting enhancers in the *Snai1* genomic landscape are active in embryonic hearts and other *Snai1*-expressing tissues. These results show that HAND2 directly regulates the molecular cascades initiating AVC cardiac valve development.

INTRODUCTION

Perturbations affecting cardiac progenitors result in embryonic lethality and severe congenital heart defects, which are a major cause of infant and even adult mortality (Bruneau, 2008). In particular, different types of congenital heart defects are caused by alterations in the progenitors of the second heart field (SHF; reviewed by Kelly, 2012). SHF progenitors migrate into the developing heart tube, where they contribute to most developing structures including the inflow pole, both atria and ventricles, and the outflow tract (OFT). The four heart chambers are formed by rapid proliferative expansion, while the cardiomyocytes in the OFT and atrioventricular canal (AVC) proliferate less and remain undifferentiated (Christoffels et al., 2010; Greulich et al., 2011). The AVC connects the left ventricle to the forming atria, is required for chamber septation, and gives rise to the atrioventricular node and mitral and tricuspid valves (Christoffels et al., 2010; Lin et al., 2012). In mouse embryos, development of the AVC valves begins at embryonic day E9.5, when endocardial cells undergo an endothelial to mesenchymal transition (EndMT or EMT) in response to signals from the myocardium. The delaminating endocardial cells migrate into the cardiac jelly and give rise to the cardiac cushion mesenchyme, which are then remodeled into the mature valve structures (MacGrogan et al., 2014). The EMT in the AVC is controlled by BMP2 signaling from the myocardium, which synergizes with myocardial TGFβ2 and endocardial NOTCH signaling to activate downstream effectors that include the *Snai1* transcriptional regulator (Luna-Zurita et al., 2010; Ma et al., 2005; Niessen et al., 2008; Timmerman et al., 2004). SNAI1 is a key EMT regulator in embryos and various diseases such as tumor metastasis (reviewed by Nieto, 2011). Its inactivation in the endothelial compartment disrupts the EMT underlying AVC cardiac cushion formation (Wu et al., 2014).

Another transcription factor essential for heart development is HAND2, which also functions in developing branchial arches and limb buds (Srivastava et al., 1997). In the developing heart, HAND2 is expressed in the myocardial compartment of the right ventricle and OFT, the epicardium, and valve progenitors in both OFT and AVC (VanDusen and Firulli, 2012; VanDusen et al., 2014b). Consistent with its complex expression pattern, genetic inactivation of *Hand2* in mice disrupts development of limb buds, branchial arches, aortic arch arteries, and the right ventricle, which causes embryonic lethality (Srivastava et al., 1997).

Specific inactivation in developing heart tissues has revealed essential *Hand2* functions in the cardiac neural crest cells that contribute to cardiac cushions in the OFT, survival of SHF progenitors, heart chamber trabeculation, and epicardial cell differentiation (Barnes et al., 2011; Holler et al., 2010; Tsuchihashi et al., 2011; VanDusen et al., 2014a). Previous studies had also pointed to HAND2 functions in cardiac cushion formation, but the potential essential functions have not been identified (Holler et al., 2010; Liu et al., 2009; VanDusen et al., 2014a). In humans, mutations in *HAND2* have been linked to congenital heart malformations that include ventricular septal defects (Shen et al., 2010; Sun et al., 2016).

We have used chromatin immunoprecipitation sequencing (ChIP-seq) to define the genome-wide interaction profile of endogenous HAND2 chromatin complexes in mouse embryonic hearts. This analysis shows that HAND2 interacts with non-coding regions associated with a large number of genes functioning during heart development. Most importantly, our analysis revealed essential HAND2 functions in the EMT underlying AVC cardiac cushion formation, which is disrupted in *Hand2*-deficient embryos. Combining transcriptome analysis of wild-type and *Hand2*-deficient AVCs with the HAND2 ChIP-seq dataset identified the HAND2 target genes that function in AVC cardiac cushion development. In particular, this analysis revealed that the EMT regulator *Snai1* is a transcriptional target of HAND2. The failure of endocardial cells to invade the cardiac jelly in *Hand2*-deficient AVCs is partially restored by re-expressing SNAI1 in mutant explants. Last but not least, we show that the two *Snai1*-associated enhancers interacting with HAND2 chromatin complexes recapitulate major aspects of *Snai1* expression in mouse embryos.

RESULTS

Genomic Regions Enriched in Endogenous HAND2 Chromatin Complexes Identify the Range of HAND2 Target Genes in Mouse Embryonic Hearts

The *Hand2*^{3xF} allele, which encodes a HAND2 protein with a 3xFLAG epitope tag inserted in its N-terminal part (Osterwalder et al., 2014), was used to profile the genomic regions enriched in HAND2-chromatin complexes. Anti-FLAG antibodies were used for chromatin immunoprecipitation, which was followed by massive parallel sequencing (ChIP-seq). To obtain sufficient chromatin for ChIP-seq, ~300 hearts per biological replicate were dissected from *Hand2*$^{3xF/3xF}$ mouse embryos at embryonic days E10.25–10.5. Two biological replicates were analyzed, and the genome-wide binding profiles using model-based analysis of ChIP-seq (MACS) (Zhang et al., 2008) identified 12,117 significantly enriched genomic regions. Genomic Regions Enrichment of Annotations Tool (GREAT) analysis (McLean et al., 2010) was used to assign these 12,117 regions to 7,792 neighboring genes, which defines the initial set of putative HAND2 targets (Table S1; Supplemental Experimental Procedures). Most of the genomic regions enriched in HAND2 chromatin complexes are located ≥10 kb away from transcriptional start sites (TSS) and encode evolutionary conserved sequences that overlap the peak summit (Figures 1A and 1B). Functional enrichment was assessed by GREAT using increasingly larger set of peaks (pool of incremental deciles; for details, see Supplemental Experimental Procedures). This analysis revealed that terms related to abnormal cardiac morphology and heart development were already enriched when using only the most enriched regions, while terms referring to specific processes such as OFT, right ventricle, and AVC development reached significance using the larger dataset (Figure S1A). In particular, 15 of the 16 most enriched Gene Ontology (GO) terms are relevant to heart development, while the remaining term identifies genes functioning in EMT processes (Figure 1C, see below). De novo motif discovery using HOMER (Heinz et al., 2010) identified the consensus *Ebox* motif (CATCTG; Dai and Cserjesi, 2002) as the most prevalent among HAND2 peaks (Figure 1D). Other significantly enriched motifs include binding sites for GATA transcription factors (Figure S1B), which are key regulators of heart development (Stefanovic and Christoffels, 2015). Indeed, computational comparison of the HAND2 (Table S1) with a published GATA4 ChIP-seq dataset (using whole mouse embryonic hearts at E12.5; He et al., 2014) shows that 28.3% of the enriched genomic regions are shared between the two datasets, which is 15-fold higher than expected by chance (data not shown). As development of *Hand2*-deficient mouse embryos results in lethality by ~E10.5 (Srivastava et al., 1997), we limited our analysis to mutant embryos at E9.0–9.5. During this early organogenic stage, no aberrant apoptosis occurs in the developing heart in contrast to branchial arches and frontonasal mass (Figure S1C). Therefore, all interactions of HAND2 chromatin complexes with candidate *cis*-regulatory modules (CRMs) associated with genes of interest were also verified by ChIP-qPCR using embryonic hearts at E9.25–9.5 (Figures 1, 4, S3, and S6).

To validate the ChIP-seq dataset as a resource for identifying HAND2 target genes, we determined which of the genes with known alterations in their expression are associated with HAND2 ChIP-seq peaks (Table S2 and references therein). This analysis showed that about half of these genes are associated with at least one HAND2 ChIP-seq peak (n = 56/114; Table S2, see also Table S1). This suggests that their altered expression in *Hand2*-deficient hearts could be a consequence of direct transcription regulation by HAND2. Next, we overlapped the HAND2 binding profiles with enhancers active in mouse embryonic hearts that were identified by a large enhancer screen (VISTA Enhancer Browser: https://enhancer.lbl.gov, Visel et al., 2007). This analysis showed that 71 of the 193 VISTA enhancers active in mouse embryonic hearts overlap with HAND2 ChIP-seq peaks (Figures 1E, 1F, and S2; Table S3). These enhancers include CRMs in the genomic landscapes of the *Gata4*, *Gata6*, *Myocd*, and *Tbx20* transcriptional regulators, which are essential for OFT and/or right ventricle development. Whole-mount in situ hybridization (WISH) showed that the expression of *Gata4*, *Myocd*, and *Gata6* is reduced in *Hand2*-deficient hearts, while the expression of *Tbx20* appears unchanged (E9.25–9.5, Figure 1G). The requirement of *Hand2* for OFT and right ventricle development is also supported by GO analysis as more than half of the genes with annotated functions in OFT and right ventricle (RV) development are associated with HAND2 ChIP-seq peaks (Figure S3A; Table S4). In particular, several ligands of the signaling pathways required for development of these structures are identified as HAND2 target genes (Figures S3B

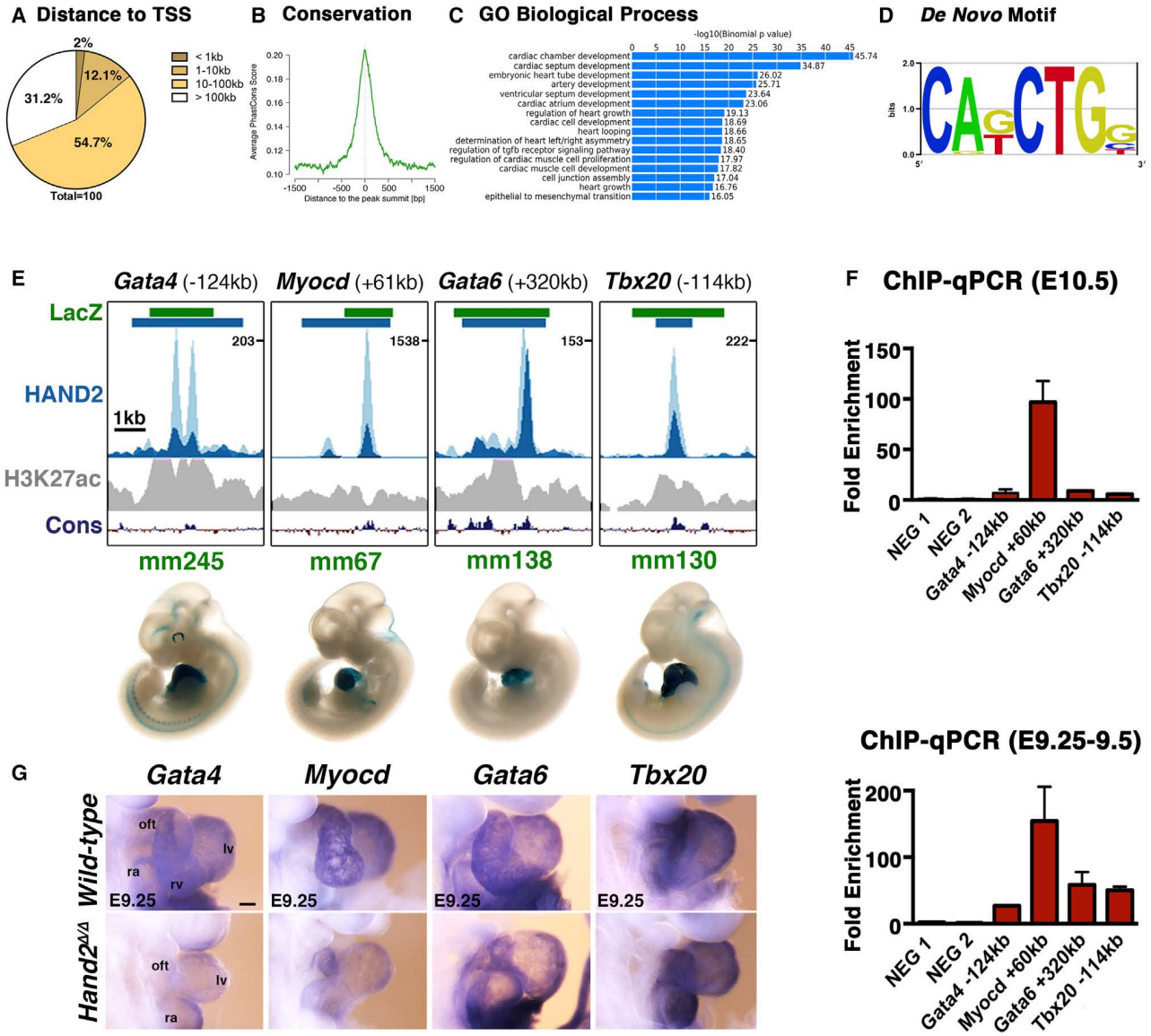


Figure 1. ChIP-Seq Analysis Using the *Hand2*^{3xF} Allele Identifies the HAND2 Cistrome in Mouse Embryonic Hearts

(A) The majority of the genomic regions enriched in HAND2 chromatin complexes from mouse embryonic hearts at E10.5 map ≥10 kb away from the closest transcription start site (TSS).

(B) In addition, the majority of the HAND2-interacting genomic regions are evolutionarily conserved.

(C) The top GO terms associated with HAND2 candidate targets reveal the preferential enrichment of genes functioning in cardiac development.

(D) The consensus *Ebox* motif is most enriched by de novo motif discovery.

(E) Selection of VISTA enhancers enriched in HAND2 chromatin complexes. Green intervals indicate the regions with enhancer activity (VISTA enhancer database); blue intervals highlight the regions enriched in HAND2 chromatin complexes (MACS peaks). Distances to the nearest TSS within the TAD are indicated on top. ChIP-seq profiles of the two biological replicates (E10.5) are shown in light and dark blue, respectively. The H3K27ac ChIP-seq profile for mouse hearts (E11.5) is shown in gray (Nord et al., 2013). The scheme at the bottom shows the placental mammal conservation (Cons) plot (PhyloP). Representative transgenic *LacZ* reporter embryos for VISTA enhancers associated with genes functioning in OFT and/or right ventricle development (*Gata4*, *Myocd*, *Gata6*, and *Tbx20*) are shown below.

(F) ChIP-qPCR validation of the ChIP-seq peaks (E) for mouse embryonic hearts at E10.5 (n = 3 biological replicates) and E9.25 (n = 2; mean ± SD).

(G) Expression of the HAND2 targets *Gata4*, *Myocd*, *Gata6*, and *Tbx20* in wild-type and *Hand2*-deficient (*Hand2*$^{\Delta/\Delta}$) embryonic hearts (E9.25).

Scale bars, 100 μm. oft, outflow tract; ra, right atrium; rv, right ventricle; lv, left ventricle; avc, atrioventricular canal; la, left atrium. See also Figures S1–S3 and Tables S1 and S2–S4.

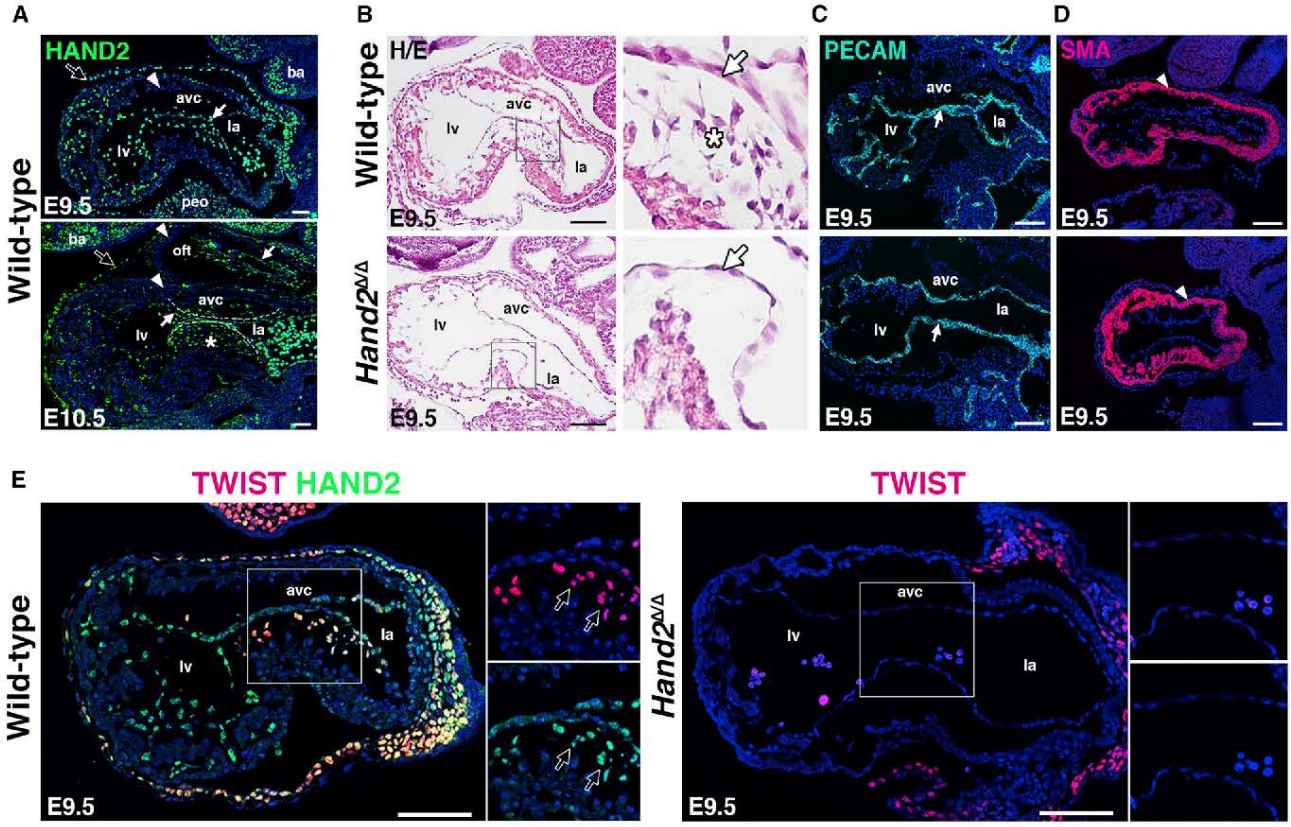


Figure 2. AVC Cardiac Cushion Agenesis in *Hand2*-Deficient Mouse Embryos

(A) Distribution of the endogenous HAND2^{3xF} protein (using anti-FLAG antibodies, green fluorescence) during in mouse embryonic hearts at E9.5 and E10.5. Representative sagittal sections are shown. White arrows, endocardium; white arrowheads, myocardium; white asterisks, HAND2 expressing AVC cardiac cushion mesenchymal cells; black arrows, epicardium. Scale bars, 50 μm.

(B) H&E staining reveals the absence of delaminating endocardial cells with mesenchymal characteristics (white asterisks) in the AVC of *Hand2*-deficient mouse embryos. White arrow, endocardium. Scale bars, 100 μm.

(C) Detection of the platelet endothelial cell adhesion molecule (PECAM) in the endocardium of wild-type and *Hand2*-deficient embryos (white arrow).

(D) Detection of smooth muscle actin (SMA) in the AVC myocardium of wild-type and *Hand2*-deficient embryos (white arrowhead). Myocardium (A–D).

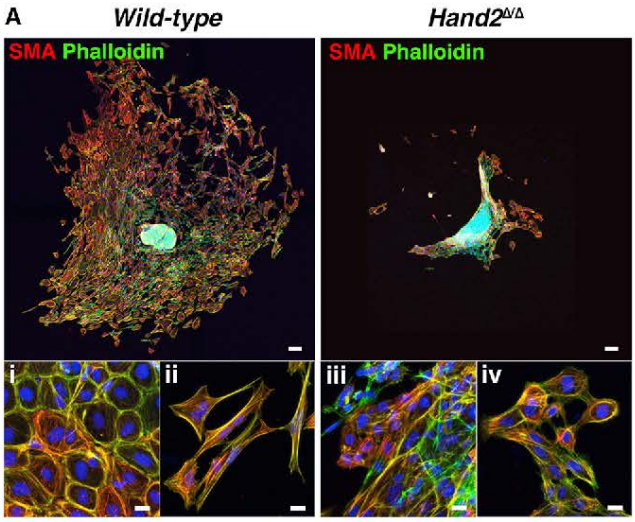
(E) Colocalization of the HAND2^{3xF} (green fluorescence) with TWIST1 transcriptional regulators (red fluorescence) in the AVC of wild-type (*Hand2*$^{3xF/3xF}$) and *Hand2*-deficient (*Hand2*$^{\Delta/\Delta}$) embryos at E9.5. Colocalization is detected in delaminating mesenchymal cells (indicated by arrows, left panel), which are missing from the mutant AVC (right panel).

Scale bars in panels (E) AND (F), 100 μm. avc, atrioventricular canal; ba, branchial arches; la, left atrium; lv, left ventricle; oft, outflow tract; peo, proepicardial organ.

and S3C). This includes *Wnt11*, *Wnt5a*, *Bmp4*, *Tgfβ2*, and *Fgf10*, whose expression is either reduced or lost from the OFT and/or RV of *Hand2*-deficient mouse embryos (E8.75–9.25, Figure S3C). Collectively, this first analysis shows that a significant fraction of genes functioning in mouse OFT and RV morphogenesis are likely HAND2 target genes (Figures 1G and S3). However, not all candidate target genes analyzed are altered in *Hand2*-deficient hearts, which points either to *cis*-regulatory redundancy or raises the possibility that the interaction of HAND2 complexes with the candidate CRMs is not essential for the adjacent genes.

HAND2 Is a Key Regulator of the EMT during AVC Cardiac Cushion Formation

Strikingly, the GO analysis identified EMT as one of the key biological processes associated with the HAND2 cistrome (Figure 1C). An essential process during heart valve development is formation of the mesenchymal compartment of the AVC cardiac cushions as endocardial cells undergo an EMT (see Introduction). In wild-type hearts, HAND2 proteins are expressed by endocardial cells in the AVC, the cells undergoing EMT and the delaminating cells forming the cushion mesenchyme continue to express HAND2 (asterisk, Figure 2A and data not shown). Furthermore, histological analysis reveals the complete absence of mesenchymal cells in the AVC cardiac cushions of *Hand2*-deficient hearts at E9.0–9.5, which points to disruption of the EMT process (Figure 2B). As the distribution of platelet endothelial cell adhesion molecule (PECAM) and smooth muscle actin (SMA) positive cells is not altered in *Hand2*-deficient hearts, the endocardial and myocardial compartments of the AVC appear to have formed normally, which underscores the specific nature of the observed cellular defect



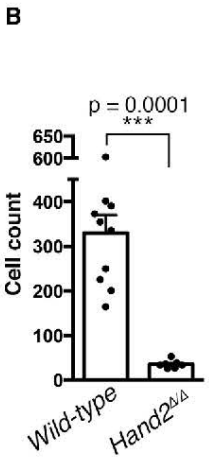


Figure 3. *Hand2*-Deficient AVC Endocardial Cells Fail to Initiate EMT and Mesenchymal Cell Migration

(A) Smooth muscle actin (red) and F-actin (green Phalloidin staining) distribution in cells that have migrated into the matrix from AVC explants of wild-type and *Hand2*-deficient embryos after 72 hr in culture. Scale bars, 100 μm. Bottom panels show cells in proximity to the explant (i and iii) and at the far edge of migration (ii and iv). Scale bars, 20 μm. (B) Quantification of the numbers of cells that migrated into the matrix from wild-type (n = 10) and *Hand2*-deficient AVC explants (n = 7). The mean ± SD (p = 0.0001, Mann-Whitney test) and all individual data points are shown.

(Figures 2C and 2D). TWIST1 forms hetero-dimeric transcriptional complexes with HAND2 and regulates cardiac cushion development (Firulli et al., 2005; VanDusen and Firulli, 2012). Therefore, the distribution of both proteins was comparatively analyzed in developing AVC cardiac cushions (Figure 2E). In wild-type embryos, HAND2 and TWIST1 are co-expressed by the delaminating mesenchymal cells that form the AVC cardiac cushions in wild-type hearts (left panels, Figure 2E). In contrast, these TWIST1-positive cells are absent in *Hand2*-deficient hearts (right panel, Figure 2E). Together, these results point to complete disruption of AVC cardiac cushion formation in *Hand2*-deficient mouse embryos.

To study this processes further, AVCs were dissected from wild-type and mutant hearts at E9.5 and cultured on collagen matrices for 72 hr (Figure 3A, Camenisch et al., 2000). Then, the mesenchymal cells that had migrated from the explant into the matrix were quantitated (Figure 3B): on average 330 ± 40 cells colonize the matrix in wild-type AVC explants (n = 10), while ~10-fold fewer mesenchymal cells (36 ± 3) are detected in AVC explants isolated from *Hand2*-deficient embryos (n = 7). In particular, wild-type endocardial cells in proximity of the AVC explant retain their cobblestone-like morphology and a cortical actin ring (Figure 3A, i), while cells that migrated further develop actin stress fibers and long filopodia characteristic of mesenchymal cells (Figure 3A, ii). In contrast, the few cells invading the matrix in cultures of *Hand2*-deficient AVCs mostly retain their cobblestone-like morphology (Figure 3A, iii and iv). This loss of mesenchymal characteristics shows that the endocardial cells of *Hand2*-deficient AVCs fail to undergo the EMT giving rise to the mesenchymal cell forming the cardiac cushions.

HAND2 Controls the Expression of Genes that Function in the EMT Underlying AVC Cardiac Cushion Formation

To identify the gene regulatory networks (GRNs) controlled by HAND2 during AVC cardiac cushion formation, the transcriptomes of dissected wild-type and *Hand2*-deficient AVCs were analyzed (E9.0–9.25: 18–23 somites; see Supplemental Experimental Procedures). Statistical analysis showed that 1,051 genes are differentially expressed (DEGs: 695 are upregulated and 365 downregulated, Figure 4A; Table S5; Figure S4A for GO analysis). Among these genes, the transcriptional targets of HAND2 were identified as those genes harboring one or more HAND2 ChIP-seq peaks in their topologically associating domains (TADs, Dixon et al., 2012; Figures 4B, 4C, and S4B; Table S5). This analysis shows that the TADS of DEGs contain on average a significantly larger number of HAND2-interacting regions (median ~5) than genes whose expression is not altered (median ~1; Figure 4B). GO analysis of these HAND2 transcriptional targets indicates that the 167 DEGs, whose expression is downregulated in *Hand2*-deficient AVCs function preferentially in heart and organ development (including cardiac EMT and mesenchyme development), while the 372 upregulated DEGs function preferentially in cardiovascular and blood vessel development (Figure 4C). In particular, this functional annotation identified a subset of 24 DEGs that function in EMT processes and/or AVC cushion formation (Figure S4C). Combining the transcriptome analysis with HAND2 ChIP-qPCR analysis establishes 19 of these DEGs as direct transcriptional targets of HAND2 in developing hearts at E9.25–9.5 (Figures 4D and 4E). Transcriptome analysis showed that the expression of seven of these HAND2 targets is downregulated, while 12 are upregulated in mutant AVCs (Figure 4D; Table S6). These differential effects are not unexpected as HAND2 transcriptional complexes are known to differentially activate or repress gene expression (see Discussion). Together, this analysis uncovers the HAND2 target GRN functioning in AVC cardiac cushion formation and reveals the differential effects of the *Hand2* deficiency on gene expression in the mutant AVC (Figures 4D and 4E).

Next, we used WISH to detect spatial alterations in the AVC of *Hand2*-deficient hearts at E9.25–9.5 (Figures 5 and S5). To uncover potential global molecular changes in the mutant AVC, we first analyzed the spatial distribution of key regulators whose transcript levels are not changed (*Bmp2*, *Hey2*, *Notch1*, *Rbpj*, *Snai2*, *Epha3*, *Tbx20*; Figures 5A and 5B; Table S5 and data not shown). The spatial distribution of all of these genes is comparable to wild-types in mutant AVCs, as exemplified by the EMT-inducer *Bmp2* and the NOTCH transcriptional target *Hey2* (Figures 5A and 5B). This indicates that the AVC domain

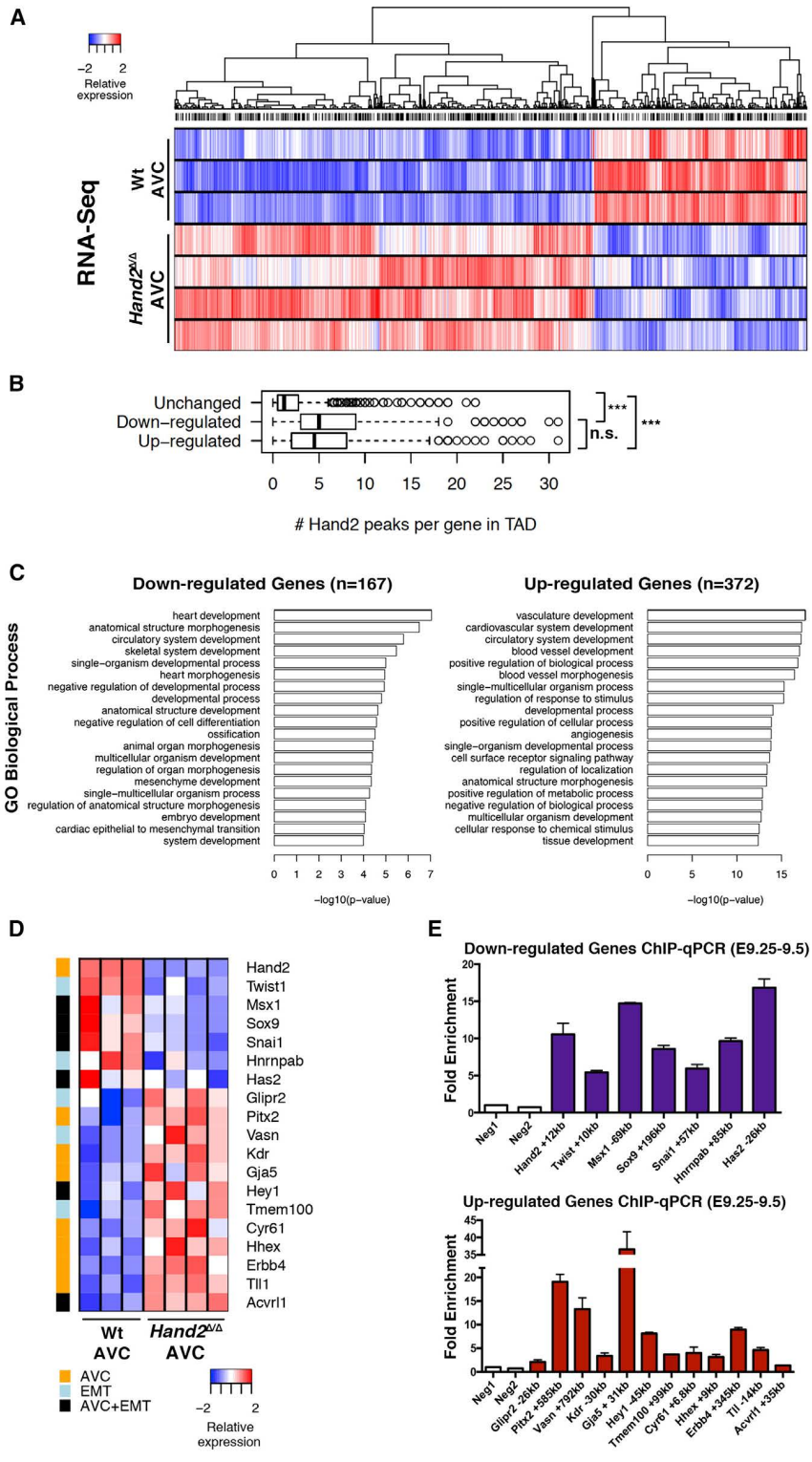


Figure 4. Transcriptome Analysis Identifies the Transcriptional Targets of HAND2 in the AVC

(A) Heatmap of the differentially expressed genes (DEGs) identified by comparing the transcriptomes of wild-type and *Hand2*-deficient AVCs from mouse embryonic hearts at E9.25–9.5 (n = 4 and n = 3 biological replicates were analyzed for mutant and wild-type AVCs, respectively). DEGs are genes whose expression is significantly changed (≥1.5-fold) between wild-type and mutant samples ($p < 0.05$).

(B) The boxplot shows the number of HAND2 ChIP-seq peaks in the TADs harboring genes with unchanged, down- or upregulated expression in mutant AVCs, respectively. The TADs of genes with altered expression encode more HAND2-interacting genomic regions. To account for the different numbers of genes and HAND2 ChIP-seq peaks per TAD, peak counts were normalized as numbers of peaks per gene for each TAD.

(C) GO enrichment analysis for biological processes for the 167 downregulated and 372 upregulated genes (in mutant AVCs) that contain HAND2 ChIP-seq peaks in their TADs.

(D) GO analysis to identify the HAND2 target genes with annotated functions in EMT processes and AV cushion morphogenesis.

(E) Mouse embryonic hearts isolated at E9.25–9.5 were used for ChIP-qPCR validation of the most prominent HAND2 ChIP-seq peaks in the TADs of the genes shown in D (n = 4 using two biological replicates, mean ± SD, $p \leq 0.05$). See also Figure S4 and Tables S5 and S6.

is correctly specified in mutant hearts and corroborates the specificity of the molecular alterations underlying the cardiac cushion agenesis (Figures 2 and 3). WISH analysis of HAND2 target genes with significantly changed transcript levels (Figure 4D; Table S6) failed to detect spatial alterations for several of them due to low and/or uniform expression (low: *Vasn*, *Tmem100*, *Erbb4*, *Tll1*, and *Acvrl*, uniform: *Cyr61* and *Hnrnapab*, data not shown). Furthermore, the small but significant differences in transcript levels for several additional DEGs could not be reliably detected by WISH (*Pitx2*, *Glipr2*, *Kdr*, *Gja5*, Figure S5A and data not shown), likely due to the qualitative nature of in situ hybridization that best detects spatial changes. This is exemplified by the fact that WISH revealed the ectopic expression of two transcriptionally upregulated HAND2 target genes in mutant AVCs, namely, the NOTCH1 transcriptional mediator *Hey1* and the homeobox transcription factor *Hhex* (Figures 5C and 5D).

In agreement with the reduced expression detected by RNA-seq (Figure 4D; Table S6), WISH corroborates the loss of the transcriptional regulators *Twist1* (Figure 5E, see also Figure 2E),

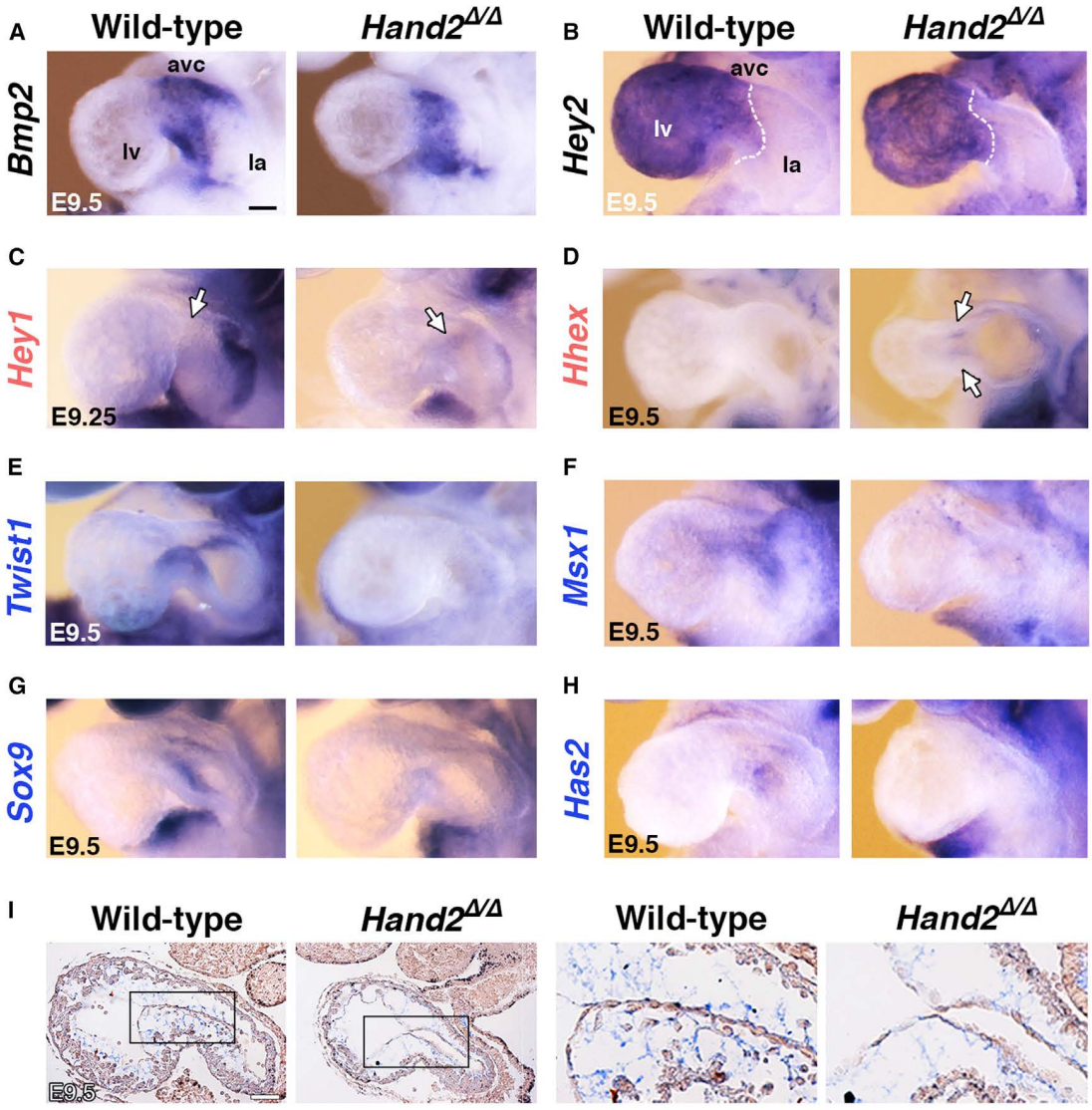


Figure 5. Whole-Mount in Situ Hybridization Reveals Spatial Changes in Some of the Differentially Expressed HAND2 Target Genes

(A and B) Unaltered expression of *Bmp2* (A) and *Hey2* (B) indicates that the AVC domain is established correctly in *Hand2*-deficient hearts.

(C and D) Ectopic expression of the *Hey1* (C) and *Hhex* transcriptional regulators (D) in the mutant AVC corroborated their upregulation detected by RNA-seq analysis.

(E–H) Loss of *Twist1* (E), *Msx1* (F), Sox9 (G), and *Has2* (H) from the mutant AVC is in agreement with their transcriptional downregulation detected by RNA-seq analysis.

(I) Alcian blue staining of glycosaminoglycans shows the reduced deposition of extra-cellular matrix in the cardiac jelly of *Hand2*-deficient mouse embryos. Right panels show the enlargements indicated by frames in the left panels. Gene names in black, unaltered; red, increased; blue, reduced transcript levels as determined by RNA-seq analysis (Figure 4).

Scale bars, 100 μm. See also Figure S5.

Msx1 (Figure 5F), and *Sox9* (Figures 5G and S5B) from mutant AVCs by E9.5. The downregulation of *Msx1* indicates that BMP signal transduction is disrupted in the mutant AVC (Figure 5F; Table S6). SOX9 regulates the proliferation of the mesenchymal progenitor cells, and its loss agrees with the lack of delaminating mesenchymal cells in mutant AVCs (Figures 5G and S5B; Akiyama et al., 2004). Most relevant to the disrupted EMT (Figure 3), the expression of *Has2* and *Snai1* is significantly downregulated in *Hand2*-deficient AVCs (Figure 4D). *Has2* encodes the enzyme that produces hyaluronic acid in the cardiac jelly (Camenisch et al., 2000). Genetic inactivation of the mouse *Has2* gene disrupts both cardiac jelly deposition and mesenchymal cell migration during cardiac cushion development. In *Hand2*-deficient hearts, the loss of *Has2* is paralleled by reduced extra-cellular matrix/cardiac jelly deposition (Figures 5H and 5I). Camenisch et al. (2000) showed that treatment of *Has2*-deficient

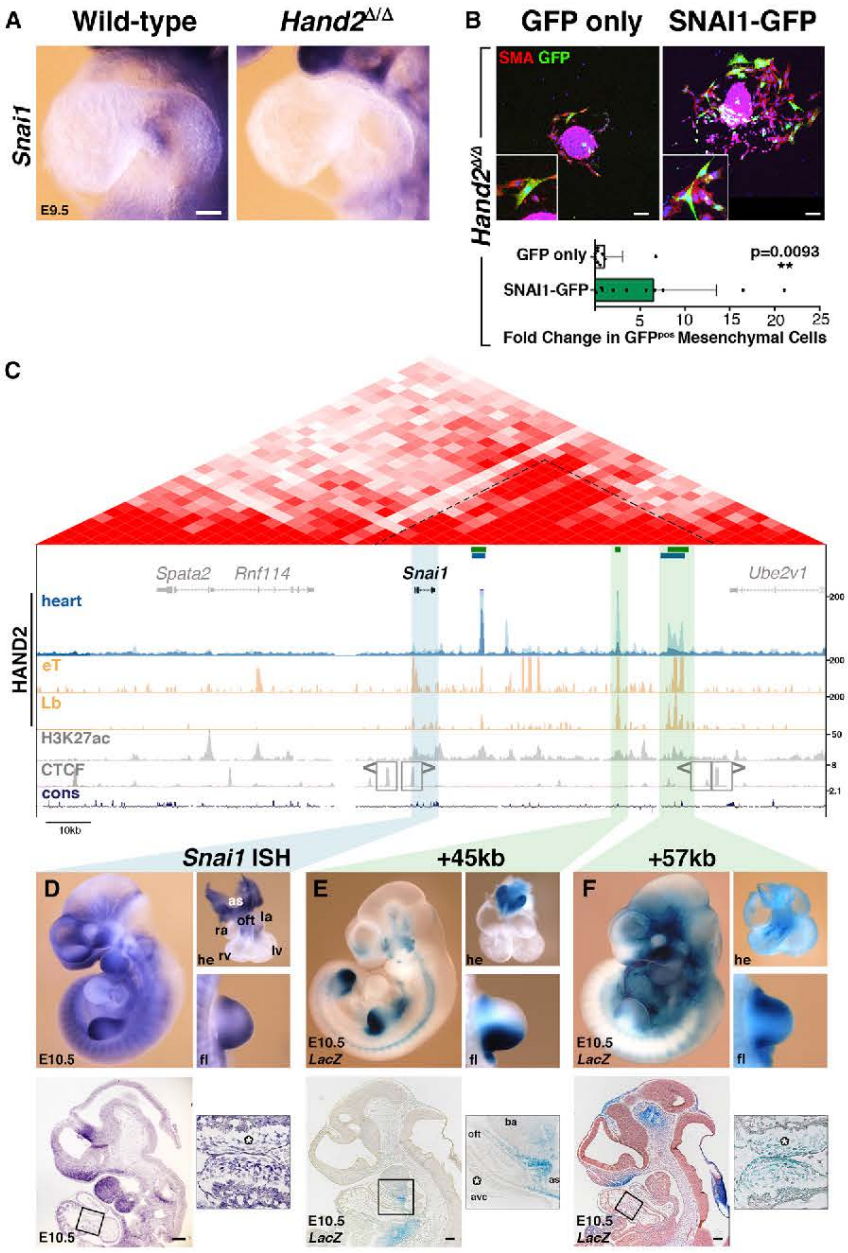


Figure 6. Re-expression of *Snai1* in *Hand2*-Deficient AVC Explants Partially Restores Mesenchymal Cell Migration and Regulation of *Snai1* Expression by Enhancers Enriched in HAND2 Chromatin Complexes

(A) *Snai1* expression is lost from the endocardium of *Hand2*-deficient hearts.

(B) Upper panels: *Hand2*-deficient AVC explants were infected either with GFP (control) or SNAI1-GFP adenovirus to re-express SNAI1. Samples were analyzed 72 hr after infection. All infected cells that migrated into the matrix are marked by GFP expression (green). Smooth muscle actin (SMA, red) was detected to reveal cellular morphology. Lower panel: quantitation of cell migration in *Hand2*-deficient AVC explants. The mean ± SD and all individual data points are shown. The observed increase in mesenchymal cells in SNAI1-GFP infected explants is significant ($p = 0.0093$, Mann-Whitney test). Scale bars, 100 μm.

(C) Scheme of the mouse *Snai1* TAD (red, HiC-data). The HAND2 ChIP-seq profiles in the heart (this study), *Hand2*-expressing tissues (eT) and limb buds (Lb) (Osterwalder et al., 2014) are shown below together with the H3K27ac profile in developing hearts (E11.5, Nord et al., 2013). The *Snai1* TAD boundaries are marked by CTCF-binding regions in opposite orientation (Gómez-Marín et al., 2015). Green bars indicate the HAND2 target CRMs located +14kb, +45kb, and +57kb that were analyzed by *LacZ* reporter assays in transgenic founder embryos.

(D) *Snai1* transcript distribution in wild-type mouse embryos (E10.5).

(E) A representative transgenic founder embryo (E10.5) shows the activity of *LacZ* reporter construct encoding the +45kb CRM. (F) A representative transgenic founder embryo (E10.5) shows the activity of *LacZ* reporter construct encoding the +57kb CRM. The upper panels in (E) and (F) depict whole embryos, dissected hearts, and forelimb buds. The lower panels show sagittal sections at the level of the heart. The boxed areas indicate the enlargements shown in the right panels.

Asterisks, AVC cardiac cushion mesenchyme; as, aortic sac; avc, atrioventricular canal; ba, branchial arch; la, left atria; lv, left ventricle; oft, outflow tract; ra, right atria; rv, right ventricle. Scale bars, 200 μm. See also Figures S5 and S6.

AVC explants with hyaluronic acid restores mesenchymal cell migration. In contrast, culturing *Hand2*-deficient AVC explants in hyaluronic acid does not suffice to restore migration (data not shown). This is in line with the fact that the genetic inactivation of *Hand2* affects the expression of multiple genes required for AVC cushion development (Figures 4 and S4).

HAND2 Regulates the Transcription of *Snai1*, a Key Regulator of the EMT and Mesenchymal Cell Migration in the AVC

The transcriptome combined with ChIP-seq/qPCR analysis (Figures 4D and 4E) and WISH (Figure 6A) establishes *Snai1* as a direct transcriptional target of HAND2. As *Snai1* is a EMT key regulator (Nieto, 2011), its loss from the mutant AVC (Figure 6A) is likely causally linked to the observed cardiac cushion agenesis (Figure 2). To test this experimentally, *Hand2*-deficient AVC explants were infected either with adenovirus producing both SNAI1 and GFP proteins (SNAI1-GFP) or control GFP virus (GFP only; Figure 6B; Tao et al., 2011). Quantitative analysis shows that infection of *Hand2*-deficient AVC explants with SNAI1-producing virus induces migration of a significantly larger fraction of GFP-positive mesenchymal cells into the collagen matrix than GFP alone (Figure 6B, $p = 0.0093$ Mann-Whitney test; see Figure S5C for wild-type controls).

This partial restoration of mesenchymal cell migration reveals the functional importance of the HAND2-*Snai1* interactions for the EMT during cardiac cushion formation. Therefore, we analyzed the potential enhancer activities of the three HAND2-interacting CRMs located in the *Snai1* TAD (Figures 6C and S6A). Our previous analysis has shown that these three candidate CRMs are also enriched in HAND2 chromatin complexes isolated from mouse limb buds (Osterwalder et al., 2014) and overlap regions of active chromatins in embryonic hearts (H3K27ac profile in Figure 6C). Their transcription enhancing potential was assessed in transgenic mouse founder embryos using *LacZ* reporter constructs (Figures 6E and 6F). In particular, *LacZ* activity reminiscent of *Snai1* expression (Figure 6D) was detected for reporters encoding the CRMs located +45kb and +57kb downstream of the *Snai1* transcription start site (Figures 6E and 6F). In contrast, no *LacZ* activity was detected using the *Snai1* +14kb genomic region, whose sequence is not well conserved in mammals (data not shown). The *Snai1* +45kb CRM is active in cells located between the OFT and aortic sac (n = 7/10), the posterior (n = 10/10) and anterior (n = 3/10) limb bud mesenchyme, and branchial arches and cranial mesenchyme (n = 5/10, Figure 6E). Most relevant with respect to the AVC, the *Snai1* +57kb CRM is active in the cardiac cushion mesenchyme of the AVC and OFT (n = 7/9, Figure 6F) and in most other embryonic tissues expressing *Snai1* (Figure 6D). Together, the activities of these two HAND2-interacting CRMs recapitulate most of the *Snai1* expression pattern in mouse embryos (Figures 6D and S6B–S6D). Indeed, in *Hand2*-deficient embryos, *Snai1* expression is not only lost from the AVC (Figure 6A), but also significantly reduced in the second branchial arch and forelimb bud mesenchyme (Figure S6E).

DISCUSSION

We show that HAND2 chromatin complexes interact with genomic regions such as enhancers located in the *cis*-regulatory landscapes of genes functioning in heart morphogenesis. Previous molecular analysis showed that the altered expression of many of these genes correlates well with the defects in right ventricle and OFT morphogenesis observed in *Hand2*-deficient mouse embryos (Cohen et al., 2012; Tsuchihashi et al., 2011; Zhao et al., 2008). We provide evidence that about half of all previously known genes with altered expression are direct transcriptional targets of HAND2. In addition, our ChIP-seq analysis reveals that a significant fraction of the genomic regions enriched in HAND2 chromatin complexes are also bound by GATA4 complexes (He et al., 2014). This is interesting in light of previous studies, which showed that HAND2 and GATA4 form transcriptional complexes regulating gene expression in developing hearts (Dai et al., 2002). In addition, it has been shown that AVC enhancers are repressed in the atrial and ventricular myocardium by complexes containing GATA4, HEY1, and/or HEY2 transcriptional repressors (Firulli et al., 2000; Stefanovic et al., 2014). These three repressors plus RUNX2 and TWIST1 are all able to form heterodimers with HAND2 (Firulli et al., 2005; Funato et al., 2009). As the expression of many HAND2 target genes is upregulated in *Hand2*-deficient AVCs (this study), HAND2-mediated transcriptional repression is likely functionally relevant to normal AVC development. For example, the HAND2 target *Hhex* is ectopically expressed in the AVC of *Hand2*-deficient embryos. Indeed, genetic inactivation of *Hhex* increases the number of mesenchymal cells in AVC cardiac cushions and causes valve dysplasia (Hallaq et al., 2004).

One key finding of our analysis is that constitutive inactivation of *Hand2* disrupts the EMT underlying cardiac cushion formation in the AVC. This disruption of AVC morphogenesis contrasts with the phenotypes resulting from specific inactivation of *Hand2* in either the endocardium or mesenchyme. Neither inactivation disrupts AVC cardiac cushion formation but specifically alters the AVC-derived tricuspid valves (tricuspid atresia; VanDusen et al., 2014a, 2014b). This discrepancy is a likely consequence of different *Hand2* inactivation kinetics. As genetic inactivation of *Hand2* in SHF progenitors also alters AVC development, we cannot formally exclude that recruitment of progenitors to the AVC is compromised in *Hand2*-deficient embryos (Tsuchihashi et al., 2011), even though the expression of early markers for AVC morphogenesis remains normal (this study).

However, our analysis shows that most genes with known functions in the EMT underlying AVC cardiac cushion formation are direct transcriptional targets of HAND2. Together with the cellular analysis, these results point to specific disruption of the EMT rather than a general arrest of AVC development and suggest that HAND2 is a very upstream regulator of AVC cardiac cushion morphogenesis. In agreement, re-expression of the HAND2 target *Snai1* in *Hand2*-deficient AVCs explants only partially restores mesenchymal cell migration, which indicates that other HAND2 target DEGs have essential functions in the EMT and/or mesenchymal cell migration during cardiac cushion development. Ingenuity pathway analysis of HAND2 target genes shows that HAND2 enhances the expression of genes such as *Msx1* and *Hnrnapab*, which, in turn, reinforce the expression of the HAND2 targets *Has2*, *Twist1*, and *Snai1* (Figure 7). This type of dual transcriptional reinforcement likely increases the robustness of the expression of HAND2 target genes with key functions in AVC morphogenesis. In fact, it is reminiscent of the dual transcriptional reinforcement seen for key genes during limb bud development. In early limb buds, HAND2 reinforces the expression of the *Shh* morphogen by directly regulating its transcription and indirectly via upregulating E26 transformation-specific (ETS) transcription factors, which also positively regulate *Shh* expression (Osterwalder et al., 2014).

Remarkably, this study identifies HAND2 as key regulator of most genes with known functions in EMT processes and cardiac cushion formation in the developing AVC including *Snai1* (Garside et al., 2013). We also provide evidence that HAND2 directly regulates *Snai1* transcription in other embryonic tissues. The notion that the direct transcriptional regulation of *Snai1* by HAND2 maybe of more general importance is supported by genetic analysis as *Hand2* and *Snai1* are both essential for the EMT of epicardial cells and morphogenesis of craniofacial structures such as the palate (Barnes et al., 2011; Murray et al., 2007; Tao et al., 2013; Xiong et al., 2009). In summary, our study identifies the HAND2 target GRN that controls the initiation of

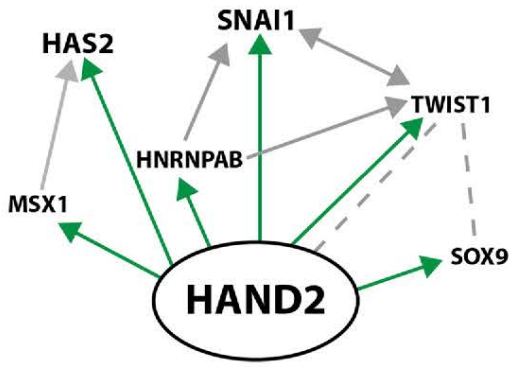


Figure 7. Scheme Depicting the Interactions among Positively Regulated HAND2 Targets Genes

The HAND2 GRN was constructed using Ingenuity pathway analysis in combination with manual annotation (Chen et al., 2008). Arrows indicate transcriptional upregulation and direct upregulation by HAND2 transcriptional complexes is indicated in green. Broken gray lines direct protein-protein interactions. This graph represents the simplest possible scheme to illustrate the relevant direct interactions.

cardiac valve formation and provides evidence for its general role in regulating the expression of *Snai1* during mouse embryogenesis. Last but not least, the identification of HAND2 as key regulator of AVC cardiac cushion morphogenesis may have important implications for regenerative medicine (see, e.g., review by Levine et al., 2015).

EXPERIMENTAL PROCEDURES

Ethics Statement, Mouse Strains, and Embryos

All experiments conducted with mice and embryos of both sexes at the developmental ages indicated (see Results and below) were performed in strict accordance with Swiss law. All animal studies were evaluated and approved by the Regional Commission on Animal Experimentation (license 1951). The 3Rs were taken into account in designing the animal studies. The procedures for generating transgenic mice at the Lawrence Berkeley National Laboratory (LBNL) were reviewed and approved by the LBNL Animal Welfare and Research Committee. The *Hand2*$^{\Delta}$ and *Hand2*^{3xF} alleles (Galli et al., 2010; Osterwalder et al., 2014) were outbred into an NMRI background as this prolongs survival of *Hand2*-deficient embryos in comparison to the previously used 129SvJ/C57BL6 background.

ChIP-Seq Analysis

To obtain sufficient material for ChIP-seq analysis, about 600 hearts had to be dissected from *Hand2*$^{3xF/3xF}$ mouse embryos at E10.5. After collection, these were split in two batches and processed as completely independent biological replicates for ChIP-seq analysis using the M2 anti-FLAG antibody (F1804; Sigma; (Osterwalder et al., 2014). Library construction and sequencing were performed by the Genome Technology Access Center using an Illumina HiSeq 2500 system. More details are included in the Supplemental Experimental Procedures.

Transcriptome Analysis

AVCs dissected from wild-type and *Hand2*$^{\Delta/\Delta}$ mutant embryos at E9.0–9.25 were flash frozen in RLT buffer (QIAGEN). Four AVCs were pooled per replicate, keeping the same gender ratio for all replicates (AVCs of two male and female embryos). RNA was extracted using the QIAGEN RNeasy mini kit. The quality of total RNA (30–60 ng) was analyzed using the Agilent 2100 Bioanalyzer, and both wild-type and mutant samples had an RNA integrity number (RIN) of 8.8–9.6. Libraries were prepared using the Clontech SMARTer kit and sequenced on a HiSeq3000 using a single-read 50 cycle protocol. More details on the computational analysis are provided in the Supplemental Experimental Procedures.

AVC Explant Cultures

AVC explant cultures were set on matrixes of rat-tail collagen type I (Luna-Zurita et al., 2010, see Supplemental Information for more details). Only wild-type and mutant AVC explants that were still beating, i.e., alive after 72 hr in culture were analyzed. Supplementation with hyaluronic acid (HA): both the collagen matrix and serum-free culture medium were supplemented with 0.75 mg/mL HA (Camenisch et al., 2000). Adenoviral infections: the titers of the SNAI1-GFP or GFP adenoviruses (Tao et al., 2011 and Vector Biolabs) were determined in mitomycin-treated mouse embryonic fibroblasts and adjusted such that equal numbers of active virus particles were used. Following attachment to the matrix, mutant and wild-type AVC explants were incubated with 6×10^6 plaque-forming units (PFUs) of either SNAI1-GFP or GFP virus for 12 hr in serum-free medium. Then, the AVC explants were cultured for 60 hr in fresh serum-free medium. After fixation in 4% paraformaldehyde (PFA) (30-min room temperature), antigens were detected using anti-SMA-Cy3 antibodies (1:250, Sigma) and Phalloidin-Alexa 488 (1:250, Life Technologies) and nuclei counterstained with Hoechst-33258 and analyzed using a Leica SP5 confocal microscope. The analysis of wild-type controls shows that GFP-virus tends to infect AVC cells more efficiently than SNAI1-GFP virus (Figure S5C). Therefore, the restoration of cell migration following infection of mutant AVCs with SNAI1-GFP virus is rather underestimated (Figure 6B).

Statistical Analysis

ChIP-seq

Following initial alignment of sequences, the genome-wide pattern of binding of HAND2 was determined using MACS (version 1.4.2) with a p value threshold of 1e-5.

ChIP-qPCR

Mean ± SD were calculated using the Prism (GraphPad Software) Student-t test.

Transcriptome

Following initial sequence alignment, *edgeR* was used to normalize the datasets (trimmed mean of M-values [TMM] normalization) and to identify the differentially expressed genes (DEGs). Only genes expressed in all samples were considered (reads per million [RPM] ≥ 1). DEGs are defined as genes with a q value ≤ 0.05 and a linear fold change ≥ 1.5.

AVC Explant Cultures

The Mann-Whitney test was used to determine significant differences in numbers of migrating mesenchymal cells. More details on statistical validation of the ChIP-seq and transcriptome analyses are included in the Supplemental Experimental Procedures.

Histology and Immunofluorescence Analysis

Embryos were collected and fixed overnight in 4% PFA at 4°C and embedded in paraffin wax. Standard protocols were used for histological staining (H&E; Alcian blue) of 7-μm paraffin sections. Minimally three biological replicates were analyzed for each stage, genotype, and antigen shown. Antibodies are listed the Supplemental Experimental Procedures.

Generation and Analysis of *LacZ* Transgenic Founder Embryos

Genomic regions were amplified by PCR from mouse genomic DNA (*Snai1* +45kb, *Snai1* +14kb) or recovered as restriction fragment (*Snai1* +57kb, BAC clone RP23-193B17) and cloned into the *Hsp68*-promoter-*LacZ* reporter vector (Osterwalder et al., 2014). Transgenic founder embryos were generated by pronuclear injection and analyzed by *LacZ* staining, and transgenic embryos were identified by genotyping.

ACCESSION NUMBERS

The accession numbers for the primary ChIP-seq and transcriptome datasets reported in this paper are GEO: GSE73368 and GSE94246, respectively.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and six tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2017.05.004.

AUTHOR CONTRIBUTIONS

F.L. performed the ChIP-seq and follow-up functional analysis, A.G. performed the transcriptome and WISH analysis for revision, J.G. participated in different aspects of the experimental studies, and I.B. performed the bioinformatics analysis. M.O. was involved in initiating this study, J.L. generated the SNAI1-GFP virus, J.A.A. generated the *LacZ* reporters, and A.V. provided the resources for the transgenic and bioinformatics analysis. The experimental study design was done and the manuscript written by F.L., J.L.-R., A.Z., and R.Z. with input from all authors.

ACKNOWLEDGMENTS

The Genome Technology Access Center (Department of Genetics at Washington University School of Medicine) is acknowledged for deep sequencing, and R. Ivanek (DBM Bioinformatics Core) performed the bioinformatics analysis of the ChIP-seq datasets. All calculations were performed using the sciCORE (https://scicore.unibas.ch/) scientific computing core facility at University of Basel. We thank D. Speziale for technical assistance, A. Offinger's team for excellent animal care, and P. Lorentz (DBM Bio-Optics Core Facility) for imaging support. V. Afzal, B. Mannion, and I. Plajzer-Frick provided assistance with transgenics, and L. Bazzani provided the GFP adenovirus preparations. T. Papoutsi, C. Rolando, and O. Pertz are acknowledged for advice on AVC explants and infections. We are grateful to V. Christoffels for helpful input and anonymous reviewers for critical input that resulted in a significantly improved study. This research was mostly supported by SNF grants 31003A_146248 and 310030B_166685 to A.Z. and R.Z. and by funds from the University of Basel. The stay of M.O. in the group of A.V. was supported by an SNSF early mobility postdoctoral fellowship. I.B., J.A., and A.V. were supported by NIH grants R24HL123879, UM1HL098166, R01HG003988, and U54HG006997, and research at LBNL was performed under Department of Energy Contract DE-AC02-05CH11231 to University of California. J.L. was supported by NIH/NHLBI grant 1R01HL127033.

Received: May 20, 2016
Revised: March 20, 2017
Accepted: April 28, 2017
Published: May 23, 2017

REFERENCES

Akiyama, H., Chaboissier, M.C., Behringer, R.R., Rowitch, D.H., Schedl, A., Epstein, J.A., and de Crombrugghe, B. (2004). Essential role of Sox9 in the pathway that controls formation of cardiac valves and septa. Proc. Natl. Acad. Sci. USA *101*, 6502–6507.

Barnes, R.M., Firulli, B.A., VanDusen, N.J., Morikawa, Y., Conway, S.J., Cserjesi, P., Vincentz, J.W., and Firulli, A.B. (2011). Hand2 loss-of-function in Hand1-expressing cells reveals distinct roles in epicardial and coronary vessel development. Circ. Res. *108*, 940–949.

Bruneau, B.G. (2008). The developmental genetics of congenital heart disease. Nature *451*, 943–948.

Camenisch, T.D., Spicer, A.P., Brehm-Gibson, T., Biesterfeldt, J., Augustine, M.L., Calabro, A., Jr., Kubalak, S., Klewer, S.E., and McDonald, J.A. (2000). Disruption of hyaluronan synthase-2 abrogates normal cardiac morphogenesis and hyaluronan-mediated transformation of epithelium to mesenchyme. J. Clin. Invest. *106*, 349–360.

Chen, Y.H., Ishii, M., Sucov, H.M., and Maxson, R.E., Jr. (2008). Msx1 and Msx2 are required for endothelial-mesenchymal transformation of the atrioventricular cushions and patterning of the atrioventricular myocardium. BMC Dev. Biol. *8*, 75.

Christoffels, V.M., Smits, G.J., Kispert, A., and Moorman, A.F. (2010). Development of the pacemaker tissues of the heart. Circ. Res. *106*, 240–254.

Cohen, E.D., Miller, M.F., Wang, Z., Moon, R.T., and Morrisey, E.E. (2012). Wnt5a and Wnt11 are essential for second heart field progenitor development. Development *139*, 1931–1940.

Dai, Y.S., and Cserjesi, P. (2002). The basic helix-loop-helix factor, HAND2, functions as a transcriptional activator by binding to E-boxes as a heterodimer. J. Biol. Chem. *277*, 12604–12612.

Dai, Y.S., Cserjesi, P., Markham, B.E., and Molkentin, J.D. (2002). The transcription factors GATA4 and dHAND physically interact to synergistically activate cardiac gene expression through a p300-dependent mechanism. J. Biol. Chem. *277*, 24390–24398.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376–380.

Firulli, B.A., Hadzic, D.B., McDaid, J.R., and Firulli, A.B. (2000). The basic helix-loop-helix transcription factors dHAND and eHAND exhibit dimerization characteristics that suggest complex regulation of function. J. Biol. Chem. *275*, 33567–33573.

Firulli, B.A., Krawchuk, D., Centonze, V.E., Vargesson, N., Virshup, D.M., Conway, S.J., Cserjesi, P., Laufer, E., and Firulli, A.B. (2005). Altered Twist1 and Hand2 dimerization is associated with Saethre-Chotzen syndrome and limb abnormalities. Nat. Genet. *37*, 373–381.

Funato, N., Chapman, S.L., McKee, M.D., Funato, H., Morris, J.A., Shelton, J.M., Richardson, J.A., and Yanagisawa, H. (2009). Hand2 controls osteoblast differentiation in the branchial arch by inhibiting DNA binding of Runx2. Development *136*, 615–625.

Galli, A., Robay, D., Osterwalder, M., Bao, X., Bénazet, J.D., Tariq, M., Paro, R., Mackem, S., and Zeller, R. (2010). Distinct roles of Hand2 in initiating polarity and posterior Shh expression during the onset of mouse limb bud development. PLoS Genet. *6*, e1000901.

Garside, V.C., Chang, A.C., Karsan, A., and Hoodless, P.A. (2013). Co-ordinating Notch, BMP, and TGF-β signaling during heart valve development. Cell. Mol. Life Sci. *70*, 2899–2917.

Gómez-Marín, C., Tena, J.J., Acemel, R.D., López-Mayorga, M., Naranjo, S., de la Calle-Mustienes, E., Maeso, I., Beccari, L., Aneas, I., Vielmas, E., et al. (2015). Evolutionary comparison reveals that diverging CTCF sites are signatures of ancestral topological associating domains borders. Proc. Natl. Acad. Sci. USA *112*, 7542–7547.

Greulich, F., Rudat, C., and Kispert, A. (2011). Mechanisms of T-box gene function in the developing heart. Cardiovasc. Res. *91*, 212–222.

Hallaq, H., Pinter, E., Enciso, J., McGrath, J., Zeiss, C., Brueckner, M., Madri, J., Jacobs, H.C., Wilson, C.M., Vasavada, H., et al. (2004). A null mutation of Hhex results in abnormal cardiac development, defective vasculogenesis and elevated Vegfa levels. Development *131*, 5197–5209.

He, A., Gu, F., Hu, Y., Ma, Q., Ye, L.Y., Akiyama, J.A., Visel, A., Pennacchio, L.A., and Pu, W.T. (2014). Dynamic GATA4 enhancers shape the chromatin landscape central to heart development and disease. Nat. Commun. *5*, 4907.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

Holler, K.L., Hendershot, T.J., Troy, S.E., Vincentz, J.W., Firulli, A.B., and Howard, M.J. (2010). Targeted deletion of Hand2 in cardiac neural crest-derived cells influences cardiac gene expression and outflow tract development. Dev. Biol. *341*, 291–304.

Kelly, R.G. (2012). The second heart field. Curr. Top. Dev. Biol. *100*, 33–65.

Levine, R.A., Hagége, A.A., Judge, D.P., Padala, M., Dal-Bianco, J.P., Aikawa, E., Beaudoin, J., Bischoff, J., Bouatia-Naji, N., Bruneval, P., et al.; Leducq Mitral Transatlantic Network (2015). Mitral valve disease–morphology and mechanisms. Nat. Rev. Cardiol. *12*, 689–710.

Lin, C.J., Lin, C.Y., Chen, C.H., Zhou, B., and Chang, C.P. (2012). Partitioning the heart: mechanisms of cardiac septation and valve development. Development *139*, 3277–3299.

Liu, N., Barbosa, A.C., Chapman, S.L., Bezprozvannaya, S., Qi, X., Richardson, J.A., Yanagisawa, H., and Olson, E.N. (2009). DNA binding-dependent and -independent functions of the Hand2 transcription factor during mouse embryogenesis. Development *136*, 933–942.

Luna-Zurita, L., Prados, B., Grego-Bessa, J., Luxán, G., del Monte, G., Benguría, A., Adams, R.H., Pérez-Pomares, J.M., and de la Pompa, J.L. (2010). Integration of a Notch-dependent mesenchymal gene program and Bmp2-driven cell invasiveness regulates murine cardiac valve formation. J. Clin. Invest. *120*, 3493–3507.

Ma, L., Lu, M.F., Schwartz, R.J., and Martin, J.F. (2005). Bmp2 is essential for cardiac cushion epithelial-mesenchymal transition and myocardial patterning. Development *132*, 5601–5611.

MacGrogan, D., Luxán, G., Driessen-Mol, A., Bouten, C., Baaijens, F., and de la Pompa, J.L. (2014). How to make a heart valve: from embryonic development to bioengineering of living valve substitutes. Cold Spring Harb. Perspect. Med. *4*, a013912.

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

Murray, S.A., Oram, K.F., and Gridley, T. (2007). Multiple functions of Snail family genes during palate development in mice. Development *134*, 1789–1797.

Niessen, K., Fu, Y., Chang, L., Hoodless, P.A., McFadden, D., and Karsan, A. (2008). Slug is a direct Notch target required for initiation of cardiac cushion cellularization. J. Cell Biol. *182*, 315–325.

Nieto, M.A. (2011). The ins and outs of the epithelial to mesenchymal transition in health and disease. Annu. Rev. Cell Dev. Biol. *27*, 347–376.

Nord, A.S., Blow, M.J., Attanasio, C., Akiyama, J.A., Holt, A., Hosseini, R., Phouanenavong, S., Plajzer-Frick, I., Shoukry, M., Afzal, V., et al. (2013). Rapid and pervasive changes in genome-wide enhancer usage during mammalian development. Cell *155*, 1521–1531.

Osterwalder, M., Speziale, D., Shoukry, M., Mohan, R., Ivanek, R., Kohler, M., Beisel, C., Wen, X., Scales, S.J., Christoffels, V.M., et al. (2014). HAND2 targets define a network of transcriptional regulators that compartmentalize the early limb bud mesenchyme. Dev. Cell *31*, 345–357.

Shen, L., Li, X.F., Shen, A.D., Wang, Q., Liu, C.X., Guo, Y.J., Song, Z.J., and Li, Z.Z. (2010). Transcription factor HAND2 mutations in sporadic Chinese patients with congenital heart disease. Chin. Med. J. (Engl.) *123*, 1623–1627.

Srivastava, D., Thomas, T., Lin, Q., Kirby, M.L., Brown, D., and Olson, E.N. (1997). Regulation of cardiac mesodermal and neural crest development by the bHLH transcription factor, dHAND. Nat. Genet. *16*, 154–160.

Stefanovic, S., and Christoffels, V.M. (2015). GATA-dependent transcriptional and epigenetic control of cardiac lineage specification and differentiation. Cell. Mol. Life Sci. *72*, 3871–3881.

Stefanovic, S., Barnett, P., van Duijvenboden, K., Weber, D., Gessler, M., and Christoffels, V.M. (2014). GATA-dependent regulatory switches establish atrioventricular canal specificity during heart development. Nat. Commun. *5*, 3680.

Sun, Y.M., Wang, J., Qiu, X.B., Yuan, F., Li, R.G., Xu, Y.J., Qu, X.K., Shi, H.Y., Hou, X.M., Huang, R.T., et al. (2016). A HAND2 loss-of-function mutation causes familial ventricular septal defect and pulmonary stenosis. G3 (Bethesda) *6*, 987–992.

Tao, G., Levay, A.K., Gridley, T., and Lincoln, J. (2011). Mmp15 is a direct target of Snai1 during endothelial to mesenchymal transformation and endocardial cushion development. Dev. Biol. *359*, 209–221.

Tao, G., Miller, L.J., and Lincoln, J. (2013). Snai1 is important for avian epicardial cell transformation and motility. Dev. Dyn. *242*, 699–708.

Timmerman, L.A., Grego-Bessa, J., Raya, A., Bertrán, E., Pérez-Pomares, J.M., Díez, J., Aranda, S., Palomo, S., McCormick, F., Izpisúa-Belmonte, J.C., and de la Pompa, J.L. (2004). Notch promotes epithelial-mesenchymal transition during cardiac development and oncogenic transformation. Genes Dev. *18*, 99–115.

Tsuchihashi, T., Maeda, J., Shin, C.H., Ivey, K.N., Black, B.L., Olson, E.N., Yamagishi, H., and Srivastava, D. (2011). Hand2 function in second heart field progenitors is essential for cardiogenesis. Dev. Biol. *351*, 62–69.

VanDusen, N.J., and Firulli, A.B. (2012). Twist factor regulation of non-cardiomyocyte cell lineages in the developing heart. Differentiation *84*, 79–88.

VanDusen, N.J., Casanovas, J., Vincentz, J.W., Firulli, B.A., Osterwalder, M., Lopez-Rios, J., Zeller, R., Zhou, B., Grego-Bessa, J., De La Pompa, J.L., et al. (2014a). Hand2 is an essential regulator for two Notch-dependent functions within the embryonic endocardium. Cell Rep. *9*, 2071–2083.

VanDusen, N.J., Vincentz, J.W., Firulli, B.A., Howard, M.J., Rubart, M., and Firulli, A.B. (2014b). Loss of Hand2 in a population of Periostin lineage cells results in pronounced bradycardia and neonatal death. Dev. Biol. *388*, 149–158.

Visel, A., Minovitsky, S., Dubchak, I., and Pennacchio, L.A. (2007). VISTA Enhancer Browser–a database of tissue-specific human enhancers. Nucleic Acids Res. *35*, D88–D92.

Wu, Z.Q., Rowe, R.G., Lim, K.C., Lin, Y., Willis, A., Tang, Y., Li, X.Y., Nor, J.E., Maillard, I., and Weiss, S.J. (2014). A Snail1/Notch1 signalling axis controls embryonic vascular development. Nat. Commun. *5*, 3998.

Xiong, W., He, F., Morikawa, Y., Yu, X., Zhang, Z., Lan, Y., Jiang, R., Cserjesi, P., and Chen, Y. (2009). Hand2 is required in the epithelium for palatogenesis in mice. Dev. Biol. *330*, 131–141.

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

Zhao, R., Watt, A.J., Battle, M.A., Li, J., Bondow, B.J., and Duncan, S.A. (2008). Loss of both GATA4 and GATA6 blocks cardiac myocyte differentiation and results in acardia in mice. Dev. Biol. *317*, 614–619.

Cell Reports, Volume *19*

Supplemental Information

HAND2 Target Gene Regulatory Networks Control Atrioventricular Canal and Cardiac Valve Development

Frédéric Laurent, Ausra Girdziusaite, Julie Gamart, Iros Barozzi, Marco Osterwalder, Jennifer A. Akiyama, Joy Lincoln, Javier Lopez-Rios, Axel Visel, Aimée Zuniga, and Rolf Zeller

Supplemental Figures

Figure S1

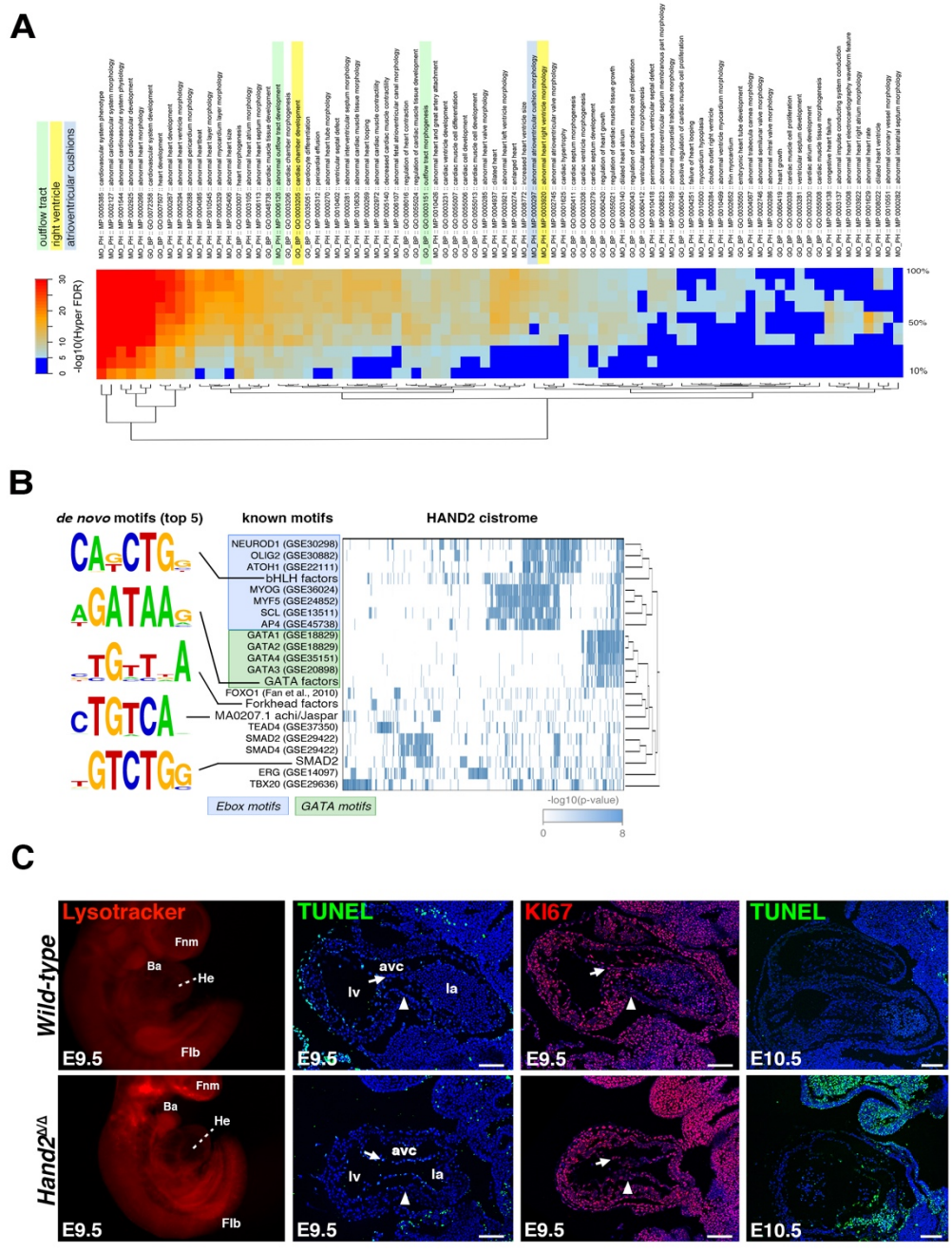


Figure S1 (related to Figure 1). Analysis of the genomic regions enriched in HAND2 chromatin complexes and associated genes.

(A) Heat map showing the enrichment of heart-related GO term categories in the list of putative HAND2 target genes defined by GREAT analysis. The columns denote ten incremental bins of HAND2-bound regions (from 10% to the complete set = 100%). For visualization, hyper-geometric p-values equal or lower to 1e-30 were set to this value. Terms were hierarchically

clustered and re-ordered according to the row-wise mean. GO terms related to the development of specific cardiac compartments are highlighted by different colors: outflow tract (green), right ventricle (yellow) and atrioventricular cushions (blue). While the general cardiac terms are systematically identified in each incremental bin (top term: cardiovascular system phenotype), GO terms related to more specific aspects of cardiac development (such as: abnormal heart right ventricle morphology) are only detected when considering an increasing number of peaks or the entire dataset. (B) Hierarchical clustering of the high-affinity matches for each of the enriched known motifs across the HAND2-contacted regions is shown. The top five binding motifs that were identified *de novo* are highlighted on the left. (C) Analysis of the patterns of cell death in *Hand2*-deficient mouse embryos. Panel Lysotracker: whole mount Lysotracker staining reveals increased apoptosis in branchial arches (Ba) and frontonasal mass (Fnm) of mutant mouse embryos at E9.5 (red fluorescence), while no aberrant apoptosis is detected in the developing heart. He: heart; Flb: forelimb bud. Panel TUNEL: analysis of serial section by TUNEL staining confirmed that apoptosis in not increased in the mutant heart at E9.5 (TUNEL positive cells fluoresce green). Panel KI67: the majority of all cells are KI67 positive (red fluorescence), which indicates that there is no major effect on cell proliferation in mutant hearts at E9.5. Right-most panel TUNEL: Only by E10.5, the apoptosis is significantly increased in mutant hearts in comparison to wild-type controls. Representative images are shown for all samples analyzed (n=3). avc: atrioventricular canal; lv: left ventricle; la: left atria; oft: outflow tract.

Figure S2

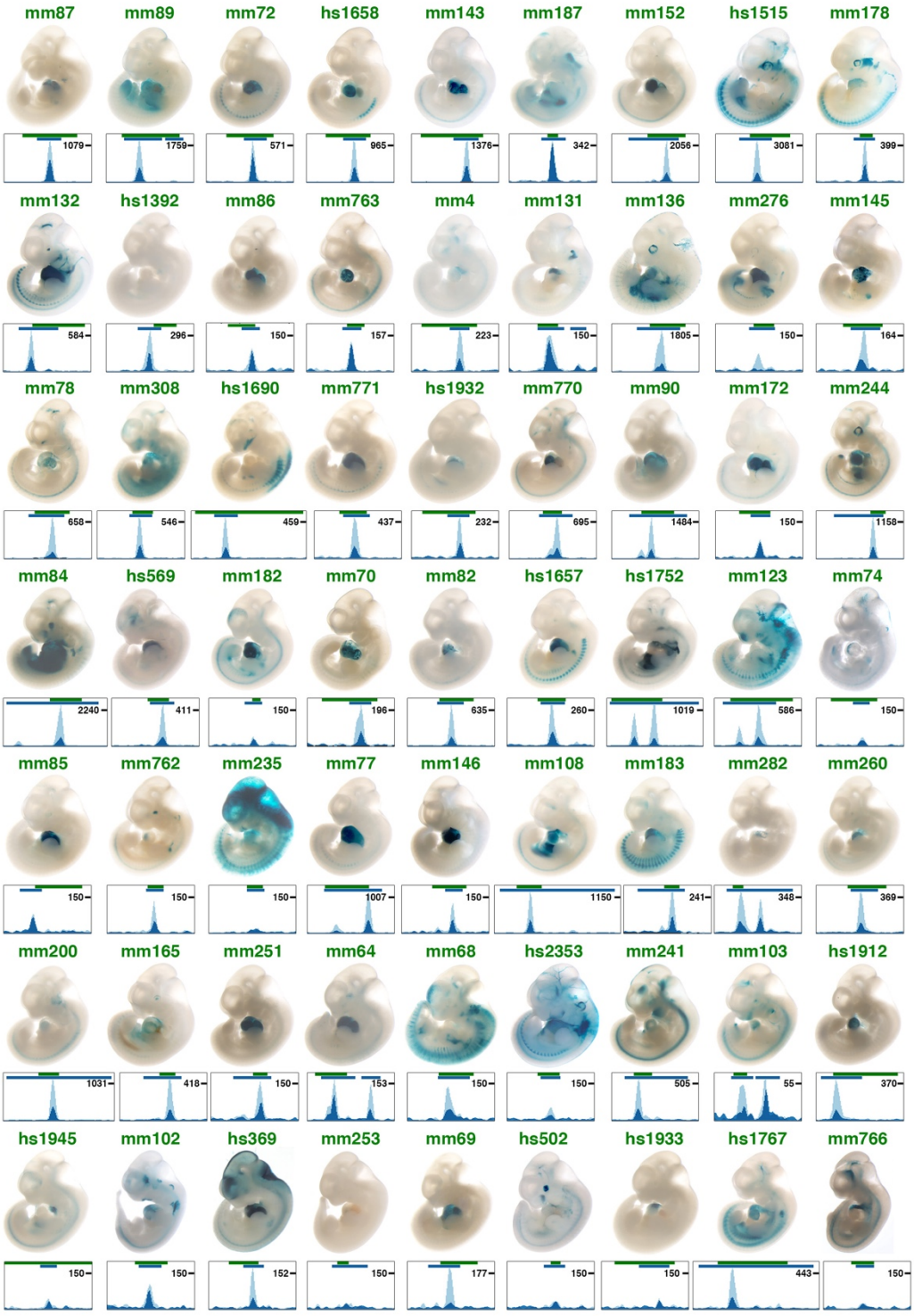


Figure S2 (related to Figure 1). Activities of the VISTA cardiac enhancers that overlap genomic regions enriched by HAND2 ChIP-Seq (E10.5).
Representative transgenic founder embryos from the public VISTA enhancer database collection (https://enhancer.lbl.gov; Visel et al., 2007) are shown. The transgenic embryos were not generated as part of this study, but images from the database collection were used for the purpose of this analysis. For each VISTA enhancer, the HAND2 ChIP-Seq peak identified by MACS analysis is indicated by a blue bar. The genomic regions used for *LacZ* reporter analysis

are indicated by a green bar. mm: mouse element; hs: human element. Nomenclature used is according to the VISTA database.

Figure S3

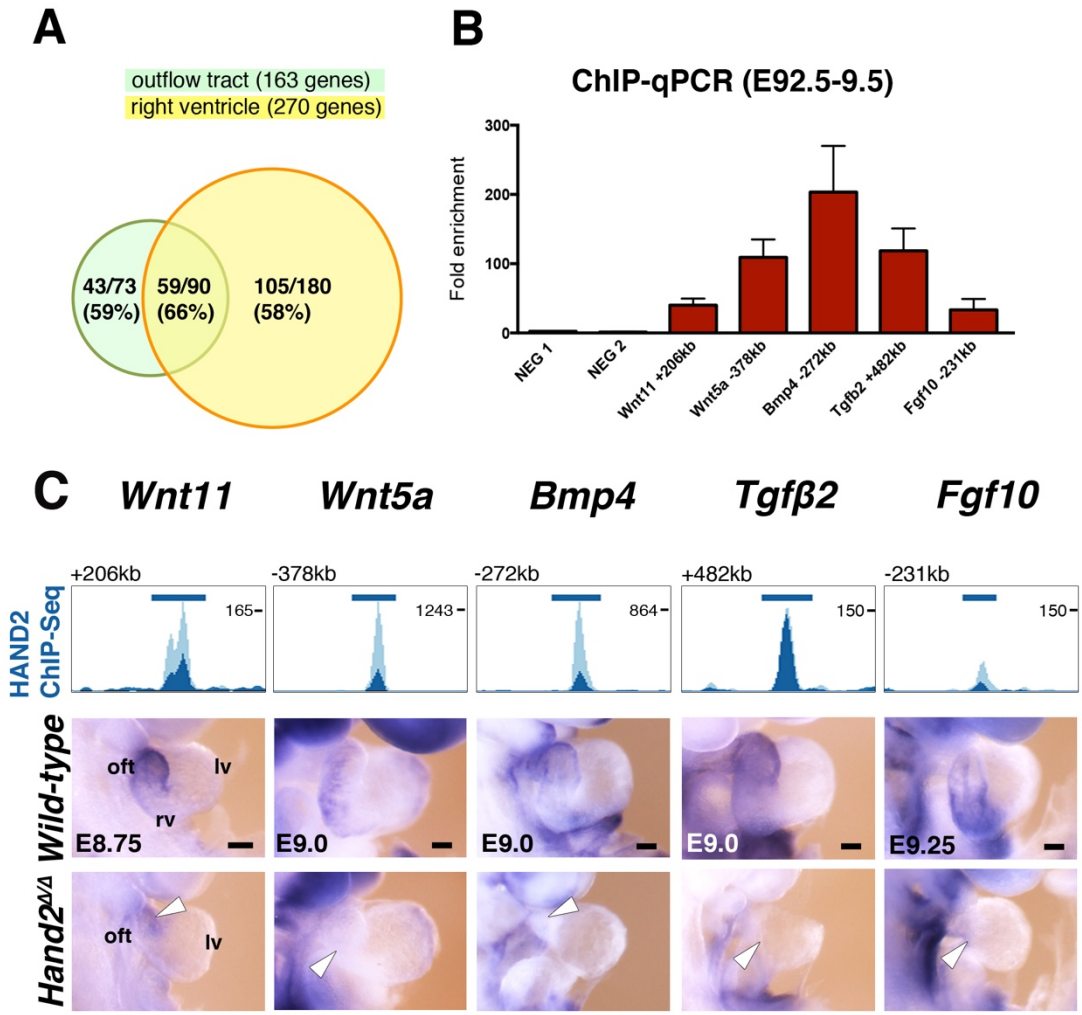


Figure S3 (related to Figure 1). HAND2 target genes encoding ligands for signaling pathways that function in OFT and/or right ventricle morphogenesis.
(A) Venn diagram shows the intersection of genes associated with the following mouse phenotype and GO terms, respectively: MP:0006126: abnormal outflow tract development; MP:0003920: abnormal heart right ventricle morphology; GO:0003151: outflow tract morphogenesis; GO:0003205: cardiac chamber development. Numbers and percentages indicate how many of the genomic landscapes associated to the terms encode regions enriched in HAND2 chromatin complexes. (B) ChIP-qPCR validation of HAND2 target regions associated to genes encoding ligands in embryonic hearts at E9.25 (n=2; mean ± SD). (C) Comparative WISH analysis of HAND2 target genes encoding signaling ligands in wild-type and Hand2-deficient mouse embryos. Graphs show the highest enriched HAND2 ChIP-Seq peaks associated with the genes analyzed. White arrowheads: reduction/loss of expression in Hand2-deficient embryos. oft: outflow tract, rv: right ventricle, lv: left ventricle. Scale bar: 100μm.

Figure S4

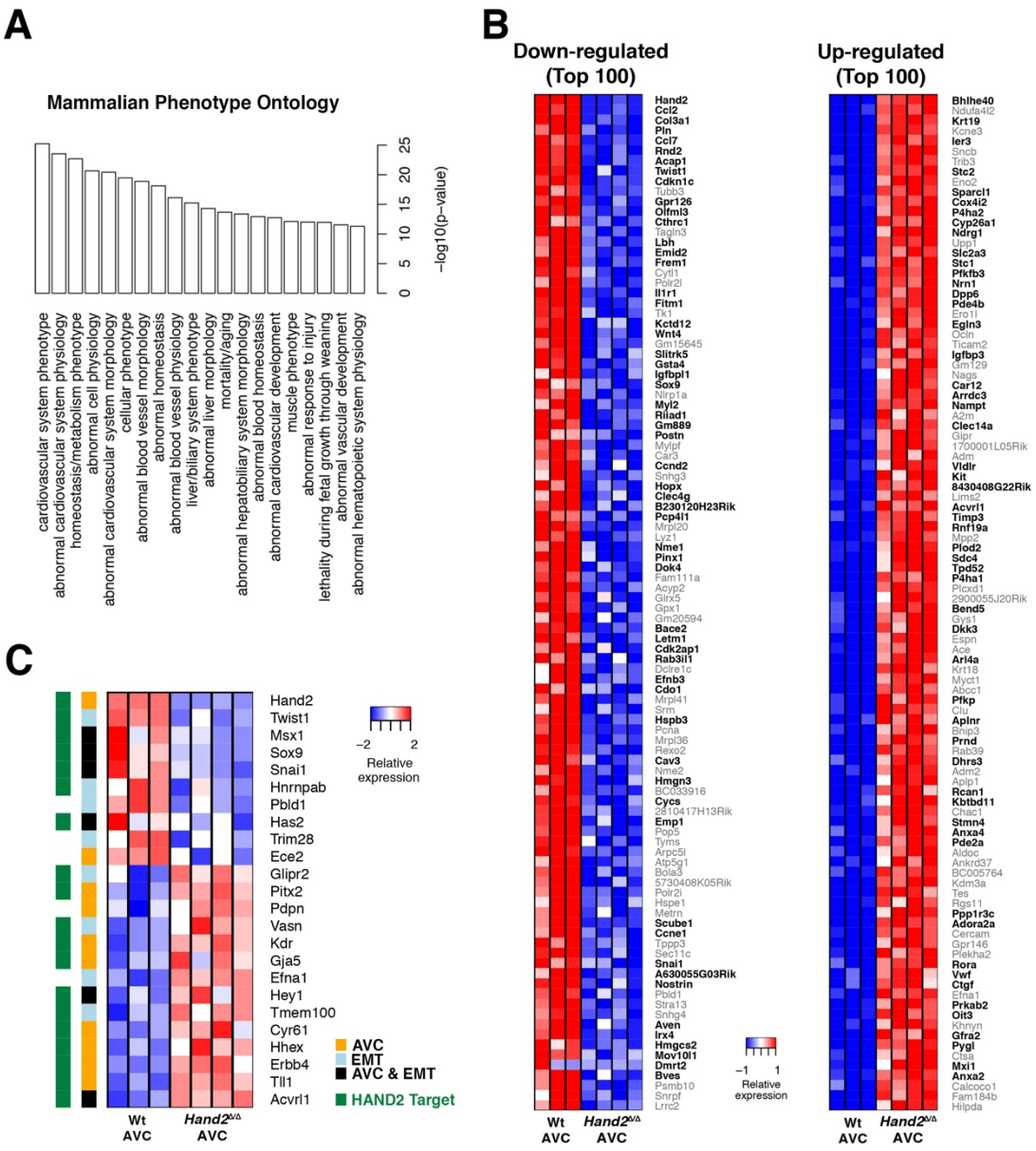


Figure S4 (related to Figure 4). Transcriptome analysis identifies the HAND2 target genes with significantly altered expression in *Hand2*-deficient AVCs.

(A) Enrichment analysis for mammalian phenotypes including all 1051 DEGs identified by comparative transcriptome analysis of *Hand2*$^{\Delta/\Delta}$ and wild-type AVCs. (B) Top 100 up- and down-regulated genes in *Hand2*-deficient AVCs in comparison to wild-type controls. Genes with regions enriched in HAND2 chromatin complexes within their TADs are indicated in bold black, others in grey. (C) Heat map of up- and down-regulated genes in *Hand2*-deficient AVCs annotated using the following gene ontology categories: MP:0000297 (abnormal AV cushion morphology + child terms) BP:0001837 (epithelial-mesenchymal transition) and BP:0010717 (regulation of EMT). Most of the DEGs in these categories are HAND2 target genes (indicated in green).

Figure S5

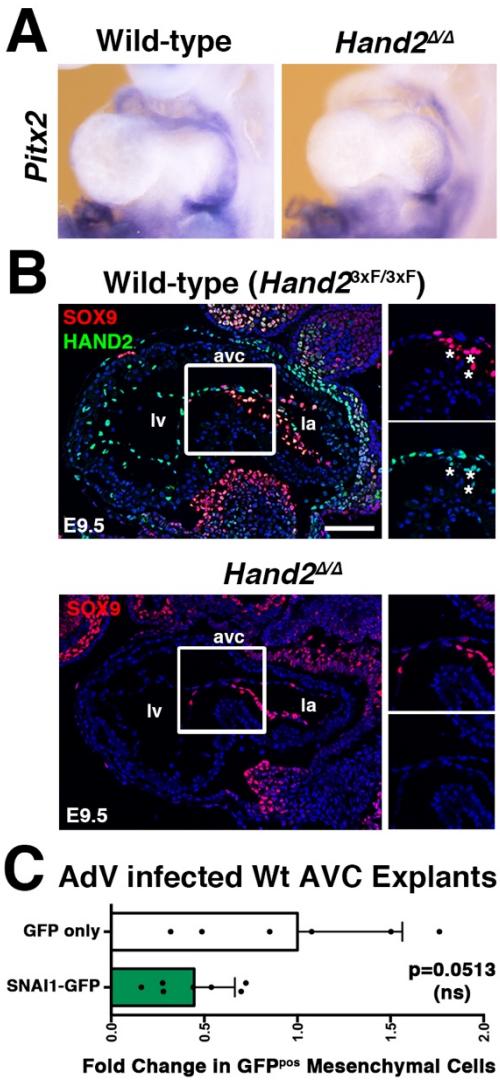


Figure S5 (related to Figures 5 and 6). Analysis of HAND2 target genes in the developing AVC.

(A) WISH analysis of the HAND2 target gene *Pitx2*, whose transcript levels are significantly altered in mutant AVCs by RNA-Seq analysis. No changes in the spatial distribution of *Pitx2* transcripts are detected. (B) Colocalization of HAND2^{3xF} proteins (green fluorescence) with the SOX9 transcriptional regulators (red fluorescence) in the AVC of wild-type (*Hand2*$^{3xF/3xF}$) and *Hand2*-deficient (*Hand2*$^{\Delta/\Delta}$) mouse embryos at E9.5. Asterisks in the enlargement (upper panels) point to SOX9-positive delaminating mesenchymal cells in the AVC, which are lacking in the *Hand2*-deficient AVC. Scale bar: 100μm. (C) Infection of wild-type AVC explants with GFP and SNAI1-GFP adenovirus (using 6x10^6 PFU for either virus per sample) indicates that GFP virus infects AVC cells slightly more efficiently than SNAI1-GFP virus. Therefore, the observed partial restoration of cell migration in *Hand2*-deficient AVCs infected with SNAI1-GFP virus is likely slightly underestimated (Figure 6B).

Figure S6

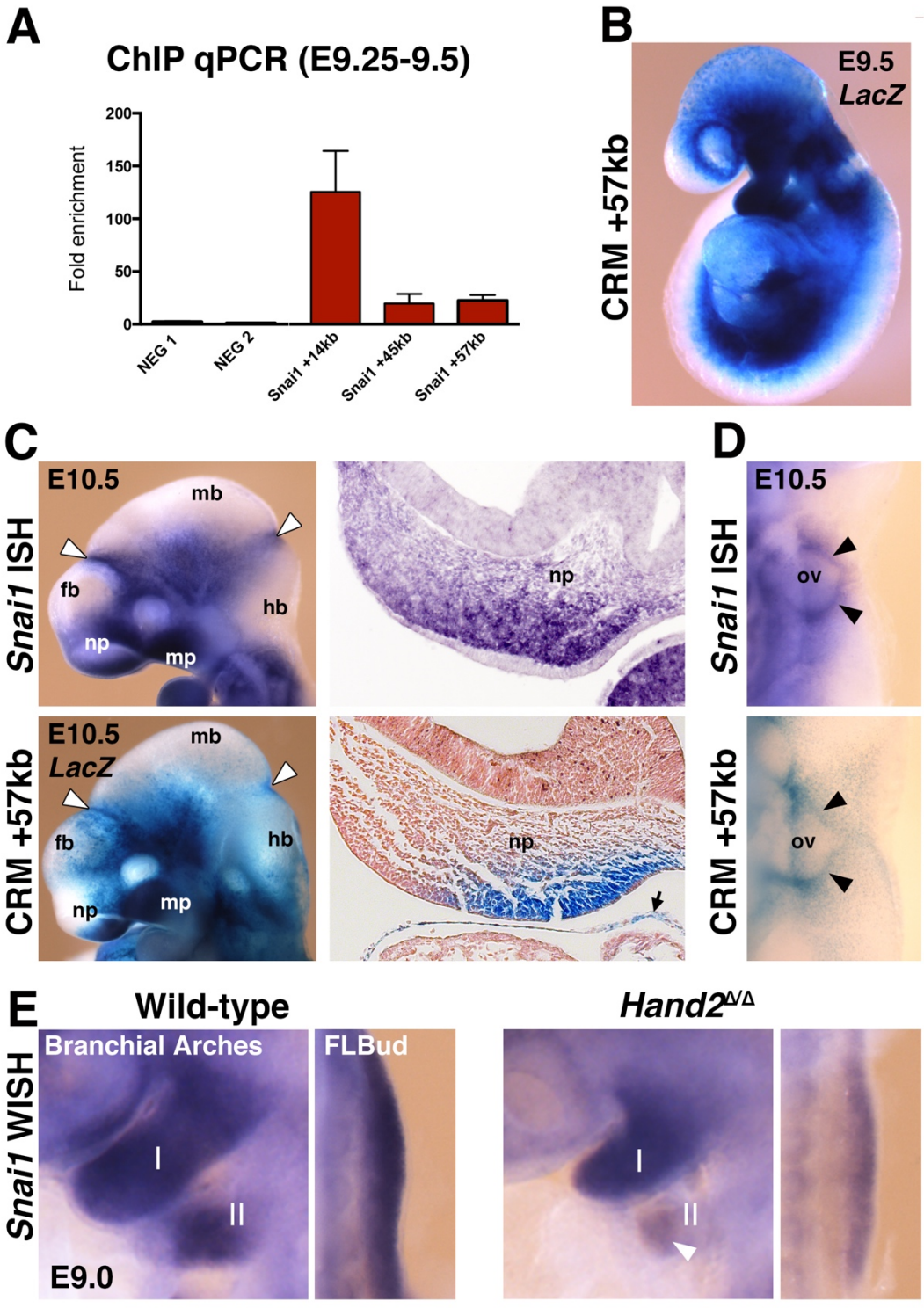


Figure S6 (related to Figure 6). The *Snai1* transcript distribution overlaps the CRM+57kb activity in craniofacial structures, branchial arches and early limb buds.
(A) ChIP-qPCR validation of the enrichment of the three CRMs in HAND2 chromatin complexes from embryonic hearts (E9.25-E9.5, n=2; mean ±SD). (B) The expression pattern of the *Snai1* CRM+57kb *LacZ* reporter transgene at E9.5. (C) *Snai1* expression and activity of the CRM+57kb *LacZ* reporter transgene in craniofacial structures. The enhancer activity overlaps well with the domain of *Snai1* transcripts in the nasal prominence (np), maxillary process (mp), fore-midbrain

and mid-hindbrain boundaries (white arrowheads). fb: forebrain, mb: midbrain, hb: hindbrain. Black arrow points to the epicardium. (D) Expression of *Snai1* and the CRM+57kb LacZ reporter transgene in migrating cardiac neural crest cells (black arrowheads) enveloping the otic vesicle (ov). (E) *Snai1* expression is reduced in the 2nd branchial arch (II) and early forelimb buds in *Hand2*-deficient mouse embryos (E9.0)

6.2. Identification of the SMAD4 targets during mouse limb bud development

As the BMP signalling pathway is crucial for early limb bud development, I have studied the impact of BMP signalling on gene expression by analysing the functions of the downstream nuclear mediator SMAD4. Recent advances in sequencing-based methodologies have made it easier to identify the genome-wide profiles underlying *cis*-regulatory networks (Pavesi, 2016). To decipher the roles of SMAD4 during limb bud development, I have used different genome-wide approaches between two embryonic stages: before the establishment of the SHH/GREM1/FGF feedback loop in E9.5-E10.0 forelimb buds and during the propagation the feedback signalling pathway in E10.5 forelimb buds. The analysis includes chromatin immunoprecipitation in combination with deep sequencing (ChIP-seq) to identify SMAD4 complexes interacting regions, ATAC-seq to map accessible chromatin and RNA-seq to describe the differentially expressed genes between WT and *Smad4*-deficent forelimb buds. This allowed us to identify the direct transcriptional targets of SMAD4 and study the SMAD4-dependent processes during early limb bud development.

6.2.1. *Smad4*^{3xFLAG} mouse generation and validation

The genome-wide profile of genomic regions bound by SMAD4 chromatin complexes by ChIP-seq required insertion of a 3xFLAG epitope tag into the endogenous SMAD4 coding region by homologous recombination in mouse embryonic stem (ES) cells. Positive ES cell clones were injected into blastocysts at the Center for Transgenic Models, Basel. To avoid unspecific effects of co-inserted Neo cassette on *Smad4* expression (Lewandoski, 2001), the floxed Neo cassette was deleted by mating chimeric males with *CMV*-CRE females (Figure 8A). Specific primer pairs were used to discriminate between *Smad4*$^{+}$, *Smad4*^{3xF} and *Smad4*$^{3xF\text{-}\Delta Neo}$ alleles (Figure 8B). The Western blot analysis shows specific detection of the Flag-tagged-SMAD4 protein in forelimb buds using FLAG antibodies (Figure 8C). Similarly, RT-qPCR analysis shows that transcription at the *Smad4* locus is unaffected in *Smad4*$^{+/+}$, *Smad4*$^{3xF/+}$ and *Smad4*$^{3xF/3xF}$ mouse limb buds (Figure 8D, left panel). As expected, the junction of exon12 and 3'UTR is fully detected in *Smad4*$^{+/+}$, partially (50%) in *Smad4*$^{3xF/+}$ and not in *Smad4*$^{3xF/3xF}$ limb buds (Figure 8D, middle panel). The junction of the FLAG epitope tag and the 3'UTR is detected specifically in *Smad4*$^{3xF/3xF}$ (100%) and *Smad4*$^{3xF/+}$ (50%) limb buds (Figure 8D, right panel). This analysis shows that the *Smad4* locus has been correctly targeted. Furthermore, we performed fluorescent immunostaining using the mouse anti-Flag antibody on frozen sections of *Smad4*$^{+/+}$ and *Smad4*$^{3xF/3xF}$ forelimb buds. A highly specific staining is observed in *Smad4*$^{3xF/3xF}$ limb buds in comparison to wild-type controls (Figure 8E, left panels). Together, this analysis demonstrates the specificity of the FLAG antibody and usefulness of the *Smad4*^{3xF} allele to study SMAD4-dependent processes. SMAD4 is uniformly expressed in the limb bud mesenchyme at E9.5 and E10.5 (Bénazet et al., 2012, Figure 8E, left panels). Our Immunohistochemical analysis shows that SMAD4 is mostly present in the cytoplasm, and only a small fraction of the protein is detected in nucleus (Figure 8E, high magnification). *Smad4*^{3xF} homozygous mice are perfectly normal, which shows that the inserted 3xFLAG epitope tag does not alter SMAD4 function nor protein localisation.

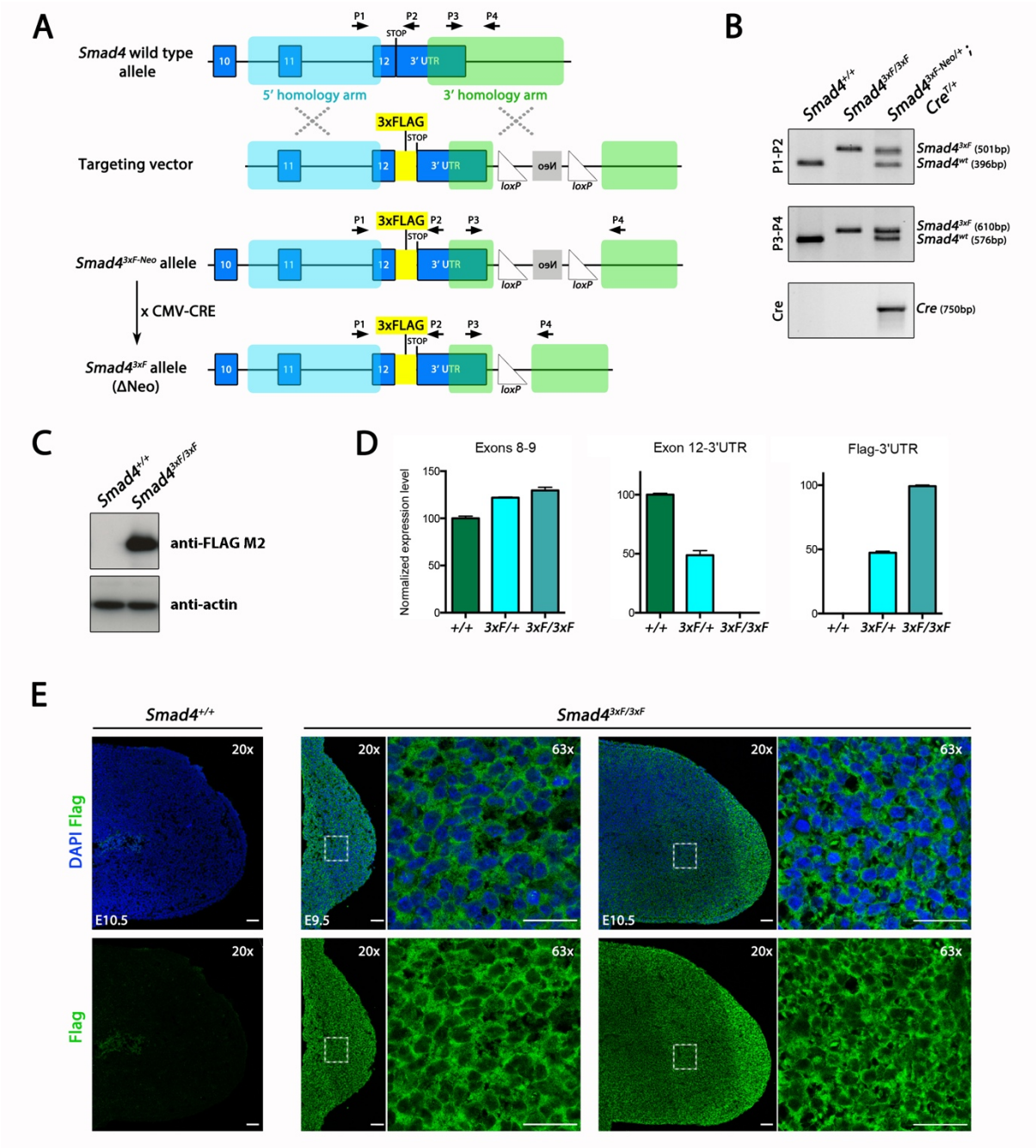


Figure 8: Generation, characterisation and validation of the *Smad4*^{3xFlag} knockin mouse.

(A) The *Smad4*^{3xF} allele was generated by homologous recombination in mouse ES cells. A vector containing two homology arms flanking the 3' end of the *Smad4* coding sequence and a 3xFLAG epitope tag was inserted in frame between the exon 12 and the 3'UTR of the *Smad4*^{wt} allele. The targeting vector also contains a floxed Neo cassette downstream of the *Smad4* sequence (targeting was done by Frédéric Laurent). Initially, the *Smad4*$^{3xF\text{-}Neo}$ mouse was crossed with *CMV-cre* mouse to remove the Neo cassette and obtain the *Smad4*^{3xF} mouse (ΔNeo). The arrows indicate the primers used for genotyping. (B) Genotyping and validation of the different *Smad4* alleles by PCR. The sizes of the PCR bands are written on the right. (C) Detection of the SMAD4^{3xF} protein by Western blot in *Smad4*$^{3xF/3xF}$ E11.75 limb buds (done by Frédéric Laurent). (D) RT-qPCR analysis of E11.75 limb buds from *Smad4*$^{+/+}$, *Smad4*$^{3xF/+}$ and *Smad4*$^{3xF/3xF}$ mouse embryos to detect

different parts of the transcript: between exon 8 and 9, exons 12 and 3'UTR and between the FLAG and 3'UTR (done by Frédéric Laurent). (E) Fluorescent immunostaining using the mouse anti-Flag antibody on frozen sections of WT and *Smad4*$^{3xF/3xF}$ limb buds. Scale bar: 100μm (50μm in high magnification). Figure adapted from Frédéric Laurent's PhD thesis.

6.2.2. SMAD4^{3xF} ChIP-seq and ATAC-seq

The genomic regions interacting with SMAD4 chromatin complexes were identified by ChIP-seq using limb buds from two embryonic stages: E9.75, characterized by essential high BMP activity and E10.5, characterized by low BMP activity (Bénazet et al., 2009, see Figure 2). Due to the small size of E9.75 forelimb buds, I have dissected these early forelimb buds along with a small part of trunk tissue from ~160 *Smad4*$^{3xF/3xF}$ mouse embryos in two replicates (Figure 9A). Furthermore, the profile of accessible/open chromatin was determined by ATAC-seq at E9.75 in wild-type (WT) forelimb buds along with a small part of adjacent flank tissue to be able to directly compare the SMAD4 ChIP-seq and ATAC-seq profiles (Figure 9A). Our bioinformatics analysis shows that 40% of the regions enriched in SMAD4 chromatin complexes are located close to transcriptional start sites (TSS, ±5kb) while 20% are located further away (±100kb) from the nearest TSS. Phastcon analysis shows that SMAD4 bound regions are highly conserved in placental mammals (Figures 9C). HOMER analysis using known motifs identified the SMAD4 consensus as the most enriched motif (Figure 9D, Kim et al., 2011). *De novo* motif search discovery also established SMAD motifs as most enriched (Figure 9D, Kim et al., 2011). To further validate the SMAD4^{3xF} ChIP-seq dataset, we performed the ChIP-qPCR using two known BMP responsive elements (BREs). The BRE near the *Id1* and *Msx2* genes show significant enrichment at the target sites and similar enrichment is observed in the SMAD4 ChIP-seq dataset (Figure 9F).

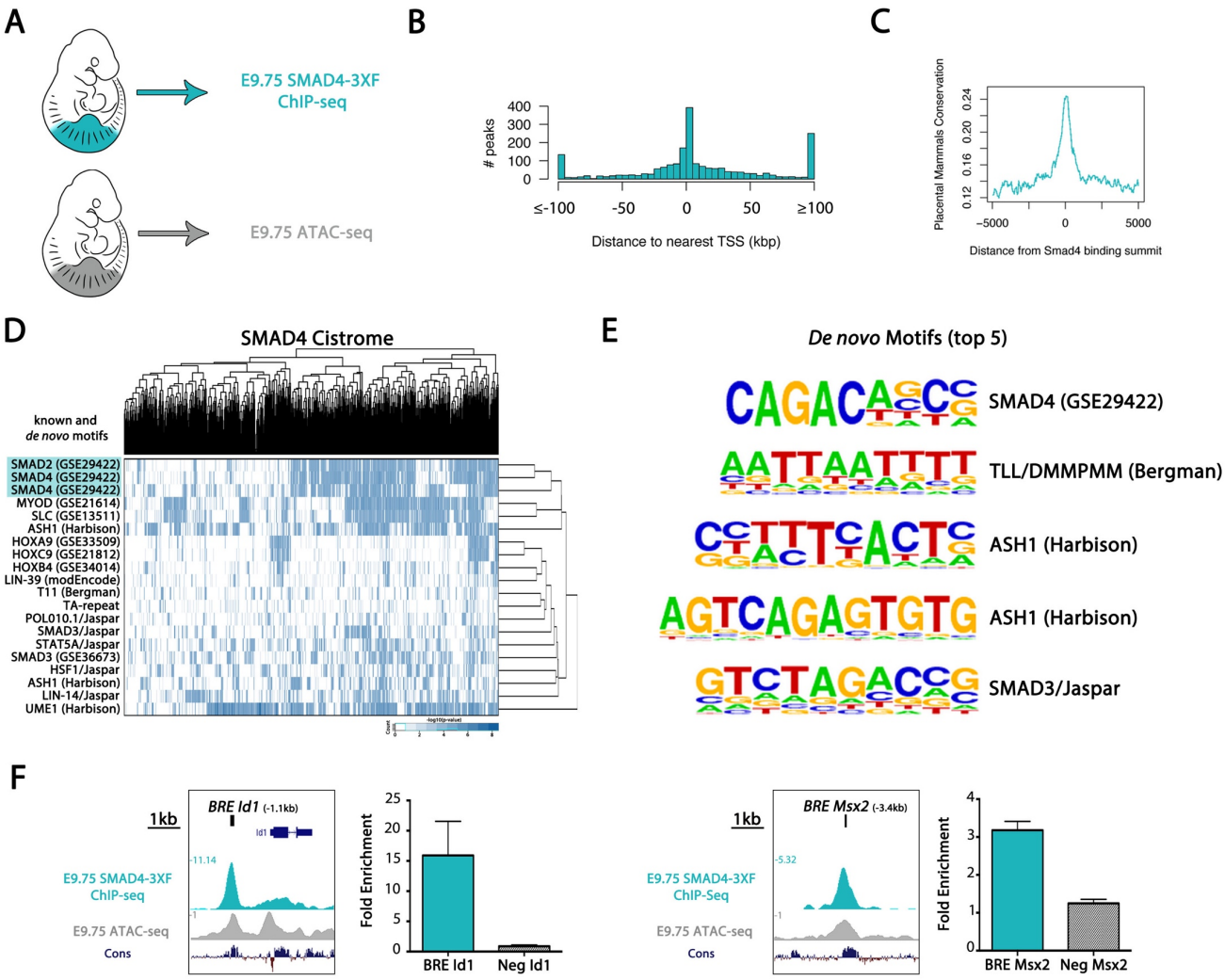


Figure 9: SMAD4^{3xF} ChIP-seq and ATAC-seq experiments using E9.75 forelimbs.
(A) E9.75 forelimbs were dissected with a portion of trunk attached to perform ChIP-seq and ATAC-seq experiments. (B) Barplot showing the distribution of SMAD4 occupied regions to the nearest TSS. (C) Histogram showing phastcon conservation of SMAD4 bound regions. SMAD4 occupied regions are conserved. (D) Hierarchical clustering of the high-affinity matches for the enriched known and *de novo* SMAD4 motifs. (E) *De novo* motif analysis showing the top 5 motifs in SMAD4 bound regions. (F) ChIP-qPCR validation for two known SMAD4 binding regions: the BMP responsive element (BRE) on the *Id1* transcription unit and the BRE of the *Msx2* gene (Korchynskyi and Dijke, 2002, Brugger et al., 2004).

At E10.5, when BMP activity is lower, fore and hindlimb buds from ~200 *Smad4*$^{3xF/3xF}$ embryos were used for ChIP-seq (in duplicate) and ATAC-seq was performed using E10.5 WT forelimb buds (ATAC-seq dataset from Javier Lopez-Rios, Figure 10A). Similar to the E9.5 ChIP-seq dataset, the regions enriched in SMAD4 chromatin complexes are located ±5kb or ≥±100kb away

from the closest TSS and are conserved in placental mammals (Figure 10B and 10C). Surprisingly, *de novo* and known motifs analysis using HOMER (Heinz et al., 2010) identified the homeobox motif PKNOX1 as the top enriched *de novo* motif (Figure 10D, E). Similar to the SMAD4 ChIP-seq at E9.5, the BREs near the *Id1* and *Msx2* target genes are also enriched at E10.5, which validates both datasets (Figure 10F).

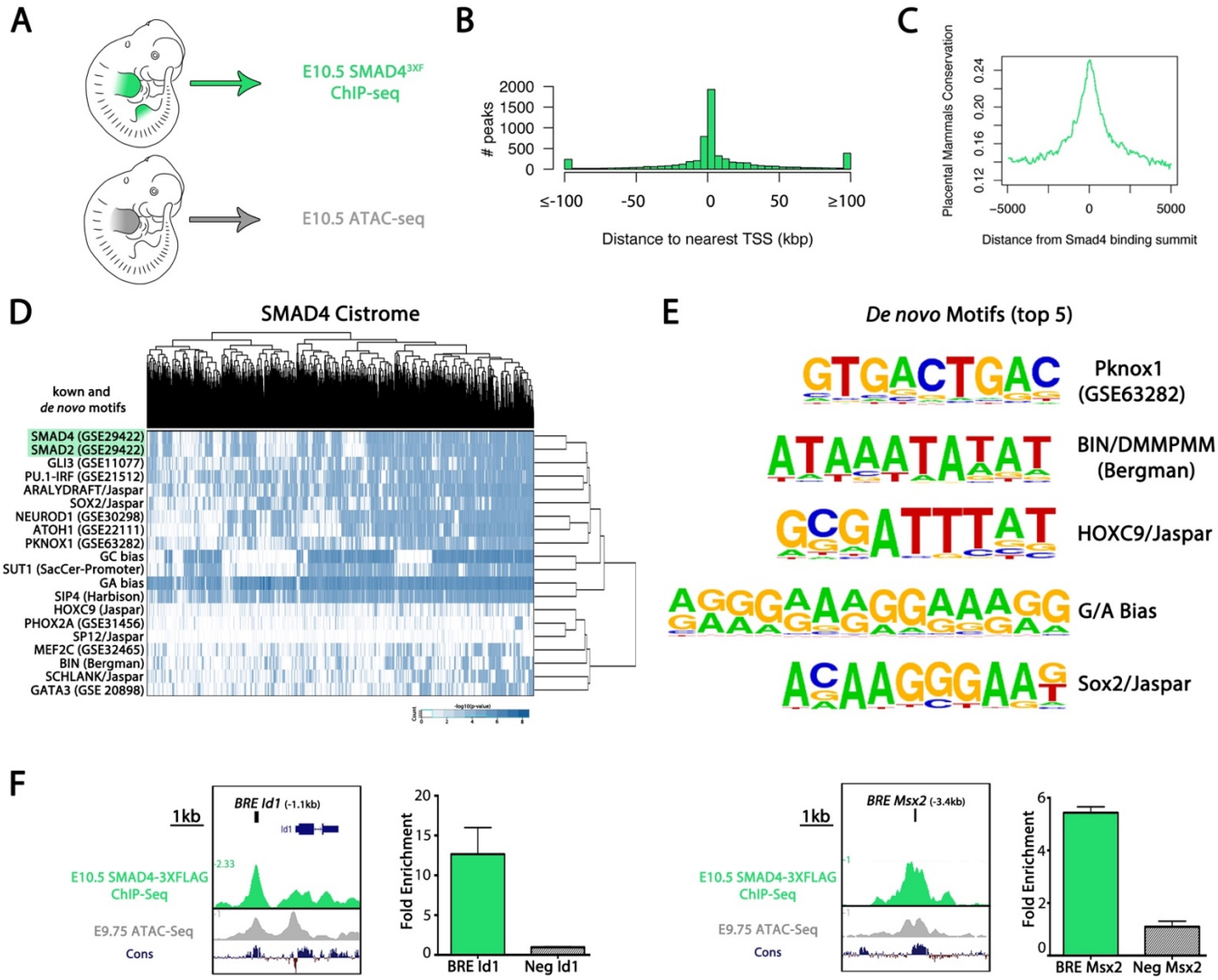


Figure 10: SMAD4^{3xF} ChIP-seq and ATAC-seq experiments using E10.5 limb buds (forelimbs and hindlimbs).
(A) E10.5 limb buds were dissected and pooled to perform ChIP-seq and ATAC-seq experiments. (B) Barplot showing the distribution of SMAD4 occupied regions to the nearest TSS. (C) Histogram showing phastcon conservation of SMAD4 bound regions. SMAD4 occupied regions are conserved. (D) Hierarchical clustering of the high-affinity matches for the enriched known and *de novo* SMAD4 motifs. (E) Results from the *de novo*

motif analysis showing the top 5 motifs in SMAD4 bound regions. (F) ChIP-qPCR validation for two known SMAD4 binding regions: the BMP responsive element (BRE) on *Id1* transcript unit and the BRE of the *Msx2* gene (Korchynskyi and Dijke, 2002, Brugger et al., 2004).

6.2.3. Comparative analysis of the transcriptome in wild-type and *Smad4*-deficient forelimb buds

To identify SMAD4-dependent gene regulatory networks, I performed RNA-seq analysis using WT and *Smad4*-deficient forelimb buds. It has been shown that constitutive *Smad4* inactivation is embryonically lethal (Chu et al., 2004). Therefore, I have conditionally inactivated *Smad4* (*Smad4*$^{\Delta/\Delta c}$) in the limb bud mesenchyme using the *Prx1*-CRE driver line (Logan et al., 2002). By E9.5, the *Prx1*-CRE transgene is activated in developing forelimb buds (Logan et al., 2002), and I estimated that CRE -mediated deletion of the *Smad4*$^{flox/flox}$ allele requires about 10hrs (see Figure 4B).

Therefore, the earliest stage used for RNA-seq analysis are WT and *Smad4*$^{\Delta/\Delta c}$ forelimb buds isolated from mouse embryos at E10.0, i.e at the time the feedback signalling system is set-up and mesenchymal BMP4 expression is high and essential for limb bud development (Bénazet et al., 2009). The RNA-seq analysis shows that 669 genes are differentially expressed (DEGs) between WT and *Smad4*$^{\Delta/\Delta c}$ E10.0 forelimb buds (fold change ≥1.2; FDR set to <0.1). Out of 667 DEGs, 306 are down- and 361 are up-regulated genes (Figure 11B). The top 25 DEGs are shown in Figure 11C. GO analysis identifies sterol biosynthesis as the top enriched term among the down-regulated DEGs. In contrast, the most enriched terms in the up-regulated DEGs are related to developmental and cellular processes (Figure 11D).

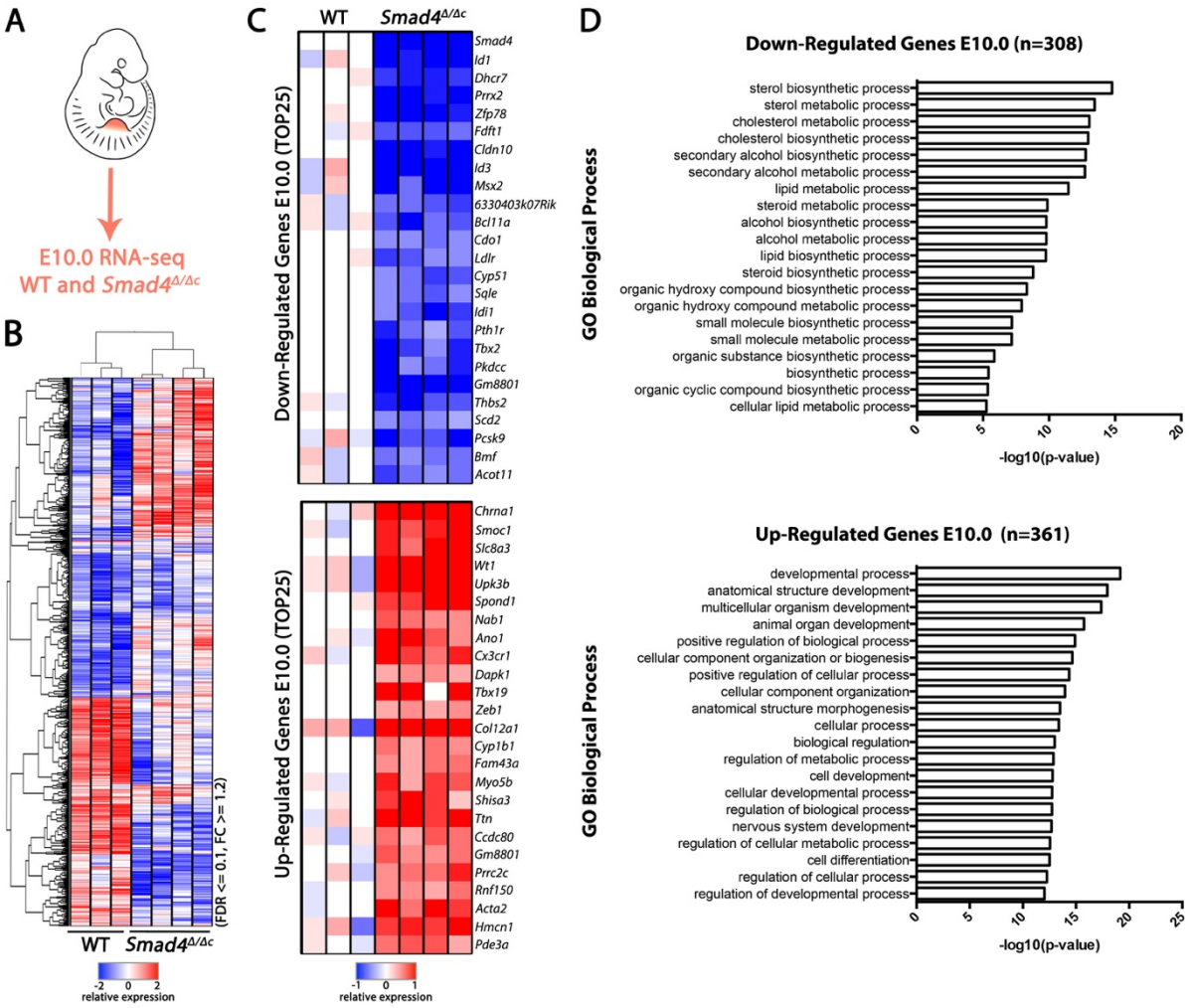


Figure 11: Transcriptome analysis in WT and *Smad4*$^{\Delta/\Delta c}$ forelimbs at E10.0
(A) Forelimbs of WT and *Smad4*$^{\Delta/\Delta c}$ E10.0 embryos were used to perform RNA-seq experiment. (B) Heatmap showing the differentially expressed genes (DEGs) identified by comparing WT and *Smad4*$^{\Delta/\Delta c}$ transcriptomes. N=3 and n=4 biological replicates for WT and *Smad4*$^{\Delta/\Delta c}$ forelimbs, respectively. DEGs showing a fold change ≥1.2 and FDR<0.1 were analysed. (C) Top 25 down-regulated (top panel) and up-regulated (bottom panel) genes in *Smad4*$^{\Delta/\Delta c}$ forelimbs, normalized to the expression mean in WT samples. (D) GO analysis of down-regulated and up-regulated genes in *Smad4*$^{\Delta/\Delta c}$ forelimbs. Note that sterol biosynthetic processes are most enriched among the down-regulated genes, while GO terms for general developmental processes are more enriched in up-regulated genes (bottom panel).

Similarly, the DEGs in WT and *Smad4*$^{\Delta/\Delta c}$ forelimbs were also identified using forelimb buds at E10.5 (Figure 12A), i.e. when the self-regulatory feedback signalling system is active and mesenchymal BMP activity is lowered due to GREM1-mediated antagonism (Bénazet et al., 2009, 2012). Using the same statistical threshold as for E10.0, 386 DEGs were identified, 143 of which are

down- and 243 up-regulated (Figure 12B). The top 25 DEGs for each category are shown in Figure 12C. In contrast to E10.0, GO analysis at E10.5 no longer detected sterol biosynthesis, but terms related to developmental processes (Figure 12D). This indicates that SMAD4-mediated BMP signal transduction controls genes functioning in developmental process at both stages, while its effect on cholesterol biosynthesis may be limited to the onset of limb bud development (see chapter 6.2.7).

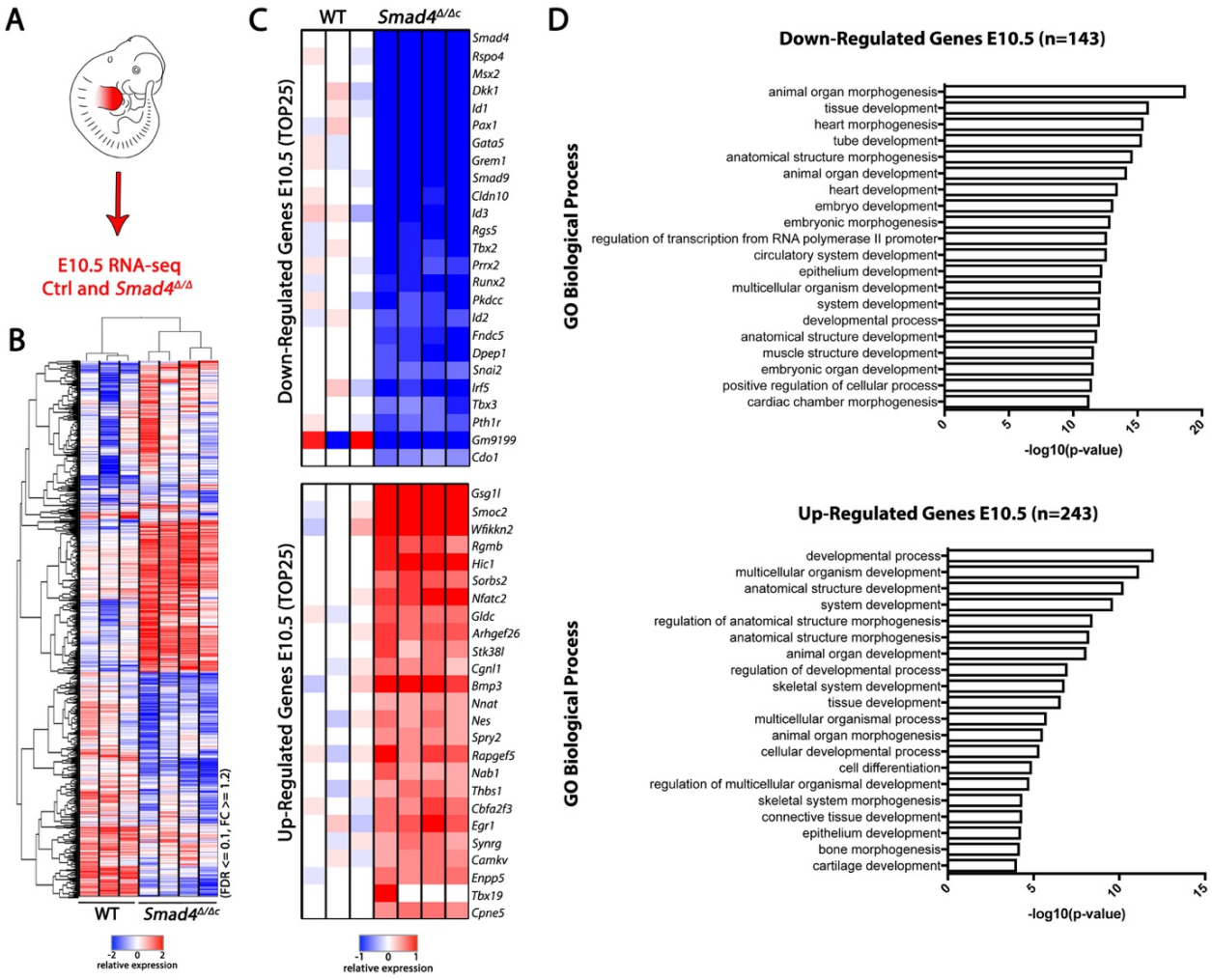


Figure 12: Transcriptome analysis in WT and *Smad4*$^{\Delta/\Delta c}$ forelimbs at E10.5
(A) Forelimbs of WT and *Smad4*$^{\Delta/\Delta c}$ E10.5 embryos were used for RNA-seq analysis. (B) Heatmap showing the differentially expressed genes (DEGs) identified by comparing WT and *Smad4*$^{\Delta/\Delta c}$ transcriptomes. N=3 and n=4 biological replicates for WT and *Smad4*$^{\Delta/\Delta c}$ forelimbs, respectively. DEGs showing a fold change ≥1.2 and FDR<0.1 were analysed. (C) Top 25 down-regulated (top panel) and up-regulated (bottom panel) genes in *Smad4*$^{\Delta/\Delta c}$ forelimbs, normalized to the expression mean in WT samples. (D) GO analysis of down-regulated and up-regulated genes in *Smad4*$^{\Delta/\Delta c}$ forelimbs. Note that global development

processes are the most enriched in both datasets. Sterol biosynthesis is no longer present in the top terms.

6.2.4. Identification of the SMAD4 target gene regulatory networks: Intersection of the ChIP-seq, ATAC-seq and RNA-seq datasets

To identify the direct transcriptional targets of SMAD4 complexes, the SMAD4^{3xF} ChIP-seq, ATAC-seq and RNA-seq datasets were combined. First, we identified the genomic regions bound by SMAD4 that overlap with ATAC-seq peaks. The regions of open chromatin that interact with SMAD4 complexes were considered as potential *cis*-regulatory modules (CRMs). Next, these CRMs were associated with DEGs located within a distance up to ±1Mb. This parameter was chosen as it has been established for several developmental regulator genes that CRMs can be located as far away as ±1Mb. For example, the CRM that regulates *Shh* expression in the posterior limb bud mesenchyme is located at around 800kb upstream of the *Shh* TSS (Lettice et al., 2003). By intersecting these three datasets, we have identified 306 candidate transcriptional targets of SMAD4 in forelimb buds at E10.0 (Figure 13A). Next, the association of SMAD4 bound CRMs and DEGs within a particular topologically associating domain (TADs, Dixon et al., 2012, appendixes 12.7 and 12.8) was analysed in comparison to genes, whose expression is not changed in *Smad4*-deficient limb buds. These results show a significant larger number of SMAD4 interacting CRMs located within the TADs of DEGs than genes whose expression is not altered in *Smad4*-deficient limb buds genes (Figure 13B). Among the 306 SMAD4 transcriptional target identified 143 are down- (Appendix 12.1) and 163 up-regulated (Appendix 12.2) in *Smad4*$^{\Delta/\Delta c}$ forelimb buds at E10.0. The GO analysis of the 143 down-regulated SMAD4 target genes is enriched for functions in sterol metabolism (Figure 13C, right bar plot), while the up-regulated SMAD4 target genes function predominantly in developmental processes (Figure 13C, left bar plot).

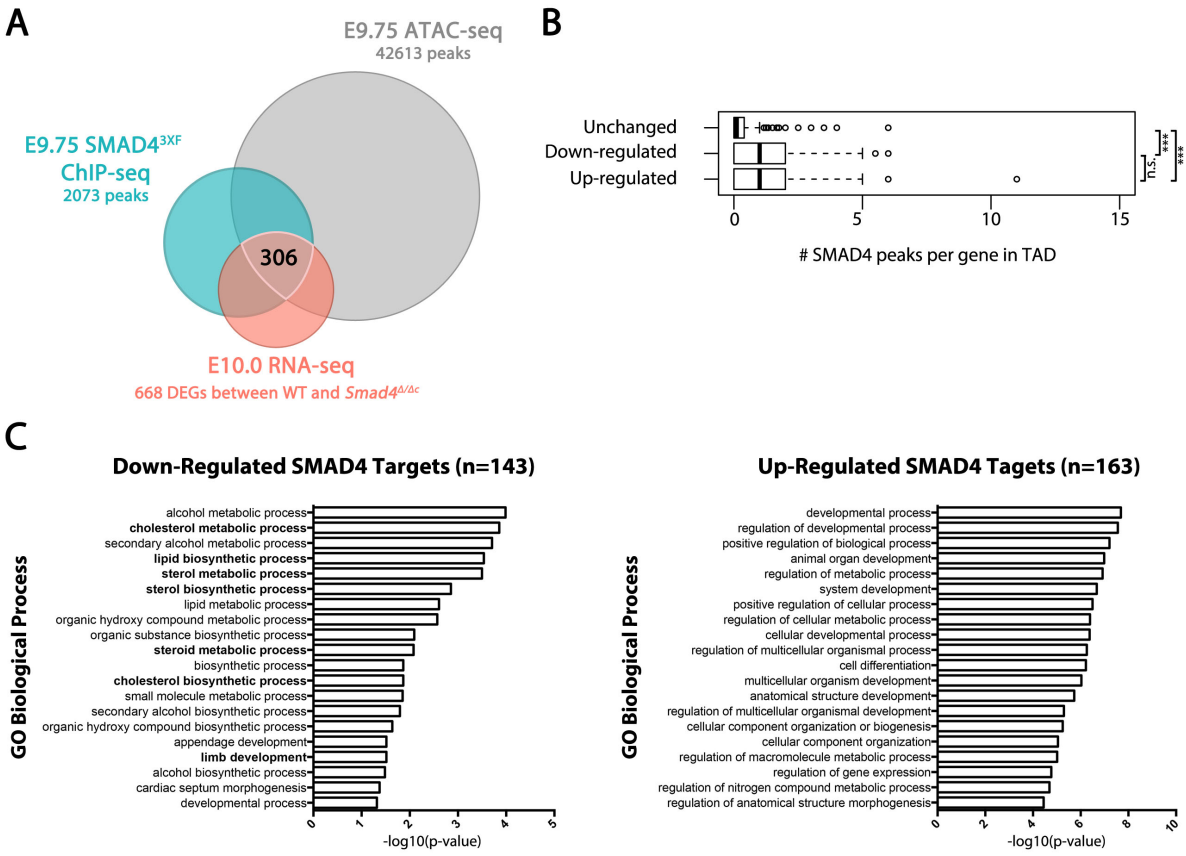


Figure 13: Identification of the putative SMAD4 targets in forelimb buds at E9.75-E10.0

(A) Schematic representation of the intersection between the ChIP-seq (E9.75), ATAC-seq (E9.75) and RNA-seq (E10.0) dataset. The total numbers of ChIP peaks, ATAC peaks and DEGs are indicated. This intersection identifies 306 candidate SMAD4 target genes in limb buds at E9.75-E10.0. (B) Boxplot representing the number of SMAD4^{3xF} ChIP-seq peaks within a topologically associated domain (TAD) harbouring genes either down- or up-regulated in *Smad4*$^{\Delta/\Delta c}$ limb buds. The differentially expressed genes are the ones located nearest to the SMAD4 interacting region coinciding with accessible chromatin as identified by ATAC-seq. (C) GO enrichment analysis for biological processes for the down- and up-regulated candidate SMAD4 target genes.

The same approach was used for the E10.5 datasets, which identified 188 candidate SMAD4 targets genes (Figure 14A). The significantly lower number of target genes identified could reflect the much lowered BMP activity. The SMAD4 ChIP-seq peaks are again more significantly abundant in TADs of DEGs than genes with unaltered expression (Figure 14B). In *Smad4*$^{\Delta/\Delta c}$ forelimb buds at E10.5, 125 SMAD4 target genes are down- (Appendix 12.3) and 63 up-

regulated (Appendix 12.4) and in both groups genes functioning in developmental and cellular processes are enriched (Figure 14C).

Taken together, this initial analysis shows that SMAD4 directly regulates the expression of genes essential for developmental process in both limb bud stages, while it appears to also regulate the expression of genes in the cholesterol biosynthesis pathway during the onset of limb bud development.

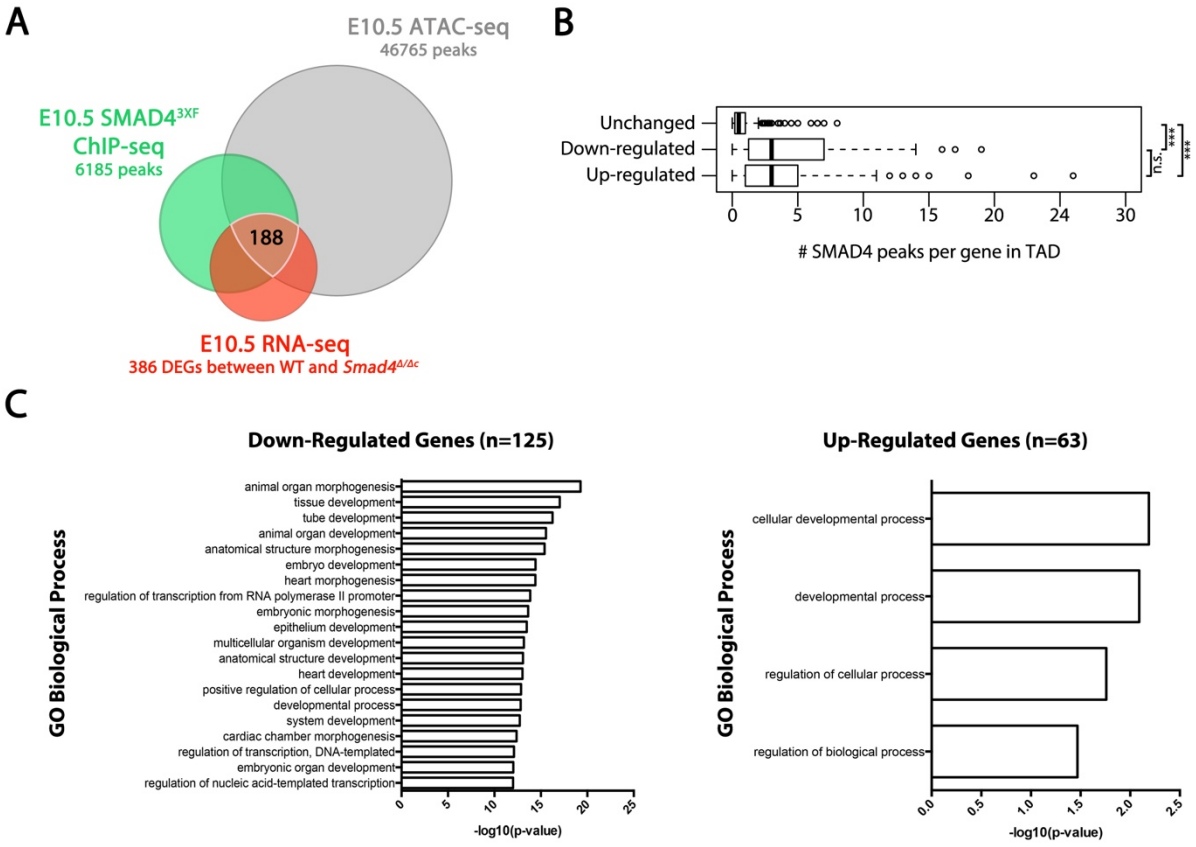


Figure 14: Identification of the putative SMAD4 targets in forelimb buds at E10.5
(A) Schematic representation of the intersection between the ChIP-seq (E10.5), ATAC-seq (E10.5) and RNA-seq (E10.5) dataset. The total numbers of ChIP peaks, ATAC peaks and DEGs indicated. This intersection identifies 188 candidate SMAD4 target genes in limb buds at E10.5. (B) Boxplot representing the number of SMAD4^{3xF} ChIP-seq peaks within a topologically associated domain (TAD) harbouring genes either down- or up-regulated in *Smad4*$^{\Delta/\Delta c}$ limb buds. The differentially expressed genes are the ones located nearest to the SMAD4 interacting region coinciding with accessible chromatin as identified by ATAC-seq. (C) GO enrichment analysis for biological processes for the down- and up-regulated candidate SMAD4 target genes at E10.5.

6.2.5. Potential roles of SMAD4 in regulating *Grem1* expression dynamics in limb buds

The cis-regulatory complexity of Grem1 expression

Previous publications from our lab have shown that several CRMs located in the *Fmn1* locus are involved in regulating *Grem1* expression in developing limb buds. Based on the evolutionary conservation, three likely functionally relevant CRMs were initially identified, called GRS1 (*Grem1* Regulatory sequence 1), HMCO2 and HMCO3 (Zuniga et al., 2012). It has been shown that deletion of these elements impairs the *Grem1* expression in developing limb buds. In addition, another element GRE1 (GLI responsive element 1) has been shown to participate in the regulation of *Grem1* expression (Li et al., 2014). The GRS1, GRE1 and HMCO2 CRMs are located within the 70kb global control region (GCR, Figure 15 blue square), which is essential for normal *Grem1* expression in mouse limb buds (Zuniga et al., 2004). Nevertheless, the combined activity of these CRMs can not explain the dynamic spatio-temporal expression of *Grem1* (see Figure 4), suggesting the existence of other CRMs.

To identify all the potential *Grem1* regulatory elements, we have used published Hi-C data from embryonic ES cells (Dixon et al., 2012), in combination with 4C-seq, and different ChIP-seq and ATAC-seq datasets to map potential additional CRMs in *Fmn1-Grem1* genomic landscape (mm9 chr2:113,143,398-113,798,025). This analysis revealed a 320kb sub-TAD within the *Grem1*-specific *cis*-regulatory landscape (Figure 15, dashed triangle). Indeed, the 4C peaks identifying the interactions of distant genomic regions with the *Grem1* proximal promoter (used as a viewpoint) are contained mostly in the *Grem1* subTAD region, which corroborates the *Grem1* specific *cis*-regulatory potential of this 320kb region (Figure 15, 4C track in red). Evolution conservation analysis of the *Grem1* sub-TAD using the ECR browser (Ovcharenko et al., 2004) to compare mouse, human, pig and bovine genomes identified eight

novel evolutionary conserved regions (ECRs), in addition to the four already known elements (GRS1, HMCO2, GRE1 and HMCO3, Figure 15, grey bars). All these ECRs are contained within ATAC-seq peaks and overlap regions decorated by H3K27ac marks indicative of active enhancers (Infante et al., 2013). In addition, they are enriched in SMAD4^{3xF} (my results), HoxA13 (Sheth et al., 2016) and/or GLI3 chromatin complexes (collaboration with Andrew P McMahon, USC, Figure 15). The chromatin accessibility, H3K27ac and interaction with transcription factor complexes points to the potential *cis*-regulatory activity of these ECR regions. The transcriptional activity of these elements has been tested by Laurène Ramos Martins using *LacZ* reporter constructs, which establishes that many of these regions are enhancers that drive limb specific *LacZ* expression, making them candidate CRMs for regulating *Grem1* expression (data not shown).

Furthermore, the binding of CTCF and Cohesin complexes as revealed by ChIP-seq (DeMare et al., 2013) identifies 6 structural elements within the *Grem1* subTAD (Figure 15, light blue bars).

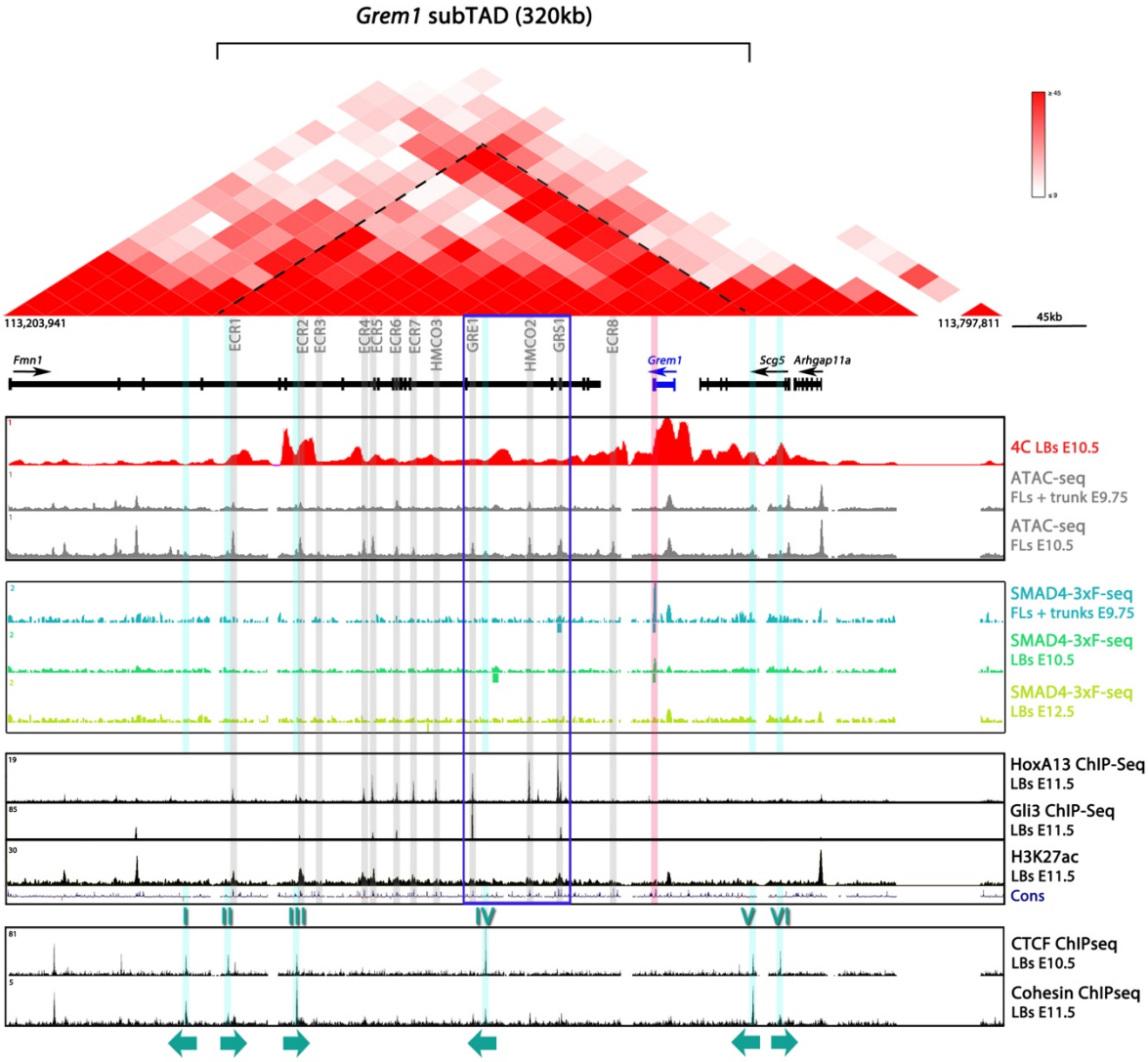


Figure 15: *Grem1* genomic landscape
UCSC genome browser view showing the Hi-C profile (top red triangle; Dixon et al., 2012), 4C-seq (with *Grem1* TSS as viewpoint), ATAC-seq, and different ChIP-seq data for the *Grem1 cis*-regulatory landscape. The sub-topologically associated subdomain (subTAD) for *Grem1* is delineated by the black dashed triangle on the Hi-C profile. Potential *cis*-regulatory modules (CRMs) interacting either with SMAD4, HoxA13 (Sheth et al., 2016) or GLI3 (collaboration with Andrew P McMahon, UCS) are highlighted by grey boxes. These are evolutionary conserved regions (ECR, from 1 to 8) and the HMCO3, GRE1, HMCO2 and GRS1 elements (Li et al., 2014; Zuniga et al., 2012). These elements are decorated by H3K27ac marks (Infante et al., 2013). The global control region (GCR) of *Grem1* is indicated by the blue square and the structural elements (I to VI) interacting with CTCF and Cohesin chromatin complexes are highlighted in light blue (DeMare et al., 2013). The arrowheads indicate the motif orientation of the CTCF regions. The *Grem1* exon 2 interacting with SMAD4 at E9.75 and E10.5 is highlighted in red.

Coding exon 2 of Grem1 is highly enriched in SMAD4 chromatin complexes

My SMAD4^{3xF} ChIP-seq analysis shows that among the *Grem1* CRMs, the GRS1 element and the *Grem1* coding exon 2 are significantly enriched (Figure 15 red bar). Alignment of SMAD4^{3xF} ChIP-seq profiles at E9.75, E10.5 and E12.5 shows that the SMAD4 complexes interactions with *Grem1* coding exon 2 are dynamic. The SMAD4 occupancy of *Grem1* coding exon 2 is highest in forelimb buds at E9.5 and decreases as limb bud development progresses (Figure 16A). As exon 2 encodes the GREM1 protein, this region is evolutionary conserved but inaccessible/closed as no corresponding ATAC-seq peak is observed (Figure 16A). The dynamics of the SMAD4-*Grem1*-exon 2 interactions were validated using ChIP-qPCR (Figure 16B). This most prominent interaction in forelimb buds at E9.5 is interesting in light of the fact that SMAD4 is required to activate *Grem1* expression during the onset of forelimb bud development (Bénazet et al., 2009, 2012). Therefore, it is possible that the interaction of SMAD4 with *Grem1* exon 2 is relevant to *Grem1* activation in limb buds. To test this hypothesis, I used whole mount *in situ* hybridisation to analyse *Grem1* and *LacZ* transcripts in *Grem1*$^{\Delta ORF/+}$ embryos, in which one copy of the *Grem1* exon2 is replaced by a *LacZ* cassette (Michos et al., 2004, Figure 16C). Interestingly, the activation of *LacZ* transcription is delayed in comparison to *Grem1* expression as *LacZ* transcripts are only detected at E10.5 instead of E9.5 (Figure 16D). Thus if coding exon 2 is replaced by *LacZ*, its expression is not activated at E9.5, which suggests that the interaction of SMAD4 with *Grem1* coding exon 2 is required for the temporally correct early activation of *Grem1* expression. The transcriptional enhancing potential of exon 2 was assessed in *LacZ* transgenic reporter assay by Laurène Ramos Martins (Figure 16E). As only 2 of 12 transgenic embryos show LacZ staining in limb buds, *Grem1* coding exon2 is unlikely to serve as a classical transcriptional enhancer.

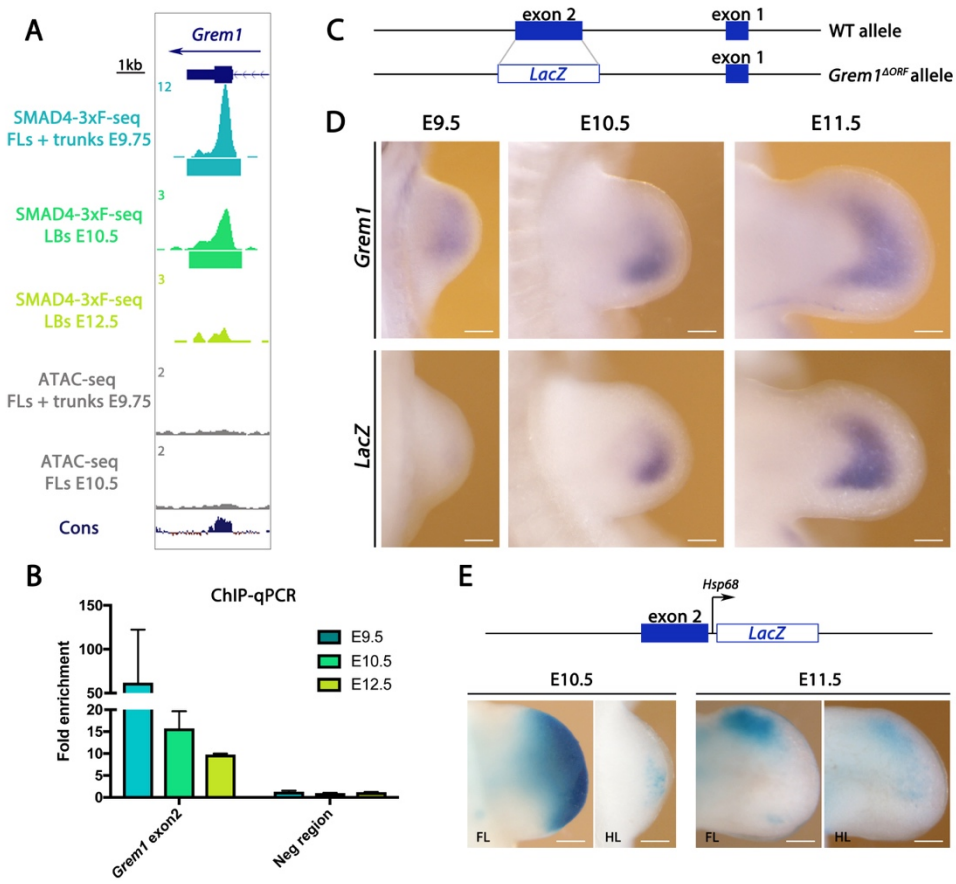


Figure 16: *Grem1* exon 2 is enriched in SMAD4 chromatin complexes and is able to drive LacZ reporter activity
(A) UCSC genome browser view of *Grem1* exon 2: SMAD4^{3xF} ChIP-seq peaks at E9.75, E10.5 and E12.5 with the corresponding ATAC-seq profiles at E9.75 and E10.5. (B) ChIP-qPCR validation of the SMAD4 interacting with *Grem1* exon 2. (C) Scheme showing the *Grem1*^{wt} and the *Grem1*$^{\Delta ORF}$ alleles (Michos et al., 2004). (D) *Grem1* and *LacZ* WISH in *Grem1*$^{\Delta ORF/+}$ limb buds at E9.5, E10.5 and E11.5. (E) Scheme showing the *Grem1 exon 2-LacZ* construct. Representative limb buds of transgenic embryos (E10.5 and E11.5) showing the *LacZ* reporter activity driven by *Grem1*-exon 2 (construct and analysis done by Laurène Ramos Martins). Scale bars: 100μm.

SMAD4 chromatin complexes interact with the GRS1 enhancer region

The interaction of SMAD4 chromatin complexes with the *Grem1* regulatory region 1 (GRS1), is only significantly enriched in forelimb buds at E9.75 (Figure 17A). The low level interaction of SMAD4 complexes with regions A and B of the GRS1 was confirmed by ChIP-qPCR (Figure 17A). The GRS1 is a robust

CRM that is able to drive strong and dynamic *LacZ* expression in forelimb buds (Figure 17C, Zuniga et al., 2012). CRISPR/*Cas9* mediated deletion of this element shows that GRS1 is required for normal levels of *Grem1* transcripts throughout limb bud development (Viviane Tschan, data not shown). The loss of GRS1 in $Grem1^{\Delta GRS1/\Delta GRS1}$ and $Grem1^{\Delta ORF/\Delta GRS1}$ forelimb buds (GRS1 deletion over a *Grem1* null allele) results in precocious termination of *Grem1* expression (Figure 17D). In addition, $Grem1^{\Delta ORF/\Delta GRS1}$ forelimbs display a large spectrum of skeletal phenotypes. About 50% of the embryos show carpal bone fusions, syndactily or soft tissue fusion (Figure 17E arrows). These results show that the GRS1 regulates *Grem1* expression likely in synergy with other CRMs. As GRS1 regulates *Grem1* expression, the low level SMAD4-GRS1 interaction may contribute to its transcriptional regulation but not specifically to *Grem1* activation.

In addition to *Grem1* locus, I also analysed the genome-wide profile of the $SMAD4^{3xF}$ interactions with genomic regions to identify the GRNs, whose expression is directly controlled by SMAD4 in E10.0 limb buds. To this aim, I have focused my analysis on the E9.75-E10.0 datasets, where the BMP activity is high, with DEGs expressed at levels detectable by whole mount RNA *in situ* hybridization (WISH) to assess their spatial distribution (Appendix 12.5). Furthermore, genes with known functions during limb bud development were given priority. Combining the unbiased GO analysis (Figure 13C) with WISH screening allowed me to identify two main GRNs directly controlled by SMAD4, namely an early limb patterning and a cholesterol biosynthesis gene network.

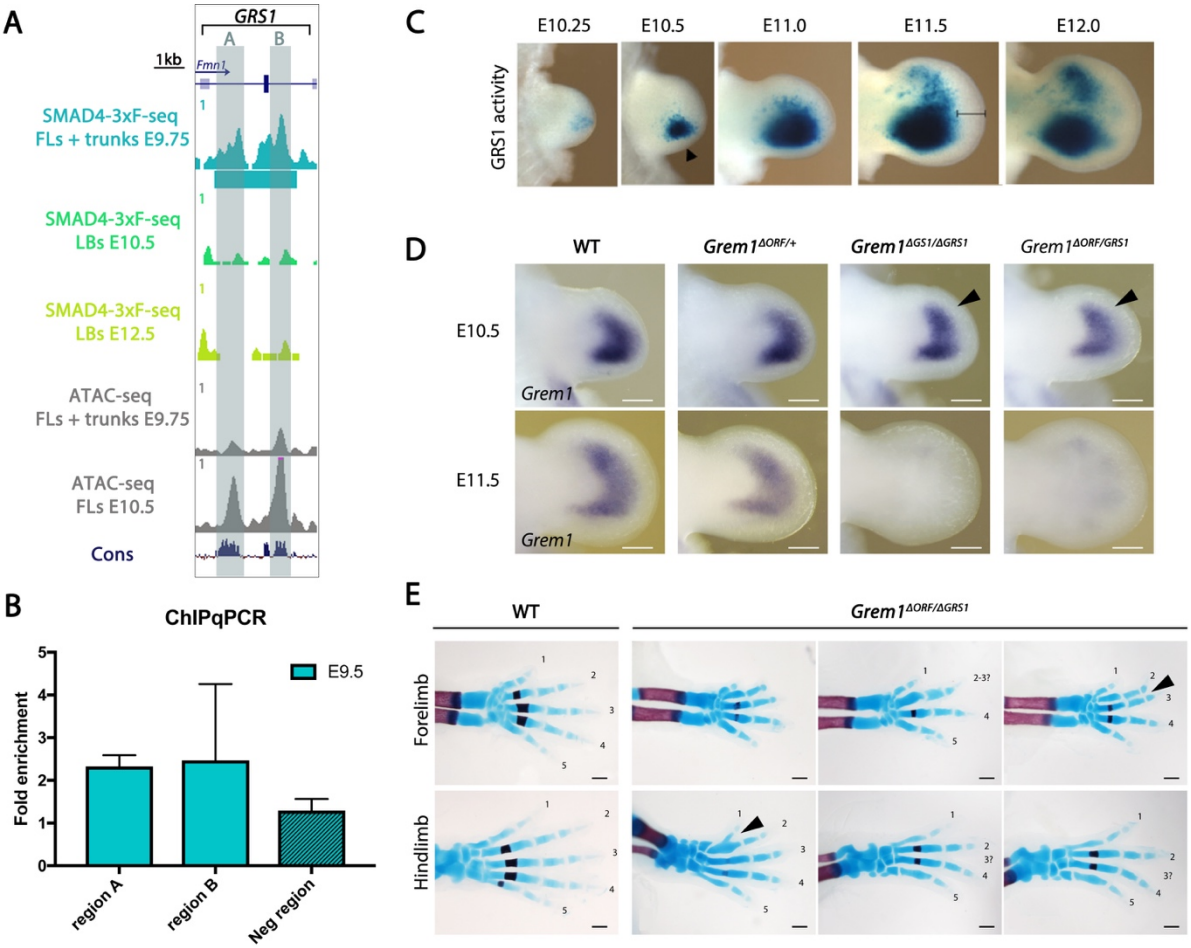


Figure 17: A *Grem1 cis*-regulatory module (CRM) is enriched in SMAD4 chromatin complexes
(A) SMAD4^{3xF} ChIP-seq and ATAC-seq profiles peaks map to the GRS1 region. (B) ChIP-qPCR validation of the SMAD4 interacting with the GRS1 element. (C) Transgenic embryonic limb buds showing the dynamic *LacZ* reporter activity driven by the *GRS1* element. (D) *Grem1* WISH in WT, *Grem1*$^{\Delta ORF/+}$, *Grem1*$^{\Delta GRS1/\Delta GRS1}$ and *Grem1*$^{\Delta ORF/\Delta GRS1}$ at E10.5 and E11.5. White scale bars: 100μm. Deletion of the GRS1 CRM results in precocious termination of *Grem1* expression. (E) Skeletal preparation of WT and *Grem1*$^{\Delta ORF/\Delta GRS1}$ E14.5 limbs. Black scale bars: 90μm.

6.2.6. Identification of a SMAD4 controlled gene regulatory network in early limb buds

Combination of the genomics analysis at E9.75-E10.0 (chapter 6.2.4) with the results of the WISH screen (appendix 12.5) identified many novel SMAD4 targets genes with functions in early limb bud patterning (Figures 18 and 19). This network of SMAD4 transcriptional targets down-regulated in the *Smad4*$^{\Delta/\Delta c}$

forelimb buds (Figures 18 and 19A) contains known targets of BMP signal transduction such as the *Id* genes (*Id1*, *Id2* and *Id3,* Hollnagel et al., 1999; Korchynskyi and Dijke, 2002), *Msx2* (Brugger et al., 2004) and the inhibitory SMADs, *Smad6* and *Smad7* (Zhao et al., 2000). Interestingly, while *Grem1* expression is down-regulated (Figure 4, Bénazet et al., 2012), the BMP antagonist *Smoc1* is up-regulated in *Smad4*$^{\Delta/\Delta c}$ forelimbs (Figure 18). In addition, several genes functioning in limb bud patterning are down-regulated in early *Smad4*$^{\Delta/\Delta c}$ forelimb buds (Figures 18 and 19A) such as *Tbx2* (Nissim et al., 2007), *Lhx2* (Tzchori et al., 2009), *Alx4* (Kuijper et al., 2005), and *Pkdcc* (Probst et al., 2011a). In particular some of the down-regulated genes have been associated with congenital limb skeletal malformations. These genes include *Sfrp2* (Morello et al., 2008), *Snai1* (Chen and Gridley, 2013) and *Prrx2* (Taher et al., 2011).

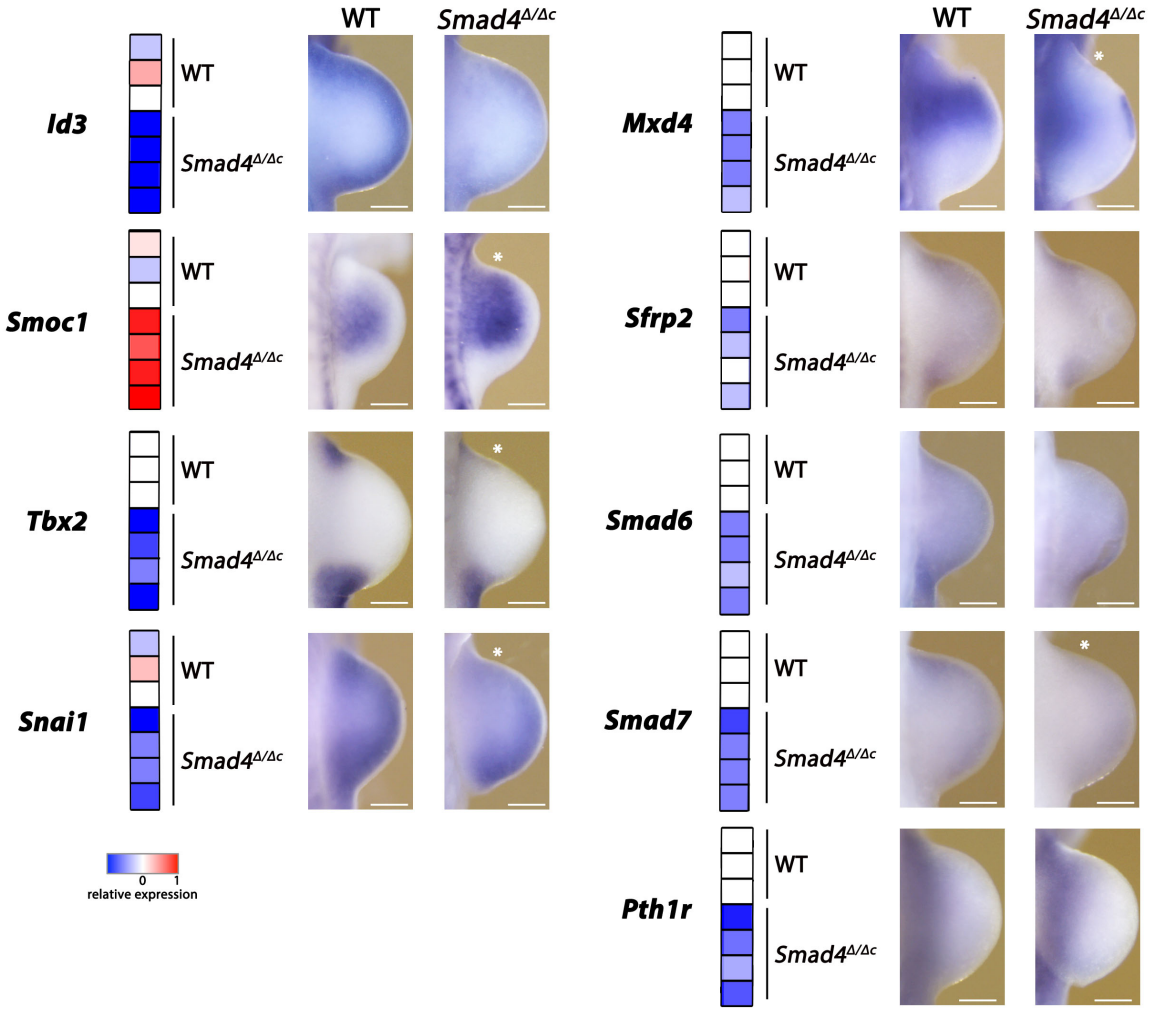


Figure 18: SMAD4 target genes involved in early limb bud patterning.
Heatmaps: blue and red boxes represent the down- and up-regulation, respectively, normalized to the expression mean in WT samples. WISH analysis of SMAD4 target genes involved in early limb bud patterning in WT and *Smad4*$^{\Delta/\Delta c}$ forelimbs. The scale bars represent 100μm. Asterisks show up or down-regulation in the anterior part of *Smad4*$^{\Delta/\Delta c}$ forelimbs.

Several of the transcriptional targets of SMAD4 are up- or down- regulated specifically in the anterior forelimb bud mesenchyme of *Smad4*-deficient forelimb buds such as *Smoc1*, *Tbx2*, *Snai1*, *Mxd4*, *Smad7*, *Id1*, *Id2*, *Msx2*, *Alx4* and *Prrx2* (Figures 18 and 19A, asterisks). This suggests that SMAD4 controls the GRNs mainly in the anterior mesenchyme, which points to potential functions in patterning the anterior limb bud. Furthermore, we have checked the regulatory potential of some of the CRMs associated with SMAD4 target genes using *LacZ* reporters in transgenic founder embryos. Interestingly, two main *LacZ* expression patterns were identified. Either the *LacZ* pattern recapitulates major aspects of the spatial expression of associated SMAD4 target gene (see e.g. *Id1* and *Msx2*) or *LacZ* expression is restricted to the anterior limb bud mesenchyme (see *Id2*, *Lhx2*, *Alx4*, *Prrx2* and *Pkdcc*, Figure 19A). This analysis corroborates the idea that SMAD4 directly controls the expression of patterning genes predominantly in the anterior limb bud mesenchyme. Most importantly, it identifies a SMAD4-controlled GRN network in the anterior mesenchyme of early forelimb buds (Figure 19B).

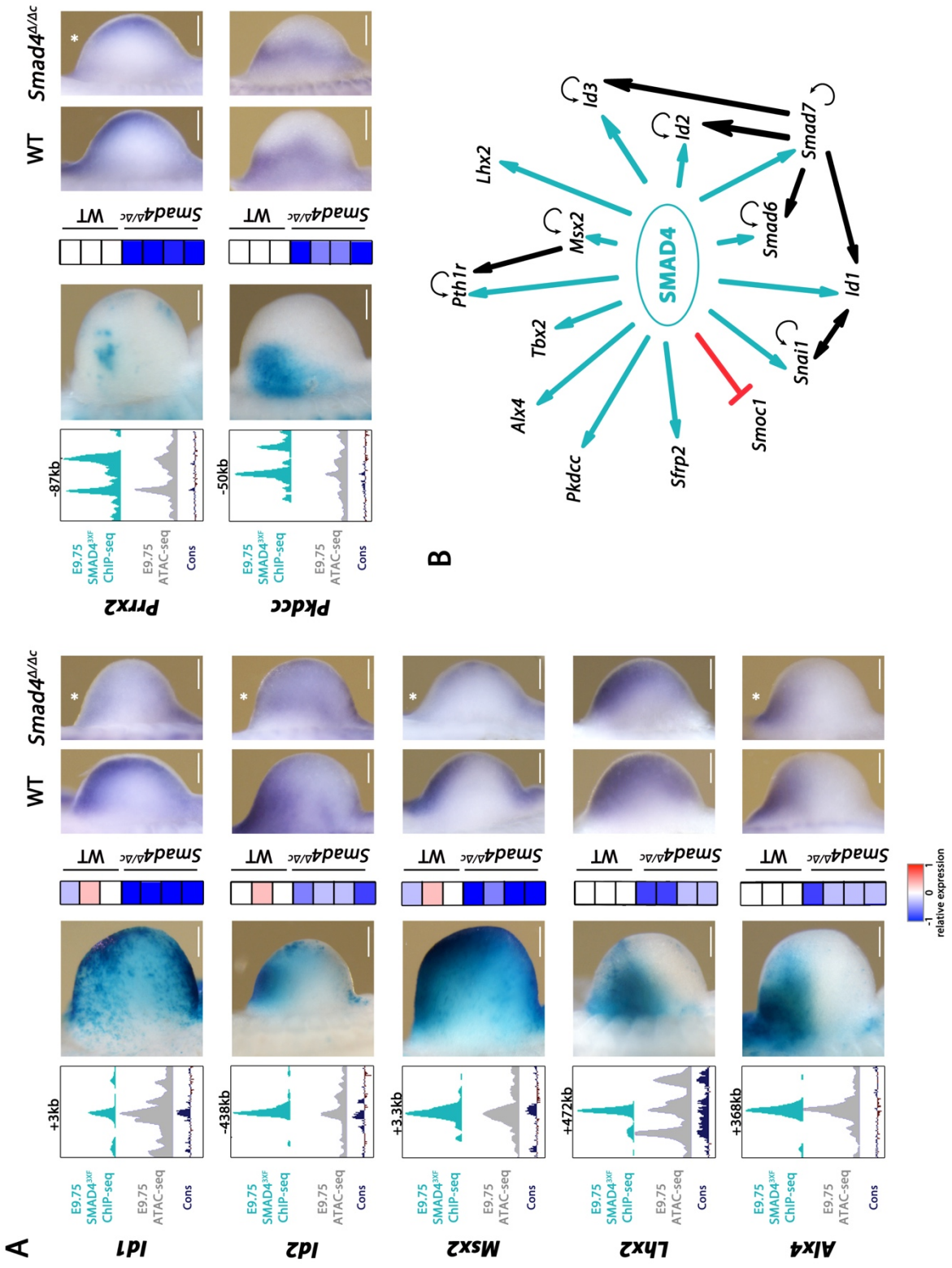


Figure 19: The limb patterning network controlled by SMAD4.
(A) Snapshots of SMAD4 bound CRMs and corresponding ATAC-seq profile are shown on the left. *LacZ* activity associated with these CRMs in embryonic limb buds at E10.5. The heatmaps show the relative expression changes as coloured boxes, normalized to the expression mean in WT samples. WISH for the associated SMAD4 target. The scale bars represent 100μm. Asterisks indicate the altered expression in the anterior part of the

Smad4$^{\Delta/\Delta c}$ forelimbs. (B) The gene regulatory network controlled by SMAD4 which is involved in early limb bud patterning.

6.2.6. SMAD4 controls the expression of enzymes in the cholesterol biosynthesis pathway

The most surprising discovery of our analysis was that SMAD4 controls the expression of many enzymes involved in cholesterol biosynthesis pathway, as almost all of them are down-regulated in *Smad4*$^{\Delta/\Delta c}$ forelimb buds at early stages (Figure 20A and 20B). Next, I checked their spatial expression patterns in WT and *Smad4*$^{\Delta/\Delta c}$ forelimbs (Figure 20C). This showed that most of these enzymes are expressed at only low levels and in a rather uniform manner during limb bud development, which precluded detection of spatial differences in expression (data not shown). However, for some genes such as *Pmvk*, *Dhcr7*, *Mvd, Pcsk9*, *Tm7sfr2* or *Sc5d* the transcriptional down-regulation can be detected in forelimb buds lacking *Smad4* expression (Figure 20C). This down-regulation of enzymes functioning in cholesterol biosynthesis indicated that the cholesterol production by *Smad4*$^{\Delta/\Delta c}$ limb bud mesenchymal cells could be altered.

To detect potential alterations in cholesterol and its intermediates, a quantitative Gas-Chromatography / Mass Spectrometry (GC/MS) analysis using WT and *Smad4*$^{\Delta/\Delta c}$ forelimb buds was done in collaboration with Prof. Dorothea Hass, University Children's Hospital, Heidelberg, Germany. This analysis showed that sterol intermediates lanosterol, demosterol, lathosterol and 7DHC (Appendix 12.6) seemed slightly increased if comparing *Smad4*$^{\Delta/\Delta c}$ to the WT forelimb buds at different embryonic stages. Unfortunately, the differences were not or barely significant with exception of the increased lanosterol and demosterol levels in forelimb buds at E11.0 (Appendix 12.6). The trend toward accumulation of sterol intermediates could be a consequence of the significantly reduced expression of the enzymes functioning in cholesterol synthesis in *Smad4*-deficient forelimb buds, as it is the case for different congenital

metabolic diseases that affect cholesterol synthesis (Porter and Herman, 2011). The quantity of cholesterol followed the same trend in the *Smad4*$^{\Delta/\Delta c}$ forelimb buds, but the differences were only significant only at E11.0 (Appendix 12.6). While these results are intriguing, they also indicate that the GC/MS method may not be sensitive enough to detect quantitative differences of sterol intermediates and cholesterol in early mouse forelimb buds as the only significant differences were seem in more advanced forelimb buds.

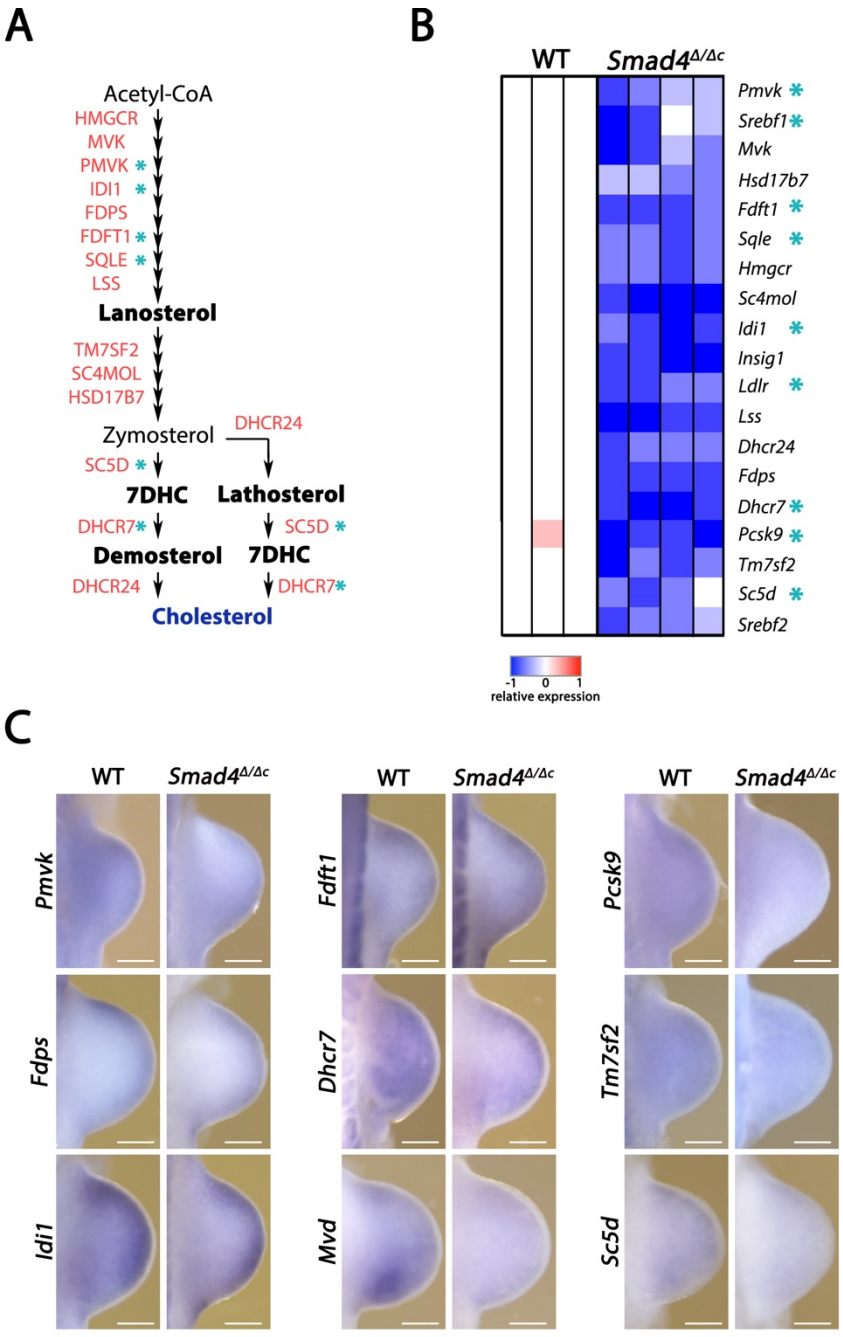


Figure 20: The enzymes essential for cholesterol synthesis are down-regulated in *Smad4*$^{\Delta/\Delta c}$ forelimbs.
(A) Schematic representation of the cholesterol synthesis pathway. The down-regulated enzymes in *Smad4*$^{\Delta/\Delta c}$ forelimbs are indicated in red and the main cholesterol intermediates in bold. The blue asterisks show the direct targets of SMAD4. (B) Heatmap showing the down-regulation of genes involved in the cholesterol biosynthesis in WT and *Smad4*$^{\Delta/\Delta c}$ forelimbs (normalized to the expression mean in WT samples). The blue asterisks indicate the direct transcriptional targets of SMAD4. (B) WISH analysis of genes encoding enzymes of cholesterol biosynthesis in WT and *Smad4*$^{\Delta/\Delta c}$ forelimbs. Scale bars: 100μm.

6.2.7. Graded SHH signalling and the expression of SHH targets in responding cells are altered in *Smad4*-deficient forelimb buds

The carboxy terminal cholesterol-modification of the SHH is crucial for its activity and spatial distribution (Hentschel et al., 2016). In limb buds, he SHH ligand that is produced by the ZPA cells in the posterior mesenchyme and the spead of its posterior to anterior graded distribution is restricted by modifying the ligand with cholesterol (Li et al., 2006). Because of the transcriptional changes that affect the cholesterol biosynthesis pathway in *Smad4*$^{\Delta/\Delta c}$ forelimb buds, I decided to check if the spread of the SHH ligand is altered in mutant limb buds. While the spatial distribution of *Shh* transcripts was not changed between E10.0 and E11.5 in *Smad4*-deficient forelimb buds (Figure 21A), immunohistochemical detection of the SHH ligand revealed a defect in the posterior to anterior spreading of the SHH ligand. While no differences are observed at E10.0, the SHH ligand remained more posteriorly restricted in *Smad4*-deficient forelimb buds at E10.5 in comparison to WT controls (Figure 21B). At E11.5, the SHH ligand remains restricted and levels are increased most posteriorly in *Smad4*$^{\Delta/\Delta c}$ forelimb buds (Figure 21B).

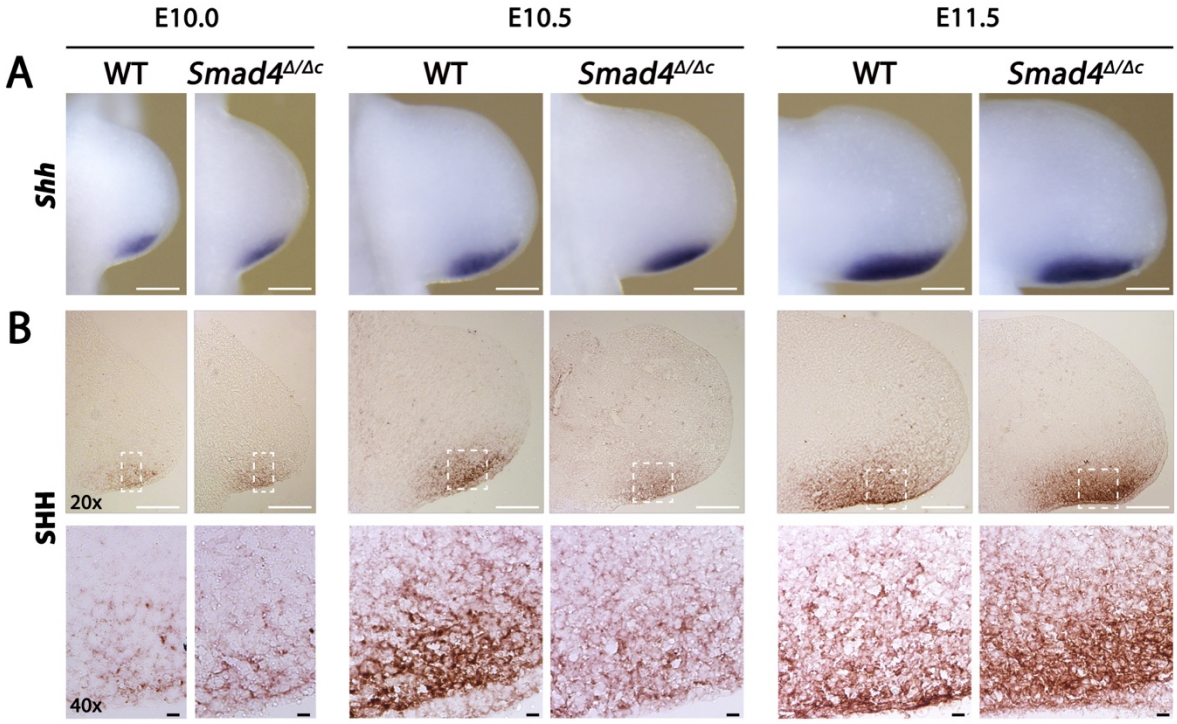


Figure 21: The spread of SHH is altered in *Smad4*$^{\Delta/\Delta c}$ forelimb buds
(A) WISH for *Shh* in WT and *Smad4*$^{\Delta/\Delta c}$ forelimbs from E10.0 to E11.5. (B) Immunohistochemical detection of SHH protein distribution in WT and *Smad4*$^{\Delta/\Delta c}$ forelimb buds. White scale bars: 100μm; black scale bars: 10μm.

The spatio-temporal expression of three direct transcriptional targets of SHH signal transduction were assessed (Figure 22). The *Gli1* expression levels, a direct effector of SHH signalling are not changed at E10.0 (Figure 22A). The *Ptch1* expression domains is expanded distal-anteriorly in mutant forelimb buds in comparison to WT controls (Figure 22B) and is up-regulated in the transcriptome of *Smad4*$^{\Delta/\Delta c}$ forelimb buds at E10.0 (left panels, Figure 22B). In addition, the *Hhip1* expression domain was enlarged and an ectopic posterior domain was observed (Figure 22C). By E10.5 and E11.5, the *Ptch1* and *Gli1* expression domains appear more posteriorly restricted (Figures 22A and 22B) in agreement with the posterior restriction of the SHH ligand (Figure 21B left panels), while *Hhip1* expression remains increased in the posterior mesenchyme at E10.5 and E11.5 (Figure 22C). As neither of these three SHH targets contains SMAD4 interacting CRM in their TADs (data not shown), the

transcriptional alterations observed in *Smad4*-deficient limb buds are caused by altered SHH signal transduction rather than being a direct effect of the *Smad4* deficiency on their transcriptional regulation. As the observed transcriptional alterations are quite dynamic (Figure 22), I decided to also assess the potential effects of altered cholesterol synthesis on the cellular response to paracrine SHH signalling using WT and *Smad4* mutant limb bud cells in culture.

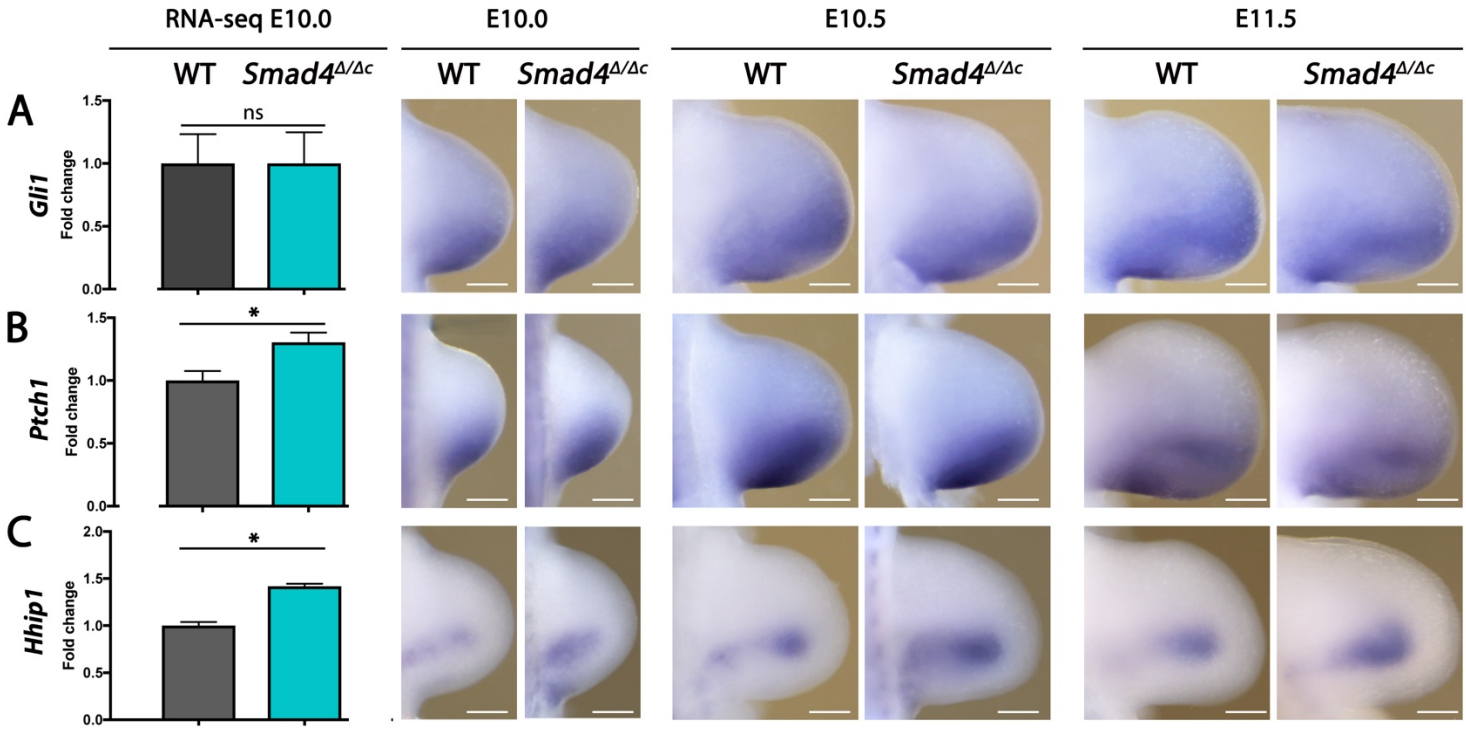


Figure 22: Molecular responsiveness to SHH signalling is altered in *Smad4*$^{\Delta/\Delta c}$ forelimb buds
Barplots show the fold change between WT and *Smad4*$^{\Delta/\Delta c}$ (from RNA-seq). Comparative WISH analysis of WT and *Smad4*$^{\Delta/\Delta c}$ forelimb buds from E10.0 to E11.5 for SHH transcriptional targets genes: *Gli1*, *Ptch1* and *Hhip1*. Scale bars: 100μm.

6.2.8. SHH signal transduction in *Smad4*$^{\Delta/\Delta c}$ limb bud cells depends on cholesterol

In addition to SHH, Smoothened (SMO), the G-protein coupled receptor that activates SHH signal transduction is cholesterol modified. This modification is crucial for activation of SHH signal transduction. Binding of SHH to the transmembrane receptor PTCH1 supresses the inhibition of SMO by PTCH.

Cholesterol modification of SMO then changes its conformation and triggers a series of intracellular events that activate signal transduction (Huang et al., 2016; Xiao et al., 2017). To study the potential effects of intracellular cholesterol concentration changes in mesenchymal cells from *Smad4*$^{\Delta/\Delta c}$ forelimb buds on the SHH signal transduction, we set-up a culture system for limb mesenchymal progenitors (LMPs). Forelimb buds were collected at E10.25 (32-34s) and following removal of the ectoderm and cell dissociation, they were plated in lipid-depleted medium in 2-wells of a 96 well plate. One well was treated with SAG, a SMO agonist which bypasses the cholesterol activation (Huang et al., 2016), and the other well served as untreated control. I assessed SHH signal transduction 24hrs later by monitoring *Gli1* and *Ptch1* transcript levels using RT-qPCR analysis (Figure 23A). If LMPs are cultured in lipid-free medium, the *Smad4*$^{\Delta/\Delta c}$ LMPs fail to maintain *Ptch1* and *Gli1* expression at levels comparable to the WT LMPs (Figure 23B). In cultures treated with SAG, the expression of both SHH transcriptional targets is comparable in LMPs from WT and *Smad4*$^{\Delta/\Delta c}$ forelimb buds (Figure 23B).

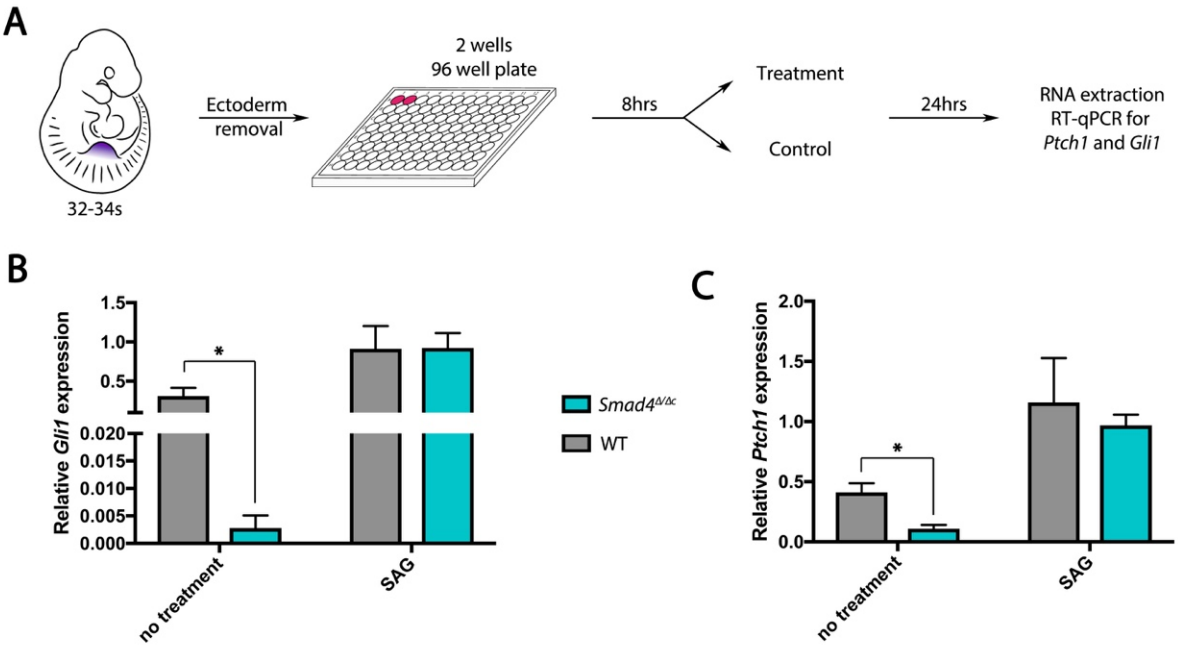


Figure 23: The response to SHH signal is altered in mesenchymal cells from *Smad4*$^{\Delta/\Delta c}$ limb buds.
(A) Experimental design using cultured limb mesenchymal progenitors (LMPs) in lipid depleted medium for analysis. (B) Expression of the SHH transcriptional targets *Gli1* and *Ptch1* in LMPs without treatment or SHH antagonist SAG treatment.

These studies show that *Smad4*$^{\Delta/\Delta c}$ LMPs are unable to maintain SHH target gene expression in lipid-depleted medium as a likely consequence of the significant down-regulation of the enzymes required for cholesterol biosynthesis. This indicates that the endogenous cholesterol synthesis is crucial for activating SHH signal transduction in LMPs from early mouse forelimb buds. Furthermore, bypassing cholesterol-dependent activation of SMO by SAG shows that the transcriptional targets *Ptch1* and *Gli1* are up-regulated to a similar extent in WT and *Smad4*$^{\Delta/\Delta c}$ LMPs, which indicates that the downstream components of the SHH signal transduction cascade are functional in *Smad4*$^{\Delta/\Delta c}$ LMPs

7. Discussion

7.1. Unbiased genome-wide analysis of the SMAD4 cistrome

As few direct transcriptional targets of the BMP/TGFβ signalling pathway were known, I combined SMAD4 ChIP-seq with open chromatin (ATAC-seq) and RNA-seq analysis to identify the SMAD4 cistrome and direct target genes during early mouse forelimb bud development. In particular, the Smad4^{3xF} allele used for ChIP-seq is a new and powerful tool to identify the CRMs interacting with the endogenous SMAD4 protein *in vivo*.

Several previous studies aimed at gaining insights into how BMP/TGFβ signalling controls transcription of its targets using SMAD4 ChIP-on-chip approaches that focused on promoter regions (Fei et al., 2010; Koinuma et al., 2009; Qin et al., 2009). These methods gave good first insights but are limited by their biased analysis that excludes SMAD4 binding sites in located in distant non-coding *cis*-regulatory regions. SMAD4 ChIP-seq studies were also done in different cell lines using commercial SMAD4 antibodies which revealed the difficulty to map the interactions of the endogenous SMAD4 protein with CRMs (Kennedy et al., 2011; Kim et al., 2011; Morikawa et al., 2011; Tsankov et al., 2015). In contrast to commercial ChIP-grade SMAD4 antibodies, which do not work in limb buds, the SMAD4^{3xF} mouse is a powerful tool to perform *in vivo* ChIP-seq experiments with the endogenous SMAD4 protein. It allows to benefit of the excellent ChIP-grade mouse anti-FLAG antibody (Laurent et al., 2017; Osterwalder et al., 2014), bypassing the lack of efficient commercial SMAD4 antibody, and avoiding any overexpression system that can potentially give rise to false positive interaction.

The nucleo-cytoplasmic shuttling and the DNA weak binding of SMAD4 itself posed two main challenges with respect to the SMAD4^{3xF} ChIP-seq analysis

that I have done. This nucleo-cytoplasmic shuttling is controlled by a nuclear localization signal (NLS) and a nuclear export signal (NES) in the SMAD4 protein (Reguly and Wrana, 2003). This dynamic sub-cellular localisation might be the cause of the dynamic functionality observed for SMAD4 by ChIP-seq in human embryonic stem cells (Kim et al., 2011; Tsankov et al., 2015) and can also contribute to the differences in the SMAD4 cistromes that I identified in forelimb buds at E9.75 and E10.5. Furthermore, the interactions of SMAD4 with DNA are rather weak, which might also be a consequence of the dynamics of nucleo-cytoplasmic shuttling (Shi et al., 1998). These SMAD4 particularities might explain that the SMAD4 cistrome is smaller in contrast to a classical transcription factor such as HAND2

Direct and indirect interactions of SMAD4 with genomic DNA

The enrichment of SMAD4 chromatin complexes in proximity to promoters is a known binding characteristic of SMAD proteins (Morikawa et al., 2013). Promoter-enhancer interactions are key to spatio-temporal regulation of transcription regulation and enhancers in developmental genes are often located very distant to the promoter. Therefore, we have decided to analyse not only proximal but also distant regions enriched in SMAD4 chromatin complexes. However, as promoter-promoter interactions are involved in transcription regulation (Joshi et al., 2015; Li et al., 2012), we have decided to considerate all the SMAD4 bound regions, even the ones located ±5kb from TSSs.

The enrichment of a genomic region in SMAD4-containing chromatin complexes does not mean that SMAD4 directly interacts with the candidate CRM. Our HOMER (Heinz et al., 2010) *de novo* and known motif analysis of the SMAD4^{3xF} ChIP-seq datasets illustrates the complexity of SMAD4 interactions with DNA. Both E9.75 and E10.5 SMAD4^{3xF} ChIP-seq are enriched for the SMAD4, SMAD2 and SMAD3 consensus motifs containing the GTCT sequence that has been identified in hECS (Kim et al., 2011). In addition it has been shown *in vitro* that SMAD3 and SMAD4 directly bind the palindromic sequence GTCTAGAC

(Zawel et al., 1998). As SMAD2/3 are the R-SMADs acting down-stream of TGFβ, these results indicate that target regions might be shared between the TGFβ and BMP signalling pathways (Lee et al., 2000). The top-enriched motif from my analysis of forelimb buds at E9.75 is the specific SMAD4 binding motif that contains the CAGA box and a CG-rich sequence identified in promoter regions of previously identified targets of BMP signal transduction (Morikawa et al., 2013).

In addition, the identified ASH1 motif is associated with the histone-lysin-N-methyltransferase that regulates the H3K36me mark, which modulates the expression of *Hox* genes (Tanaka et al., 2011). The MYOD binding motif enriched in the SMAD4 cistromes at both limb bud stages is associated with local histone acetylation and the PU.1 identified in the E10.5 forelimb bud data set is also associated with regions enriched in monomethylation of H3K4 (Morikawa et al., 2013). Taking together these results suggest that the SMAD4 chromatin complexes might contain or be associated with co-factors that modulate chromatin accessibility at the level of histone modifcations. This could explain why a fraction some $SMAD4^{3xF}$ ChIP peaks do not overlap with ATAC-seq peaks. In addition, SMAD4 might interact with pioneer factors that function in opening chromatin to make it accessible for the transcription machinery and/or transcriptional enhancing complexes at SMAD binding sites (Morikawa et al., 2013).

In the SMAD4 cistrome of forelimb buds at E10.5 the PKNOX1 (or PREP1) motif is also significantly enriched. PREP1 is a member of the TALE (three amino acid loop extension) superfamily together with the PBX transcription factors. PREP1 dimerizes with PBX1 to control the FGF and the Wnt signalling pathway in stem cells and is known to interact with SMAD4 (Laurent et al., 2015, Bailey et al., 2004). As PBX transcription factors and PREP1 are crucial for the onset of limb bud development (Capellini et al., 2011), my results suggest that SMAD4, PREP1 and PBX1 interact directly during the onset of

limb bud development and establishment of the SHH/GREM1/FGF signalling feedback loop.

These interactions of SMAD4 with different co-factors likely determines its functions in possibly altering chromatin states during activation of transcription in proximity to promoters or/and to regulate spatio-temproal aspects of gene expression by interacting with CRMs. This associations are likely cell-type/lineage-specific and provides the dynamic specificity (Morikawa et al., 2013) also detectable by comparing the SMAD4 cistromes in forelimb buds at E9.75 to E10.5. The differences in the enrichment of SMAD4 binding motifs between both limb bud stages overlap with the change from high to low BMP activity that occurs between initiation (≤E10.0) and propagation (≥E10.5) of the SHH/GREM1/FGF feedback signaling system. Furthermore, as phospho-SMAD2 bind to its targets in a dose-dependent manner (Lee et al., 2011), we can speculate that this decrease in BMP activity triggers the change in the range SMAD4 target genes observed (see below).

7.2. Transition from high to low BMP activity causes a change in the range of SMAD4 target genes in early limb buds

The change from high to BMP low activity observed when the SHH/GREM1/FGF feedback signalling system is established (Bénazet et al., 2009; Pignatti et al., 2014) is also clearly apparent in the changes of differentially expressed genes between WT and *Smad4*-deficient for limb buds within the 10hr time window analysed. With this short time window, we expected minimal fold changes in differential gene expression. Nevertheless, at E10.0, when the BMP activity is still high, we observe many more DEGs (669 genes) than at E10.5 (346 genes), and only 151 DEGs are shared between the two

limb bud stages. This drop of SMAD4 dependent gene regulation is apparent when focusing on the direct SMAD4 target genes. There are 306 direct SMAD4 targets in forelimb buds at E9.75-10.0 in contrast to only 190 genes at E10.5. The two stages share only 26 direct SMAD4 target genes. This divergence fits with the two phases of BMP activity: i) during initiation of limb bud outgrowth (E9.0 to E10.0), high BMP results in high levels of SMAD4-dependent transcriptional regulation of target genes, while ii) the activity of SHH/GREM1/FGF feedback signalling system (E10.5) lowers BMP activity, which in turn decreases SMAD4 dependent regulation of target genes. In addition, the range of SMAD4 target genes is different when BMP4 activity is low. The transition between these two phases occurs as a consequence of the BMP-dependent transcriptional activation of *Grem1,* which lowers BMP activity and antagonising BMP ligands (Bénazet et al., 2009)

BMP activates Grem1 expression via SMAD4

Grem1, a critical node in the SHH/GREM1/FGF feedback signalling system, is regulated in a complex but robust manner by the different CRMs located in the *Grem1* sub-TAD. My analysis identifies an unexpected interaction of SMAD4 chromatin complexes in *Grem1* coding exon 2 and the corresponding SMAD4 binding site overlaps a ASH1 motif (Figure 24). As discussed before ASH1 motif is associated with the histone-lysine-N-methyltrasnferase which regulates the H3K36me marks (Dorighi and Tamkun, 2013). Indeed, there is an enrichment of H3K36 methylation in the *Grem1* coding exon 2 at E10.5, which is characteristic for transcribed exons (Wagner and Carpenter, 2012). During transcriptional elongation, RNA polymerase II (Pol-II) recruits the Rpd3S histone deacetylase to remove histone acetylation marks in transcribed regions in a H3K36me dependant manner (Lee et al., 2013). Lee et al. (2013) shows that chromatin remodelling by modulating nucleosomal spacing with H3K36me, is crucial for Rsp3S functions and Pol-II mediated transcriptional elongation. Taken together, these results indicate that SMAD4 might be part of the chromatin complexes that recruit ASH1 to coding exon 2 of the *Grem1* transcription unit to rearrange

the chromatin and induce H3K36 methylation. By this mechanism the *Grem1* coding region would be deacetylized by Rpd3S, which then allows the continuation of *Grem1* transcription in a SMAD4-ASH1 independent manner. This hypothesis would explain the dynamics of the SMAD4 interactions with *Grem1* coding exon 2 and the lack of open chromatin and histone acetylation. Unfortunately, the H3K36me ChIP-seq profile for E9.5 forelimb buds is not available, which would be needed to understand if the H3K36 methylation of the *Grem1* coding exon 2 is indeed higher at E9.5 than E10.5 similar as seen for its dynamic interactions with SMAD4 complexes. Transcriptional regulation through fine-tuning Pol-II activity would constitute a novel SMAD4 function, but as the current analysis focused on the interaction of SMAD4 complexes with regions of open chromatin, I would have missed other target genes whose expression initiation could be SMAD4-H3K36me sensitive. This chance observation of the binding to coding exon 2 of *Grem1* paves the way for an interesting follow-up analysis.

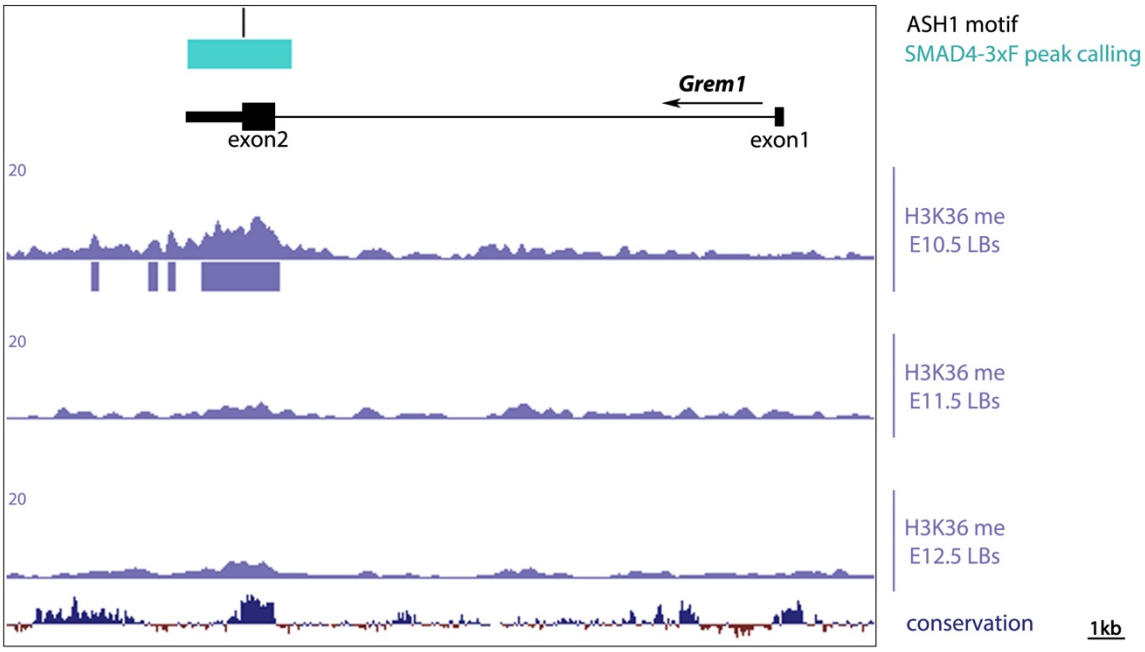


Figure 24: *Grem1* coding exon 2 is decorated with H3K36me marks
UCSC genome browser view of the *Grem1* locus (mm10): alignment of the ASH1 motif, the SMAD4-3xF ChIP-seq peak (E9.75 and E10.5) and the H3K36me ChIP-seq profiles (E10.5, E11.5 and E12.5, ENCODE). Note that the *Grem1* coding exon 2 is enriched in $SMAD4^{3xF}$ chromatin complexes encompassing the ASH1 motif and decorated by H3K36 methylation only at E10.5.

After initiation, the expression of *Grem1* is regulated in a dynamic spatio-temporal fashion as limb bud development progresses. This fine regulation needs to be robust as loss of *Grem1* expression results in drastic phenotypes (Zuniga et al., 2004). It appears that this dynamic expression is orchestrated by the many CRMs present in the *Grem1* genomic landscape, which likely have differential and/or redundant activities. Deletion of the GRS1 element uncovered its role during up-regulation and termination of *Grem1* expression. However, we can not exclude a more general role that is not revealed by its deletion as several CRMs might function in a similar and/or redundant manner (Cannavò et al., 2016). We are deleting other CRMs individually or in combination to test the robustness of the system and to understand how the dynamics of *Grem1* expression is regulated (this study is done by Laurène Ramos Martins).

However, these candidate CRMs can only contribute to the spatiotemporal regulation if they can interact with the *Grem1* promoter. Many studies have identified CRMs regulating gene expression that located many kilobases away from the relevant promoter and TSS (Lettice et al., 2003; Zuniga et al., 2012). The 3D architecture of chromatin is important for these interactions. In particular, chromatin loops within TADs allow specific interactions of CRMs with the respective promoter. This looping depends on CTCF and Cohesin binding sites that drive the 3D conformation of the chromatin and organize the TADs (Rao et al., 2014). The orientation of the CTCF binding sites is responsible for the intra-TAD chromatin configuration and establishing boundaries between TADs. A chromatin loop can be formed between two converging CTCF and a TAD boundary is formed between two diverging CTCF (Gómez-Marín et al., 2015; Rao et al., 2014). Six such structural elements are present in the *Grem1* genomic landscape. Using these data, I have been able to come up with two possible models for the 3D architecture of the *Grem1* genomic landscape: i) the first is the “loop in the loop” model (Figure 25A), ii) the second is the “loop and loop” model (Figure 25B). In both cases, the CRMs interacting with either with HoxA13, GLI3 or SMAD4 in closer proximity and able to interact with the *Grem1*

promoter. However, the *in situ* Hi-C profile from B-lymphoblastoid cells (Rao et al., 2014) points to the formation of a small loop in a bigger loop (Figure 25C), which is in favour of the “loop in the loop” model.

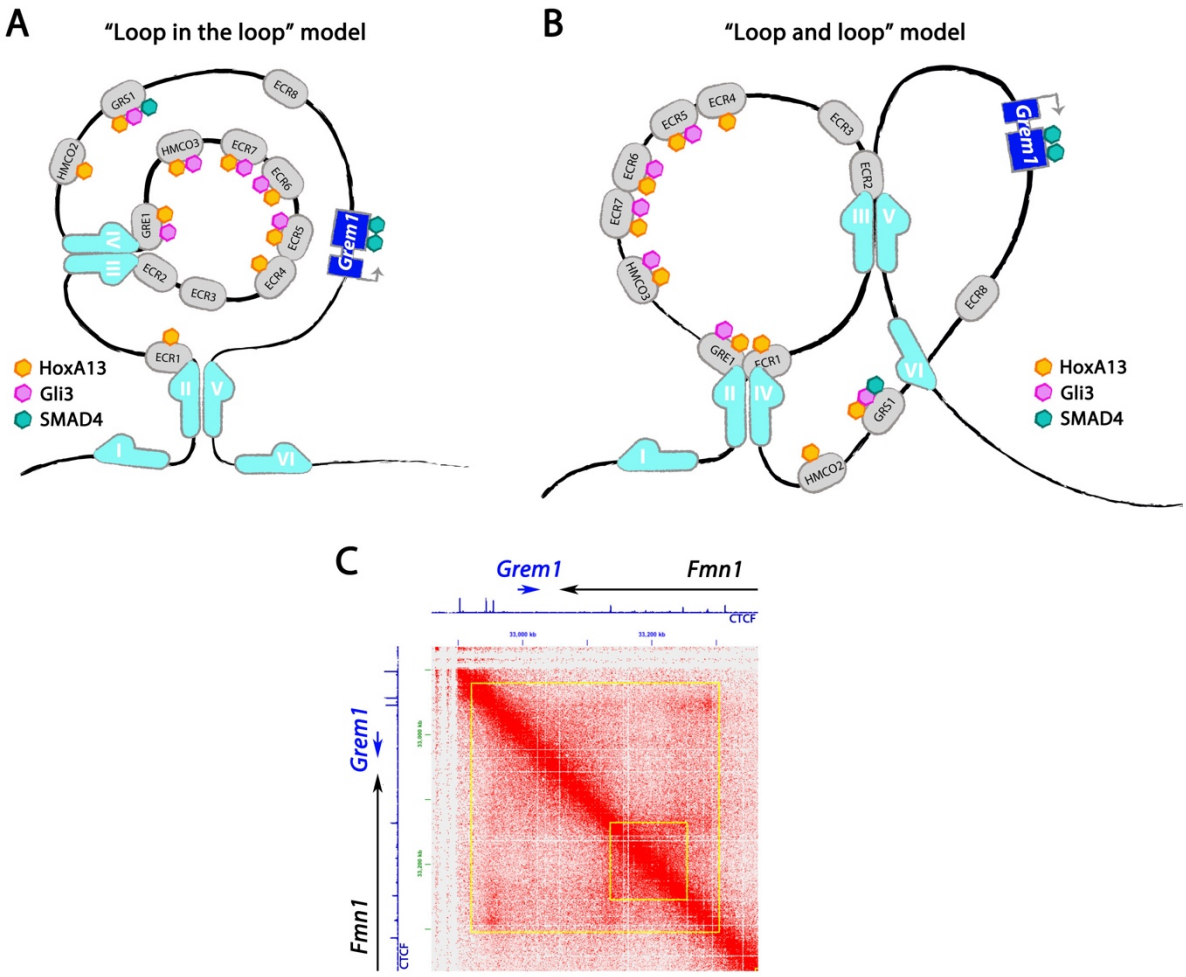


Figure 25: Models for the likely chromatin configuration of the *Grem1* genomic landscape

Two basic chromatin configurations are possible at the *Grem1* locus. (A) Scheme of the “Loop in the loop” model for the 3D chromatin architecture. The CTCF/Cohesin binding sites are indicated by the blue arrows (I to VI), the functionally relevant CRMs by grey boxes. Regions enriched in HoxA13 chromatin complexes are indicated in yellow, the GLI3 interaction regions in pinks and SMAD4 binding regions in green. The *Grem1* exons are represented by the dark blue boxes. (B) “Loop and loop” model. (C) *In situ* Hi-C profile for the *Grem1* genomic landscape in human B-lymphoblastoid cells (Rao et al., 2014). The loop calling is represented by the yellow squares. The smaller loop is within the bigger loop, which is in support of the “Loop in the loop” model.

As there is good evidence that many chromatin interactions and contacts are conserved between cell types and during evolution (Rao et al., 2014; Vietri Rudan et al., 2015), I propose the “loop in the loop” model as more feasible model for the 3D chromatin architecture at the *Grem1* locus. Within this conceptual framework, the different CRMs detect the information from signalling inputs and interact with the *Grem1* promoter in a dynamic and temporally controlled manner as limb bud development progresses.

7.3. SMAD4 has multiple functions during early limb bud development

The unbiased WISH screen and genome wide-analysis of the SMAD4 transcriptional targets has allowed me to identify the limb bud GRNs controlled by SMAD4. These SMAD4 GRNs include known BMP targets such as *Id1*, *Id2*, *Id3* (Hollnagel et al., 1999), *Msx2* (Brugger et al., 2004), *Smad6* and *Smad7* genes (Benchabane and Wrana, 2003; Ishida et al., 2000), which validate the unbiased analysis and identified GRNs. More importantly, my analysis identified many novels SMAD4 targets, the majority of which are linked to limb defects and have established functions in different processes during mouse limb bud development.

First, my analysis identifies an additional BMP antagonist called *Smoc1* (Rainger et al., 2011) as a direct SMAD4 target in early limb buds. Interestingly, *Smoc1* expression is negatively regulated by SMAD4, which contrasts with the positive regulation of the BMP antagonist *Grem1*. This balance between SMAD4-mediated activation and repression of BMP antagonists likely reflects the fine-tuned auto-regulation of the spatial BMP activity in early limb buds. During the onset of limb bud development, BMP signalling via SMAD4 inhibits *Smoc1* in the anterior and induces *Grem1* in the posterior limb bud mesenchyme. This indicates that the anterior part remains free of BMP antagonists, which enables BMP/SMAD4 to regulate the expression of their

anterior target genes, which is corroborated by the early down-regulation of many SMAD4 target genes in the anterior of *Smad4*-deficient limb buds. In contrast, the BMP/SMAD4-mediated induction of *Grem1* in the posterior mesenchyme results in the establishment of the SHH/GREM1/AER-FGF feedback signalling system.

The Wnt inhibitor *Sfrp2* is also a direct SMAD4 target, whose expression is down-regulated in the $Smad4^{\Delta/\Delta c}$ forelimb buds as other Wnt antagonists like *Dkk1* or *Wif1* (my study and Klaus and Birchmeier, 2008). The inhibition of Wnt signalling decreases in *Smad4* mutant limb buds, which indicates a direct and indirect regulation by SMAD4. Wnt signalling participates in limb bud initiation and outgrowth by interfering with the FGFs that induce the AER and promote proliferation of the limb bud mesenchyme (Geetha-Loganathan et al., 2008). Therefore, the reduced Wnt inhibition in *Smad4* deficient limb buds might explain their slightly increased size (Bénazet et al., 2012).

A significant fraction of SMAD4 targets are related to the establishment of axes polarity during initiation of limb bud outgrowth. The coordinated establishment and outgrowth of the three axes depends on feedback interactions and SMAD4 could be one of common transcriptional regulators involved in controlling their coordinated growth and patterning. For example, *Tbx2* and *Alx4* are SMAD4 targets that are part of the limb patterning GRN and function in AP limb bud axis patterning by restricting *Shh* expression in the ZPA (Kuijper et al., 2005; Nissim et al., 2007). Furthermore, *Pkdcc* participates to the PD axis patterning by it being also regulated by retinoic acid signalling, which proximalizes the early limb bud (Probst et al., 2011b). Members of the LIM homeobox TF family, such as the SMAD4 target *Lhx2,* function in the coordinated control of all three axes during limb bud outgrowth (Tzchori et al., 2009). Finally, the SMAD4 transcriptional targets *Snai1* and *Prrx2* have important functions during early limb bud development. However, as there is functional redundancy with *Snai2* and *Prrx1* respectively, only double mutants in which both paralogues have been inactivated show strong limb phenotypes (Chen and Gridley, 2013; Taher

et al., 2011). In summary, my analysis identifies the limb bud GRN of SMAD4 target genes, which reveals how the direct targets have diverse roles in key processes such as AER formation, LMP proliferation and axis patterning

7.4. SMAD4 controls the embryonic cholesterol biosynthesis and thereby modulates SHH signalling

The analysis of the SMAD4 target genes identified a second, rather unusual limb bud GRN that controls cholesterol biosynthesis by LMPs.

The embryonic cholesterol biosynthesis is enhanced by SMAD4 before the initiation of the SHH/GREM1/FGF feedback signalling system

My analysis establishes that SMAD4 directly and positively regulates the expression of several enzymes in the cholesterol synthesis pathway in early limb buds. In addition, SMAD4 also controls the expression of the sterol regulatory element binding transcription factor 1 (*Srebf1*), which a master regulator of these enzymes (see chapter 4.4.). These genes are significantly down-regulated in early *Smad4*-deficient limb buds (E10.0) as by E10.5 levels in wild-type become similarly to mutant limb buds (Figure 26). This reveals the early precocious down-regulation of several of these enzymes involved in cholesterol synthesis in *Smad4*-deficient forelimb buds. This analysis shows that SMAD4 is directly required to up-regulate several of the enzymes of the cholesterol synthesis during initiation of limb bud development prior to establishment of the SHH/GREM/FGF feedback signalling system (Figure 26). This shift from SMAD4-dependent enhancement of cholesterol synthesis during the onset of limb bud development to its decrease during feedback signalling points to specific roles during limb bud development.

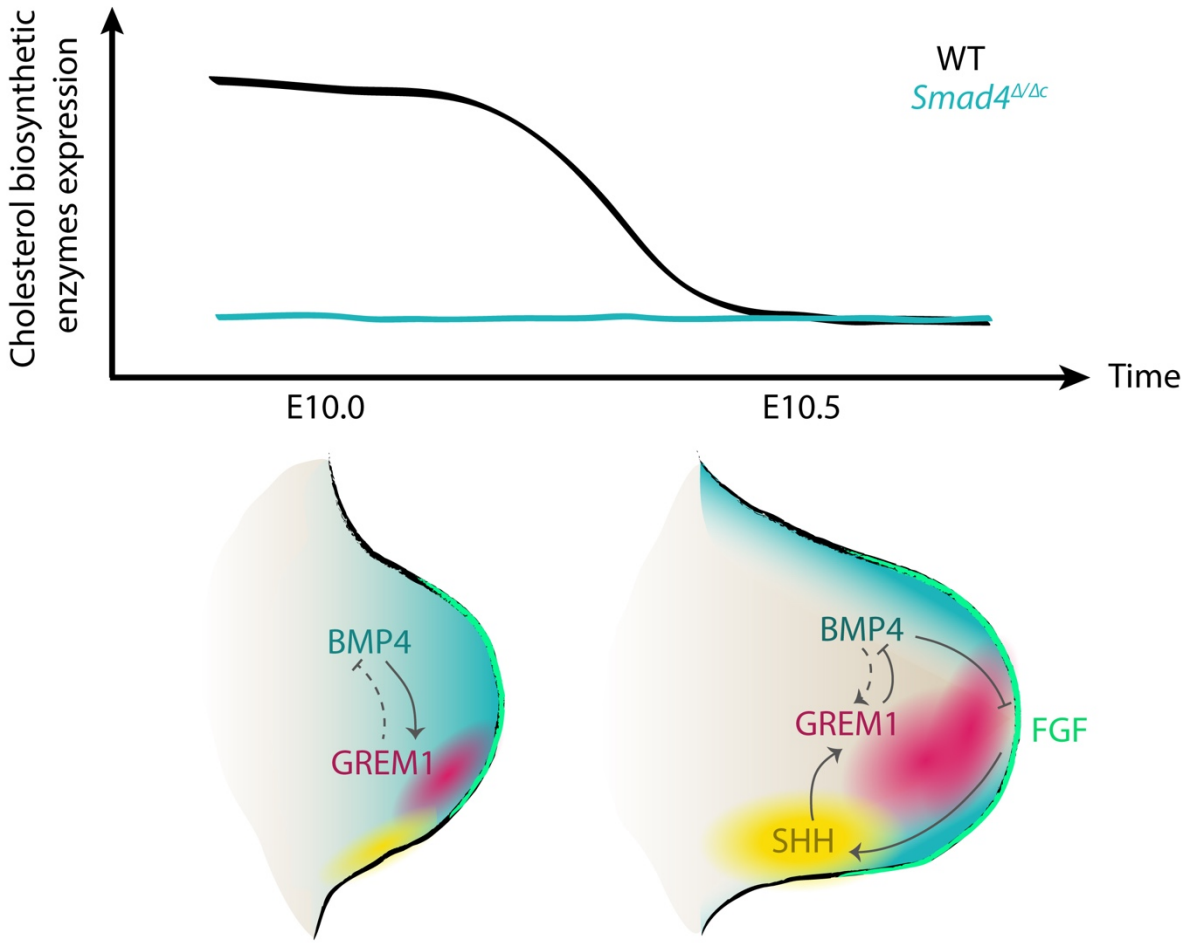


Figure 26: The expression of the enzymes for cholesterol biosynthesis in early limb buds is increased prior to establishing the feedback signalling system
Scheme showing the dynamic transcriptional regulation of the cholesterol biosynthesis enzymes in WT (black line) and *Smad4*$^{\Delta/\Delta c}$ (blue line) forelimb buds between E10.0 and E10.5. The scheme below shows the signalling interactions (E10.0) before and after establishing the SHH/GREM1/AER-FGF feedback signalling system. Blue: BMP activity; yellow: SHH; red: GREM1; green: AER-FGF signalling. Between the WT and the *Smad4*$^{\Delta/\Delta c}$ forelimb buds, the differences in expression of cholesterol biosynthesis enzymes is only apparent during the onset of limb bud development.

Balancing embryonic cholesterol synthesis with uptake of maternal cholesterol

Postnatally, the adjustment of plasma cholesterol levels is crucial as an excess is linked to an increased risk of cardiovascular diseases. The strict regulation of plasma cholesterol levels depends on the intracellular balance between cholesterol synthesis and cholesterol uptake. It has been shown that treatment with statins (a HMG-CoA reductase) directly inhibits cholesterol synthesis but

indirectly increases the cholesterol uptake by cells (Cohen, 2008, Kostner, 2007).

Cholesterol levels of cells and tissues that synthesize and take up cholesterol are controlled by regulatory feedback, which provides a possible straight forward explanation to the observed trend that cholesterol accumulation in *Smad4*$^{\Delta/\Delta c}$ forelimb buds in spite of the decreased expression of cholesterol synthesis enzymes (GC/MS appendixes 12.6, Figure 20B). As they are down-regulated in the *Smad4*-deficient forelimb buds, the cholesterol uptake may simply increase to balance the deficiency. During the developmental stages relevant to my analysis, the mouse embryo starts to synthetize its own cholesterol but still depends on maternal cholesterol (Woollett and Heubi, 2000, 2008). In *Smad4*$^{\Delta/\Delta c}$ forelimb buds the cholesterol uptake is likely enhanced further as the expression of another direct SMAD4 target, *Pcsk9* is also down-regulated. The PCSK9 enzyme modulates the amount of LDL-receptors (LDL-R) by controlling their degradation (see Chapter 4.4.) and its down-regulation may result in an increase of LDL-R levels in membranes of LMPs, which would increase their cholesterol uptake (Figure 27). As this is a working hypothesis and the situation may be more complex, I plan to directly assess the potential higher levels of LDL-R on cell membranes of *Smad4*-deficient LMPs using FACS analysis.

Altered cholesterol levels affect SHH signal transduction

One of the main functions of cholesterol during development is the regulation of SHH signalling. It was shown that cholesterol modification of the SHH signalling peptide restricts its posterior to anterior spread in the limb bud. If the cholesterol modification is disrupted, the graded distribution of SHH gradient expands anteriorly in limb buds (Li et al., 2006a). My analysis shows that while the *Shh* expression is not changed, the SHH ligand remains more posteriorly restricted during progression of *Smad4* deficient limb bud development. This is a likely

consequence of the tendency to increased cholesterol accumulation. Hence, the expression domains of the SHH transcription targets *Gli1* and *Ptch1* remains also more posteriorly restricted. One possibility is that the SHH protein is “hyper-cholesterol-modified” in *Smad4* mutant limb buds due to the increased uptake. To directly evidence the potential hyper-cholesterylation of the SHH ligand I plan to assess the cholesterol modification of SHH by comparative Western blot analysis of WT and *Smad4*$^{\Delta/\Delta c}$ forelimbs.

However, the cholesterol modification is also involved in SMO activation and thereby in the stimulation of the SHH signal transduction in responding cells (Huang et al., 2016; Xiao et al., 2017). In early *Smad4*$^{\Delta/\Delta c}$ forelimb buds, SHH signal transduction could be enhanced by increasing SMO activity in agreement with the observed transcriptional up-regulation of the SHH target genes *Hhip1* and *Ptch1*. However, *Gli1* is not up-regulated at these early stages, but the differential effect on *Gli1*, *Ptch1* and *Hhip1* expression in *Smad4*$^{\Delta/\Delta c}$ forelimbs can be explained by differences in GLI binding sites in the CRMs regulating the differential response of these three genes. It is indeed known that low and high affinity GLI binding sites are responsible for differential responsiveness to SHH signalling (Cohen et al., 2014; Oosterveen et al., 2012; Parker et al., 2011; Peterson et al., 2012). Furthermore, we have noticed that *Ptch1* is more sensitive to changes in SHH activity than *Gli1*. Hence, *Ptch1* is a better readout for SHH signal transduction than *Gli1* (data not shown). In particular, in early mouse forelimb buds at E10.0, the kinetics of *Ptch1* and *Hhip1* up-regulation might be faster than for *Gli1* in early mouse forelimb buds. Furthermore, in contrast to the GLI1 transcriptional activator, both PTCH1 and HHIP1 are strong SHH ligand antagonists (Jeong and McMahon, 2005). The observed early enhancement of SHH activity in *Smad4*-deficient limb buds will up-regulate the expression of both antagonists PTCH1 and HHIP1, which in turn will inhibit SHH signalling. This provides a straightforward explanation for the decrease in *Ptch1* expression and the up-regulation of *Hhip1* during *Smad4*$^{\Delta/\Delta c}$ limb bud development progresses. Finally, as PTCH1 and HHIP1 limit the spread and SHH signal transduction in the neural tube (Ribes and Briscoe, 2009), the up-

regulation of *Ptch1* and *Hhip1* in *Smad4*$^{\Delta/\Delta c}$ forelimbs provides an alternate explanation for the more posteriorly restricted SHH ligand in mutant limb buds (Figure 27).

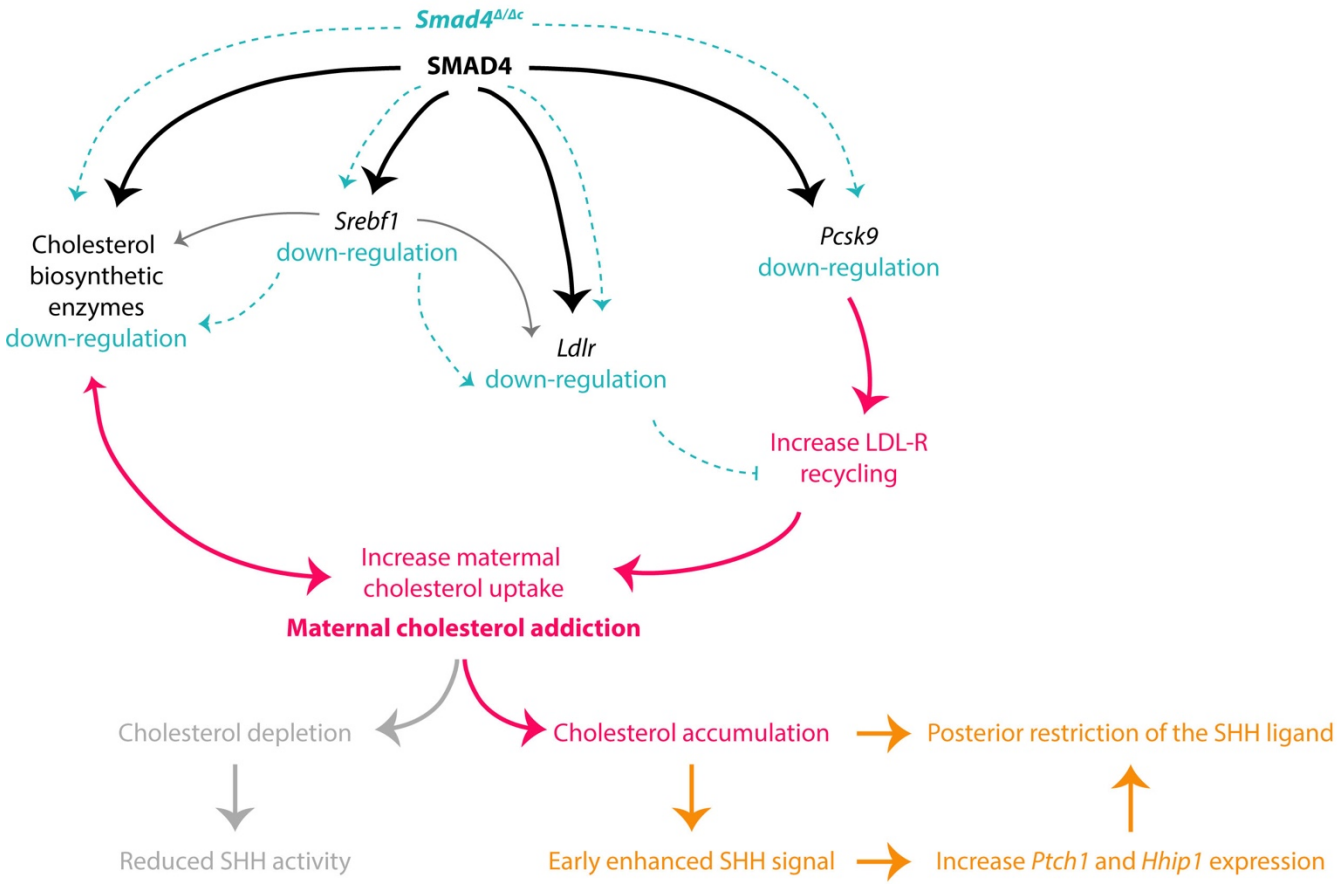


Figure 27: Causes and consequences of maternal cholesterol addiction.
SMAD4 directly regulates the cholesterol biosynthesis and uptake by positive regulation of the expression of cholesterol biosynthetic enzymes, *Srebf1*, *Ldlr* and *Pcsk9* (black arrows), which are involved in embryonic intra-cellular cholesterol synthesis. Cross-regulation between these enzymes is indicated by grey arrows. In the E10.0 *Smad4*$^{\Delta/\Delta c}$ forelimb buds, these genes are down-regulated (dashed blue arrows). This results in increased LDL-R recycling, which in turn increases uptake of extra-cellular maternal cholesterol (=maternal cholesterol addiction; red arrows). This enhances both SHH signal transduction and increased posterior retention of the SHH ligand (orange arrows). Cholesterol depletion results in reduction of SHH activity (light grey arrows).

Are Smad4$^{\Delta/\Delta c}$ LMPs addicted to maternal cholesterol?

In addition to observed alteration *in vivo*, I have been able to show that the culture of LPMs from *Smad4*$^{\Delta/\Delta c}$ and WT limb buds in lipid free medium disrupts the maintenance of SHH signal transduction in *Smad4* deficient LMPs. These results indicate that *Smad4*-deficient LMPs depend on enhanced cholesterol uptake to maintain SHH signal transduction. As E10.0 mouse embryos are mostly depending on the maternal cholesterol (Woollett and Heubi, 2000), the *Smad4*$^{\Delta/\Delta c}$ LPMs might be dependent of the maternal cholesterol (Figure 27). Therefore, our results show that SMAD4 promotes the autonomous cholesterol production of the limb bud cells. This task might be crucial for a robust and fine regulation of SHH gradient and activity, independently of the maternal diet.

8. Conclusions and Outlook

The major aim of my PhD research was to investigate of the roles of BMP signalling during the early limb bud development through the activity of the nuclear mediator SMAD4. Therefore, I have identified and functionally analysed the transcriptional gene regulatory networks controlled by SMAD4 during the onset of limb bud development.

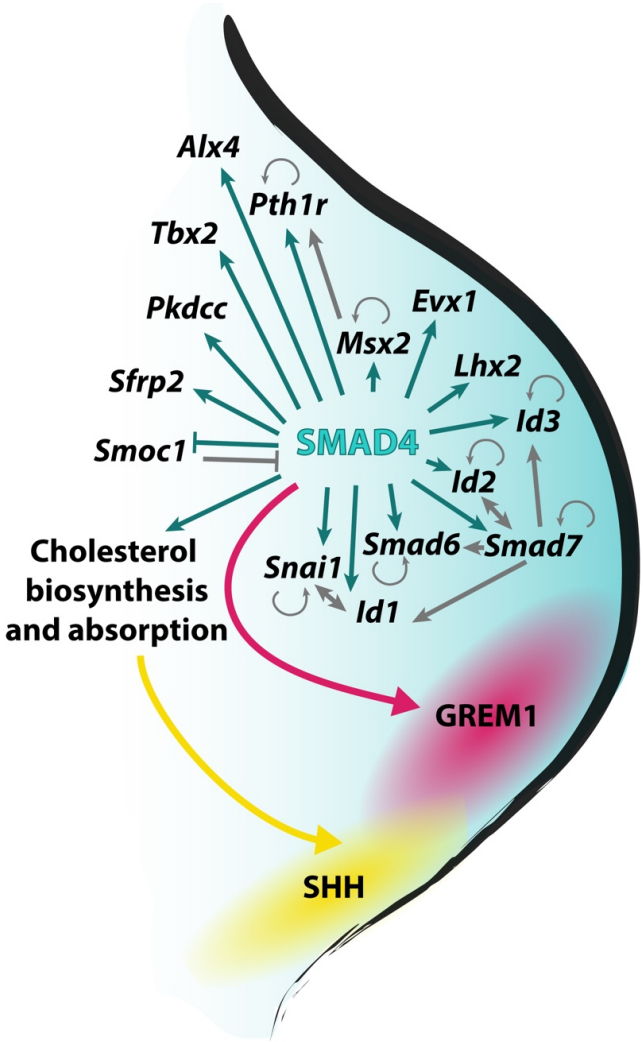


Figure 28: The gene regulatory network controlled by SMAD4 during the onset of limb bud development
SMAD4 regulates limb bud developmental genes in the anterior limb bud (in blue), initiates *Grem1* expression (in red) and regulates the SHH gradient and signal transduction by directly controlling cholesterol biosynthesis and uptake (in yellow).

Using different genome-wide approaches (ChIP-seq, ATAC-seq, RNA-seq), mouse molecular genetics and cell-biochemistry has allowed me to uncover and study the facets of SMAD4 functions during initiation of limb bud development. The different key functions of SMAD4 during early limb bud development are driven by high mesenchymal BMP activity. First, SMAD4 initiates *Grem1* expression, which is a key step in inaugurating the SHH/GREM1/FGF feedback signalling system that regulates limb bud outgrowth and proliferative expansion and patterning of mesenchymal progenitor. By activating *Grem1* in the posterior and restraining *Smoc1* in the anterior, SMAD4 creates an BMP antagonist-free region specifically in the anterior limb bud mesenchyme. This in turn allows BMPs to regulate the expression of many genes involved in the anterior part of the early limb bud by SMAD4-mediated transcriptional regulation. Most interestingly, SMAD4 directly fine-tunes the spatial distribution of the SHH ligand, reception and transduction of the SHH signal in early limb buds by directly controlling the expression of genes involved in embryonic cholesterol biosynthesis and uptake of maternal cholesterol. These multi-functional control of the onset of limb bud development by SMAD4 is crucial to correct establishment of the feedback signalling system and orderly progression of subsequent limb bud development.

I have discovered *in vivo* this novel role of BMP/SMAD4 mediated signal transduction in regulating SHH signalling via controlling cholesterol synthesis and levels during the onset of limb bud development. We are now wondering whether the BMP/SMAD4 mediated control of cholesterol metabolism and SHH signal transduction is a limb bud-specific or general mechanism. The developing spinal cord is another excellent model to study the opposing effects of the BMP and SHH signalling pathways (Jessell, 2000). Therefore, we have established a collaboration with Dr. Anna Kicheva (IST, Vienna, Austria) to investigate the effects of the *Smad4* deficiency in the neuronal progenitors of the spinal cord using an inducible *Sox2*-CRE transgene for spatio-temporally controlled inactivation. A recent paper shows that inactivating the cholesterol modification of SMO alters the identity of the ventral neuronal progenitors in the

spinal cord (Xiao et al., 2017). This study encourages us to investigate if SMAD4-mediated regulation of cholesterol biosynthesis also impacts on SHH signal transduction and neuronal patterning in the developing spinal cord.

9. Materials and Methods

9.1. Mouse husbandry and embryo analysis

9.1.1. Ethics statement

All experiments with mice were performed with strictly respecting Swiss laws, the 3R principles and of the Basel Declaration. They were classified as grade 0, which implies no or only minimal suffering. The researchers are licensed for animal experimentation by successfully completing the LTK-1 course and fulfilling the continuing animal education course requirements.

9.1.2. Mouse strains

For this study, the following mouse strains were used: the *Grem1* null (Michos et al., 2004), *Smad4*^{3xF} (generated by Frédéric Laurent) and *Prx1-Cre* transgene (Logan et al., 2002) alleles, which are maintained in a NMRI background. The *Smad4 conditional* (*Smad4*$^{flox/flox}$, Yang et al., 2002) allele is maintained in a mixed background. To inactivate *Smad4* in the mesenchyme, *Prx1-Cre*$^{Tg/+}$ mice were crossed with mice carrying the *Smad4*$^{\Delta}$ allele to generate the *Prx1-Cre*$^{Tg/+}$*Smad4*$^{\Delta/+}$ mice. *Prx1-Cre*$^{Tg/Tg}$ *Smad4*$^{\Delta/+}$ males were crossed with *Smad4*$^{flox/+}$ females to obtain experimental embryos that carry a constitutive *Smad4* null allele and a conditionally inactivated *Smad4* allele (*Prx1-Cre*$^{Tg/+}$*Smad4*$^{\Delta/\Delta c}$, referred as *Smad4*$^{\Delta/\Delta c}$), and control embryos *(Prx1-Cre*$^{Tg/+}$*Smad4*$^{+/+}$, referred as WT). All mice and embryos were genotyped by PCR using the primers listed in Table 9.7.2.

9.1.3. Generation of LacZ reporter transgenic embryos

Candidate CRM regions were amplified by PCR from mouse genomic DNA (for primers see Table 9.7.1). They were then cloned into a Hsp68-LacZ reporter vector using the Gibson Assembly® Method (performed by Jens Stolte). Transgenic embryos were generated by pronuclear injection at the Center for Transgenic Models, Basel. Founder embryos were stained for *LacZ* activity as described below.

9.1.4. Whole-mount *LacZ* staining of mouse embryos

Embryos were isolated in ice-cold PBS and staged by counting their somite numbers, before fixation in 1% formaldehyde, 0.2% glutaraldehyde, 0.02% NP40, 0.01% sodium deoxycholate in PBS for 20 min at 4°C. Embryos were then washed 3 times in PBS for 5min at room temperature (RT), and incubated for at least 4hrs in the dark in 1mg/mL X-Gal in dimethyl formamide, 0.25mM K3Fe(CN6), 0.25mM K4Fe(CN6), 0.01% NP40, 0.4mM $MgCl_2$. Cells expressing β-galactosidase activity turn blue. The reaction was stopped by washing the embryos 3 times in PBS for 5min at RT.

9.1.5. Whole-mount *in situ* hybridization (WISH)

Embryos were isolated in cold-PBS and transferred in a 2mL Eppendorf tube to be fixed overnight in 4% paraformaldehyde (PFA) in PBS at 4°C. The next day they were rinsed twice in PBS-0.1% Twen-30 (PBT) and dehydrated in a graded series of 25%, 50%, 75% MetOH/PBT and 2 times in 100% MetOH (5min each). The dehydrated embryos were stored at -20°C in 100% MetOH. For the first day of the WISH, embryos were rehydrated in reverse gradient series of 75%, 50%, 25% MethOH/PBT and washed twice in PBT (5min each). Embryos were bleached in 6% hydrogen peroxide (H_2O_2)/PBT for 15min and washed 3 times in PBT (5min each). This step was followed by treatment with 10μg/mL

proteinase K in PBT for 15min. To inactivate the proteinase K, embryos were washed with fresh 2mg/mL glycine in PBT for 5min. After 2 washes in PBT (5min), embryos were re-fixed in fresh 0.2% glutaraldehyde, 0.1% Tween-20 in 4% PFA (in PBS) for 20min, and rinsed again twice in PBT (5min). Then, they were equilibrated in 2mL of prewarmed prehybridization buffer (50% deionized formamide, 5xSSC pH 4.5, 2% BCl blocking powder, 0.1% Tween-20, 0.5% CHAPS (Sigma C3023), 50μg/mL yeast RNA (Sigma R8759), 5mM EDTA, 50μg/mL heparin (Sigma H5515) at 70°C for at least 1hr. The prehybridization buffer was then replaced by 1mL of fresh prewarmed prehybridization buffer containing 10μl/mL of digoxigenin-labelled RNA riboprobe (see below) and incubated overnight at 70°C.

On the second day, the riboprobe-containing prehybridization buffer was recovered and replaced by pre-warmed prehybridization buffer, followed by a series of 75%, 50% and 25% prehybridization buffer in 2xSSC (0.3M NaCl, 0.03M sodium citrate pH 4.5). All these steps were performed for 5min at 70°C. Embryos were then washed twice in 2xSSC, 0.1% CHAPS for 30min at 70°C in a rotating wheel. To remove unbound RNA probe, embryos were treated with 20μg/mL RNase A (Roche 10109169001) in 2xSSC,0.1% CHAPS for 45min at 37°C. This was followed by 2 washes at RT of 10min using 100mM maleic acid disodium, 150mM NaCl pH 7.5. Then 2 additional washes at 70°C for 30min were done. Next, the embryos were washed for 3 times 5min each in fresh TBST (140mM NaCl, 2.7mM KCl, 25mM Tris-HCl pH 7.5, 1% Tween-20). Embryos were then blocked in 10% lamb serum/TBST for 1hr or more, at RT. This blocking solution was replaced with a solution containing anti-digoxigenin-Alkaline Phosphatase Fab Fragments (Roche 11093274910) diluted 1:5000 in 1% lamb serum in TBST. Embryos were incubated overnight at 4°C with gentle rocking.

On the third day, embryos were washed 3 times in TBST for 5min and 5 times for 1hr to 1.5hrs at RT. The last wash was overnight at 4°C.

On the fourth day, the embryos were equilibrated 3 times in NTMT (100mM NaCl, 100mM Tris-HCl pH 9.5, 50mM $MgCl_2$, 1% Tween-20) for 10min each, and transferred into 1mL of BM purple Alkaline Phosphatase substrate (Roche) at RT. The *in situs* were developed in the dark and checked roughly every hour until reaching appropriate probe-specific signal intensity. The development was stopped by washing the stained embryos 5 times in PBT (10min) and twice in PBS. Pictures were taken using a Leica MZ FLII stereomicroscope and the Leica Application Suite V3 software. Embryos were stored in 4% PFA (in PBS) at 4°C.

9.1.6. Digoxigenin-labelled RNA probe preparation

The plasmids containing the cDNA of interest were linearized and transcribed using the appropriate T3, T7 or SP6 RNA polymerase with the DIG RNA labelling kit (Roche). The dig-UTP riboprobes were purified using the mini Quick Spin RNA columns (Roche). For the WISH screen, all cDNAs of interest were generated by PCR amplification (see Table 9.7.4.) using primers that include the binding site for SP6. After transcription, riboprobes were also purified. Before use, the riboprobes were heated at 85°C for 5min and equilibrated in prehybridization buffer at 70°C. Probes in prehybridization buffer were stored at -20°C and re-used several times.

9.1.7. Culture of limb mesenchymal progenitors (LMPs)

Forelimbs (32-34 somites) were collected in ice-cold PBS and incubated in cold 2% Trypsin (Gibco 15090-046)/PBS at 4°C for 30min. The reaction was stopped by adding an excess of EMFI medium containing serum: DMEM (Gibco 41966029) containing 4.5g/L Glucose (Gibco 41966-029), 10% FCS (PanBiotech P30.3302), 100U Penicillin, 0.1mg/mL Streptomycin (Sigma P-0781) and 200mM L-Glutamine (Sigma G-7513). The limb bud ectoderm was removed using forceps and a sharp dissection needle. Limb buds were transferred into a 1.5mL tube containing 500μL EMFI medium and dissociated as single cells by pipetting up-and-down. Limb bud cells were centrifuged at

2.5rpm for 5min at 4°C. After resuspending the pellet into 500mL EMFI medium using lipid depleted FCS (see protocol below), cells were plated in two wells of a 96 well plate. After 8-9hrs of culture, one well of each duplicate was treated for 24hrs 0.5μM SAG (Millipore 566660) to activate SHH signal transduction. The day after, cells were harvested, flash frozen in RLT buffer (Quiagen) and stored at -80°C for subsequent RNA extraction and RT-qPCR analysis.

Lipid depleted FCS preparation: 500mL FCS were stirred overnight at 4°C with 10g Cab-osil M-5 (ACROS Organics 7631-86-9). The mix was then centrifuged for 10min at 3000rpm and the supernatant filtrated under sterile conditions. Insulin-Transferrin-Sodium-Selenite media supplement (Sigma I-1884-1) was dissolved in 50mL water, acidified by adding 250μL HCl and filtrated under sterile conditions. 25mL of the Insulin-Transferrin-Sodium-Selenite solution was added to 500mL of lipid depleted FCS. Aliquots of 30mL were frozen at -20°C and used for preparing lipid depleted EMFI medium.

9.1.8. Limb bud collection for GC/MS

E10.0 (30 somites), E10.5 (35 somites) and E11.0 (40 somites) forelimbs were collected in ice-cold PBS. Two forelimbs from one embryo were recovered with a P1000 pipette and flash frozen in a drop of PBS into a 2mL Eppendorf tube containing liquid nitrogen. After evaporation of the liquid nitrogen forelimbs were stored at -80°C.

9.1.9. Skeletal preparations

The skeletal preparations were done by Nathalie Riesen. E14.5 embryos were isolated in ice-cold PBS and euthanized by cutting the head. They were incubated overnight in tap water at room temperature. The day after, embryos were scalded in hot water (65-70°C) for 20-30 seconds, eviscerated (liver biopsies were taken for genotyping), and the skin of the limbs were peeled off. Embryos were then fixed in 95% ethanol for a minimum of 24 hrs. Alcian blue staining (30mg Alcian Blue 8GX (Sigma), 85% ethanol, 20% glacial acetic acid)

was performed overnight and followed by 3 washes in 95% ethanol on the second day. Embryos were then cleared in 1% potassium hydroxide (KOH) for 10min, counterstained with Alizarin red (50mg Alizarin red (Sigma) in 1% KOH) for 1hr and cleared in 1% KOH for an additional hour. Embryos were then moved through progressively higher ratios of Glycerol:1% KOH (80:20, 60:40, 40:60 and 20:80) according to their speed of clearing and stored in 80% glycerol in water. Pictures were taken using a Leica MZ FLII stereomicroscope.

9.2. Generation of transgenic mice using CRISPR/*Cas9* genome editing

9.2.1. ES cell targeting

The design of the CRISPR strategy and the targeting of ES cells were performed by Nathalie Riesen and Aimée Zuniga. SgRNA (see Table 9.7.6.) were designed using the website http://crispr.mit and were cloned in the px459 vector. For transfection, 300000 G4 ES cells were plated in a 6cm dish on a monolayer of EMFIs in ES medium: DMEM (Gibco 41966029), 15% FCS (PanBiotech P30.3302), 100U Penicillin, 0.1mg/mL Streptomycin (Sigma P-0781), 200mM L-Glutamine (Sigma G-7513), 1mM βMercapto-Ethanol (Gibco 31350-010), 10^7U/mL EsGRO LIF (Gibco 13275-029), 1x Non-Essential Amino Acids (Gibco 11140-035), 100mM Sodium Pyruvate (Gibco 11360-39). The next morning, the medium was changed for 1.75mL ES medium per plate, and the evening cells were transfected by using the FuGENE method (Promega TM-328). Plasmids (5μg) were diluted in 125μL OptiMEM (Gibco 51985-026) and 25μL FuGENE were mixed with 100μL of OptiMEM. 125μL DNA-mixture was combined with 125μL of diluted FuGENE reagent and incubated for 15min at RT. 250μL of this mix were then added per dish. After 12hrs, the transfection was stopped by changing the medium, and 24hrs later cells were split into new 6cm dishes (1:3) containing DR4 resistant feeders. Cells were selected using

puromycin (2μg/mL, Sigma P8833) for 48hrs, and then the medium was replaced by ES medium. The positive clones were grown for 4 to 6 days and picked to be expanded and screened for the correct deletion. The positive clones were used to generate transgenic mice by aggregation chimera.

9.2.2. Superovulation of oocyte donor females

On day 0, female NMRI mice of 13 weeks were injected intraperitoneally using 25G needle with 5 IU PMSG (Pregnant Mare Serum Gonadotropin – *Pregnyl* from Organon). On day 2, mice were injected with 5 IU hCG (*Folligon* from Intervet). PMSG and hCG were dissolved in sterile PBS to 50 IU/mL and aliquots of 1mL were stored at -20°C. The hCG injection was done 46 to 47hrs after the PMSG injection and matings were set using these primed females and NMRI males.

9.2.3. Pseudo-pregnant females

On day 1, bedding from a NMRI male was added to each cage of NMRI females. Mating between these females and vasectomized males were set on day 3.

9.2.4. ES cell preparation for aggregation

ES cells for aggregation were prepared by Nathalie Riesen. On day 5, ES cells from a confluent 6 cm dish were trypsinized and dissociated by pipetting up and down in ES medium. Cells were then pre-plated for 30min to remove excess EMFIs. ES cells were collected and counted. After centrifugation (1200rpm, 5min), ES cells were re-suspended to $1.5x10^6$cells/mL in aggregation medium: 18mL DMEM (Gibco 41966029), 66mg Ca-lactate (Sigma 21185) and 4% FCS. Drops of 50μL ES cells were made on a 10cm Petri dish, covered by mineral oil (Sigma M5310) and incubated 10min before aggregation.

9.2.5. Embryo collection for aggregation

All the solutions must be pre-warmed at 37°C. All the steps of the aggregation were performed by mouth pipetting.

On day 5, Embryos (E2.5, morula stage) were flushed out of uteri using M2 medium injected through the infundibulum. They were pooled in a 3cm dish in M2 Medium. At this stage, a selection was performed to keep only round embryos with a nicely formed *zona pellucida*, and 8 to 16 blastomeres and without black residue inside. These embryos were then treated with Tyrode's acid solution (Sigma T1788) for few seconds to remove the *zona pellucida*. Then they were washed 3 times in M2 medium.

Composition of the M2 Medium: 6.4mg/mL NaCl (Merck 1064041000), 350μg/mL $NaHCO_3$ (Merck 1063290500), 36μg/mL Na-Pyruvate (Merck 1066190050), 50μg/mL Streptomycin. Sulf. (Sigma 56501) 160μg/mL KH2PO4 (Merck 1048731000), 465μg/mL Ca-Lactate.3$H_2$0 (Sigma 44388), 356μg/mL KCl (Merck 1049361000), 294μg/mL MgSO4.7$H_2$0 (Merck 1058860500), 1mg/mL Glucose (Sigma G8270), 621μg/Ml HEPES (Sigma 54457), 75μg/mL K-PenG (Sigma P7794) and 4mg/mL BSA (Sigma A3311) in Aqua ad inject (Braun/Aichele Medico 530108).

9.2.6. ES cell-embryos aggregation

10 embryos were gently placed in a drop of ES cells; without touching the monolayer of ES cells and incubated for 2hrs at 37°C 5% CO_2. During this time, Petri dishes with 50μL drops of KSOM (Millipore MR-106-D) covered with mineral oil were prepared and prewarmed at 37°C. Aggregated embryos were then gently removed, the excess of cells detached by pipetting up and down. For generating good chimeras, an embryo must be aggregated with 5 to 10 ES

cells. Ten aggregated embryos were placed in a drop of KSOM, well separated from each other and incubated overnight at 37°C in 5% CO_2.

9.2.7. Preparation for transfer

Around 20hrs after aggregation, embryos were mostly blastocysts and were ready for transfer. They were washed once in prewarmed M2 medium before to be kept in M2 at 37°C during the transfer.

9.2.8. Embryo transfer

Pseudo-pregnant females were anesthetized with a mix of Ketamine (100mg/mL), Xylazine (20mg/mL) and Acepromazine (10mg/mL). Subsequently they were shaved on the back, and a protective gel was added on the eyes. For the surgery, a first incision was made in the skin, on the middle line of the back, just posterior to the ribs. This first incision was used for transferring embryos to both uterus horns. The ovary was visible by transparency and a second incision was made in the peritoneum. The ovary was pulled out by the attached fatpad and clamped to keep the uterus out of the body wall. Ten embryos were loaded into a glass transfer capillary between two air bubbles. Using a needle, an incision was made in the uterus proximal to the oviduct and used to transfer the embryos. The uterus horn was placed back into the body cavity and the transfer was repeated from the other uterus horn. Finally, the incision on the mouse back was closed using sterile suture clips. These clips were removed 10 days after surgery.

9.3. Molecular biology

9.3.1. Chromatin Immunoprecipitation (ChIP)

For ChIP-qPCR experiments, each duplicate contains dissected tissues from 45 *Smad4*$^{3xF/3xF}$ embryos at E9.75 (forelimbs with attached trunk) or E10.5

(forelimbs/hindlimbs) or E12.5 (forelimbs/hindlimbs). For ChIP-seq experiments, each duplicate contains dissected tissues from 80 *Smad4*$^{3xF/3xF}$ embryos at E9.75 (forelimbs with attached trunk) or 100 *Smad4*$^{3xF/3xF}$ embryos E10.5 (forelimbs/hindlimbs) or 60 *Smad4*$^{3xF/3xF}$ embryos E12.5 (forelimbs/hindlimbs).

The ChIP protocol for embryonic tissues was setup by Marco Osterwalder in the group (Osterwalder et al., 2014). *Smad4*$^{3xF/3xF}$ embryos were collected in ice-cold PBS and dissected tissues were pooled in a 2mL Eppendorf tubes. After 2 washes in ice-cold DPBS containing Ca^{2+} and Mg^{2+} (Gibco), the dissected tissues were transferred to a 2mL glass douncer (Tissue Grind Tube Size 2mL, Kimble-Chase) on ice. Cells were disaggregated applying 25 stokes with the pestle A (Tissue Grind Pestle LC 2mL, Kimble-Chase) and 25 stokes with the pestle B (Tissue Grind Pestle SC 2mL, Kimble-Chase). The nuclei solution was transferred back to the 2mL Eppendorf tube and the douncer was rinsed with 300μL cold DPBS w/ $Ca^{2+}Mg^{2+}$ that were added to the sample. The solution was then centrifuged at 3000rpm for 3min (4°C) and the supernatant was discarded. The nuclei pellet was resuspended in 1.5mL RT DPBS w/ $Ca^{2+}Mg^{2+}$ containing 150μL 11X crosslinking buffer (0.1M NaCl, 1mM EDTA, 0.5mM EGTA, 50mM HEPES pH 8.0; with 11% formaldehyde added before using). The nuclei solution is crosslinked for 5min at RT on a horizontal shaker. 75μL 2.5M Glycine were added to stop the reaction, followed by 5min incubation at RT with horizontal shaking. The sample was then centrifuged for 3min at 3000rmp (4°C) and the pellet was re-suspended in 1.5mL ice-cold DPBS w/ $Ca^{2+}Mg^{2+}$. After re-centrifugation for 3min at 3000rpm (4°C), the supernatant was discarded and the nuclear pellet was flash-frozen in liquid nitrogen and stored at -80°C.

The day before the ChIP, antibodies were coupled to magnetic beads. For one ChIP sample, 20μL of Dynabeads® Protein G (Invitrogen) were rinsed 6 times in 1mL freshly prepared ice-cold BSA (5mg/mL in DPBS w/ $Ca^{2+}Mg^{2+}$) using a magnetic rack. Beads were re-suspended in a volume corresponding to 2.5x the original volume and transferred to a 2mL screw cap tube (Sarstedt). 2μg of anti-FLAG M2 antibody (Sigma, F1804) were added and the mixture was

incubated overnight on a rotating wheel at 4°C to couple the antibody with the beads. The following day, the frozen nuclear pellets were thawed on ice and re-suspended in 6mL (for ChIP-qPCR) or 30mL (for ChIP-seq) cold Lysis buffer (50mM HEPES pH 7.5, 140mM NaCl, 1mM EDTA pH 8.0, 10% Glycerol, 0.5% NP40, 0.25% Triton X-100, 1x Complete Mini protease inhibitor cocktail from Roche). Several frozen nuclear pellets were pooled in Lysis buffer to reach the correct number of cells. Lysates were incubated at 4°C for 10min with rocking and centrifuged at 2500rpm for 10min (4°C). The supernatant was discard and the pellet resuspended in 5mL (for ChIP-qPCR) or 24mL (for ChIP-seq) of Protein extraction buffer (0.2M NaCl, 1mM EDTA pH 8.0, 0.5mM EGTA pH 8.0, 10mM Tris-HCl pH 8.0, 1x Complete Mini protease inhibitor cocktail). After 10min of incubation at 4°C on a rocking platform, the sample was centrifuged at 2500rpm for 10min (4°C). The supernatant was discarded and the nuclear pellet was re-suspended in 1mL (for ChIP-qPCR) or 5mL (for ChIP-seq) of chromatin extraction buffer (1mM EDTA pH 8.0, 0.5mM EGTA pH 8.0, 10mM Tris-HCl pH 8.0, 0.1% Na-Deoxycholate, 0.5% N-lauroylsarcosine, 3x Complete Mini protease cocktail). The sample was transferred into Covaris TC 12x12 tubes (5 tubes of 1mL for ChIP-seq) and DNA was sheared using the Covaris Ultra-sonicator S220 for 15min (5% duty cycle, 140 watts Peak Incident Power and 200 cycles per Burst). During the time of sonication, Dynabeads®-antibody coupled complexes were rinsed 6 times in 1mL cold BSA/PBS and finally re-suspended in the original volume (20μL per sample) and kept on ice.

After sonication, the sheared chromatin was transferred to a new 1.5mL Eppendorf tube and centrifuged at 1300rpm for 10min (4°C). The supernatant was transferred into a 2mL screw cap tube to adjust the volume to 1.060mL for ChIP-qPCR or 5.060mL for ChIP-seq with Chromatin extraction buffer. Two aliquots of 30μL were taken and kept on ice: one for input control and one to run on an agarose gel to check at the quality of sonication. For the immunoprecipitation (IP), 300μL of Cocktail mix were added (for one sample 130μL Triton X-100 10%, 3μL Na-Deoxycholate 10%, 26μL Complete protease inhibitor solution (from a 50x tablet dissolved in 1mL mQ H_20), 131μL TE buffer

(100mM Tris HCl pH 7.4, 10mM EDTA pH 8.0) and 10μL mQ H_20) to 1mL sample aliquots (1 aliquot for ChIP-qPCR and 5 aliquots for ChIP-seq). For each tube, 20μL of the freshly rinsed Dynabeads®-antibody coupled complexes were added and incubated for 6hrs on a rotating wheel (4°C). After IP, the samples were recovered in a 1.5mL Eppendorf tube and 1mL of fresh cold RIPA buffer was added (50mM HEPES pH 8.0, 1mM EDTA pH 8.0, 1% NP40, 0.7% Na-Deoxycholate, 0.5M LiCl, 1x Complete Mini protease inhibitor cocktail), followed by 6 washes with 1mL of cold RIPA buffer on a magnetic rack. The beads were rinsed once with 1mL TE-plus (100mM Tris HCl pH 8.0, 10mM EDTA pH 8.0, 50mM NaCl, 1x Complete Mini protease inhibitor cocktail) and centrifuged at 1000rpm for 3min (4°C). On the magnetic rack, the TE-plus was removed and the beads were re-suspended in 100μL of fresh elution buffer (1mM EDTA pH 8.0, 1% SDS, 10mM Tris-HCl pH 8.0). Elution of chromatin-protein complexes from the Dynabeads® was processed for 15min at 65°C, with a strong shacking at 1300rpm. After centrifugation at 13000rpm for 1min, the elution buffer (containing the chromatin-protein complexes) was removed on the magnetic rack and transferred in a PCR microtube. The 30μL aliquots (for input control and gel verification) were also transferred into PCR microtubes with 120μL of Elution buffer. The reverse crosslinking was performed in a PCR machine at 65°C overnight (±19hrs).

The next day, 5 aliquots for ChIP-seq were pooled, and transferred to a new 1.5mL Eppendorf tube. For ChIP-qPCR, the samples were also individually transferred to a 1.5mL Eppendorf tube. They were treated with 0.2μg/μL RNase A (from bovine pancreas, Sigma) in TE (100mM Tris HCl pH 8.0, 10mM EDTA pH 8.0) for 1hr at 37°C. Then they were treated with 0.2μg/μL Proteinase K for 2hrs at 55°C. Finally, the chromatin was purified using the QIAquick Gel Extraction Kit (Qiagen). The DNA was eluted in two steps of 30μL EB buffer. All samples (ChIP, input and for gel verification) were aliquoted and stored at -20°C. To verify the quality of the sonication, 30μL of the sample for gel verification was loaded on a 1.5% agarose gel to estimate the range of DNA fragments sizes (with the maximal intensity between 100 and 300bp).

9.3.2. ChIP-seq library construction and sequencing

Library construction was performed using the KAPA Hyper Prep Kit (ref KK8502). The samples were sequenced by the Genomics Facility Basel - ETH Zürich using an Illumina NextSeq 500 system (SR75).

9.3.3. ChIP-qPCR

To determine the enrichment of SMAD4 chromatin complexes on regions of interest (ROI), duplicates of ChIP and input samples were analysed using qPCR. The reaction was performed in 20μL containing 0.3μM of each primer diluted in EB buffer (10mM Tris-HCl pH 8.5), 50% of SYBR green and 1μL of chromatin (for primers see Table 9.7.7.). The program for qPCR is the same as the one used for RT-qPCR. A region in the *β-actin* locus (Galli et al., 2010) was used as normalizing control region (NCR). Fold-change enrichments between input and ChIP samples were calculated as follows:

(CqROI ChIP) - (CqROI input) = ΔCqROI

(CqNCR ChIP) – (CqNCR input) = ΔCqNCR

Fold enrichment = $2^{-(\Delta CqROI - \Delta CqNCR)}$

ChIP-qPCR results are shown as mean ± standard deviation (SD) and are based on the analysis of 2 biological replicates. To avoid artefact bias of fold enrichment, a Cq of 32 was defined as minimum background value threshold.

9.3.4. ATAC-seq

The ATAC-seq was performed by following the protocol of Buenrostro et al., 2013, for 2 replicates. One replicate consists of a pair of forelimbs from a E9.5 wild-type (WT) embryo with the attached portion of trunk (for carrier DNA as in the ChIP-seq experiment). After two washes in ice-cold PBS, dissected tissues

were transferred to a 2mL glass douncer (Tissue Grind Tube Size 2mL, Kimble-Chase) on ice. Cells were disaggregated by applying 20 stokes with the pestle A (Tissue Grind Pestle LC 2mL, Kimble-Chase) and 20 stokes with the pestle B (Tissue Grind Pestle SC 2mL, Kimble-Chase). After cell counting, the volume corresponding to 75000 cells was centrifuged for 5min at 2300rpm (4°C). The supernatant was gently removed and the samples were rinsed with 100μL ice-cold PBS and centrifuged for 5min at 2300rpm (4°C). The PBS was discarded and the samples were lysed in 50μL Lysis buffer (10mM Tris-HCl pH 7.5, 10mM NaCl, 3mM $MgCl_2$ and 0.1% NP40) by pipetting them gently up-and-down and centrifuging for 10min at 2300rpm (4°C). After discarding the supernatant, the pellet of nuclei was transposed for 30min at 37°C using the Nextera DNA Prep Kit (Illumina 15028212) in a 50μL reaction mix (25μL TD (2x reaction buffer), 2.5μL TDE1 (Nextera Tn5 Transposase) and 22.5μL nuclease-free H_20). Immediately following the transposition, the samples were purified using a Qiagen MinElute PCR Purification Kit and the transposed DNA was eluted in 13μL EB buffer (Tris-HCl pH 8.5) and stored at -20°C. The second day, 10μL transposed DNA fragments were amplified by PCR using 25μL KAPA HiFi HotStart Ready Mix (Roche 07958927001) with 25μM PCR Primer 1 and 25μM Barcoded PCR Primer 2.1 for the first replicate and Primer 2.2 for the second replicate in a total volume of 50μL.

PCR Primer 1	5'-AATGATACGGCGACCACCGAGATCTACACTCGTCGGCAGCGTCAGATGTG-3'
Barcoded PCR Primer 2.1	5'-CAAGCAGAAGACGGCATACGAGATTCGCCTTAGTCTCGTGGGCTCGGAGATGT-3'
Barcoded PCR Primer 2.2	5'-CAAGCAGAAGACGGCATACGAGATCTAGTACGGTCTCGTGGGCTCGGAGATGT-3'

PCR program details:

72°C	5min	
98°C	30sec	
98°C	10min	13 cycles
63°C	30sec	
72°C	1min	
4°C	∞	

The amplified library was purified using QIAquick PCR purification kit and eluted in 30μL EB buffer. Sequencing was done using an Illumina NextSeq 500 system (SR75) by the Genomics Facility Basel - ETH Zürich.

9.3.5. RNA-seq

Dissected WT and *Smad4*$^{\Delta/\Delta c}$ forelimbs from E10.0 embryos (30 somites) and 10.5 embryos (35 somites) were collected in RNAlater® (Sigma R0901), incubated overnight at 4°C and then stored at -80°C. Both forelimbs of an embryo were pooled for one replicate. After genotyping, 4 age-matched replicates of *Smad4*$^{\Delta/\Delta c}$ and 3 replicates of WT forelimb buds per stage were sequenced. RNA was extracted using the Qiagen RNeasy micro kit. For each replicate, the quality of total RNA was analysed using the RNA 6000 Pico kit (Agilent 2100 bioanalyzer), which was followed by polyA RNA library preparation. Sequencing was done on the HiSeq 2500 machine using the single-read 50 cycles protocol. Library construction and sequencing were performed by the Genomics Facility Basel - ETH Zürich.

9.3.6. RNA extraction and Real Time-quantitative PCR (RT-qPCR)

For RT-qPCR analyses of cultured limb bud mesenchymal cells (LMPs), total RNA was extracted using the RNeasy Micro Kit (Qiagen) and eluted in 14μL RNAse-free water. To synthetize cDNA, 10μL RNA were mixed with 1μL oligo(dT)$_{12\text{-}18}$ (500ng) and 1μL dNTP mix (10mM each of dATP, dTTP, dGTP and dCTP). The reaction mixture was denatured at 65°C for 5 min to allow annealing and left on ice for at least 1min. Samples were transferred in PCR microtubes and the following reagents were added: 8μL 5x First Strand Buffer, 1μL 0.1M DTT, 1μL RNAseOUT™ 40U/μL, 1μL Superscript™ III RT 200U/μL (all from Invitrogen). Samples were incubated for 1hr at 50°C and then 15min at 70°C (in a PCR machine). cDNA samples were stored at -20°C.

The RT-qPCR reaction was done using the Bio-Rad CFX96 Real-Time PCR system using the iQ SYBR Green Supermix (Bio-Rad). Each PCR reaction was done in 20μL containing 0.3μM each primer, 50% SYBR green and either 0.5μL cDNA, or 1μL DNA diluted 1:100 in MiliQ H2O (for *Cre* genotyping) or 1μL ChIP/Input samples (for ChIP-qPCR). The program for qPCR is:

95°C	3min	
95°C	10sec	13 cycles
60°C	1min	
95°C	10sec	
65°C	5sec	
95°C	5sec	

For gene expression analysis, the *ribosomal protein L19* (*RPL19*) transcript was used as normalizer. Relative Cq values of the target transcripts were normalised to the Cq values of *RPL19*, and normalized fold expression level ($2^{-\Delta\Delta Cq}$) are shown as mean ± SD.

For genotyping, the *Dopamine Beta-Hydroxylase (DBH)* genomic region was used as a normalizer. Results are reported as mean ± SD of normalized expression level ($2^{-\Delta\Delta Cq}$).

9.3.7. Gas Chromatography / Mass Spectrometry (GC/MS)

The Gas Chromatography / Mass Spectrometry (GC/MS) was performed by Dr. Dorothea Haas, MD, in the University Children's Hospital Heidelberg and the following protocol was provided to me. Samples (forelimb bud pairs from 12 embryos each) were washed with PBS and hydrolyzed with ethanolic KOH. 5α-cholestane was added as internal standard. The extraction was performed with water and n-hexane and the sample was derivatized with 50 μl MSHFBA to form the trimethylsilyl derivatives. For GC/MS analysis the quadrupole mass spectrometer MSD 5972A (Agilent) was run in the selective ion monitoring mode. The following characteristic mass fragments were used for quantification: *m/z* 217/357 (5α-cholestane, internal standard), *m/z* 329/368 (cholesterol), *m/z*

325/351 (7- and 8-DHC), *m/z* 343/372 (desmosterol), *m/z* 255/458 (lathosterol), *m/z* 393/498 (lanosterol), *m/z* 213/229 (8(9)cholestanol). We also quantify cholestanol, campesterol, beta-sitosterol and stigmastanol. Gas chromatographic separation was achieved on a capillary column (DB-5MS, 30 m x 0.25 mm; df: 0.25; J & W Scientific) using helium as a carrier gas. The initial oven temperature of 100°C was raised after 2 minutes to 300°C at a rate of 35°C per minute. The injector was held at 280°C and the transfer line at 290°C. 1 μl of the derivatized sample was injected in a splittless mode.

9.4. Histology

9.4.1. Paraffin embedding of mouse embryos and tissues

Embryos were collected in ice-cold PBS and fixed overnight in cold 95% ethanol containing 1% glacial acetic acid (Sainte Marie's fixation). After 3 washes of 30min in cold 95% ethanol, embryos were stored in 70% ethanol at 4°C. For the paraffin embedding, they were dehydrated through an ascending ethanol washes series: once 30min in 95% ethanol (10min if embryos are younger than E11.0) and 3 times 30min in absolute ethanol (10min if younger than E11.0). They were then cleared with xylene 3 times for 15min each and transferred to a 50:50 (v/v) xylene/paraffin mixture for 30min at 60°C. Tissues were incubated with freshly melted paraffin 4 times for 1hr at 60°C. Finally, samples were embedded in a fresh paraffin wax (60°C) using a stereomicroscope for proper orientation. Paraffin blocks were incubated at RT for 1hr and then stored at 4°C. The Microm HM 355 microtome was used to cut sections of 6μm that were mounted on Superfrost Plus slides (Thermo Scientific) and stored at 4°C until use.

9.4.2. Immunohistochemistry (IHC) using paraffin sections

The first day, sections of embryonic tissues were deparaffinised for 2x5min in xylene, followed by 2x5min in absolute ethanol, 2x5min in 95% ethanol and 1x5min in PBS. Endogenous peroxidases were blocked by incubating the slides in 0.3% H_2O_2 (Sigma H1009) in methanol for 40min. After 3x5min washes in PBS, the sections were first blocked for 30min at RT in TNB solution. TNB solution consists of 0.1M Tris-HCl pH 7.5, 0.15M NaCl, 0.5% NEN blocking reagent from the TSA Biotin System kit (PerkinElmer NEL700). Sections were then blocked a second time in 5% goat serum, 0.2% BSA and 0.1% Triton X-100 mixture in PBS for 40min at RT. Finally, the sections were incubated overnight at 4°C with anti-SHH Ab80 rabbit antibody (Gritli-Linde et al., 2001) diluted 1:600 in PBS containing 0.2% BSA and 0.1% Trition-X100. The second day, sections were first washed 3x5min in TNT solution (0.1M Tris-HCl pH 7.5, 0.15M NaCl, 0.5%, 0.025% Tween-20) and then incubated for 45min at RT using biotinylated goat anti-rabbit antibodies (Vector Laboratories BA100), which was diluted 1:300 in TNT solution containing 2% non-fat dry milk (Sigma M7409). Sections were then washed 3x 5min in TNT and incubated 30min in the dark with Streptavidin-HRP from the TSA Biotin System kit (diluted 1:250 in TNB). After washing for 3x5min in TNT, sections were incubated exactly 9min in Biotinyl Tyramide diluted 1:50 in Amplification buffer (TSA Biotin System kit). Sections were washed again 3x5min in TNT and then incubated in Streptavidin-HRP from the TSA Biotin System kit (diluted 1:250 in TNB) for 30min in the dark. To develop the sections, they were first washed for 3x5min in TNT and developed in DAB solution for 5 to 10min. The DAB solution is prepared as follows: 900μL of stable peroxide solution are mixed with 100μL of DAB substrate (Metal Enhanced DAB Substrate Kit, Thermofisher 34065). WT and *Smad4*$^{\Delta/\Delta c}$ sections were developed in parallel for exactly the same time. Slides were then washed 3 x 5min in PBS and dried at 42°C for 1hr. Then they were mounted in Histomount medium (Life Technologies 008030).

9.4.3. Optimum Cutting Temperature (O.C.T) embedding of embryos to prepare frozen sections

Embryos were collected in ice-cold PBS and fixed for 2hrs at 4°C in 4% PFA/PBS and washed 3 times in PBS (5min). Samples were then cryoprotected using a gradient of sucrose: 10% sucrose/PBS (w/v), 20% sucrose/PBS, 30% sucrose/PBS (1hr each) at 4°C. Then, the embryos were transferred into embedding molds, the 30% sucrose/PBS was carefully removed and replaced by 50:50 (v/v) O.C.T/30% sucrose. The orientation was checked using a stereomicroscope and the samples frozen by dipping and holding the bottom of the embedding mold into Isopentane cooled with dry ice. The frozen blocks were stored at -80°C and 10μm sections prepared using a Leica CM3050S Cryostat (-20°C). Frozen sections were mounted on Superfrost Plus slides and stored at -80°C.

9.4.4. IHC using frozen sections

Smad4$^{3xF/3xF}$ or WT sections were washed 3x5min in PBS, once 30min in PBT (PBS with 0.2% Triton X-100) and again 5min in PBS. They were then blocked in 1% BSA in PBT for 1hr at RT and incubated overnight at 4°C with the monoclonal mouse anti-FLAG M2 antibody (Sigma, F1804) diluted 1:500 in 1% BSA/PBS. The second day, the sections were washed 3x5min in PBS, once in PBT and were incubated in the dark for 1hr at RT with the goat anti-mouse Alexa 488 secondary antibody diluted 1:500 in 1% BSA/PBS. Sections were finally washed 3x10min PBS, once in PBT (5min), nuclei counterstained in 1μg/mL Hoechst-33258/PBS (5min) and rinsed again 3x5min in PBS. Then they were mounted in Mowiol 4-88 and dried overnight at RT in the dark.

9.5. Genomics online resources

All the sequences alignments were retrieved from UCSC (Flicek et al., 2014). Alignments between species were performed using the ECR browser (Ovcharenko et al., 2004) or VISTA alignment browser (Dubchak et al., 2000;

Frazer et al., 2004). TADs analysis was performed using the 3D Genome Browser (Wang et al., 2017). Other useful online resources used include the Galaxy browser (Sloggett et al., 2013), the VISTA enhancer browser (Visel et al., 2007) and the Mouse Genome Informatics website (http://informatics.jax.org).

9.6. Bioinformatics Analysis

All bioinformatics analysis for the ChIP-seq, ATAC-seq and RNA-seq experiments were performed by Iros Barozzi (Lawrence Berkeley National Laboratory, Berkeley, CA, USA). This part of the material and methods section was written by Iros Barozzi and provided to me for this thesis.

9.6.1. ChIP-seq raw data analyses and annotation

Short reads obtained from Illumina HiSeq were aligned to the mm9 genome using Bowtie v1.1.0 (Langmead et al., 2009). Only those reads with a unique match to the genome with two or fewer mismatches (*-m 1 -v 2*) were retained. In order to make different runs comparable, the 3' of reads were trimmed to 63 bp before alignment. This step was performed using fastx_trimmer (*-l 63*), a tool part of the FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/) (v0.0.13). Peak calling was performed using MACS v1.4 (Zhang et al., 2008) with the following parameters: *--gsize=mm --bw=300 --nomodel --shiftsize=100 --pvalue=1e-2*. Matched input DNA was used as control. Wiggle tracks were also generated with MACS; these were then re-scaled linearly according to sequencing depth (RPM, Reads Per Million sequenced reads). MACS was run with a permissive threshold (*p*-value 0.01) in order to identify a larger list of sub-significant regions across biological replicates. Evidences from these replicates were combined using MSPC (Jalili et al., 2015), with the following parameters *-r biological -s 1E-5 -W 1E-2*. The confirmed peaks were assigned the best *p*-value (as defined by MACS) among the overlapping peaks across replicates. Only the peaks showing reproducibility were retained for further analysis (we

termed this set as *golden*; one golden set per developmental stage). These lists of peaks were annotated to the TSS of the nearest RefSeq genes using the script *annotatePeaks.pl* available in HOMER (Heinz et al., 2010). A region was considered as promoter-proximal if annotated within 2.5 kbp from a RefSeq promoter. The remaining regions were divided into intragenic and intergenic, whether the region overlapped the body of an annotated gene or not.

9.6.2. Motif enrichment and de novo motif discovery analyses

The script *findMotifsGenome.pl* available in HOMER (Heinz et al., 2010) was used to perform enrichment analysis for known transcription-factor binding sites and motif discovered *de novo*. The script was run with the following arguments: *-size -150,150 -len 6,7,8,9,10,12,14*, using the summit of the peaks in the golden set as reference. The top ten most significant, over-represented known matrices along with the top ten motif discovered *de novo* were then used to scan every single region for high-affinity sites using FIMO (Grant et al., 2011) (v4.10.0). The following parameters were used: *--thresh 1e-4 --no-qvalue*. The resulting list of sites was transformed into a matrix in which each region was represented as a vector of *p*-values, one for each different motif, corresponding to the *p*-value of the highest-scoring site identified (*p*-value = 1 if no significant match was found). *P*-values were then log10-transformed and their sign inverted, then hierarchically clustered (Euclidean distance, complete linkage).

9.6.3. Evolutionary conservation analysis of genomic regions enriched in SMAD4 chromatin complexes

The genome-wide track of base-pair *phastcons* (Siepel et al., 2005) conservation scores in placental mammals was downloaded from the UCSC genome browser (Tyner et al., 2017) (track name: *mm10.60way.phastCons60wayPlacental.bw*). The coordinates of the peaks in the golden sets were converted from mm9 to mm10 using *liftOver* (Tyner et al., 2017) (*-minMatch=0.95*). The base-pair scores for the 300 bp centered on the summit of the peaks were then extracted using *bwtool* (Pohl and Beato, 2014).

9.6.4. ATAC-seq raw data analysis and annotation

Short reads obtained from Illumina HiSeq were aligned to the mm9 genome using Bowtie v1.1.0 (Langmead et al., 2009) (-m 1 -v 2, see *"Chip-seq raw data analyses and annotation"*). Accessible regions were identified using MACS v1.4 (Zhang et al., 2008) with the following parameters: *--gsize=mm --bw=150 --nomodel --nolambda --shiftsize=75*. Genome-wide profiles were generated using MACS and re-scaled linearly according to sequencing depth (RPM). Gene annotation was performed using HOMER (Heinz et al., 2010), as described in *"Chip-seq raw data analyses and annotation"*. Evidences from biological replicates were combined using MSPC (Jalili et al., 2015), using the following parameters *-r biological -s 1E-10 -W 1E-6*. The confirmed regions were assigned the best *p*-value (as defined by MACS) among the overlapping regions across replicates.

9.6.5. RNA-seq data analysis

Single-end reads obtained from Illumina HiSeq were aligned to the mm9 reference genome and to the *Mus Musculus* transcriptome (iGenome refGene GTF) using TopHat v2.0.13 (Kim et al., 2013). The option *--no-coverage-search* was specified, while all the other parameters were left to default. Only uniquely mapped reads were considered for the analysis. Tracks for the UCSC genome browser (Tyner et al., 2017) were produced using *genomeCoverageBed* from BedTools v2.17.0 (Quinlan and Hall, 2010); these were linearly re-scaled according to sequencing depth (RPM). Gene-wise counts were computed using *htseq-count* from the HTSeq package (Anders et al., 2015) with *-s* set to *no*. Genes on chromosomes X, Y and M were excluded from further analysis. edgeR (Robinson et al., 2010) was used to identify differentially expressed genes (DEGs). Only genes showing expression (in terms of Fragments Per Million sequenced reads equal or higher than 1) in at least three samples were considered for further analyses. Libraries were normalized according to TMM normalization. Tag-wise estimation of dispersion was evaluated using *prior.df = 10*. Differential expression between pairs of conditions was evaluated using the

exactTest R function. False discovery rates were estimated using Benjamini-Hochberg correction (Benjamini and Hochberg, 1995). DEGs were defined as those genes showing a *q*-value <= 0.1 and a linear fold-change equal or higher than 1.2. Functional enrichment analyses were conducted using DAVID (Huang et al., 2009).

9.6.6. Hierarchical clustering, plots and statistical testing

Clustering, plots, heat maps and statistics were handled in the statistical computing environment R v3.

9.7. Tables

9.7.1. LacZ reporter primers table

CRM	Forward primer (5'-3')	Reverse primer (5'-3')	Size (bp)
Alx4 +368kb	CACTAAAGGGAACAAAAGCTGGTACCTC CTGAGTGCTGTCAGTTCG	TTTGGATGTTCCTGGAGCTCGGTACCC CCAGCGTAAATGCTGTGTA	885
Id1 +3kb	CACTAAAGGGAACAAAAGCTGGTACGGT ACCGGATTCCCACTGGGGCTAAAG	TTTGGATGTTCCTGGAGCTCGGTACCG GTAGCCAATCACTCTTATTGC	1441
Id2 -438kb	CACTAAAGGGAACAAAAGCTGGTACCTG TCCTTTCTTTAATGTGATCCAA	TTTGGATGTTCCTGGAGCTCGGTACCG AGGGCTGCCATTAACACTC	842
Lhx2 +472kb	CACTAAAGGGAACAAAAGCTGGTACCCC TACTTCTGCCTCTCAAATGC	TTTGGATGTTCCTGGAGCTCGGTACCT GGAGAGAGAGCCTTTTATGTTTC	1771
Msx2 -3.3kb	CACTAAAGGGAACAAAAGCTGGTACCTA ATGTTTATGCGCCGTGAA	TTTGGATGTTCCTGGAGCTCGGTACCC ACTGTCAGGAAGCCAACAA	1180
Pkdcc -50kb	CACTAAAGGGAACAAAAGCTGGTACGGT ACCTGCAATTTGCAGCTTGCC	TTTGGATGTTCCTGGAGCTCGGTACCG GGCAGCGTGATATAGGAG	1238
Prrx2 -87kb	CACTAAAGGGAACAAAAGCTGGTACGGT ACCAAGAAGAGTCTGAATCTCCC	TTTGGATGTTCCTGGAGCTCGGTACCG GTCTTGAACACCTCAGAAG	3228

9.7.2. Genotyping primers table

Locus	Forward primer (5'-3')	Reverse primer (5'-3')	Size (bp)	Allele
LacZ	5'-GATCCCGTCGTTTTACAACG-3'	5'-AATGTGAGCGAGTAACAACCCG-3'	362	LacZ
PrxCre	5'-GGCTCTCTCCTTAGCTTCCC-3'	5'-CCTGGCGATCCCTGAACATGTCC-3'	400	Tg

Smad4	5'-ACAGCCTCCACACTTGTGCT-3'	5'-TGTCTGCTAAGAGCAAGGCA-3'	396 + 501	WT + 3xFlag
	5'-AAGAGCCACAGGTCAAGCAG-3'	5'-CCTGACCCAAACGTCACCTTC-3'	500	Δ
	5'-GGGCAGCGTAGCATATAAGAC-3'	5'-CCTGACCCAAACGTCACCTTC-3'	390 + 450	WT + Fluxed

9.7.3. qPCR primers table

Locus	Forward primer (5'-3')	Reverse primer (5'-3')	Size (bp)
Cre	5'-ATACCGGAGATCATGCAAGC-3'	5'-TTGCCCCTGTTTCACTATCC-3'	88
DBH	5'-AGGACATCAGCCACTCTGCT-3'	5'AATTGTCTTGGTGGCCCTC-3'	117
Gli1	5'-CAAGTGCACGTTTGAAG-3'	5'-CAACCTTCTTGCTCACACATGTAAG-3'	76
Hhip1	5'-GCTCTGTCGAAACGGCTACT-3'	5'-GGTGTCTGTGTCAGACCGAA-3'	129
Ptch1	5'-GGCAAGTTTTTGGTTGTGGGTC-3'	5'-TGCTGCTGATGGATGGGAAC-3'	157
Rpl19	5'-ACCCTGGCCCGACGG-3'	5'-TACCCTTTCCTCTTCCCTATGCC-3'	53

9.7.4. WISH screen primers table

Locus	Forward primer (5'-3')	ATAT-SP6 sequence- Reverse primer (5'-3')	Size (bp)
1500004A13Rik	AGAACCAGAGAAGCCTGAGGTC	ATATATTTAGGTGACACTATAGAAGAGGCTACGGGTGAGGTTATCT	746
4933404O12Rik	CCCGTGAGAGACATGAGCTAAG	ATATATTTAGGTGACACTATAGAAGGGCTACAGAGTGAGATTCAGG	513
5830417I10Rik	GAGGGTGAGTCTGAGGAAGAGA	ATATATTTAGGTGACACTATAGAAACTTACGAGTCCACACGACAGG	594
Abhd8	GCCCAGCCTATGCTCCATCTTCA	ATATATTTAGGTGACACTATAGAAGCACTGAGTCTCCAACACTGAC	638
Acot11	GGAGACTCTCTGTTCAGGCTTC	ATATATTTAGGTGACACTATAGAAACTGAGTCAGAGTCCAGCCAAC	538
Acta2	CCTGGAGAAGAGCTACGAACTG	ATATATTTAGGTGACACTATAGAAGCTAGGCCAGGGCTACAAGTTA	588
Adat3	GAAGACAAGCAGGTGACCAGTG	ATATATTTAGGTGACACTATAGAACTGCAGTCTGGAAACTGAGAGG	738
Adora2a	ACCCCTTCATCTACGCCTACAG	ATATATTTAGGTGACACTATAGAACTGCTGTAGCCCTGACCTAACT	609
Akap9	TCTGGACTATCCTCGGTCTCTC	ATATATTTAGGTGACACTATAGAAAAGGTACTCGTCCAGCAGCACT	540
Ano1	CCTTCACGTCTGACTTCATCCC	ATATATTTAGGTGACACTATAGAACTCATCAGAGTGCTCCTCCACAC	687
Apba3	CCTGCAGACCATCTCCTACATC	ATATATTTAGGTGACACTATAGAACGGTGGGCAGTGGATAATACTC	577

Arhgef19	CTCAGAGAGTACAAGAGGGTCTCC	ATATATTTAGGTGACACTATAGAAGGACGA GGACACGTAGTAGGAT	592
Atoh8	CTCTCCAAACTGGCCATCCTGA	ATATATTTAGGTGACACTATAGAAGTATCG GGAACAGGAGTCAGAG	748
B4galnt2	GGCCATCTCACAGGTGACTACT	ATATATTTAGGTGACACTATAGAAACTTCA GGGTCCTACCCTAGCTC	700
Bbc3	CTGGGTGCACTGATGGAGATAC	ATATATTTAGGTGACACTATAGAACAGACT CCTCCCTCTTCTGAGAC	653
Bcl11a	CGGGATGAGTGCAGAATATGCC	ATATATTTAGGTGACACTATAGAAGGTCAG GGGTCATGCTCATTTC	600
Bcl9l	AAGCAGCACCCTCCAGTACTTC	ATATATTTAGGTGACACTATAGAAGGCCTG CTAGAAGGGTAGATTG	502
Bnc2	GTCCTTGACCTGAGTACCACCT	ATATATTTAGGTGACACTATAGAACACACA CTGGCTATGGCAGTTC	500
Cacng7	CTACACCCTTCCCTATGGTCAG	ATATATTTAGGTGACACTATAGAAGAGATG GTCCGGGTACTTGATG	506
Cc2d1a	GTACCAGGACGTAGTACAGCGTAG	ATATATTTAGGTGACACTATAGAAGTATCTT TGGAGAGGGCTGTGG	524
Ccdc124	GGAGCTGGAAGATGCTTACTGG	ATATATTTAGGTGACACTATAGAAGACTCA GCCTGGTTCTGTCACT	576
Cldn10	GGTGCTAGTGTCTTCCACACTG	ATATATTTAGGTGACACTATAGAAGTGGGT CCGTTGTATGTGTAGC	546
Clstn3	GACCTAGACCCTGAACGAGAGAG	ATATATTTAGGTGACACTATAGAAAGTCAC TGCTGTCCTCTTCCTC	658
Canter	CAGCACACACCATCACAGATGCC	ATATATTTAGGTGACACTATAGAATGCATG GTCCTCCTCTCTGGATG	564
Col12a1	CATCATACAGGAGGGTGGAGAC	ATATATTTAGGTGACACTATAGAACTAGAG ACTGAACTCGGGGTGA	623
Col1a1	CTCAACCCCGTCTACTTCCCTAC	ATATATTTAGGTGACACTATAGAAACAGAC CAAGAGAGAGGCAGAG	766
Dhcr24	GACATCCACGTCTACCCCATCT	ATATATTTAGGTGACACTATAGAAGACTAG CCCAGAGTCAGTCACAC	563
Dhcr7	CCATCGACATCTGCCATGACCAC	ATATATTTAGGTGACACTATAGAAGGCTCT CCATTCTCCAGATGAGC	671
Dyrk1b	CGACAACAGAGCCTACCGATAC	ATATATTTAGGTGACACTATAGAAGTAGCA GCAAGTCCAGTCAAGG	557
Fam171a2	GCAGGCTCTAACCGAGAAGAAG	ATATATTTAGGTGACACTATAGAACCCAGC ACATACCACAGATGAG	713
Fam189a2	GCTACAGTGACTCTGAGGAGAGG	ATATATTTAGGTGACACTATAGAAGATACT GTGAGACCCAGCAGAC	711
Fam189b	ACTCTCTAGGGGACCTGAAAGG	ATATATTTAGGTGACACTATAGAACTGAGG TCATGCAGCTAGAGGT	741
Fdft1	ACGTCCTCACCTACCTGTCAAG	ATATATTTAGGTGACACTATAGAACTCTGC CCAGTTCCTAAAGGTC	577
Fops	CTGGTGTGTAGAACTGCTCCAG	ATATATTTAGGTGACACTATAGAACTCGTA GCAGACACTGAACCAC	566
Fras1	GGAGGACACTCTGGAGGAATAC	ATATATTTAGGTGACACTATAGAAGGGGAC AGTGCACAGATGTCTA	750
Glis1	CAGGGGGACAGTCATTCTCTAC	ATATATTTAGGTGACACTATAGAAAGGTAG GGAAGTACCCTGTGGT	527
Gm5779	GAGGACCTCACTGAGATTAGGG	ATATATTTAGGTGACACTATAGAAGCAGCT GGGAAAGTGTACTCAG	506
Gng8	GACTCGGTCTCTGAGTGTCTGTC	ATATATTTAGGTGACACTATAGAAGTCACC AGTGGGTCATCCTTAG	716
Gpr162	CCTGTGACGACTACACAGATGG	ATATATTTAGGTGACACTATAGAAAGTGGT CTAGGCCTACGAGGAGA	529

H2-D1	CTCCTCCGTCCACTGACTCTTAC	ATATATTTAGGTGACACTATAGAACAGACA CTGCAGAAGAGGGTGT	624
H2-K1	CTGGTGAAGCAGAGAGACTCAG	ATATATTTAGGTGACACTATAGAACTGGGA GAGACAGATCAGAGGT	558
Hdac5	TCCCTCTGTGCTCTACATCTCC	ATATATTTAGGTGACACTATAGAACCTCCT CTTTCTCACCTGTCTG	596
Hop	GGTGGAGATCCTGGAGTACAAC	ATATATTTAGGTGACACTATAGAAGGGAAG TGAACTCAGGAGGTGT	588
Idea	ATCAACTCAGCAGCCCCTACAG	ATATATTTAGGTGACACTATAGAAGACTCC ATTCACCAGTCAGCAC	556
Ier5l	AAGAACCTCCTGGTGTCCTACG	ATATATTTAGGTGACACTATAGAAGTAGCC GTTTTCTACCGTGGTC	729
Inf2	GATCTCTCAGCTCCCAGTACTCC	ATATATTTAGGTGACACTATAGAAGTACCT GCTGTCTCATGTCCAG	760
Irf5	CTTCTACAGCCAGCTAGAGGCTAC	ATATATTTAGGTGACACTATAGAACTTCACT ACCACCAGACCATCC	708
Lhx2	GATGCTGAACACCTGGATCGTG	ATATATTTAGGTGACACTATAGAACTCTGTT TCCAGGCGAGATCCT	615
Ltbp4	GTCTGAGACTCCTGACCCACCTA	ATATATTTAGGTGACACTATAGAACCCTCT GTGTCTGTCCATCCTA	552
Mamdc2	GGAGATCACACTACTGGGGTAGG	ATATATTTAGGTGACACTATAGAAGAAAAC AGAGGAGAGCCCTGAG	618
Mafia	GCTCTGTAACATCCTCCTGGAC	ATATATTTAGGTGACACTATAGAAGGAGCA GAGGTAGAGAAGGACAC	519
Mdn1	AGACAGAGTCCTAGCAGCAGTCC	ATATATTTAGGTGACACTATAGAAGTGACC TCTACCTGGTCCCTAGT	547
Mfap4	GGATGGCTATACCCTCTACGTG	ATATATTTAGGTGACACTATAGAAAGCTGT CGTGAAGGGGTAGAAG	555
Mib2	GCTGAGTACTCCTAACACCGTGAC	ATATATTTAGGTGACACTATAGAAGTTCCT GATCTCCAGGCAGTCT	646
Mt2	TGCTGGCCATATCCCTTGAGCCA	ATATATTTAGGTGACACTATAGAATTGTGG AGAACGAGTCAGGGTTG	478
Mxd4	CTGACAGCACACGTCACACTAC	ATATATTTAGGTGACACTATAGAAGTAAGA GCCTGACACTGCTTCC	621
Naif1	CCTACCCAGTACCACAGAGATCC	ATATATTTAGGTGACACTATAGAAAGCAGA GTACTGGCTTGACCTC	583
Nckap1l	CTGCTGCCCTCTTCACTCTGTA	ATATATTTAGGTGACACTATAGAAGAAGTG CAGGAGACAATAGCAGCTC	716
Ndufs7	CAGCTGCGCAGAGTTCATCAGA	ATATATTTAGGTGACACTATAGAAAGATGT CCACTGGCACAATGCG	502
Nfatc4	TATAGCACACCCGCTCTGTACC	ATATATTTAGGTGACACTATAGAAAAGCTA GAAGCTGGGCTAGGAG	588
Nfkbil1	GGAGAAGGAACTGTGTGAGAGC	ATATATTTAGGTGACACTATAGAAAGTCCA GAGCAGTGGTTAGAGC	569
Nup210l	GTGTGGCAGACACCTCAGTCTAC	ATATATTTAGGTGACACTATAGAAGTGTTGT AGCCGACTCTGCACT	647
Pard6a	GATGCTCACGCAGAAGGTGACT	ATATATTTAGGTGACACTATAGAAGGTGGC GATTCTCAATGACCAG	598
Pcsk9	GATCTCAGGTCCTTCAGAGCAG	ATATATTTAGGTGACACTATAGAAGAACCT CCAAGGATCTCTGTCC	560
Pdk2	GGAAGATCGAGAGGCTCTTCAG	ATATATTTAGGTGACACTATAGAACACTTA GGGACCACCACAGTTG	501
Phactr3	GAGCACAGGACTATGACAGGAG	ATATATTTAGGTGACACTATAGAAAGTCCC CAGCTATACCTCATCC	754
Pave	GGTCCACTCAAGGAGGAGTATG	ATATATTTAGGTGACACTATAGAAGAGAGC GTACACACAGGGAAAC	771

Rec8	CCAGAGATCAGTCGAGGAGACT	ATATATTTAGGTGACACTATAGAACTTCTGC TGCTTCTAGGGAGAG	781
Rfx1	CCTCAAGTGGTCCTTCTACAGC	ATATATTTAGGTGACACTATAGAAAGACTC TCCTAGCCTTCCCTGT	549
Rgl3	CTCACCTACCTCCAGTGTGTCTC	ATATATTTAGGTGACACTATAGAAACCTAA GCTAAGCTCCCAGCTC	784
Rhpn1	ACATGACTCGAGGAGAGGGAAG	ATATATTTAGGTGACACTATAGAAGTCCCC AAGAGTAGAGGACAGAC	555
Rspo4	GGAGTCCCTGCATACACAATGG	ATATATTTAGGTGACACTATAGAATGCAGA GAGCCACATGGAGGAT	572
Rusc1	AGTAAGTGTGCTGGCTCTGGTG	ATATATTTAGGTGACACTATAGAAAGGGCT AGAGACAGAGTGAGGAG	637
Scd1	GACTACTCTGCCAGTGAGTACCG	ATATATTTAGGTGACACTATAGAAGGGGAC TTGCTCTATCCCTAGT	732
Serinc2	CTGATGCAGACAGAGGAGTGTC	ATATATTTAGGTGACACTATAGAACTTCTG GAGACCCTGAGAACAG	548
Slc27a1	GTCAGACCTCAGACAGGCTCTT	ATATATTTAGGTGACACTATAGAAGACAGC CAGGGCTATACAGAGA	661
Srebf1	GGACAGCTTAGCCTCTACACCA	ATATATTTAGGTGACACTATAGAAAGCCAC TAAGGTGCCTACAGAG	751
Thbs3	GAAACAGACGGAGCAGACCTAC	ATATATTTAGGTGACACTATAGAACCAGGG TCTCCAGAACAGAGTT	515
Tm7sf2	GGTATGAGGAGTCTGTCCTCACC	ATATATTTAGGTGACACTATAGAAAGGTGG ACCTGCTTCAGTAGAC	555
Tmem119	ACAGAGTCTCCCCCAGTGTCTA	ATATATTTAGGTGACACTATAGAAAGCAGG AACCCTTAGAGTAGCC	679
TN	CTACCTCTGGCCTCTACACCAT	ATATATTTAGGTGACACTATAGAAGCCCTG ACTGTGGTTATTGTCC	521
Tnnt1	GAGGCAAAGAAGAGAGCAGAGG	ATATATTTAGGTGACACTATAGAACTCTCC CAGGCAGTATGGAGAT	528
TN	GAAGCGTCTCGTCTCAGTCAGT	ATATATTTAGGTGACACTATAGAACCCTGG CCCTCTTACATAGAAC	550
Upk3b	CCACAGACTGCTGCTAAGATCC	ATATATTTAGGTGACACTATAGAAAAGCTA CAGAGACCAGGCTACG	593
Uqcr11	TACGTCTGCACATGCGTAGTGC	ATATATTTAGGTGACACTATAGAAAGGCAG CCCTAGTGTCTGTCAA	391
Wnt2b	TGGAGAGCACTCTCAGACTTCC	ATATATTTAGGTGACACTATAGAACTCAGC CTCCTAAATCCATCCC	634
Wt1	GACTTCCAAGACAGCACACCTG	ATATATTTAGGTGACACTATAGAAAGGAAG GCTCCTCTCTGTCCTA	536
Zfp628	CTAATGGAGGAGCTGGCACTAC	ATATATTTAGGTGACACTATAGAAGTACCA GTTGGACTGCAGGAAG	540
Zfp78	GGTCAGAGGCCCTATGTATGTG	ATATATTTAGGTGACACTATAGAACTGAGA TGACCTCCCAGAACTC	777

9.7.5. WISH probe cloning primers table

Locus	Forward primer (5'-3')	Reverse primer (5'-3')	Size (bp)
Cyp51	GGGATAGAGCCCATCGAGAGAT	CTCTCTCCACACTGGCTTCTTG	786
Dhcr24	GACATCCACGTCTACCCCATCT	GACTAGCCCAGAGTCAGTCACAC	563
Dhcr7	CCATCGACATCTGCCATGACCAC	GGCTCTCCATTCTCCAGATGAGC	671

Fdft1	ACGTCCTCACCTACCTGTCAAG	CTCTGCCCAGTTCCTAAAGGTC	577
Fops	CTGGTGTGTAGAACTGCTCCAG	CTCGTAGCAGACACTGAACCAC	566
Id3	GGTGTCTCTTTTCCTCCCTCTC	CCACCCAAGTTCAGTCCTTCTC	598
Idi1	CTCAGGAGACAGCTCAGTGTACC	CTAACTCAAGCCAGACCCCCTAC	710
Lidl	GATCTCTCAGCTCCCAGTACTCC	GTACCTGCTGTCTCATGTCCAG	760
Lhx2	GATGCTGAACACCTGGATCGTG	CTCTGTTTCCAGGCGAGATCCT	615
Mad	CCTACCTCAATGACACCTCCAG	GTCAGCTAGCCCTTCAGAAACC	585
Mxd4	CTGACAGCACACGTCACACTAC	GTAAGAGCCTGACACTGCTTCC	621
Pcsk9	GATCTCAGGTCCTTCAGAGCAG	GAACCTCCAAGGATCTCTGTCC	560
Pave	GGTCCACTCAAGGAGGAGTATG	GAGAGCGTACACACAGGGAAAC	771
Prrx2	GCAAGAACTTCTCGGTGAGC	GAGGACAGAGACAGACAAGGCTA	735
Sc5d	GGTTAGCGTCGTATCCTTCCTC	TACTGTCAGGACCATCCCTCAG	671
Sfrp2	CCTGGAGACAAAGAGCAAGACC	CCCTCGGGTCAAAGTACTACAG	591
Srebf1	GGACAGCTTAGCCTCTACACCA	AGCCACTAAGGTGCCTACAGAG	751
Tm7sf2	GGTATGAGGAGTCTGTCCTCACC	AGGTGGACCTGCTTCAGTAGAC	751

9.7.6. CRISPR/*Cas9* deletion strand table

CRM	strand 1 (5'-3')	strand 2 (5'-3')	Size (bp)
GRS1	AGCGGCAGTTCGGCTTCCGG	TCTCATACGATCCAGGAGAA	9909

9.7.7. ChIP-qPCR primer table

Locus	Forward primer (5'-3')	Reverse primer (5'-3')	Size (bp)
BRE Id1	AGAATGCTCCAGCCCAGTTT	TGACGTCACCCATTCATAAAA	93
BRE Msx2	CCATTAGGGCGAATTGTCAT	GAGCCGCGTTAATTGCTCT	90
Grem1 exon 2	TCATTGTGCTGAGCCTTGTC	CAACAGCTGAGGGGAAAAAG	70
GRS1-A	ATGAGACTTGGAGCAGCAGTT	TCCTCGGTTATTGGGCAATA	123
GRS1-B	ACACACACAGAGCAAGAGGCT	AACCGGTGCAGTCTCAGAAG	115
Neg Grem1	ATCCAATCACGTGCATAACAA	AAGCAAGGATCCATGGAACA	92
Neg Id1	TTCTTCTCTGGCTGCCAGTG	AACTGAGCCTTGCATCATGC	140
Neg Msx2	GACTAGGGCTCTCTTTTCCTGA	CATTTCTCCACCCCAGCTTA	84
β-actin	GATCTGAGACATGCAAGGAGTG	GGCCTTGGAGTGTGTATTGAG	115

10. Acknowledgements

First, I thank Prof. Rolf Zeller and PD Dr. Aimée Zuniga for giving me the opportunity to do my PhD in their laboratory, for the training, their patience and for giving me the independence and freedom to develop my ideas.

I would like specially to thank Frédéric Laurent for being my mentor during the first years of my PhD and for still being one of my best friend. Thank you for our great scientific (or no scientific) discussions, for your availability, your blessed jokes and your enthusiasm that has motivated me so much. And of course, thank you for generating the *Smad4*^{3xFlag} mouse line with what the project could start. I would like also to thank Marco Osterwalder for his excellent training for ChIP experiments, his patience during my first year of PhD and for giving me still support and ideas from San Francisco.

I deeply thank Iros Barozzi, without whom this project would not lead to this great story. Thank you for your excellent bioinformatics skills, for your reactivity, your availability and your great scientific ideas.

I thank Prof. Verdon Taylor, Prof. Gerhard Christofori and Dr. Vanessa Ribes for being members of my PaC and my thesis committee. Thanks for your advices and your guidance that helped me to take the good decisions at the right moment. I also thank Prof. Markus Affolter for accepting to be the Chair of my PhD committee.

I am indebted to Rushikesh Sheth for the good scientific discussions and his help for the thesis corrections.

Many thanks to our collaborators Prof. Dorothea Hass for the GC/MS analysis and Dr. Anna Kicheva for her great input to improve the LMP culture protocol and for giving me the opportunity to come back to the spinal cord development field to test our hypothesis.

I am deeply thankful to Gretel Nusspaumer and Javier Lopez-Rios for our amazing professional and friend relationship, for your support, your help and for making us feeling the laboratory as a family during these four years.

Big big big thanks to Dario Speziale for being the best neighbour. Thank you for your kindness, your smile, your hugs and your happiness everyday during three amazing years.

I thank my girl team Laurène Ramos Martins, Virginie Tissières, Maëva Luxey and Fabiana Gullotta for listening to my complaints, cheering me up and for our shared crazy laughs and good moments. I want to thank Nathalie Riesen for our great work together and our complicity for three years.

I thank Jens Stolte for his help concerning the cloning for the LacZ reporters, and also Laurène for the WISH. Furthermore, I want to thank Aline Baur, Robert Reinhardt, Ausra Girdziusaite, Jonas Malkmus, Shalu Jhanwar and Artal Moreno-Fortuny for sharing good time and interesting discussions. I would like also to thank very much the previous members of the laboratory Sumit Jaiswal, Erkan Ünal, Emanuele Pignatti and Jorge Dorado. I also thank people from Verdon Taylor's laboratory, mainly Chiara Rolando, Anna Engler, Andrea Erni and Zahra Ehsaei for sharing their good mood on the first floor and for our good and constructive discussions during coffee breaks.

It is important for me to thank Christian Beisel, Ina Nissen and Katja Eschbach from the Genomics Facility, Pawel Pelczar, Alain Brülhart and Andreas Siegrist from the Transgenic Mouse Core Facility. Special thanks to Angelika Offinger and Emilia Stanislawa Terszowska for taking care so well of our mice.

Thanks to Susie Boudebaba and Saléha Rabehi who always provided us autoclaved solutions and glassware, and thanks Saléha for our daily breakfasts. I am thankful to Chris Müller, Barbara Widmer, Debora Neyer and Sonia Calzascia for their administrative and translation helps.

Last, but definitively not the least, I want to thank my little Mum who always cheered me up during my studies and who gave me the motivation to continue. I thank my brother and my sister in law for helping me to unwind for many weekends during these four years and for giving me my little niece who brought me happiness during the last months of my PhD. And thanks to my friends from my Master for our great European weekends that recharged me in energy when I needed it.

11. Bibliography

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166–169.

Andrey, G., Schöpflin, R., Jerković, I., Heinrich, V., Ibrahim, D.M., Paliou, C., Hochradel, M., Timmermann, B., Haas, S., Vingron, M., et al. (2017). Characterization of hundreds of regulatory landscapes in developing limbs reveals two regimes of chromatin folding. Genome Res. *27*, 223–233.

Bailey, J.S., Rave-Harel, N., McGillivray, S.M., Coss, D., and Mellon, P.L. (2004). Activin Regulation of the Follicle-Stimulating Hormone β-Subunit Gene Involves Smads and the TALE Homeodomain Proteins Pbx1 and Prep1. Mol. Endocrinol. *18*, 1158–1170.

Barolo, S. (2012). Shadow enhancers: Frequently asked questions about distributed cis-regulatory information and enhancer redundancy. BioEssays *34*, 135–141.

Bastida, M.F., Sheth, R., and Ros, M.A. (2009). A BMP-Shh negative-feedback loop restricts Shh expression during limb development. Development *136*, 3779–3789.

Benazet, J.-D., and Zeller, R. (2013). Dual requirement of ectodermal Smad4 during AER formation and termination of feedback signaling in mouse limb buds. Genesis *51*, 660–666.

Bénazet, J.-D., Bischofberger, M., Tiecke, E., Gonçalves, A., Martin, J.F., Zuniga, A., Naef, F., and Zeller, R. (2009). A Self-Regulatory System of Interlinked Signaling Feedback Loops Controls Mouse Limb Patterning. Science *323*, 1050–1053.

Bénazet, J.-D., Pignatti, E., Nugent, A., Unal, E., Laurent, F., and Zeller, R. (2012). Smad4 is required to induce digit ray primordia and to initiate the aggregation and differentiation of chondrogenic progenitors in mouse limb buds. Development *139*, 4250–4260.

Benchabane, H., and Wrana, J.L. (2003). GATA- and Smad1-Dependent Enhancers in the Smad7 Gene Differentially Interpret Bone Morphogenetic Protein Concentrations. Mol. Cell. Biol. *23*, 6646–6661.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. R. Stat. Soc. Ser. B Methodol. *57*, 289–300.

Blassberg, R., Macrae, J.I., Briscoe, J., and Jacob, J. (2016). Reduced cholesterol levels impair Smoothened activation in Smith–Lemli–Opitz syndrome. Hum. Mol. Genet. *25*, 693–705.

Briscoe, J., and Thérond, P.P. (2013). The mechanisms of Hedgehog signalling and its roles in development and disease. Nat. Rev. Mol. Cell Biol. *14*, nrm3598.

Briscoe, J., Chen, Y., Jessell, T.M., and Struhl, G. (2001). A Hedgehog-Insensitive Form of Patched Provides Evidence for Direct Long-Range Morphogen Activity of Sonic Hedgehog in the Neural Tube. Mol. Cell *7*, 1279–1291.

Brugger, S.M., Merrill, A.E., Torres-Vazquez, J., Wu, N., Ting, M.-C., Cho, J.Y.-M., Dobias, S.L., Yi, S.E., Lyons, K., Bell, J.R., et al. (2004). A phylogenetically conserved cis-regulatory module in the Msx2 promoter is sufficient for BMP-dependent transcription in murine and Drosophila embryos. Development *131*, 5153–5165.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat. Methods *10*, 1213–1218.

Buenrostro, J.D., Wu, B., Chang, H.Y., and Greenleaf, W.J. (2015). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide: ATAC-seq for Assaying Chromatin Accessibility. In Current Protocols in Molecular Biology, F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl, eds. (Hoboken, NJ, USA: John Wiley & Sons, Inc.), p. 21.29.1-21.29.9.

Bulger, M., and Groudine, M. (2011). Functional and Mechanistic Diversity of Distal Transcription Enhancers. Cell *144*, 327–339.

Büscher, D., Bosse, B., Heymer, J., and Rüther, U. (1997). Evidence for genetic control of Sonic hedgehog by Gli3 in mouse limb development. Mech. Dev. *62*, 175–182.

Cannavò, E., Khoueiry, P., Garfield, D.A., Geeleher, P., Zichner, T., Gustafson, E.H., Ciglar, L., Korbel, J.O., and Furlong, E.E.M. (2016). Shadow Enhancers Are Pervasive Features of Developmental Regulatory Networks. Curr. Biol. *26*, 38–51.

Capdevila, J., Tsukui, T., Esteban, C.R., Zappavigna, V., and Belmonte, J.C.I. (1999). Control of Vertebrate Limb Outgrowth by the Proximal Factor Meis2 and Distal Antagonism of BMPs by Gremlin. Mol. Cell *4*, 839–849.

Capellini, T.D., Giacomo, G.D., Salsi, V., Brendolan, A., Ferretti, E., Srivastava, D., Zappavigna, V., and Selleri, L. (2006). Pbx1/Pbx2 requirement for distal limb patterning is mediated by the hierarchical control of Hox gene spatial distribution and Shh expression. Development *133*, 2263–2273.

Capellini, T.D., Zappavigna, V., and Selleri, L. (2011). Pbx homeodomain proteins: TALEnted regulators of limb patterning and outgrowth. Dev. Dyn. *240*, 1063–1086.

Chen, Y., and Gridley, T. (2013). Compensatory regulation of the Snai1 and Snai2 genes during chondrogenesis. J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res. *28*, 1412–1421.

Chen, Y., and Struhl, G. (1996). Dual Roles for Patched in Sequestering and Transducing Hedgehog. Cell *87*, 553–563.

Chen, M.-H., Li, Y.-J., Kawakami, T., Xu, S.-M., and Chuang, P.-T. (2004). Palmitoylation is required for the production of a soluble multimeric Hedgehog protein complex and long-range signaling in vertebrates. Genes Dev. *18*, 641–659.

Chiang, C., Litingtung, Y., Lee, E., Young, K.E., Corden, J.L., Westphal, H., and Beachy, P.A. (1996). Cyclopia and defective axial patterning in mice lacking Sonic hedgehog gene function. Nature *383*, 407–413.

Chiang, C., Litingtung, Y., Harris, M.P., Simandl, B.K., Li, Y., Beachy, P.A., and Fallon, J.F. (2001). Manifestation of the Limb Prepattern: Limb Development in the Absence of Sonic Hedgehog Function. Dev. Biol. *236*, 421–435.

Chu, G.C., Dunn, N.R., Anderson, D.C., Oxburgh, L., and Robertson, E.J. (2004). Differential requirements for Smad4 in TGFβ-dependent patterning of the early mouse embryo. Development *131*, 3501–3512.

Chuang, P.-T., Kawcak, T., and McMahon, A.P. (2003). Feedback control of mammalian Hedgehog signaling by the Hedgehog-binding protein, Hip1, modulates Fgf signaling during branching morphogenesis of the lung. Genes Dev. *17*, 342–347.

Cohen, D.E. (2008). Balancing Cholesterol Synthesis and Absorption in the Gastrointestinal Tract. J. Clin. Lipidol. *2*, S1–S3.

Cohen, M., Page, K.M., Perez-Carrasco, R., Barnes, C.P., and Briscoe, J. (2014). A theoretical framework for the regulation of Shh morphogen-controlled gene expression. Development *141*, 3868–3878.

Cooper, M.K., Wassif, C.A., Krakowiak, P.A., Taipale, J., Gong, R., Kelley, R.I., Porter, F.D., and Beachy, P.A. (2003). A defective response to Hedgehog signaling in disorders of cholesterol biosynthesis. Nat. Genet. *33*, 508–513.

DeMare, L.E., Leng, J., Cotney, J., Reilly, S.K., Yin, J., Sarro, R., and Noonan, J.P. (2013). The genomic landscape of cohesin-associated chromatin interactions. Genome Res. *23*, 1224–1234.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376–380.

Dorighi, K.M., and Tamkun, J.W. (2013). The trithorax group proteins Kismet and ASH1 promote H3K36 dimethylation to counteract Polycomb group repression in Drosophila. Development *140*, 4182–4192.

Dubchak, I., Brudno, M., Loots, G.G., Pachter, L., Mayor, C., Rubin, E.M., and Frazer, K.A. (2000). Active Conservation of Noncoding Sequences Revealed by Three-Way Species Comparisons. Genome Res. *10*, 1304–1306.

Fei, T., Xia, K., Li, Z., Zhou, B., Zhu, S., Chen, H., Zhang, J., Chen, Z., Xiao, H., Han, J.-D.J., et al. (2010). Genome-wide mapping of SMAD target genes reveals the role of BMP signaling in embryonic stem cell fate determination. Genome Res. *20*, 36–44.

Flicek, P., Amode, M.R., Barrell, D., Beal, K., Billis, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fitzgerald, S., et al. (2014). Ensembl 2014. Nucleic Acids Res. *42*, D749–D755.

Frankel, N., Davis, G.K., Vargas, D., Wang, S., Payre, F., and Stern, D.L. (2010). Phenotypic robustness conferred by apparently redundant transcriptional enhancers. Nature *466*, 490–493.

Frazer, K.A., Pachter, L., Poliakov, A., Rubin, E.M., and Dubchak, I. (2004). VISTA: computational tools for comparative genomics. Nucleic Acids Res. *32*, W273–W279.

Galli, A., Robay, D., Osterwalder, M., Bao, X., Bénazet, J.-D., Tariq, M., Paro, R., Mackem, S., and Zeller, R. (2010). Distinct Roles of Hand2 in Initiating Polarity and Posterior Shh Expression during the Onset of Mouse Limb Bud Development. PLOS Genet. *6*, e1000901.

Geetha-Loganathan, P., Nimmagadda, S., and Scaal, M. (2008). Wnt signaling in limb organogenesis. Organogenesis *4*, 109–115.

Gómez-Marín, C., Tena, J.J., Acemel, R.D., López-Mayorga, M., Naranjo, S., Calle-Mustienes, E. de la, Maeso, I., Beccari, L., Aneas, I., Vielmas, E., et al. (2015). Evolutionary comparison reveals that diverging CTCF sites are signatures of ancestral topological associating domains borders. Proc. Natl. Acad. Sci. *112*, 7542–7547.

Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a given motif. Bioinformatics *27*, 1017–1018.

Gritli-Linde, A., Lewis, P., McMahon, A.P., and Linde, A. (2001). The Whereabouts of a Morphogen: Direct Evidence for Short- and Graded Long-Range Activity of Hedgehog Signaling Peptides. Dev. Biol. *236*, 364–386.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. Mol. Cell *38*, 576–589.

Hentschel, A., Zahedi, R.P., and Ahrends, R. (2016). Protein lipid modifications—More than just a greasy ballast. PROTEOMICS *16*, 759–782.

Hollnagel, A., Oehlmann, V., Heymer, J., Rüther, U., and Nordheim, A. (1999). Id Genes Are Direct Targets of Bone Morphogenetic Protein Induction in Embryonic Stem Cells. J. Biol. Chem. *274*, 19838–19845.

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

Huang, P., Nedelcu, D., Watanabe, M., Jao, C., Kim, Y., Liu, J., and Salic, A. (2016). Cellular Cholesterol Directly Activates Smoothened in Hedgehog Signaling. Cell *0*.

Infante, C.R., Park, S., Mihala, A.G., Kingsley, D.M., and Menke, D.B. (2013). Pitx1 broadly associates with limb enhancers and is enriched on hindlimb cis-regulatory elements. Dev. Biol. *374*, 234–244.

Ishida, W., Hamamoto, T., Kusanagi, K., Yagi, K., Kawabata, M., Takehara, K., Sampath, T.K., Kato, M., and Miyazono, K. (2000). Smad6 Is a Smad1/5-induced Smad Inhibitor CHARACTERIZATION OF BONE MORPHOGENETIC PROTEIN-RESPONSIVE ELEMENT IN THE MOUSE Smad6 PROMOTER. J. Biol. Chem. *275*, 6075–6079.

Jalili, V., Matteucci, M., Masseroli, M., and Morelli, M.J. (2015). Using combined evidence from replicates to evaluate ChIP-seq peaks. Bioinformatics *31*, 2761–2769.

Jeong, J., and McMahon, A.P. (2005). Growth and pattern of the mammalian neural tube are governed by partially overlapping feedback activities of the hedgehog antagonists patched 1 and Hhip1. Development *132*, 143–154.

Jessell, T.M. (2000). Neuronal specification in the spinal cord: inductive signals and transcriptional codes. Nat. Rev. Genet. *1*, 20–29.

Joshi, O., Wang, S.-Y., Kuznetsova, T., Atlasi, Y., Peng, T., Fabre, P.J., Habibi, E., Shaik, J., Saeed, S., Handoko, L., et al. (2015). Dynamic Reorganization of Extremely Long-Range Promoter-Promoter Interactions between Two States of Pluripotency. Cell Stem Cell *17*, 748–757.

Karamboulas, K., Dranse, H.J., and Underhill, T.M. (2010). Regulation of BMP-dependent chondrogenesis in early limb mesenchyme by TGF signals. J. Cell Sci. *123*, 2068–2076.

Kennedy, B.A., Deatherage, D.E., Gu, F., Tang, B., Chan, M.W.Y., Nephew, K.P., Huang, T.H.-M., and Jin, V.X. (2011). ChIP-seq Defined Genome-Wide Map of TGFβ/SMAD4 Targets: Implications with Clinical Outcome of Ovarian Cancer. PLOS ONE *6*, e22606.

Khokha, M.K., Hsu, D., Brunet, L.J., Dionne, M.S., and Harland, R.M. (2003). Gremlin is the BMP antagonist required for maintenance of Shh and Fgf signals during limb patterning. Nat. Genet. *34*, 303–307.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. *14*, R36.

Kim, S.W., Yoon, S.-J., Chuong, E., Oyolu, C., Wills, A.E., Gupta, R., and Baker, J. (2011). Chromatin and transcriptional signatures for Nodal signaling during endoderm formation in hESCs. Dev. Biol. *357*, 492–504.

Klaus, A., and Birchmeier, W. (2008). Wnt signalling and its impact on development and cancer. Nat. Rev. Cancer *8*, nrc2389.

Koinuma, D., Tsutsumi, S., Kamimura, N., Imamura, T., Aburatani, H., and Miyazono, K. (2009). Promoter-wide analysis of Smad4 binding sites in human epithelial cells. Cancer Sci. *100*, 2133–2142.

Korchynskyi, O., and Dijke, P. ten (2002). Identification and Functional Characterization of Distinct Critically Important Bone Morphogenetic Protein-specific Response Elements in the Id1 Promoter. J. Biol. Chem. *277*, 4883–4891.

Krivega, I., and Dean, A. (2012). Enhancer and promoter interactions—long distance calls. Curr. Opin. Genet. Dev. *22*, 79–85.

Kuijper, S., Feitsma, H., Sheth, R., Korving, J., Reijnen, M., and Meijlink, F. (2005). Function and regulation of Alx4 in limb development: Complex genetic interactions with Gli3 and Shh. Dev. Biol. *285*, 533–544.

Kvon, E.Z., Kamneva, O.K., Melo, U.S., Barozzi, I., Osterwalder, M., Mannion, B.J., Tissières, V., Pickle, C.S., Plajzer-Frick, I., Lee, E.A., et al. (2016). Progressive Loss of Function in a Limb Enhancer during Snake Evolution. Cell *167*, 633–642.e11.

Lagace, T.A. (2014). PCSK9 and LDLR degradation: regulatory mechanisms in circulation and in cells. Curr. Opin. Lipidol. *25*, 387–393.

Lagace, T.A., Curtis, D.E., Garuti, R., McNutt, M.C., Park, S.W., Prather, H.B., Anderson, N.N., Ho, Y.K., Hammer, R.E., and Horton, J.D. (2006). Secreted PCSK9 decreases the number of LDL receptors in hepatocytes and inlivers of parabiotic mice. J. Clin. Invest. *116*, 2995–3005.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

Laurent, A., Calabrese, M., Warnatz, H.-J., Yaspo, M.-L., Tkachuk, V., Torres, M., Blasi, F., and Penkov, D. (2015). ChIP-Seq and RNA-Seq Analyses Identify Components of the Wnt and Fgf Signaling Pathways as Prep1 Target Genes in Mouse Embryonic Stem Cells. PLOS ONE *10*, e0122518.

Laurent, F., Girdziusaite, A., Gamart, J., Barozzi, I., Osterwalder, M., Akiyama, J.A., Lincoln, J., Lopez-Rios, J., Visel, A., Zuniga, A., et al. (2017). HAND2 Target Gene Regulatory Networks Control Atrioventricular Canal and Cardiac Valve Development. Cell Rep. *19*, 1602–1613.

Lee, C.-H., Wu, J., and Li, B. (2013). Chromatin Remodelers Fine-Tune H3K36me-Directed Deacetylation of Neighbor Nucleosomes by Rpd3S. Mol. Cell *52*, 255–263.

Lee, K.L., Lim, S.K., Orlov, Y.L., Yit, L.Y., Yang, H., Ang, L.T., Poellinger, L., and Lim, B. (2011). Graded Nodal/Activin Signaling Titrates Conversion of Quantitative Phospho-Smad2 Levels into Qualitative Embryonic Stem Cell Fate Decisions. PLOS Genet. *7*, e1002130.

Lee, K.-S., Kim, H.-J., Li, Q.-L., Chi, X.-Z., Ueta, C., Komori, T., Wozney, J.M., Kim, E.-G., Choi, J.-Y., Ryoo, H.-M., et al. (2000). Runx2 Is a Common Target of Transforming Growth Factor β1 and Bone Morphogenetic Protein 2, and Cooperation between Runx2 and Smad5 Induces Osteoblast-Specific Gene Expression in the Pluripotent Mesenchymal Precursor Cell Line C2C12. Mol. Cell. Biol. *20*, 8783–8792.

Lee, M.P., Ratner, N., and Yutzey, K.E. (2014). Genome-wide Twist1 occupancy in endocardial cushion cells, embryonic limb buds, and peripheral nerve sheath tumor cells. BMC Genomics *15*, 821.

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

Lettice, L.A., Williamson, I., Wiltshire, J.H., Peluso, S., Devenney, P.S., Hill, A.E., Essafi, A., Hagman, J., Mort, R., Grimes, G., et al. (2012). Opposing Functions of the ETS Factor Family Define Shh Spatial Expression in Limb Buds and Underlie Polydactyly. Dev. Cell *22*, 459–467.

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

Lewandoski, M. (2001). Conditional control of gene expression in the mouse. Nat. Rev. Genet. *2*, 743–755.

Li, G., Ruan, X., Auerbach, R.K., Sandhu, K.S., Zheng, M., Wang, P., Poh, H.M., Goh, Y., Lim, J., Zhang, J., et al. (2012). Extensive Promoter-centered Chromatin Interactions Provide a Topological Basis for Transcription Regulation. Cell *148*, 84–98.

Li, Q., Lewandowski, J.P., Powell, M.B., Norrie, J.L., Cho, S.H., and Vokes, S.A. (2014). A Gli silencer is required for robust repression of gremlin in the vertebrate limb bud. Development *141*, 1906–1914.

Li, Y., Zhang, H., Litingtung, Y., and Chiang, C. (2006a). Cholesterol modification restricts the spread of Shh gradient in the limb bud. Proc. Natl. Acad. Sci. *103*, 6548–6553.

Li, Y., Zhang, H., Litingtung, Y., and Chiang, C. (2006b). Cholesterol modification restricts the spread of Shh gradient in the limb bud. Proc. Natl. Acad. Sci. *103*, 6548–6553.

Logan, M., Martin, J.F., Nagy, A., Lobe, C., Olson, E.N., and Tabin, C.J. (2002). Expression of Cre recombinase in the developing mouse limb bud driven by a Prxl enhancer. Genesis *33*, 77–80.

Lopez-Rios, J., Speziale, D., Robay, D., Scotti, M., Osterwalder, M., Nusspaumer, G., Galli, A., Holländer, G.A., Kmita, M., and Zeller, R. (2012). GLI3 Constrains Digit Number by Controlling Both Progenitor Proliferation and BMP-Dependent Exit to Chondrogenesis. Dev. Cell *22*, 837–848.

Lopez-Rios, J., Duchesne, A., Speziale, D., Andrey, G., Peterson, K.A., Germann, P., Ünal, E., Liu, J., Floriot, S., Barbey, S., et al. (2014). Attenuated sensing of SHH by Ptch1 underlies evolution of bovine limbs. Nature *511*, 46–51.

Matharu, N., and Ahituv, N. (2015). Minor Loops in Major Folds: Enhancer–Promoter Looping, Chromatin Restructuring, and Their Association with Transcriptional Regulation and Disease. PLOS Genet. *11*, e1005640.

Michos, O., Panman, L., Vintersten, K., Beier, K., Zeller, R., and Zuniga, A. (2004). Gremlin-mediated BMP antagonism induces the epithelial-mesenchymal feedback signaling controlling metanephric kidney and limb organogenesis. Development *131*, 3401–3410.

Mora, A., Sandve, G.K., Gabrielsen, O.S., and Eskeland, R. (2016). In the loop: promoter–enhancer interactions and bioinformatics. Brief. Bioinform. *17*, 980–995.

Morello, R., Bertin, T.K., Schlaubitz, S., Shaw, C.A., Kakuru, S., Munivez, E., Hermanns, P., Chen, Y., Zabel, B., and Lee, B. (2008). Brachy–syndactyly caused by loss of Sfrp2 function. J. Cell. Physiol. *217*, 127–137.

Morikawa, M., Koinuma, D., Tsutsumi, S., Vasilaki, E., Kanki, Y., Heldin, C.-H., Aburatani, H., and Miyazono, K. (2011). ChIP-seq reveals cell type-specific

binding patterns of BMP-specific Smads and a novel binding motif. Nucleic Acids Res. *39*, 8712–8727.

Morikawa, M., Koinuma, D., Miyazono, K., and Heldin, C.-H. (2013). Genome-wide mechanisms of Smad binding. Oncogene *32*, 1609–1615.

Moustakas, A., and Heldin, C.-H. (2009). The regulation of TGFβ signal transduction. Development *136*, 3699–3714.

Nelson, A.C., and Wardle, F.C. (2013). Conserved non-coding elements and cis regulation: actions speak louder than words. Development *140*, 1385–1395.

Nissim, S., Hasso, S.M., Fallon, J.F., and Tabin, C.J. (2006). Regulation of Gremlin expression in the posterior limb bud. Dev. Biol. *299*, 12–21.

Nissim, S., Allard, P., Bandyopadhyay, A., Harfe, B.D., and Tabin, C.J. (2007). Characterization of a novel ectodermal signaling center regulating Tbx2 and Shh in the vertebrate limb. Dev. Biol. *304*, 9–21.

Ong, C.-T., and Corces, V.G. (2011). Enhancer function: new insights into the regulation of tissue-specific gene expression. Nat. Rev. Genet. *12*, 283–293.

Oosterveen, T., Kurdija, S., Alekseenko, Z., Uhde, C.W., Bergsland, M., Sandberg, M., Andersson, E., Dias, J.M., Muhr, J., and Ericson, J. (2012). Mechanistic Differences in the Transcriptional Interpretation of Local and Long-Range Shh Morphogen Signaling. Dev. Cell *23*, 1006–1019.

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

Ovcharenko, I., Nobrega, M.A., Loots, G.G., and Stubbs, L. (2004). ECR Browser: a tool for visualizing and accessing data from comparisons of multiple vertebrate genomes. Nucleic Acids Res. *32*, W280–W286.

Parker, D.S., White, M.A., Ramos, A.I., Cohen, B.A., and Barolo, S. (2011). The cis-Regulatory Logic of Hedgehog Gradient Responses: Key Roles for Gli Binding Affinity, Competition, and Cooperativity. Sci Signal *4*, ra38-ra38.

Pavesi, G. (2016). ChIP-Seq Data Analysis to Define Transcriptional Regulatory Networks. In Network Biology, (Springer, Cham), pp. 1–14.

Peters, C., Wolf, A., Wagner, M., Kuhlmann, J., and Waldmann, H. (2004). The cholesterol membrane anchor of the Hedgehog protein confers stable membrane association to lipid-modified proteins. Proc. Natl. Acad. Sci. U. S. A. *101*, 8531–8536.

Peterson, K.A., Nishi, Y., Ma, W., Vedenko, A., Shokri, L., Zhang, X., McFarlane, M., Baizabal, J.-M., Junker, J.P., Oudenaarden, A. van, et al.

(2012). Neural-specific Sox2 input and differential Gli-binding affinity provide context and positional information in Shh-directed neural patterning. Genes Dev. *26*, 2802–2816.

Pignatti, E., Zeller, R., and Zuniga, A. (2014). To BMP or not to BMP during vertebrate limb bud development. Semin. Cell Dev. Biol. *32*, 119–127.

Pizette, S., and Niswander, L. (2001). Early steps in limb patterning and chondrogenesis. Novartis Found. Symp. *232*, 23-36-46.

Pohl, A., and Beato, M. (2014). bwtool: a tool for bigWig files. Bioinformatics *30*, 1618–1619.

Porter, F.D. (2008). Smith–Lemli–Opitz syndrome: pathogenesis, diagnosis and management. Eur. J. Hum. Genet. *16*, 535–541.

Porter, F.D., and Herman, G.E. (2011). Malformation syndromes caused by disorders of cholesterol synthesis. J. Lipid Res. *52*, 6–34.

Porter, J.A., Young, K.E., and Beachy, P.A. (1996). Cholesterol Modification of Hedgehog Signaling Proteins in Animal Development. Science *274*, 255–259.

Probst, S., Kraemer, C., Demougin, P., Sheth, R., Martin, G.R., Shiratori, H., Hamada, H., Iber, D., Zeller, R., and Zuniga, A. (2011). SHH propagates distal limb bud development by enhancing CYP26B1-mediated retinoic acid clearance via AER-FGF signalling. Dev. Camb. Engl. *138*, 1913–1923.

Qin, H., Chan, M.W., Liyanarachchi, S., Balch, C., Potter, D., Souriraj, I.J., Cheng, A.S., Agosto-Perez, F.J., Nikonova, E.V., Yan, P.S., et al. (2009). An integrative ChIP-chip and gene expression profiling to model SMAD regulatory modules. BMC Syst. Biol. *3*, 73.

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

Rainger, J., Beusekom, E. van, Ramsay, J.K., McKie, L., Al-Gazali, L., Pallotta, R., Saponari, A., Branney, P., Fisher, M., Morrison, H., et al. (2011). Loss of the BMP Antagonist, SMOC-1, Causes Ophthalmo-Acromelic (Waardenburg Anophthalmia) Syndrome in Humans and Mice. PLOS Genet. *7*, e1002114.

Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., et al. (2014). A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. Cell *159*, 1665–1680.

Reguly, T., and Wrana, J.L. (2003). In or out? The dynamics of Smad nucleocytoplasmic shuttling. Trends Cell Biol. *13*, 216–220.

Ribes, V., and Briscoe, J. (2009). Establishing and Interpreting Graded Sonic Hedgehog Signaling during Vertebrate Neural Tube Patterning: The Role of Negative Feedback. Cold Spring Harb. Perspect. Biol. *1*.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140.

Sagai, T., Masuya, H., Tamura, M., Shimizu, K., Yada, Y., Wakana, S., Gondo, Y., Noda, T., and Shiroishi, T. (2004). Phylogenetic conservation of a limb-specific, cis-acting regulator of Sonic hedgehog (Shh). Mamm. Genome *15*, 23–34.

Scherz, P.J., Harfe, B.D., McMahon, A.P., and Tabin, C.J. (2004). The Limb Bud Shh-Fgf Feedback Loop Is Terminated by Expansion of Former ZPA Cells. Science *305*, 396–399.

Sheth, R., Barozzi, I., Langlais, D., Osterwalder, M., Nemec, S., Carlson, H.L., Stadler, H.S., Visel, A., Drouin, J., and Kmita, M. (2016). Distal Limb Patterning Requires Modulation of cis-Regulatory Activities by HOX13. Cell Rep. *17*, 2913–2926.

Shi, Y., Wang, Y.-F., Jayaraman, L., Yang, H., Massagué, J., and Pavletich, N.P. (1998). Crystal Structure of a Smad MH1 Domain Bound to DNA: Insights on DNA Binding in TGF-β Signaling. Cell *94*, 585–594.

Shlyueva, D., Stampfel, G., and Stark, A. (2014). Transcriptional enhancers: from properties to genome-wide predictions. Nat. Rev. Genet. *15*, 272–286.

Siepel, A., Bejerano, G., Pedersen, J.S., Hinrichs, A.S., Hou, M., Rosenbloom, K., Clawson, H., Spieth, J., Hillier, L.W., Richards, S., et al. (2005). Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res. *15*, 1034–1050.

Sloggett, C., Goonasekera, N., and Afgan, E. (2013). BioBlend: automating pipeline analyses within Galaxy and CloudMan. Bioinformatics *29*, 1685–1686.

Stevens, T.J., Lando, D., Basu, S., Atkinson, L.P., Cao, Y., Lee, S.F., Leeb, M., Wohlfahrt, K.J., Boucher, W., O'Shaughnessy-Kirwan, A., et al. (2017). 3D structures of individual mammalian genomes studied by single-cell Hi-C. Nature *544*, 59–64.

Taher, L., Collette, N.M., Murugesh, D., Maxwell, E., Ovcharenko, I., and Loots, G.G. (2011). Global Gene Expression Analysis of Murine Limb Development. PLOS ONE *6*, e28358.

Tanaka, Y., Kawahashi, K., Katagiri, Z.-I., Nakayama, Y., Mahajan, M., and Kioussis, D. (2011). Dual Function of Histone H3 Lysine 36 Methyltransferase ASH1 in Regulation of Hox Gene Expression. PLOS ONE *6*, e28171.

Tint, G.S., Yu, H., Shang, Q., Xu, G., and Patel, S.B. (2006). The use of the Dhcr7 knockout mouse to accurately determine the origin of fetal sterols. J. Lipid Res. *47*, 1535–1541.

Tsankov, A.M., Gu, H., Akopian, V., Ziller, M.J., Donaghey, J., Amit, I., Gnirke, A., and Meissner, A. (2015). Transcription factor binding dynamics during human ES cell differentiation. Nature *518*, 344–349.

Tukachinsky, H., Kuzmickas, R.P., Jao, C.Y., Liu, J., and Salic, A. (2012). Dispatched and Scube Mediate the Efficient Secretion of the Cholesterol-Modified Hedgehog Ligand. Cell Rep. *2*, 308–320.

Tyner, C., Barber, G.P., Casper, J., Clawson, H., Diekhans, M., Eisenhart, C., Fischer, C.M., Gibson, D., Gonzalez, J.N., Guruvadoo, L., et al. (2017). The UCSC Genome Browser database: 2017 update. Nucleic Acids Res. *45*, D626–D634.

Tzchori, I., Day, T.F., Carolan, P.J., Zhao, Y., Wassif, C.A., Li, L., Lewandoski, M., Gorivodsky, M., Love, P.E., Porter, F.D., et al. (2009). LIM homeobox transcription factors integrate signaling events that control three-dimensional limb patterning and growth. Development *136*, 1375–1385.

Verheyden, J.M., and Sun, X. (2008). An Fgf/Gremlin inhibitory feedback loop triggers termination of limb bud outgrowth. Nature *454*, 638–641.

Vietri Rudan, M., Barrington, C., Henderson, S., Ernst, C., Odom, D.T., Tanay, A., and Hadjur, S. (2015). Comparative Hi-C Reveals that CTCF Underlies Evolution of Chromosomal Domain Architecture. Cell Rep. *10*, 1297–1309.

Visel, A., Minovitsky, S., Dubchak, I., and Pennacchio, L.A. (2007). VISTA Enhancer Browser—a database of tissue-specific human enhancers. Nucleic Acids Res. *35*, D88–D92.

Visel, A., Rubin, E.M., and Pennacchio, L.A. (2009). Genomic views of distant-acting enhancers. Nature *461*, 199–205.

Wagner, E.J., and Carpenter, P.B. (2012). Understanding the language of Lys36 methylation at histone H3. Nat. Rev. Mol. Cell Biol. *13*, nrm3274.

Wang, Y., Zhang, B., Zhang, L., An, L., Xu, J., Li, D., Choudhary, M.N., Li, Y., Hu, M., Hardison, R., et al. (2017). The 3D Genome Browser: a web-based browser for visualizing 3D genome organization and long-range chromatin interactions. bioRxiv 112268.

Whalen, S., Truty, R.M., and Pollard, K.S. (2016). Enhancer–promoter interactions are encoded by complex genomic signatures on looping chromatin. Nat. Genet. *48*, ng.3539.

Woollett, L., and Heubi, J.E. (2000). Fetal and Neonatal Cholesterol Metabolism. In Endotext, L.J. De Groot, G. Chrousos, K. Dungan, K.R.

Feingold, A. Grossman, J.M. Hershman, C. Koch, M. Korbonits, R. McLachlan, M. New, et al., eds. (South Dartmouth (MA): MDText.com, Inc.), p.

Xiao, X., Tang, J.-J., Peng, C., Wang, Y., Fu, L., Qiu, Z.-P., Xiong, Y., Yang, L.-F., Cui, H.-W., He, X.-L., et al. (2017). Cholesterol Modification of Smoothened Is Required for Hedgehog Signaling. Mol. Cell *0*.

Yang, X., Li, C., Herrera, P.-L., and Deng, C.-X. (2002). Generation of Smad4/Dpc4 conditional knockout mice. Genesis *32*, 80–81.

Ye, J., and DeBose-Boyd, R.A. (2011). Regulation of Cholesterol and Fatty Acid Synthesis. Cold Spring Harb. Perspect. Biol. *3*.

Zawel, L., Le Dai, J., Buckhaults, P., Zhou, S., Kinzler, K.W., Vogelstein, B., and Kern, S.E. (1998). Human Smad3 and Smad4 Are Sequence-Specific Transcription Activators. Mol. Cell *1*, 611–617.

Zeller, R., and Zuniga, A. (2007). Shh and Gremlin1 chromosomal landscapes in development and disease. Curr. Opin. Genet. Dev. *17*, 428–434.

Zeller, R., López-Ríos, J., and Zuniga, A. (2009). Vertebrate limb bud development: moving towards integrative analysis of organogenesis. Nat. Rev. Genet. *10*, 845–858.

Zeng, X., Goetz, J.A., Suber, L.M., Scott, W.J., Schreiner, C.M., and Robbins, D.J. (2001). A freely diffusible form of Sonic hedgehog mediates long-range signalling. Nature *411*, 716–720.

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

Zhao, J., Shi, W., Chen, H., and Warburton, D. (2000). Smad7 and Smad6 Differentially Modulate Transforming Growth Factor β-induced Inhibition of Embryonic Lung Morphogenesis. J. Biol. Chem. *275*, 23992–23997.

Zhu, A.J., and Scott, M.P. (2004). Incredible journey: how do developmental signals travel through tissue? Genes Dev. *18*, 2985–2997.

Zhu, J., and Mackem, S. John Saunders' ZPA, Sonic hedgehog and digit identity – How does it really all work? Dev. Biol.

Zhu, J., Nakamura, E., Nguyen, M.-T., Bao, X., Akiyama, H., and Mackem, S. (2008). Uncoupling Sonic Hedgehog Control of Pattern and Expansion of the Developing Limb Bud. Dev. Cell *14*, 624–632.

Zuniga, A. (2015). Next generation limb development and evolution: old questions, new perspectives. Development *142*, 3810–3820.

Zuniga, A., and Zeller, R. (1999). Gli3 (Xt) and formin (ld) participate in the positioning of the polarising region and control of posterior limb-bud identity. Development *126*, 13–21.

Zuniga, A., Michos, O., Spitz, F., Haramis, A.-P.G., Panman, L., Galli, A., Vintersten, K., Klasen, C., Mansfield, W., Kuc, S., et al. (2004). Mouse limb deformity mutations disrupt a global control region within the large regulatory landscape required for Gremlin expression. Genes Dev. *18*, 1553–1564.

Zuniga, A., Laurent, F., Lopez-Rios, J., Klasen, C., Matt, N., and Zeller, R. (2012). Conserved cis-regulatory regions in a large genomic landscape control SHH and BMP-regulated Gremlin1 expression in mouse limb buds. BMC Dev. Biol. *12*, 23.

12. Appendixes

12.1. List of down-regulated E10.0 SMAD4 targets

Name	FC≥1.2	FDR≤0.1
0610007P14Rik	1.223213211	0.056799258
2010300C02Rik	1.504834891	0.073984364
4933404O12Rik	2.004332684	0.044882612
5730559C18Rik	1.290858343	0.019139858
6330403K07Rik	1.508664379	3.89524E-07
Abhd16a	1.24608185	0.09352162
Abhd8	1.444853433	0.031398748
Acot11	1.510298434	0.000706461
Adck2	1.269849919	0.053510024
Ak1	1.288264513	0.084434247
Alx4	1.276624344	0.048800519
Arhgef19	1.685088225	0.010136361
Asphd2	1.361954393	0.035797172
Atoh8	1.58651375	0.006026576
B4galnt2	2.941716269	0.002239574
Bcl11a	1.641380458	1.70703E-06
Bmf	1.404071466	0.000262814
Camkk1	1.415235228	0.09032415
Capn10	1.245119445	0.069714202
Ccdc114	1.503556462	0.080130779
Cdc42ep4	1.230723108	0.026268054
Cdo1	1.364654367	1.33791E-05
Cldn10	3.204662217	5.47353E-08
Cnr1	1.654817759	0.065354795
Col9a1	1.397206172	0.073609136
Coq2	1.322902689	0.029650811
Cpm	1.297272519	0.092481427
Cyp51	1.461316389	2.9263E-05
Ddt	1.266929421	0.096916279
Dennd4b	1.26803366	0.096293979
Dhcr7	1.687168612	3.69244E-12
Doc2a	1.311881765	0.049983445
Dpysl2	1.224445424	0.033878106
Dtx3	1.281518615	0.09231527
Efs	1.296565145	0.074977557
Map1lc3b	1.208616112	0.046004698

Name	FC≥1.2	FDR≤0.1
Ehd1	1.22030945	0.039005303
Elovl1	1.222853191	0.07258805
Emilin1	1.352773594	0.038229041
Enho	1.346124066	0.090965994
Entpd6	1.241480606	0.098154326
Evx1	2.115921936	0.016857195
Fam109a	1.339684573	0.013832231
Fam189a2	1.685234454	0.005669977
Fam73b	1.235010985	0.086811535
Fdft1	1.508039858	1.13E-08
Fstl1	1.25480529	0.056799258
Gm8801	73.48224447	0.000130513
Gorasp1	1.248886959	0.083061136
Gpc1	1.342314791	0.017366866
Gyltl1b	1.387249006	0.012330766
H2-K1	4.538026704	0.000805342
Hdac5	1.327257845	0.066331597
Hoxb5	1.342350441	0.075793457
Id1	2.342791413	8.72E-14
Id2	1.323485236	0.006665039
Id3	2.343132461	1.86E-07
Idi1	1.567668537	3.19E-05
Idua	1.621408573	0.001754518
Irf5	1.828965929	0.03063142
Isyna1	1.289655106	0.067052356
Itgb5	1.250836931	0.024837689
Ivd	1.210725265	0.056799258
Jag2	1.40250283	0.012385673
Kctd15	1.275631279	0.003800507
Klc4	1.322491584	0.039005303
Krt5	1.310752469	0.047302977
Ldlr	1.481010476	1.57E-05
Lhx2	1.377767616	0.001699451
Lmbr1l	1.299074779	0.050344764
Lpcat3	1.227511788	0.063115812
Sfrp2	1.296565145	0.074977557

Mfap4	1.367496175	0.015928149
Mfsd2a	1.305163384	0.073752477
Mib2	1.392808946	0.0313346
Mknk2	1.221946653	0.063023398
Mnt	1.272643152	0.09352162
Mogat2	1.313646569	0.008677396
Mpped2	1.223213211	0.056799258
Msx2	1.504834891	0.073984364
Mt1	2.004332684	0.044882612
Mt2	1.290858343	0.019139858
Mtch1	1.508664379	3.90E-07
Mxd4	1.24608185	0.09352162
Ndrg1	1.444853433	0.031398748
Nfatc4	1.510298434	0.000706461
Nfs1	1.269849919	0.053510024
Pafah1b3	1.288264513	0.084434247
Palmd	1.276624344	0.048800519
Pank1	1.685088225	0.010136361
Pck2	1.361954393	0.035797172
Pcsk9	1.58651375	0.006026576
Pkdcc	2.941716269	0.002239574
Pmepa1	1.641380458	1.71E-06
Pmvk	1.404071466	0.000262814
Prrx2	1.415235228	0.09032415
Pth1r	1.245119445	0.069714202
Ptpdc1	1.503556462	0.080130779
Ptpru	1.230723108	0.026268054
Ptx3	1.364654367	1.34E-05
Pycr2	3.204662217	5.47E-08
Rab36	1.654817759	0.065354795
Rab38	1.397206172	0.073609136
Rhpn1	1.322902689	0.029650811
Rilpl2	1.297272519	0.092481427
Rnh1	1.461316389	2.93E-05
Rspo4	1.266929421	0.096916279
Rusc1	1.26803366	0.096293979
Samd1	1.687168612	3.69E-12
Sc5d	1.311881765	0.049983445
Scd2	1.224445424	0.033878106
Scn4b	1.281518615	0.09231527

Sgsm1	1.22030945	0.039005303
Sh3bp5l	1.222853191	0.07258805
Shkbp1	1.352773594	0.038229041
Slc26a6	1.346124066	0.090965994
Slc2a4	1.241480606	0.098154326
Slc38a6	2.115921936	0.016857195
Smad4	1.339684573	0.013832231
Smad6	1.685234454	0.005669977
Smad7	1.235010985	0.086811535
Snai1	1.508039858	1.13E-08
Sqle	1.25480529	0.056799258
Srebf1	73.48224447	0.000130513
Stard4	1.248886959	0.083061136
Stk36	1.342314791	0.017366866
Suox	1.387249006	0.012330766
Tbx2	4.538026704	0.000805342
Thbs2	1.327257845	0.066331597
Tmco6	1.342350441	0.075793457
Toe1	2.342791413	8.72E-14
Tsen34	1.323485236	0.006665039
Ucp2	2.343132461	1.86E-07
Unc45a	1.567668537	3.19E-05
Wbp1	1.621408573	0.001754518
Wdyhv1	1.828965929	0.03063142
Wif1	1.289655106	0.067052356
Wnt5b	1.250836931	0.024837689
Zbtb8a	1.210725265	0.056799258
Zfp14	1.40250283	0.012385673
Zfp78	1.275631279	0.003800507
Zfyve27	1.322491584	0.039005303
Shkbp1	0.739221999	0.038229041

12.2. List of up-regulated E10.0 SMAD4 targets

Name	FC≥1.2	FDR≤0.1
0610010F05Rik	1.273105085	0.069714202
4932438A13Rik	1.470140636	0.020535029
9430020K01Rik	1.209416141	0.094621319
Acta2	2.881635984	0.001214181
Adamts9	1.487338344	0.005981323
Ahctf1	1.257322348	0.080130779
Akap12	1.40910242	0.040021864
Akap9	1.55211704	0.003476453
Ankhd1	1.29782512	0.026223827
Ano1	1.818360084	2.36053E-05
Apc	1.298197037	0.016437025
Arhgdib	1.338648021	0.037066172
Ascc3	1.220495361	0.056918764
Ash1l	1.325395395	0.022636805
Aspm	1.310442695	0.053604355
Asxl2	1.250567225	0.052275677
Atg2b	1.229051648	0.089607264
Atp6v0e2	1.410200138	0.035091533
Baz1a	1.213954453	0.063023398
Baz2a	1.247188844	0.07127768
Baz2b	1.293154627	0.018426838
Birc6	1.519544554	0.066331597
Bmp4	1.303693552	0.010187961
Bnc1	1.400389972	0.076022368
Bnc2	1.508321253	0.000971332
Bptf	1.312476885	0.014565125
Cadm4	1.301536212	0.074977557
Casc5	1.349435091	0.040138936
Ccdc80	1.496872439	0.00057127
Ccdc88a	1.242155564	0.045146856
Cenpe	1.392559841	0.008569638
Cep170	1.247979739	0.026409533
Chd1	1.275838644	0.01665887
Chd7	1.34722048	0.065354795
Chd9	1.273240321	0.07258805
Chrna1	4.302732456	6.39271E-18

Name	FC≥1.2	FDR≤0.1
Col12a1	2.269966443	0.000158317
Cx3cr1	2.531228175	6.4546E-05
Cyp1b1	1.459490179	0.000158317
D630045J12 Rik	1.235482401	0.09536828
D830031N03 Rik	1.408348094	0.01302066
Dapk1	1.347587138	6.45938E-05
Ddx21	1.232108575	0.031753565
Dgkh	1.502180399	0.037056449
Dip2c	1.273581841	0.047302977
Dnajc13	1.222884788	0.09231527
Dnajc2	1.238032937	0.007119986
Dock4	1.337828867	0.03063142
Dock7	1.320985414	0.015299074
Dync1h1	1.658706362	0.014285017
E130114P18 Rik	2.242280347	0.002361324
E330013P04 Rik	1.636553183	0.023249153
Eftud1	1.215564887	0.053822519
Ep400	1.332176341	0.073455313
Fam43a	1.45761561	0.000168566
Fez1	1.564071452	0.017366866
Flnc	1.639152012	0.01329156
Fryl	1.343152566	0.016613589
Gkap1	1.278909471	0.035565632
Gldc	1.284840271	0.049412771
Gm5860	2.027088716	0.090455686
Gm8994	18.33196422	0.02900654
H2-T24	2.033852422	0.006665039
Herc1	1.499037309	0.042921494
Hic1	1.785760887	0.067052356
Hipk2	1.307452722	0.056799258
Hmcn1	1.945679564	0.001266265
Hopx	2.000106888	0.09276796
Hoxd13	1.353435296	0.043642618
Igfbp3	1.366188867	0.003543566
Ism1	1.447962226	0.037889567
Itga8	1.500719406	0.013289505

Itgb3	1.420416371	0.042274448
Itpr1	1.337880133	0.062830909
Itsn2	1.271206004	0.017460401
Jmjd1c	1.27819505	0.02273063
Kit	1.271872823	0.071456821
Larp1b	1.628684433	0.03785712
Lats2	1.213096551	0.073455313
Lcor	1.286326594	0.09231527
Loxl1	1.60227637	0.083061136
Lrrfip1	1.209524224	0.035473289
Luc7l3	1.257194474	0.045664283
Magi3	1.227180964	0.077799941
Mamdc2	1.524765082	0.083061136
Mdn1	1.680598047	0.039823034
Med13	1.266791398	0.078245833
Meox1	1.32673576	0.089583765
Myf5	1.29782512	0.026223827
Myh10	1.818360084	2.36053E-05
Myl9	1.298197037	0.016437025
Myo5b	1.338648021	0.037066172
Nab1	1.220495361	0.056918764
Ncapd2	1.325395395	0.022636805
Nckap1l	1.310442695	0.053604355
Ncor1	1.211792717	0.065650592
Ngfr	1.431169375	0.047273752
Nin	1.410200138	0.035091533
Nlrp1a	1.61428141	0.071456821
Nol7	1.247188844	0.07127768
Pcm1	1.262827826	0.081895568
Pcsk5	1.249835766	0.044414578
Pde3a	1.303693552	0.010187961
Pim1	1.221057981	0.075613128
Pkhd1l1	1.843030246	0.076329069
Plagl1	1.312476885	0.014565125
Plekha6	1.301536212	0.074977557
Plxna4	1.349435091	0.040138936
Polq	1.496872439	0.00057127
Polr3a	1.251546203	0.036354469
Prpf8	1.392559841	0.008569638
Prrc2c	1.247979739	0.026409533
Ptch1	1.275838644	0.01665887
Qser1	1.34722048	0.065354795

Rgl1	1.210888172	0.075613128
Rnasel	4.302732456	6.39271E-18
Rnf150	1.328741552	0.015908128
Rock2	1.213467697	0.091346996
Rragd	1.406355677	0.054954981
Rsf1	1.287593941	0.036354469
Sash1	1.243549639	0.063115812
Setbp1	1.263940721	0.055280462
Setd2	1.408348094	0.01302066
Sfrp1	1.323811506	0.098429423
Shisa3	1.232108575	0.031753565
Shisa7	1.479200646	0.097869431
Six1	1.247869873	0.066427079
Slc8a3	1.222884788	0.09231527
Smchd1	1.238032937	0.007119986
Smoc1	1.337828867	0.03063142
Snrnp200	1.275154004	0.098354292
Sorbs2	1.658706362	0.014285017
Spon1	2.242280347	0.002361324
Srek1	1.201616904	0.084518133
Srrm2	1.31127167	0.065354795
Ssh2	1.262088185	0.066670988
Stk38l	1.45761561	0.000168566
Stk39	1.564071452	0.017366866
Sulf1	1.397904395	0.042274448
Tbx19	1.343152566	0.016613589
Tcf15	1.435320008	0.084842802
Tcf4	1.284840271	0.049412771
Thsd4	1.307885144	0.062232478
Tnc	3.228952512	0.034316441
Tnnt1	2.033852422	0.006665039
Topors	1.499037309	0.042921494
Trio	1.785760887	0.067052356
Trip11	1.242842119	0.090596344
Ttn	1.945679564	0.001266265
Ubr5	1.298612923	0.0560448
Upk3b	1.353435296	0.043642618
Usp53	1.268446323	0.096098197
Vax2	1.447962226	0.037889567
Vps13d	1.435002121	0.05838564
Wdfy3	1.334215627	0.086811535
Wisp1	1.337880133	0.062830909

Xrn1	11.27819505	0.02273063
Wt1	1.271206004	0.017460401
Ylpm1	1.271872823	0.071456821
Zeb1	1.628684433	0.03785712
Zfml	1.216144599	0.077194163
Zfp292	1.254141464	0.08078979
Zfp74	1.259487643	0.056452159

12.3. List of down-regulated E10.5 SMAD4 targets

Name	FC≥1.2	FDR≤0.1
2610305D13Rik	4.619970705	0.001991779
6330403K07Rik	1.353588273	0.000546566
Abhd14a	1.355427296	0.029986532
Acta2	2.400462392	0.005196229
Aldh1a2	2.283367421	0.000289412
Alx1	1.513382553	0.018069258
Alx4	1.357309563	0.006144064
Angpt1	1.353412474	0.083394116
Asb4	1.494894253	5.28019E-05
Asphd2	1.332305714	0.09503196
Atoh8	1.444588287	0.075301781
Baiap2l1	1.404523376	0.016173776
Bcat1	1.322356258	0.005967889
Bcl11a	1.346759125	0.017406184
Bmp5	1.523089692	0.05309409
Ccdc141	1.65466589	0.005193884
Cdk5r1	1.373865966	0.078104418
Cdo1	1.344156424	3.83944E-05
Cdx1	1.824085654	0.09503196
Cldn10	4.251051577	4.01232E-10
Coq2	1.410782225	0.003688684
Crb2	1.641471701	0.01014109
Csf2rb	1.697824353	0.036049917
Cyp1b1	1.297407324	0.031093575
Des	2.311314246	0.000519718
Diras2	1.914453053	0.000862856
Dkk1	3.725613563	6.81559E-16
Dkk2	1.364160227	0.072729724
Dpep1	1.897850231	7.28145E-07
Efna3	1.633570021	0.003734445

Name	FC≥1.2	FDR≤0.1
Elovl2	1.239404666	0.055784591
Erbb3	1.333604808	0.048572688
Fgfr2	1.397455482	0.000179595
Fgfr3	1.42736082	0.008990632
Fndc5	1.965394259	6.6217E-07
Fzd10	1.27113332	0.013500803
Gabrb3	1.680025665	0.004992025
Gata2	1.945211276	0.000936015
Gata3	1.698673545	0.0175389
Gata5	5.783389317	5.47101E-12
Gata6	1.860073566	0.075744724
Gchfr	2.016478848	0.003715917
Glis1	1.760618466	0.043329953
Gm6682	3.728412699	0.029986532
Gm9199	100.8708348	2.89903E-05
Gpx3	1.317074994	0.036843284
Grem1	2.118345336	2.14702E-11
Gsc	2.091532905	0.002758797
Hey1	1.397900999	0.008653745
Hoxb8	1.497658975	0.016173776
Hr	1.602195775	0.036395698
Id1	2.496708281	9.93875E-16
Id2	1.551940362	4.23804E-07
Id3	2.582655579	1.83175E-09
Igfbp4	1.296772072	0.003980822
Irak1bp1	1.717902655	0.062933466
Irf5	2.91059218	2.24822E-06
Itga3	1.382776097	0.086584833
Jph2	2.006502376	0.000223032
Kcnip3	1.628902274	0.012251099

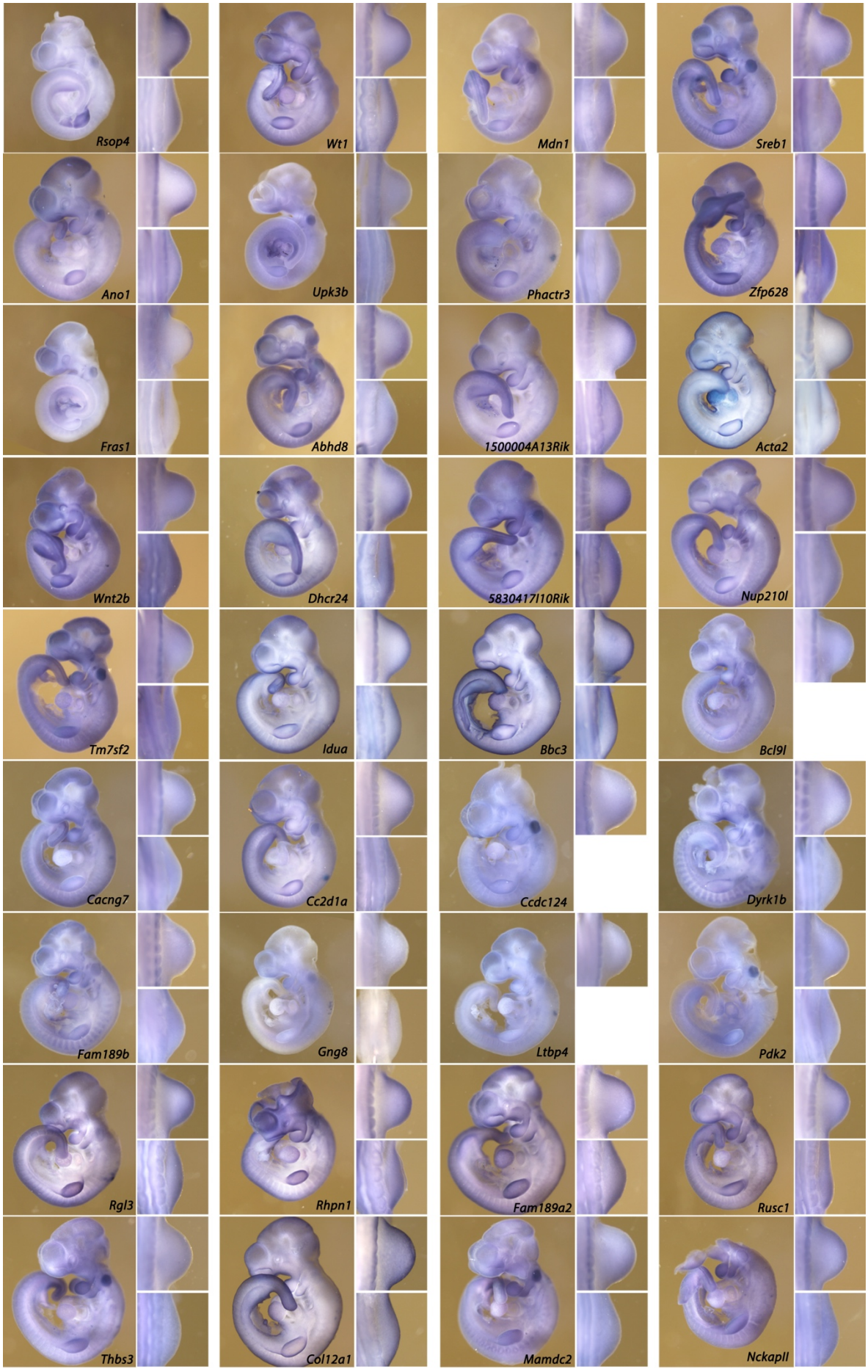
Kctd13	1.363316763	0.03678417
Kctd15	1.201186624	0.079937738
Lhfp	1.245270492	0.09630261
Lhx2	1.448922718	0.00012372
Meis3	1.735009285	0.001724936
Mogat2	1.366304714	0.001677719
Mpp2	1.259832148	0.092164588
Msx2	3.133139691	3.21169E-18
Mtmr11	1.340822746	0.023415401
Mycn	1.295475301	0.010649228
Myh3	4.513264386	0.00051811
Mylpf	2.032415074	0.021511832
Myog	7.214930774	0.000253668
Nfs1	1.248895975	0.085359385
Pax1	6.076301242	2.83811E-15
Pcdh7	1.36163166	0.036898113
Pkdcc	1.747796183	2.65803E-07
Pmepa1	1.359416637	0.082861523
Pou4f1	2.25929917	0.027751185
Ppm1m	1.348496256	0.064836824
Ppp1r14c	1.58925043	0.086797957
Ppp2r1b	1.244210313	0.022759408
Prdm6	2.183499488	0.000267873
Prrx2	1.764758136	6.68235E-08
Psd2	1.603287481	0.006520976
Ptges	1.31442431	0.046901341
Pth1r	1.522504475	8.75461E-06
Ptpru	1.360930887	0.057436675
Rarb	1.326699637	0.05394564
Rasl12	1.48732778	0.037738494
Rbpms	1.214265015	0.087697872
Rgs5	3.446046865	1.83175E-09
Rspo4	2.672974934	8.78744E-23
Runx2	2.244776832	1.78088E-07
Rxfp2	1.738181157	0.07116529
Scd1	1.295704812	0.002758797
Scd2	1.225502847	0.029986532
Scn4b	1.622396896	0.001203892
Slc16a3	1.43717567	0.027455416
Slc38a3	1.797917334	0.016134579
Smad4	4.457697015	2.94151E-94
Smad6	1.418101899	0.056234884
Smad7	1.364379553	0.005075381
Smad9	4.704523045	6.70383E-11
Smyd3	1.28943657	0.055784591
Snai1	1.625943442	0.00044489
Snai2	1.478269605	1.82476E-06
Snhg6	1.41528784	0.055076437
Sowahb	1.772763629	0.015870524
Spire1	1.222103383	0.070422456
St3gal1	1.390079179	0.004389587
Sv2a	1.53403362	0.005193884
Syt7	1.757984855	4.1323E-05
Tbx2	2.005274655	1.45684E-08
Tbx3	1.472322155	6.24825E-06
Thbs2	1.878628799	0.000783597
Tnni1	1.965492685	0.047236802
Tshz1	1.365437427	0.00012372
Ttn	3.304995298	0.011149265
Unc5b	1.233824807	0.048365955
Upk3b	3.369708068	0.034545596
Ush1c	1.566590844	0.068563798
Wdyhv1	1.43330712	0.011575155
Wt1	2.667628433	0.007918727
Zfp503	1.216830312	0.073484866

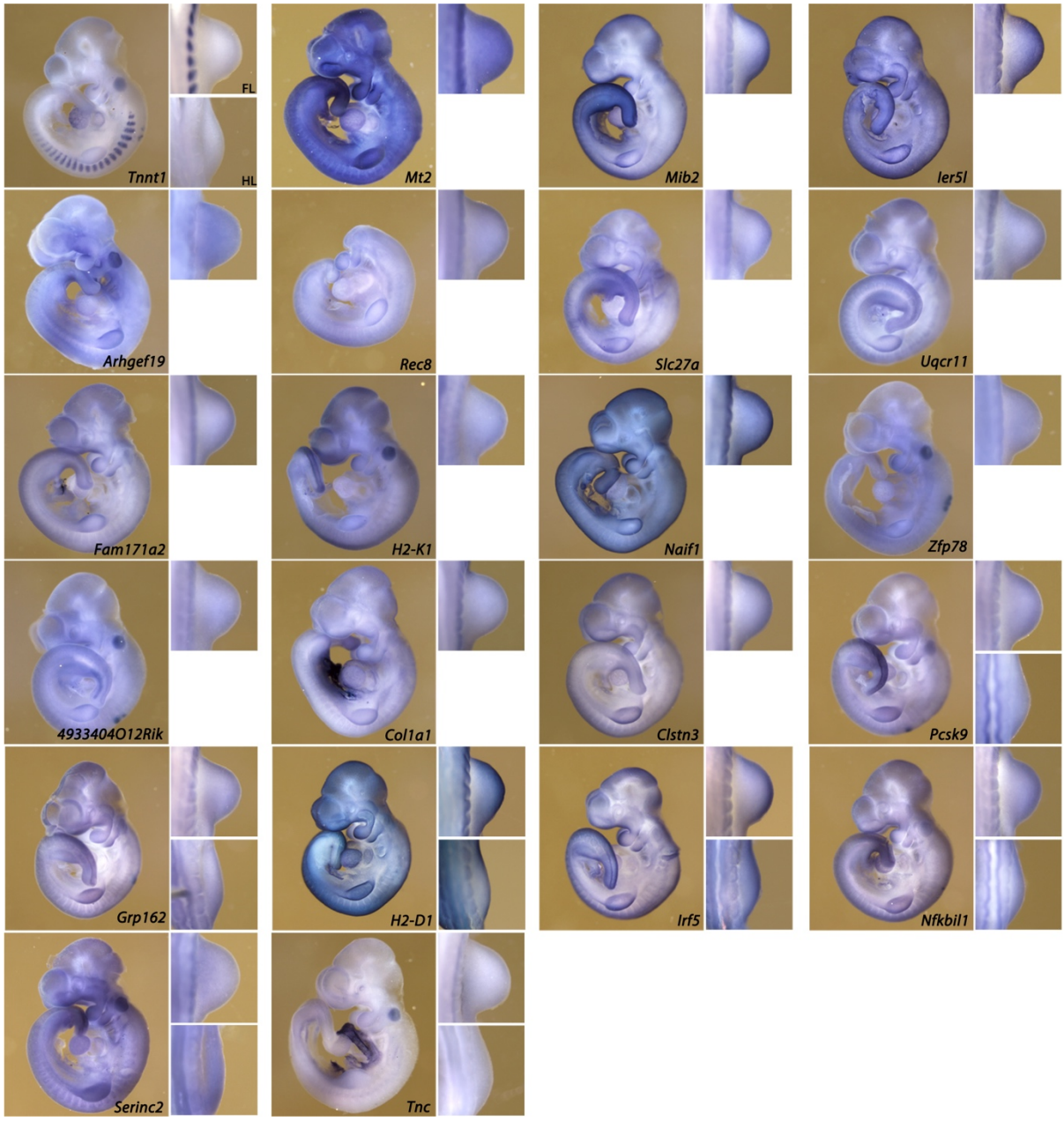
12.4. List of up-regulated E10.5 SMAD4 targets

Name	FC≥1.2	FDR≤0.1
2310030G06Rik	1.650118041	0.059807184
Ankrd6	1.301570513	0.016450526
Arhgef26	1.662223047	2.91013E-05
Bmp3	3.005347102	0.000253668
Cacna1g	1.329255179	0.03678417
Camkv	1.493711991	0.002758797
Cbfa2t3	1.610057198	0.000948731
Cdh4	1.598253052	0.041265602
Cgnl1	1.386480188	0.000208984
Cln6	1.260096116	0.052335746
Cpa2	1.402517053	0.013128303
Cpne5	1.552670699	0.004926138
Egr1	1.914656361	0.001271825
Elf3	1.354580856	0.090004931
Elovl7	1.805183656	0.032054834
Enpp5	1.570888219	0.003715917
Fam69c	1.783714696	0.035569151
Gldc	1.531116837	2.89903E-05
Gsg1l	5.445416132	3.83863E-17
Hemk1	2.122968991	0.036898113
Hic1	3.202513174	1.15411E-07
Hist2h2bb	6.838020283	0.070234156
Hmcn1	1.728292144	0.016134579
Itsn2	1.247887278	0.043311275
Jmjd7	1.575642141	0.056609109
Lrig3	1.253999376	0.059807184
Mertk	1.321740003	0.035569151
Mtss1	1.274045271	0.008540741
Myl12b	1.237308862	0.088064112
Myo5b	1.622816836	0.006649154
Nab1	1.397221384	0.000546566
Nes	1.415448776	0.000340858
Nfatc2	2.15559656	6.24825E-06
Ngfr	1.562271502	0.007828651
Nnat	1.354111416	0.000289412
Pea15a	1.257700158	0.01942771
Pmaip1	2.156528457	0.020820366

Name	FC≥1.2	FDR≤0.1
Ptgis	1.429125032	0.01014109
Rapgef5	1.733292157	0.000472245
Rgmb	1.727604332	9.52748E-08
Sash1	1.294005146	0.018927654
Sema6a	1.297386967	0.054586591
Sfmbt2	1.523483113	0.034088767
Ska1	1.284429623	0.05915533
Smoc2	2.826609617	1.34424E-13
Sorbs2	1.634033672	2.73057E-07
Spry2	1.410420365	0.000370853
Srgap1	1.227995797	0.092512731
Srms	1.910879732	0.015431625
Stambpl1	1.255539157	0.08799457
Stc1	1.49821295	0.029178832
Stk38l	1.499392347	5.96723E-05
Syne2	1.386775436	0.099079413
Synrg	1.429148997	0.001866157
Syt13	1.513325361	0.01151969
Tbx19	212.395487	0.004039331
Tbx4	1.744224272	0.098563338
Tcf15	1.465162912	0.077695558
Tdrkh	1.568022679	0.016173776
Thbs1	1.412328419	0.00083572
Tubb4a	1.778693913	0.013500803
Wfikkn2	3.4247874	3.14784E-11
Xrn1	1.310146706	0.096993443

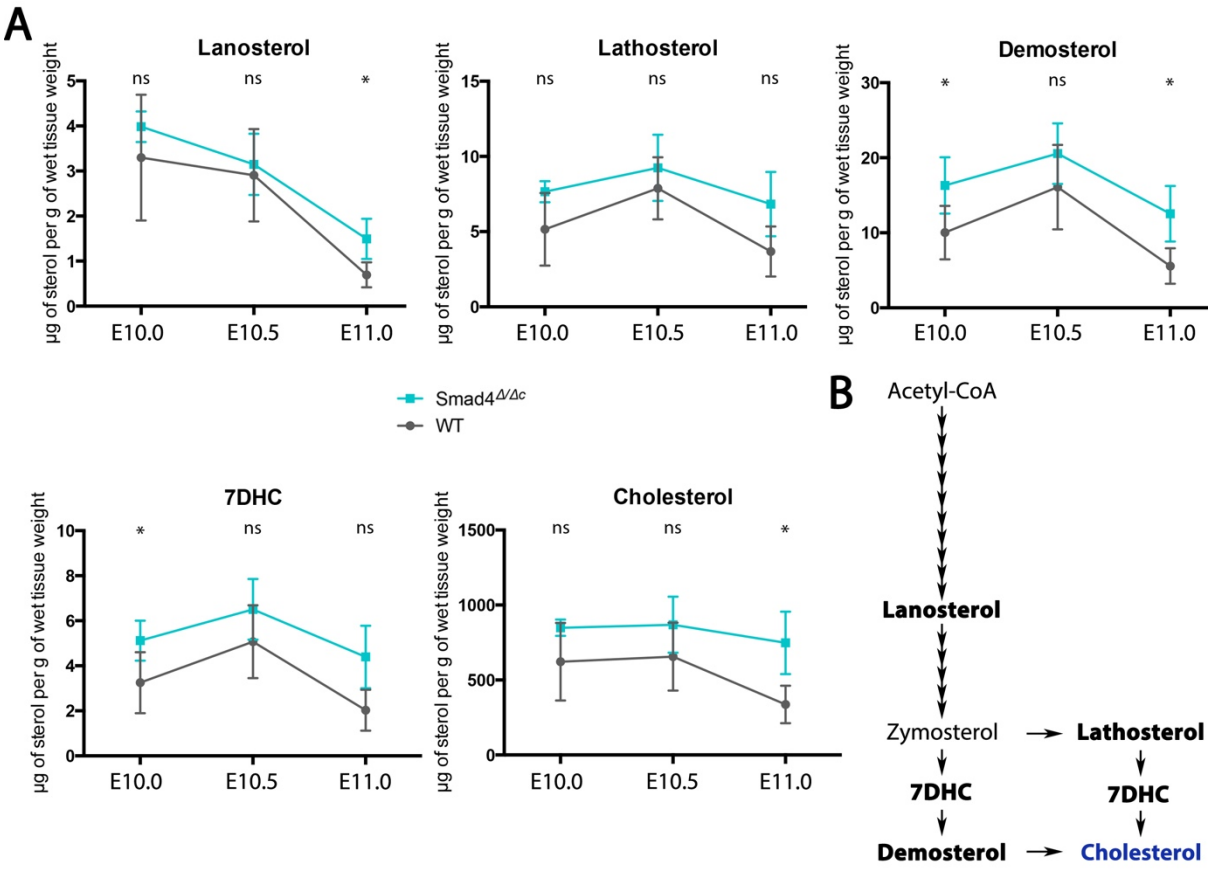
12.5. WISH screen


Rsop4
Wt1
Mdn1
Sreb1
Ano1
Upk3b
Phactr3
Zfp628
Fras1
Abhd8
1500004A13Rik
Acta2
Wnt2b
Dhcr24
5830417I10Rik
Nup210l
Tm7sf2
Idua
Bbc3
Bcl9l
Cacng7
Cc2d1a
Ccdc124
Dyrk1b
Fam189b
Gng8
Ltbp4
Pdk2
Rgl3
Rhpn1
Fam189a2
Rusc1
Thbs3
Col12a1
Mamdc2
Nckapll



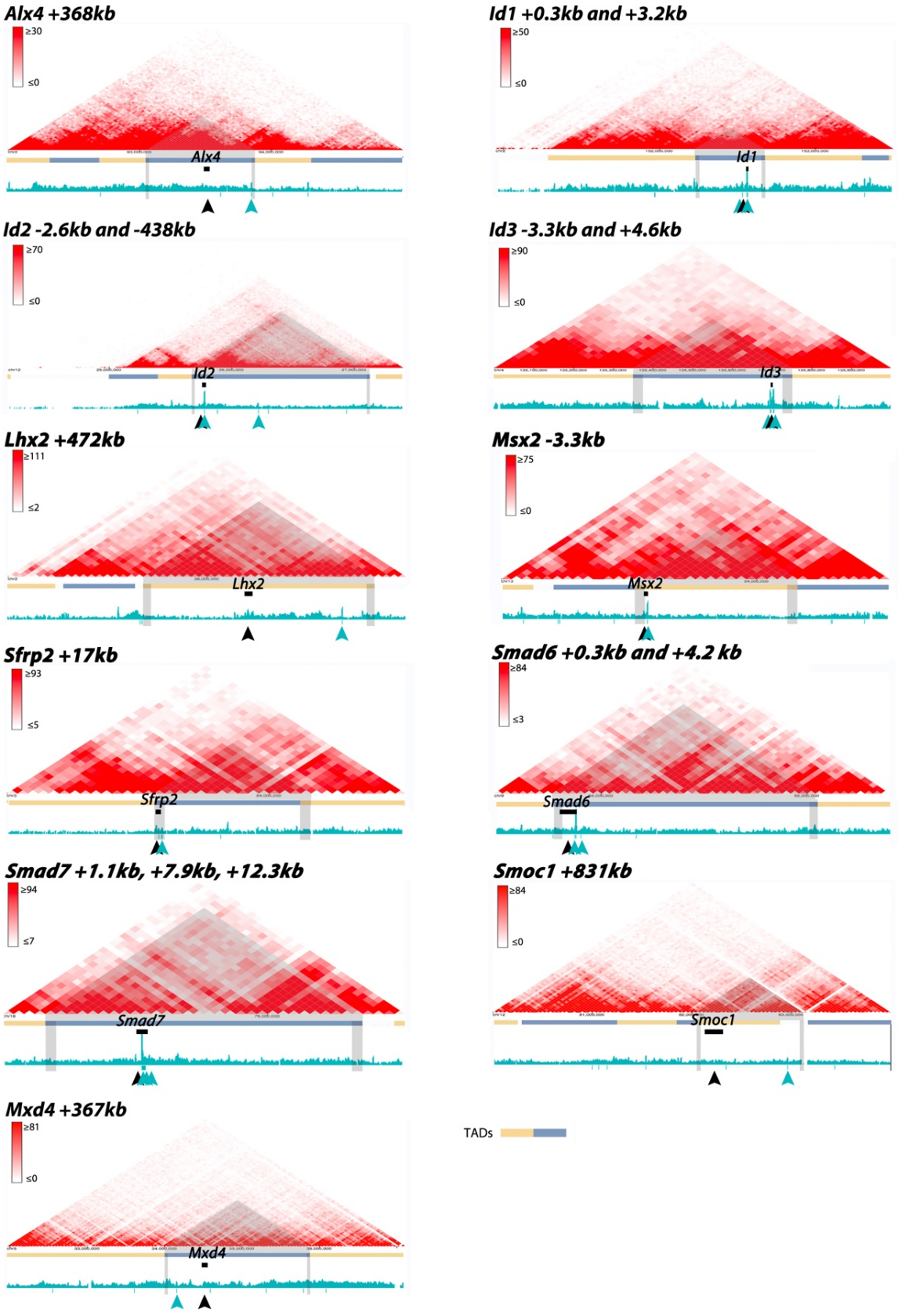
Expression pattern of 90 potential SMAD4 direct targets tested by WISH in E10.0 WT embryos.

12.6. Sterol intermediates and cholesterol quantification by GC/MS in WT and *Smad4*-deficient forelimb buds



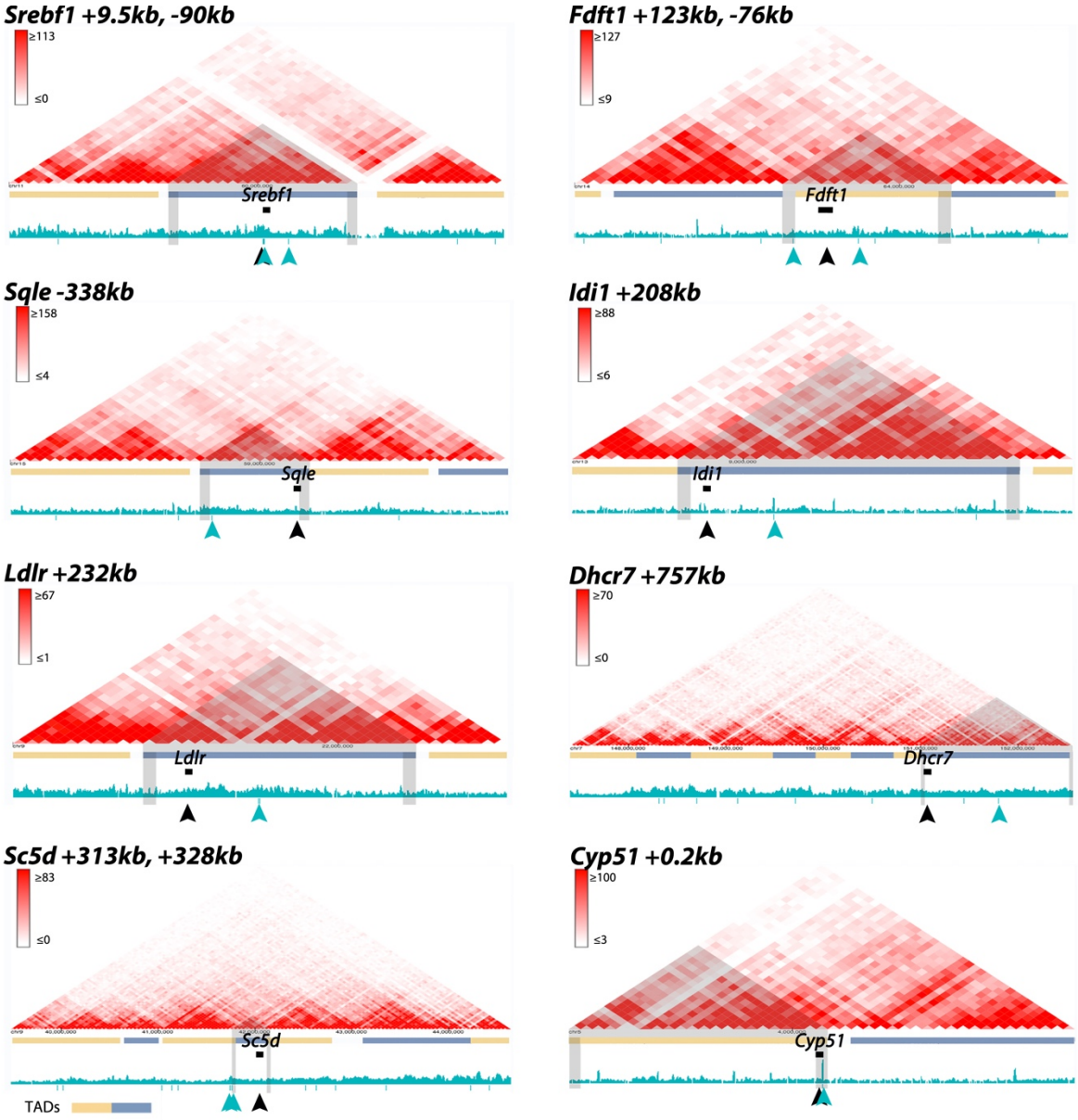
(A) Quantification of Lanosterol, Lathosterol, Demosterol, 7-dehydro-cholesterol (7DHC) and cholesterol by GC/MS in forelimb buds at E10.0, E10.5 and E11.0. (B) Simplified cholesterol synthesis pathway with the sterol intermediates quantified by GC/MS.

12.7. Analysis of SMAD4-interacting regions located in TADs of target genes involved in limb patterningx



The E9.75 SMAD4^{3xF} ChIP-seq track (in blue) is aligned with the Hi-C track from ES cells in red (Dixon et al., 2012). TADs are represented by the yellow and blue lines. The TAD of the SMAD4 target genes is highlighted in grey. The black arrow shows the SMAD4 target gene and the blue arrows show the SMAD4 bound regions within the TAD. *Pth1r, Snai1, Pkdcc* and *Tbx2* genes are not associated with a TAD in ES cells.

12.8. Analysis of SMAD4-interacting regions located in TADs of target genes involved in the cholesterol synthesis



The SMAD4^{3xF} ChIP-seq track (in blue) is aligned with the Hi-C track from ES cells in red (Dixon et al., 2012). TADs are represented by the yellow and blue lines. The TAD of the SMAD4 target genes is highlighted in grey. The black arrow shows the SMAD4 target gene and the blue arrows show the SMAD4 bound regions within the TAD.

12.9. Curriculum Vitae

Julie Gamart

Date and place of birth: 30th December 1987, Saint-Martin-Boulogne, France
Nationality: French
Address: DBM, Mattenstrasse 28, 4058 Basel, Switzerland
Phone number: +41 61 695 30 40
Email: julie.gamart@unibas.ch

EDUCATION AND DIPLOMAS:

- **PhD student in Genetics**, Department of Biomedicine, University of Basel, Switzerland (2013-Present)
- **European Master of Genetics**, with Honors, Paris-Diderot University, Paris, France (2011-2013)
- **Pasteur Courses**: Mouse Genetics, with Honors, Pasteur Institute, Paris, France (January to February 2013)
- **Bachelor of Cell Biology and Physiology**, with Honors, University of Lille, France (2008-2011)
- **Baccalauréat**, French national secondary-school diploma giving right of entry to university, Major subjects: Mathematics and Biology, with Honors, Lille, France (2003-2006)

RESEARCH TRAINING:

- **PhD. in Developmental Genetics (since September 2013)**
  Supervisors: Prof. Dr Rolf Zeller and PD Dr Aimée Zuniga
  Department of Biomedicine, University of Basel, Switzerland
  "SMAD4: a Multifunctional Regulator in Limb Bud Initiation and Outgrowth"
- **Master Thesis (December 2012-July 2013)**
  Supervisor: Dr Vanessa Ribes
  Institut de Myologie Pitié-Salpêtrière, UMRS787, Paris, France
  "In Vitro Generation of Spinal Progenitor Subtypes with Distinct Transcriptional Activities of Pax3 & Pax7"

⇨ **Research Internship (April-September 2012)**

Supervisor: Dr. Prashiela Manga

Ronald O.Perelman Department of Dermatology, NYU, New York, USA

"Endoplasmic Reticulum Maturation of Tyrosinase: Roles of TRYP1 and OCA2 Proteins"

⇨ **Research Introduction Internship (April 2011)**

Supervisor: Dr. Marie-Christine Chartier-Harlin

Institut de Recherche sur le Cancer de Lille; UMR837, Lille, France

"Transfection Conditions Establishment for Analysis of Mutations Associated with Parkinson's Disease"

PUBLICATION:

⇨ Laurent, F., Girdziusaite, A., **Gamart, J.**, Barozzi, I., Osterwalder, M., Akiyama, J.A., Lincoln, J., Lopez-Rios, J., Visel, A., Zuniga, A., et al. (2017). HAND2 Target Gene Regulatory Networks Control Atrioventricular Canal and Cardiac Valve Development. Cell Rep. 19, 1602–1613.

CONFERENCES PARTICIPATION:

⇨ **1st September 2017:** 4th Tri-Regio Developmental and Stem Cell Biology Meeting, Basel, Switzerland - Selected speaker

⇨ **23th-26th July 2017:** 14th International Limb Development and Regeneration Conference, Edinburg, UK - Selected speaker

⇨ **27th-30th September 2016:** 12th Transcription and Chromatin Meeting, EMBL Heidelberg, Germany - Poster.

⇨ **30th September 2015:** 2nd Tri-Regio Developmental and Stem Cell Biology Meeting, Freiburg, Germany - Selected speaker

⇨ **16th-19th September 2015:** 28th Mouse Molecular Genetics Meeting, Hinxton, UK - Poster, Selected for the Regeneron Bursary

TECHNICAL SKILLS:

- ⇨ **Animal handling:** mouse care, husbandry and colony management, embryo collection, and transgenic mice generation by Aggregation Chimera.
- ⇨ **Molecular biology:** RNA and DNA extraction, PCR, qPCR, RT-qPCR, molecular cloning with plasmids, whole-mount *in situ* hybridization, immunoblotting, chromatin immuno-precipitation (ChIP-seq) with library construction for sequencing, ATAC-seq, RNA-seq.
- ⇨ **Cell culture:** HEK293, melanocytes, mouse embryonic stem cells, limb mesenchymal progenitors culture. Gene targeting by cassette exchange, transfection, luciferase reporter assays, establishment of neuronal differentiation protocols and XGal assay.
- ⇨ **Histology:** Tissue preparation and sectioning with microtome cryostat, and immunohistochemistry
- ⇨ **Microscopy:** light, fluorescence and confocal microscopy, immuno-staining and fluorescence intensities quantification (FIJI)

IT (INFORMATION TECHNOLOGY) SKILLS:

- ⇨ Microsoft Office Suite: Word, Excel, and PowerPoint
- ⇨ Graphic Packages: Adobe Photoshop, Adobe Illustrator
- ⇨ ChIP-Seq data analysis: Galaxy, Nebula, IGV and UCSC Browser
- ⇨ Sequence conservation analysis: ECR and Vista Browsers

TEACHING EXPERIENCE:

⇨ 2015 and 2016: Supervision of two 1st year PhD students and a Master student regarding molecular biology (Cloning, PCR, RT-qPCR, RNA-seq)

⇨ 2016: Assistant in Developmental Biology Practical for medicine students, University of Basel

LANGUAGES

- ⇨ French: mother tongue
- ⇨ English: fluent (spoken and written)