The *Gremlin1 cis*-Regulatory Landscape: A Paradigm to Study Enhancer Cooperation in Regulation of Transcription Dynamics

Inauguraldissertation

Zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch Naturwissenschaftlichen Fakultät

der Universität Basel

Von

Laurène Ramos Martins

aus Frankreich

2020

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von

Prof. Dr. R. Zeller (Fakultätsverantwortlicher), PD Dr. A. Zuniga (Dissertationsleiterin), Prof. Dr. M. Affolter (Korreferent).

Basel, den 21.04.2020

Dekan Prof. Dr. Martin Spiess

1. Table of Contents

2. Summary

The transcriptional regulation of developmental genes expression patterns in time, space and levels, is governed by *cis*-regulatory modules (CRMs). The activity of CRMs is controlled by transcription factor complexes that act as downstream mediators of signaling inputs. CRMs are associated with their target genes in chromatin domains with enhanced contact frequency, the so-called topologically associating domains (TADs). The incoming signaling cues are integrated into specific transcriptional outputs, which orchestrate development and differentiation. Limb bud development is one of the main molecular and cellular paradigms to study the roles of gene expression regulation during embryonic development. Limbs are external organs, easily accessible, largely dispensable for embryonic and postnatal survival and have adapted to numerous specific functions during vertebrate evolution, resulting in the high level of morphological diversity among vertebrates.

The molecular pathways and morphogenetic events that govern limb patterning are largely conserved, reflecting their crucial roles in gene regulation during limb development. Our group previously identified and functionally analyzed the SHH/GREM1/AER-FGF epithelial-mesenchymal (e-m) self-regulatory signaling system that controls early limb bud outgrowth and patterning. The BMP antagonist *Gremlin1* (*Grem1*) is one of the functionally most essential nodes in this system. Its spatio-temporal expression is regulated by the converging *trans*-acting inputs of the major limb bud signaling pathways. These inputs are integrated into the dynamic regulation of *Grem1* expression by its 310 kb *cis*-regulatory landscape.

For my Ph.D. research, I used the mouse *Grem1 cis*-regulatory landscape as a paradigm to study gene transcriptional regulation in the context of embryonic limb bud development. I identified and genetically analyzed the functionally relevant *Grem1*-associated CRMs. To this end, I initially used reporter assays in transgenic mouse embryos to assess their potential enhancer activity. CRMs with established enhancer activities were then functionally studied by generating CRISPR/*Cas9*-engineered loss of enhancer function mutant mice. This, in combination with molecular analysis, was

used to assess their role(s) in the *Grem1* transcriptional regulation. In addition, I used 4C-seq assays to study the physical interactions among CRMs and the *Grem1* promoter, in wild-type and mutant mouse limb buds. I also addressed the question of the downstream consequences of enhancer deletions on limb bud development by tracking apoptosis and quantifying limb buds' cellular proliferation.

My studies revealed that the enhancer redundancy and diversity that regulates the *Grem1* expression dynamics during mouse limb bud development was much more complex than the one-to-one correlation often described by others. None of the CRMs characterized was essential on its own for limb development. The transcriptional activities of different CRMs were additive in levels and partially redundant in regulating the spatial and temporal dynamics of the *Grem1* expression. The spatio-temporal changes in *Grem1* expression levels, caused by the loss of different enhancers alone, were not sufficient to explain the observed phenotypes. Therefore, additional mouse strains lacking several CRMs were generated and analyzed. In light of these results, I performed a comparative molecular analysis of key genes in the self-regulatory SHH/GREM1/AER-FGF signaling system, which provided a better molecular understanding of how these *cis*-regulatory alterations affect the limb bud outgrowth and patterning. This analysis showed that the *cis*-regulatory alterations affecting levels and spatio-temporal kinetics of the *Grem1* expression are accompanied by specific changes in the self-regulatory feedback loops in mutant limb buds. In addition, I investigated potential effects on the structure of the *Grem1* TAD and revealed that alterations in the interactions among CRMs and the *Grem1* promoter contributed to the transcriptional regulation of *Grem1* expression.

This extensive genetic analysis led to the following major conclusion: the control of transcript levels by the *Grem1*-associated CRMs is additive, while they function in a cooperative manner to regulate the spatial dynamics of the *Grem1* expression in mouse limb buds. In particular, deleting several of the CRMs that regulate spatial aspects of the *Grem1* expression disrupts this cooperativity. This, in turn, weakens the robustness of the limb patterning system and results in the loss of pentadactyly. It appears that the observed limb skeletal deformity phenotypes strongly correlate with

reduced cell proliferation. Structural analyses reveal that intra-TAD rearrangements play a major role in the robustness of the *Grem1* expression.

In summary, my research provides novel insights into the transcriptional regulation of the *Grem1* gene expression, both at the *cis*-regulatory and at the signaling systems level.

3. List of Abbreviations

4C	Circular Chromatin Conformation Capture	**hr**	hour
AER	Apical Ectodermal Ridge	**HR**	Homologous recombination
AP	Antero-posterior	***Id***	*Inhibitor of Differentiation*
ArhGap11a	*Rho GTPase Activating Protein 11A*	***Ihh***	*Indian hedgehog*
ATAC	Assay for transposase-accessible chromatin	**kb**	kilobase
BA	Branchial arches	**KO**	Knock out
BAC	Bacterial artificial chromosome	***LacZ***	*Beta galactosidase coding gene*
BMP	Bone morphogenetic protein	**LB**	Lucia Broth
bp	base pair	**LCR**	Locus control region
BSA	Bovine serum albumin	**ld**	Limb deformity
C-L	Cenani-Lenz	**LIF**	Leukemia Inhibitory Factor
Cas	CRISPR-associated protein	**LMP**	Limb mesenchymal progenitor
cDNA	Complementary DNA	**LOF**	Loss of function
ChIP	Chromatin immunoprecipitation	**M-phase**	Mitotic phase
Chr	chromosome	**Mb**	Megabase
CLM	Congenital limb malformations	**MEF**	Mouse embryonic fibroblast
Cq	Quantification cycle	**min**	minute
CRC	colorectal carcinoma	**n**	number
CRM	*cis*-regulatory module	**NRE**	Negative regulatory element
CTCF	CCCTC-binding Factor	**O/N**	overnight
DIG	Digoxygenin	**ORF**	Open reading frame
DMEM	Dulbecco's modified Eagle medium	**PAM**	Protospacer adjacent motif
DNA	Deoxyribonucleic acid	**PBS**	Phosphate Buffered Saline
DV	Dorso-ventral	**PBT**	PBS-0.1% Tween 20
E	Embryonic day	**PCR**	Polymerase chain reaction
e-m	epithelial-mesenchymal	**PD**	Proximo-distal
EC	Enhancer cluster	**PFA**	Paraformaldehyde
ECh	Enhancer chain	**PIC**	Pre-initiation complex
EDTA	Ethylenediaminetetraacetic acid	**polyA**	Polyadenylation
Eff.	Efficiency	**ref.**	Reference(s)

EGTA	ethylene glycol-bis (β-amino ethyl ether)-N,N,N',N'-tetra acetic acid
EMT	Epithelial-to-mesenchymal transition
ERCE	Early replication control elements
ESC	Embryonic stem cell
EUCOMM	European Conditional Mouse Mutagenesis Program
FBS	Fetal bovine serum
FGF	Fibroblast growth factor
FL	forelimb
Fmn1	*Formin1*
G1-phase	Gap phase 1
GBR	GLI binding region
GCR	Global control region
Gli	*GLI-Kruppel family member*
GOF	Gain of function
GRE1	Gli responsive element 1
Grem1	*Gremlin 1*
GRN	Gene regulatory network
GRS1	Gremlin1 regulatory sequence 1
GTF	General transcription factor
H3K27ac	Histone 3 lysine 27 acetylation
H3K4me1	Histone 3 lysine 4 monomethylation
H3K4me3	Histone 3 lysine 4 trimethylation
Hi-C	High-throughput Chromosome Conformation Capture (*Hi-C*)
HL	Hindlimb
HMBS	Hydromethylbilane synthase
HMCO	Human mouse chicken oppossum
HMPS	Hereditary mixed polyposis syndrome
RNA	Ribonucleic acid
Rpl19	*Ribosomal Protein L19*
rpm	Revolutions per minute
RT	Room Temperature
RT-qPCR	Real-time quantitative PCR
S-phase	Synthesis phase
Scg5	*Secretogranin V*
SDS	Sodium Dodecyl Sulfate
SE	Super enhancer
SEM	Standard error of the mean
sgRNA	Single guide RNA
Shh	*Sonic hedgehog*
Smad	*Small Mothers Against Decapentaplegic*
Smoc1	*Secreted modular calcium-binding protein 1*
Sox9	*Sex determining region Y box 9*
TAD	Topologically associated domain
Tbx2	*T-box transcription factor 2*
TF	Transcription factor
Tmco5	*Transmembrane and Coiled-Coil Domains 5A*
TSS	Transcription Start Site
UCSC	University of California Santa Cruz
UTR	untranslated region
WISH	Whole-mount *in situ* Hybridization
Wnt	*Wingless-type MMTV integration site*
WT	Wild-type
Xgal	5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside
ZPA	Zone of polarizing activity
ZRS	ZPA regulatory sequence

4. Introduction

4.1. Promoter and *cis*-regulatory modules (*CRM*s) communication

4.1.1. CRMs are the cornerstones for the integration of signaling inputs to generate a transcriptional output

4.1.1.1. Basic concepts: what are CRMs?

During embryonic development, a unicellular zygote will follow genetic instructions and go through a series of cellular divisions to give rise to the various cell-types that comprise functional tissues and organs. While every single cell of a multicellular organism possesses the exact same genome, they regulate gene expression in disparate and cell type-specific spatio-temporal programs that control development and cellular differentiation. Such diversity in gene expression is supported by essential short non-coding DNA elements that activate, refine or prevent gene transcription (Alexander et al., 2019; Bartman et al., 2016; Benabdallah et al., 2019; Bolt and Duboule, 2020; Bulger and Groudine, 1999; Furlong and Levine, 2018; Gaszner and Felsenfeld, 2006; Gerstein et al., 2010; Levine, 2010; Ong and Corces, 2014; Petrykowska et al., 2008; Rusche et al., 2003; Sandelin et al., 2004; Schoenfelder and Fraser, 2019; Stathopoulos and Levine, 2005). Although transcription initiates at promoters, where the basal transcriptional machinery is recruited, promoters have little to no part in transcriptional regulation (Banerji et al., 1981). Promoters can be considered as a specific type of CRMs. They rely on other categories of CRMs that are in charge of the transcriptional control of gene expression in time, space and levels during development. CRMs are key regulators of cell fate decisions. Identifying and understanding CRM functions is essential to expand our knowledge of the complexity

in transcriptional coordination of developmental genes. CRMs – excluding promoters – belong to three categories: transcriptional enhancers, transcriptional silencers and structural insulator elements.

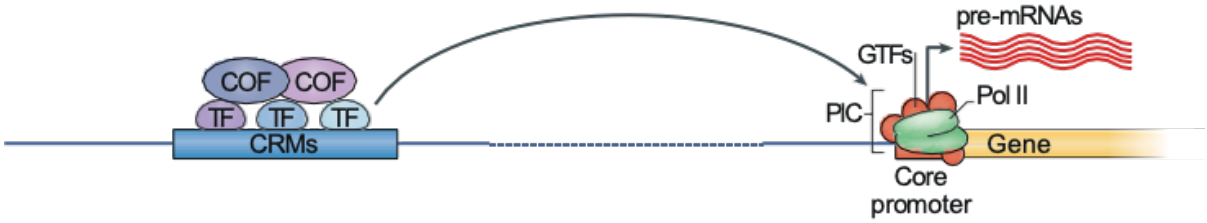


Figure 1. Schematic of a *cis*-regulatory module (CRM)-Promoter synergy in gene transcriptional regulation. RNA polymerase II (Pol II) and general transcription factors (GTFs) form a complex, the pre-initiation complex (PIC), that is recruited at the core promoter and under the control of CRMs that locally bind transcription factors (TFs) and convey regulatory cues to the PIC, modulating gene expression in precise spatio-temporal pattern and levels. *Adapted from Haberle et al. (2018).*

4.1.1.2. Basic concepts: how to identify CRMs

Functional CRMs correspond to open chromatin regions that are bound by transcription factor (TF) complexes (Felsenfeld, 1996; Krebs et al., 2017; Thurman et al., 2012). TFs are the nuclear mediators relaying cell-intrinsic and cell-extrinsic signaling cues to modulate gene transcription (Hardison and Taylor, 2012; Long et al., 2016). The first indication to identify putative CRMs is to look for open chromatin loci. Open chromatin regions are sensitive to DNases (Gross and Garrard, 1988) and transposases (Buenrostro et al., 2013). DNase hypersensitivity footprints (Consortium et al., 2007; Lu et al., 2011) or assay for transposase-accessible chromatin followed by deep sequencing (ATAC-seq, Buenrostro et al., 2013) have been commonly used to detect open and accessible chromatin regions. Although promoters and structural elements also overlap with chromatin accessible regions (Boyle et al., 2008; Felsenfeld and Groudine, 2003; Gross and Garrard, 1988; Stalder et al., 1980), they are not involved in the tissue-specific transcriptional control of developmental genes and can be removed from the analyses. To refine the selection, chromatin immune precipitation (ChIP) followed by deep sequencing is commonly performed. ChIP-seq

experiments performed *in vivo*, targeting the TFs regulating the transcription of gene(s) of interest, provide other criteria to identify CRMs. Those experiments have to be performed *in vivo* because *in vitro* assays will identify TF binding motifs and therefore millions of DNA loci while *in vivo* only about 1 in 500 motifs interact with the corresponding TFs (Zang et al., 2009; Zhang et al., 2008). In addition to binding TFs, CRMs can recruit histone-modifying enzymes that decorate histones with different marks. Detectable with ChIP-seq, those marks form a "code" that helps to predict the chromatin state. Active CRMs are decorated with monoacetylation on histone 3 lysine 27 (H3K27ac) and monomethylation of histone 3 on lysine 4 (H3K4me1, Creyghton et al., 2010; Rada-Iglesias et al., 2011; Shlyueva et al., 2014; Zentner et al., 2011, Figure 2). Poised enhancers overlap with the H3K4me1 mark only (Bonn et al., 2012, Figure 2). Promoters are recognized by the presence of H3K27ac and the trimethylation of histone 3 on lysine 4 (H3K4me3, Heintzman et al., 2007, Figure 2). The trimethylation of histone 3 on lysine 27 (H3K27me3) is characteristic of an inactive or repressed state of the chromatin (Bernstein et al., 2006; Schuettengruber et al., 2007). Chromatin epigenetic profiling is a reliable method to identify different classes of putative CRMs. ChIP-seq targeting known insulator proteins such as the CCCTC-binding factor (CTCF, Dekker and Mirny, 2016; Nichols and Corces, 2018) are used to separate them from the putative CRMs that actively regulate transcription in space, time and levels (Phillips-Cremins et al., 2013; Rao et al., 2014; Sanyal et al., 2012, Figure 2). It is important to perform these experiments in the organ/tissue of interest and at the developmental stage and time where the genes of interest are expressed as CRMs activity depends on their cell-specific genomic environment. An active CMR in a specific organ may be inactive and undetectable in another organ and different environmental conditions.

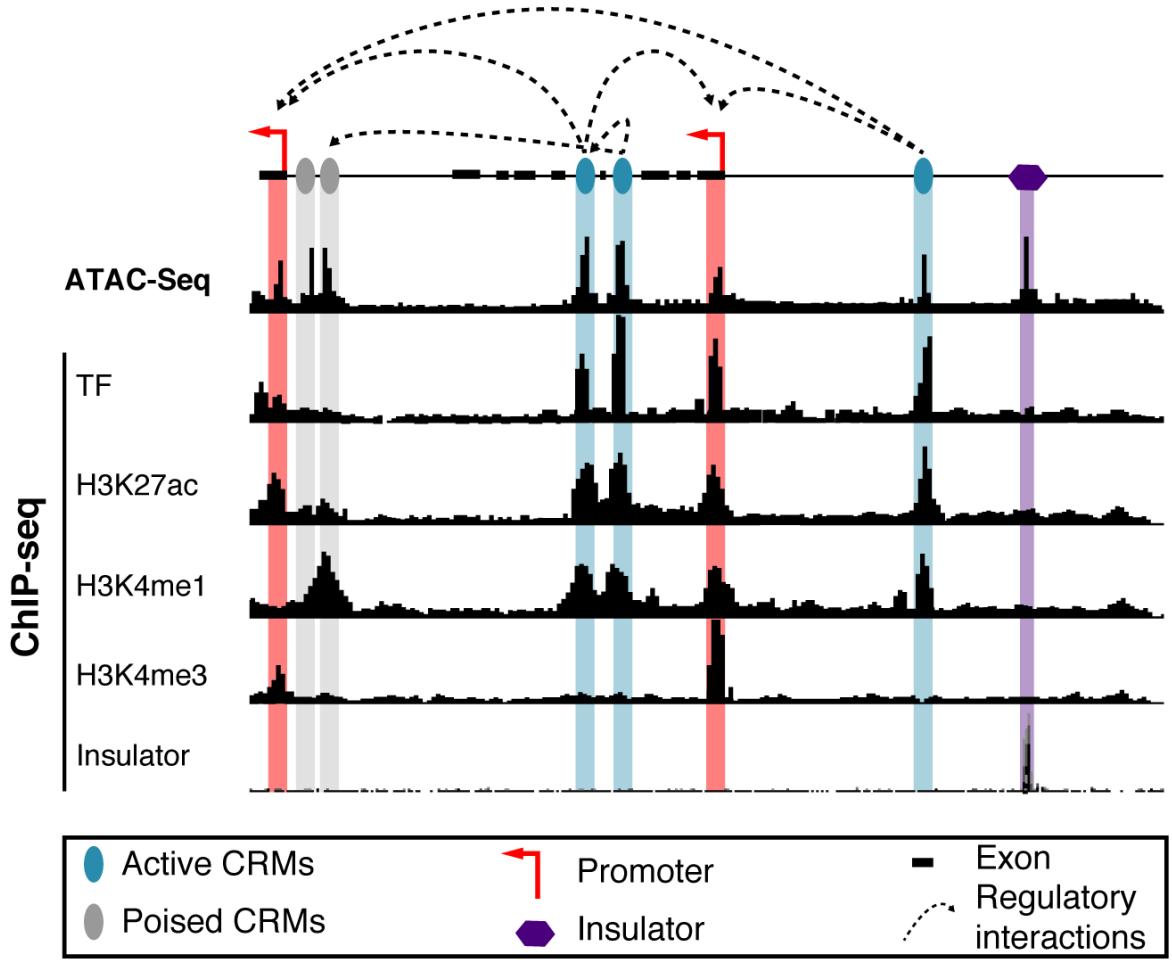


Figure 2. Schematic representation of CRM identification strategy. Putative CRMs (blue ovals) map with open chromatin regions uncovered by ATAC-seq assays. ChIP-seq experiments show that they bind at least one transcription factor (TF) and are enriched in monoacetylated histone 3 on its lysine 27 for the active ones (H3K27ac, vertical blue bars) and in monomethylated histone 3 on lysine 4 for the poised ones (H3K4me1, vertical grey bars). The trimethylation of histone 3 on lysine 4 (H3K4me3) allows the discrimination between promoters (red ovals, vertical red bars) and CRMs. A ChIP-seq experiment targeting a structural insulator (in purple) identifies an open chromatin region that is not a CRM. Dashed arrows show the possible regulatory connections between CRMs and their associated gene(s).

Another remarkable property of developmental CRMs is their high level of sequence and/or ancestral function conservation throughout evolution (Bhatia et al., 2016; Boffelli et al., 2004; Harmston et al., 2013; Long et al., 2016; Navratilova et al., 2009; Rada-Iglesias et al., 2011; Shen et al., 2012), even though sequence conservation does not always correlate with CRM functions (Nelson and Wardle, 2013). CRMs with silencing functions have not been studied systematically (Petrykowska et al., 2008)

but even poised enhancers (Banerji et al., 1981) have been able to drive a cell-type and stage-specific reporter expression despite poor sequence conservation from one species to another one (Rada-Iglesias et al., 2011; Ritter et al., 2010). Transient or transgenic reporter assays represent a quick and robust strategy to functionally characterize enhancers as their activity is known to overlap with that of their associated gene (Kvon, 2015; Levine, 2010; Stanojevic et al., 1991; Visel et al., 2009b).

After identifying putative CRMs, one has to be able to associate them with their corresponding gene(s). The linear distance (nucleotide count) between a promoter and its enhancers is not a criterion as many studies have reported that CRMs can be found in inter- or intragenic regions, in the vicinity of their associated promoter or at great linear distances, even up to 3 Mb away from their associated gene (Krivega and Dean, 2012; Lettice et al., 2003; Zeller and Zuniga, 2007; Zuniga et al., 2012a; Zuniga et al., 2004). Vakoc and colleagues have demonstrated that promoter and enhancer overcome linear distances as they are brought in close proximity through chromatin looping (Vakoc et al., 2005). Enhancers can be associated with genes far away in linear distances forming so-called “long-range” interactions (Lettice et al., 2003; Montavon et al., 2011; Sagai et al., 2005). Impressive technological progress in the molecular methods based on proximity ligation (Chromatin Conformation Capture-(3C-) based technology) have been made this past ten years (Dekker et al., 2002).

4.1.2. Developmental genes are organized into Topologically Associated Domains (TADs)

In 2012, the establishment of DNA-DNA interaction profiles with High-throughput Chromosome Conformation Capture (Hi-C) has uncovered the existence of chromatin domains with a high frequency of internal physical interaction: the Topologically Associating Domains (TADs, Dixon et al., 2012; Nora et al., 2012; Sexton et al., 2012). TADs are chromatin domains found in eukaryotes that exhibit increased probability of internal physical interactions (Dixon et al., 2012; Nora et al., 2012; Rao et al., 2014;

Sexton et al., 2012; Stam et al., 2019) and are delimited by boundary loci (i.e. insulators) occupied, in most cases, by CTCF (Dekker and Mirny, 2016; Nichols and Corces, 2018). CRM activities seem to be directed toward the gene(s) present within the same TAD, with limited contacts with genes or CRMs outside of these domains (Dixon et al., 2016; Dixon et al., 2012, Figure 3). According to the theory supporting the contact model (Fulco et al., 2019), physical contacts between CRM-TF complexes and the promoter-pre-initiation complexes (PICs) are a prerequisite to CRMs functions. HiC-seq with high-resolution (Barutcu et al., 2018) but also Circular Chromosome Conformation Capture followed by high throughput sequencing (4C-seq, Simonis et al., 2006) are appropriate methods to identify putative CRMs within *cis*-regulatory landscapes when compared with chromatin profiling. 4C-seq is an assay enabling the identification of all contacts with a promoter when it is set as the viewpoint. Nevertheless, 4C-seq, unlike HiC-seq, lacks the ability to identify TAD boundaries. However, knowing that CTCF-binding sites in converging orientation often surround chromatin regions with high contact frequencies (Dixon et al., 2012; Rao et al., 2014), while the boundaries between two adjacent TADs are characterized by divergent CTCF-motifs (Gomez-Marin et al., 2015), one can approximate *cis*-regulatory landscapes thanks to publicly available datasets (Figure 3).

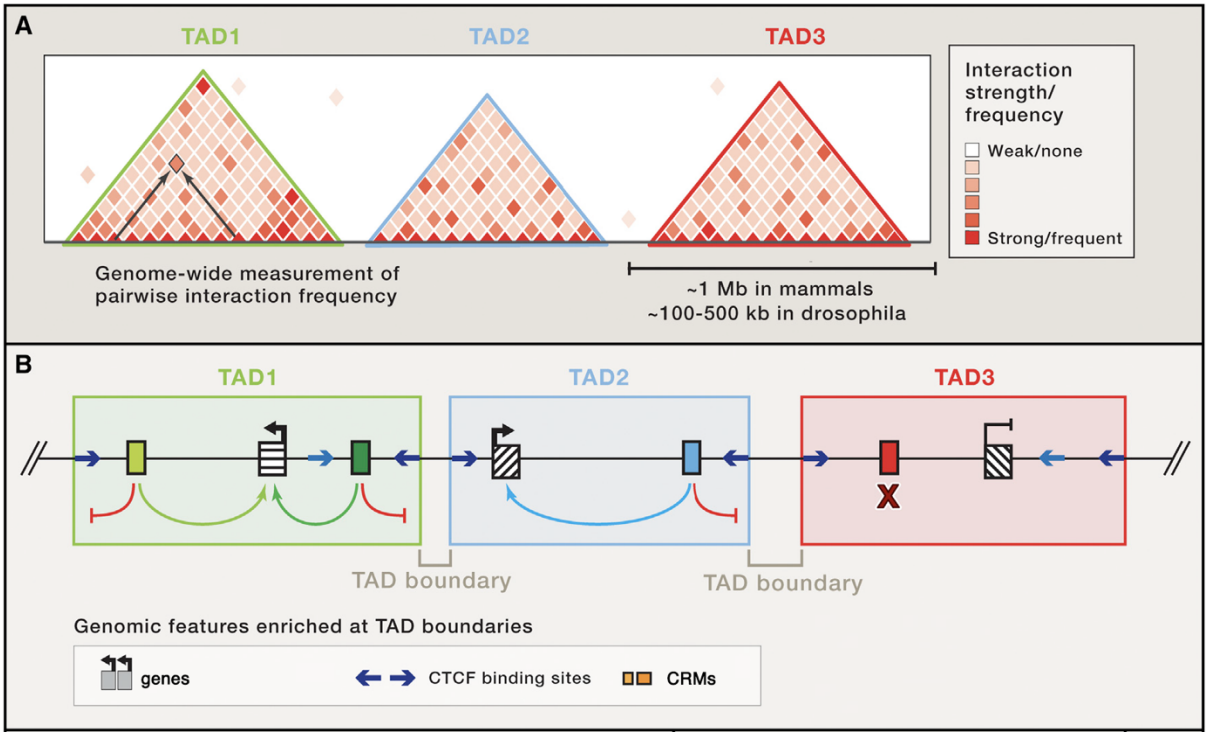


Figure 3. Schematic representation of the organization of chromatin into Topologically Associated Domains (TADs). (A) Visualization of chromatin compartmentalization into TADs with Hi-C heatmaps, interaction frequencies are represented in triangular blocks of increased color intensity. (B) CRMs influence is limited within TAD boundaries that also restrict the spread of chromatin modifications. The TAD boundary loci are associated with CTCF binding sites. TAD1 and TAD2 include actively transcribed genes while TAD3 is transcriptionally inactive. *Adapted from Long et al. (2016a).*

There is no established causal effect between TAD structures and enhancer-promoter interactions (Mir et al., 2019). When TAD formation is disrupted by depletion of its boundary-establishing protein, CTCF or cohesin, only a limited subset of genes activity is affected (Haarhuis et al., 2017; Nora et al., 2017; Rao et al., 2017; Schwarzer et al., 2017; Sofueva et al., 2013; Wutz et al., 2017). Although genomic rearrangements that suppress or shuffle TAD boundaries have more impressive phenotypes and can lead to "TADopathies" in mice and humans (Harmston et al., 2017; Lupianez et al., 2015; Matharu and Ahituv, 2015; Rao et al., 2014; Symmons et al., 2014), the relative importance of TADs for gene expression regulation remains ambiguous (Mir et al., 2019). Indeed, some transcribed loci do not present a TAD organization and multiple studies have uncovered mild consequences of TAD disruption (Despang et al., 2019; Ghavi-Helm et al., 2019; Kragesteen et al., 2018; Rodriguez-Carballo et al., 2019; Williamson et al., 2019, Figure 4). Besides, TADs rearrangements can be highly dynamic, resulting in great cell to cell heterogeneity (Cattoni et al., 2017; Finn et al., 2019; Flyamer et al., 2017; Nagano et al., 2013; Stevens et al., 2017).

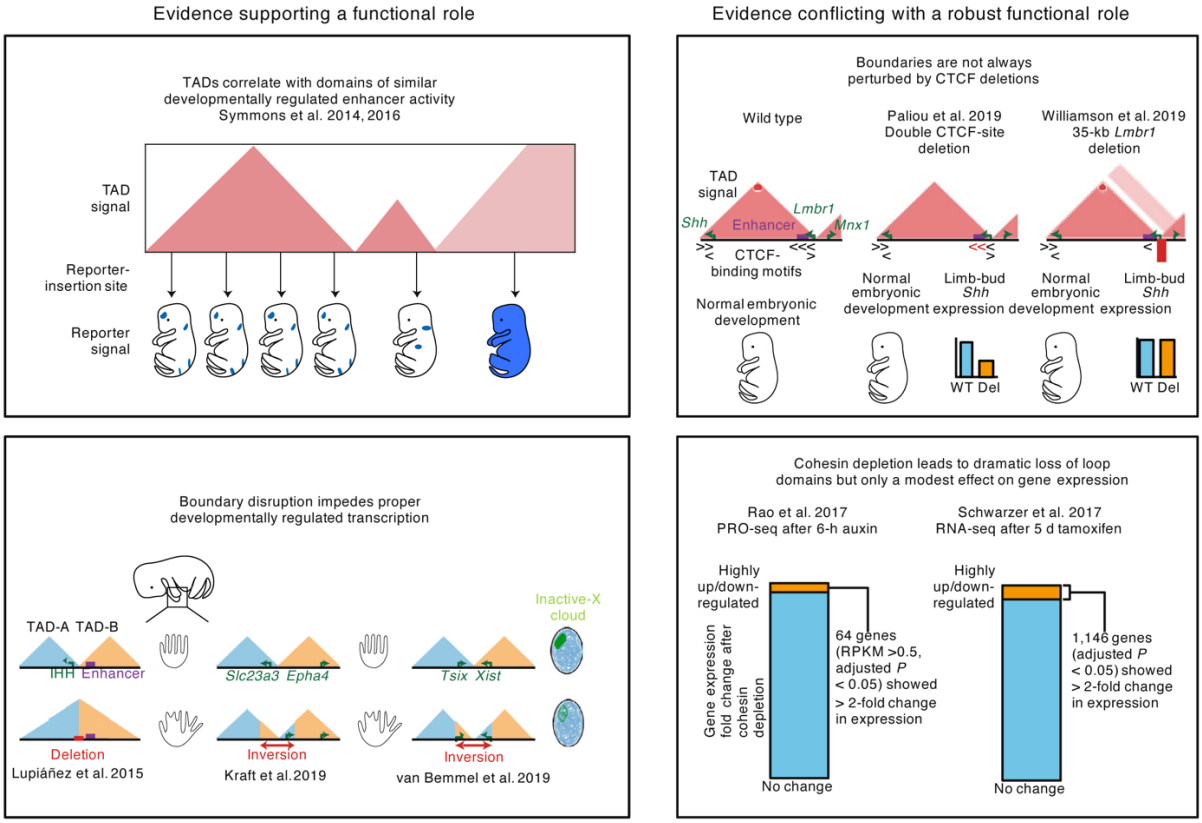


Figure 4. The functional relevance of TAD in gene transcriptional regulation is open for debate. Left panel on top: schematic of a reporter assay exposing the activity of randomly integrated CRMs in genomic loci (Symmons et al., 2016; Symmons et al., 2014). Left panel on the bottom: some published cases of boundary disruption leading to developmental issues. Right panel on top: schematic of uneventful TAD CTCF boundaries perturbation for the *Shh* gene during mouse embryonic development (Paliou et al., 2019; Williamson et al., 2019). Right panel on the bottom: schematic representation of the relatively mild transcriptional changes measured after cohesin depletion, rPKM, reads per kilobase per million mapped reads (Rao et al., 2017; Schwarzer et al., 2017). *Adapted from Beagan and Phillips-Cremins (2020).*

Nevertheless, the observation that CRMs appear to confine their activity toward the genes present in the same TAD and that TAD coordinates are often a good estimation of regulatory landscapes – a defined genomic region including all CRMs and their associated genes that are transcriptionally regulated in a coordinated fashion (Grosveld et al., 1987; Spitz et al., 2003) – is commonly observed for most studied developmental genes (Bolt and Duboule, 2020; Harmston et al., 2017; Lupianez et al., 2015; Rao et al., 2014; Symmons et al., 2014). CRMs within the same TAD and regulating the same genes have overlapping activities (Symmons et al., 2016; Symmons et al., 2014, Figure 3, Figure 4).

4.1.3. Overview of the current state of the knowledge about CRMs function in gene transcriptional regulation

4.1.3.1. Insulator functions

Insulators, or boundary elements, inhibit inappropriate contacts between adjacent TADs (Gaszner and Felsenfeld, 2006; Petrykowska et al., 2008; Valenzuela and Kamakaka, 2006; Wallace and Felsenfeld, 2007). The most studied protein associated with insulator regions so far are the CTCFs (Bell et al., 1999; Hark et al., 2000). CTCF proteins together with upstream stimulatory factors recruit histone-modifying enzymes and prevent the spread of repressive heterochromatin state (Huang et al., 2007). They also limit irrelevant action of enhancers ("enhancer blockers" role) or silencers ("barriers" role, Valenzuela et al., 2008; Wallace and Felsenfeld, 2007). Additionally, CTCFs associate with cohesins at boundary regions (Wendt et al., 2008) and shape chromatin into TADs. CTCF/cohesin dynamics are the foundation for polymer simulations and genomic analyses proposing the loop extrusion model as the mechanism allowing TADs formation (Fudenberg et al., 2016). CTCFs protect cohesins from chromatin release factors and stop their linear diffusion. Cohesin complexes extrude chromatin forming loops. Cohesins can get trapped at CTCF binding sites. A steady loop domain, a TAD, happens when flanked by CTCFs, cohesin proteins cannot extrude or release chromatin anymore, creating stable TAD boundaries (Li et al., 2020, Figure 5). Besides, structural complexes such as CTCFs/cohesin or Mediator/cohesin (Berlivet et al., 2013; Kagey et al., 2010) help to connect regulatory sequences to their associated promoters (Majumder and Boss, 2010; Ong and Corces, 2014; Sanyal et al., 2012) ensuing intra-TAD looping. Nevertheless, so far, they have not been described as functional modulators of transcription levels or dynamics.

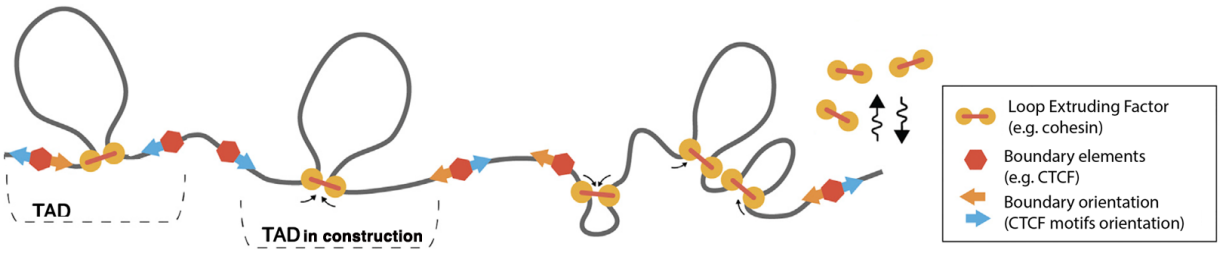


Figure 5. Schematic of the interphase loop extrusion model. Dashed lines represent a TAD region, bordered by converging CTCF binding sites, established because of a loop extrusion mechanism. *Adapted from Fudenberg et al. (2016).*

4.1.3.2. Enhancer and silencer functions

Enhancers (Banerji et al., 1981; Fromm and Berg, 1983; Gillies et al., 1983) and silencers (Rusche et al., 2003) are defined by their positive or negative effects on gene transcriptional regulation. Their activities are often tested with reporter assays as they can act independently of their orientation or even position in the genome. Enhancer and silencer activities depend on their *trans*-environment. One module can assume a positive or negative function depending on the TF complexes they recruit (Istrail and Davidson, 2005; Jing et al., 2008; Li et al., 2014; Martowicz et al., 2005).

Enhancers form the CRM category considered as a key player in the spatio-temporal control of gene expression during embryonic development (Bolt and Duboule, 2020; Long et al., 2016). The first described example of long-distance gene regulation by enhancers is the *ß-globin* locus (Carter et al., 2002; Tolhuis et al., 2002). A Locus Control Region (LCR) containing six DNaseI hypersensitive sites was identified more than 60 kb upstream of the *ß-globin* locus in humans (Grosveld et al., 1987) and in mice (Moon and Ley, 1990). The LCR is required for high-level transcription of the *ß-globin* genes (Bender et al., 2000; Bulger and Groudine, 1999; Epner et al., 1998; Reik et al., 1998; Schubeler et al., 2001). These studies proposed an "active chromatin hub" concept in which close physical proximity with an associated gene was a prerequisite for enhancer functions (Carter et al., 2002; Tolhuis et al., 2002). The stage- and tissue-specific expression of developmental genes can rely on the activity of multiple CRMs, but this is not true for all genes. During limb bud development, the *Sonic hedgehog* gene (*Shh*) transcription is entirely regulated by a single enhancer the Zone of

polarizing activity Regulatory Sequence (ZRS, Lettice et al., 2003). The ZRS is located within another gene (*Lmbr1*) 1 Mb upstream of *Shh*, both present at the opposite ends of the same TAD. The ZRS deletion phenocopies the *Shh* loss of function (LOF) limb phenotype, demonstrating the importance of this single enhancer for gene expression (Lettice et al., 2003; Sagai et al., 2005). Only few other studies have described the predominance of a single enhancer over several tissues/time-specific ones in the regulation of a developmental gene (loss of one enhancer function followed by partially penetrant phenotypes) but none were as striking as the removal of the ZRS (Gonen et al., 2017; Kragesteen et al., 2019; Kragesteen et al., 2018). These studies revolving around one single enhancer dominance seem to be the exception and not the rule as many more studies present instances where one single gene is under the control of multiple enhancers with various regulatory crosstalk.

Enhancers can cooperate in an additive and simultaneous manner (Osterwalder et al., 2018; Will et al., 2017, Figure 6a). In such contexts, the transcriptional levels of a gene are equal to the sum of each enhancer activity within the same *cis*-regulatory landscape. Thus, the number of transcripts and severity of phenotypes scales to the number of enhancers present in the *cis*-regulatory landscape. Synergistic interactions have also been documented. Two or more enhancers work in synergy whereby their combined action is stronger than the sum of their individual ones (Maekawa et al., 1989; Stine et al., 2011, Figure 6b). When enhancers working in an additive or cooperative fashion are genetically altered without major changes in gene expression, functional redundancy between enhancers has been proposed as an explanation (Ahituv et al., 2007; Dickel et al., 2018; Lam et al., 2015; Osterwalder et al., 2018). Functional redundancy has been extensively studied, in particular by Mark Levine and colleagues who introduced the concept of shadow enhancers (Hong et al., 2008). They described shadow enhancers as “remote secondary enhancers mapping far from the target gene and mediating activities overlapping with the primary enhancer”. Shadow enhancers have overlapping patterns of activity. They bind the same TFs as the primary enhancer but may assume different roles in the spatio-temporal fine-tuning of gene expression (Barolo, 2012). They may not be essential in normal context but when the “primary” enhancer is inactivated, they could compensate and ensure the robustness of gene expression (Barolo, 2012; Frankel et al., 2010). This mechanism

could be the consequence of competitive interactions between enhancers (Bothma et al., 2015; El-Sherif and Levine, 2016, Figure 6c). Enhancer associations with promoters can prevent the association of other enhancers to the same promoter. In competitive situations, one enhancer could only reach its full potential in the absence of its competitor(s). Hierarchical regulations between enhancers are a non-additive regulatory crosstalk between one essential enhancer activating gene expression on its own and other(s), under the control of the main enhancer, that can bring a specific trait to gene expression (e.g. tissue specificity, Iampietro et al., 2010; Leddin et al., 2011; Maeda and Karch, 2011; Mihaly et al., 2006, Figure 6d).

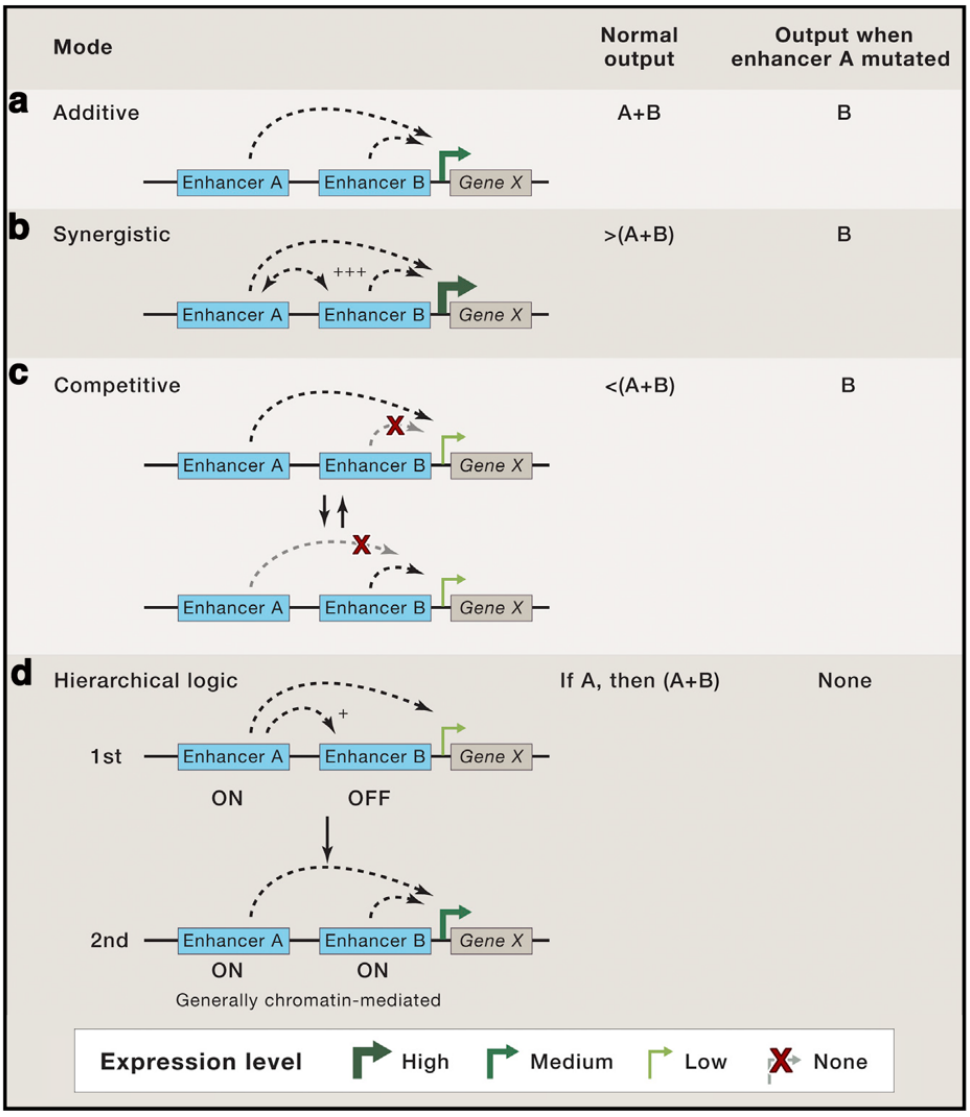


Figure 6. Schematic of known regulatory crosstalk between enhancers. (a-d) Possible enhancer interactions and resulting transcriptional output in wild-type or mutant contexts. *Adapted from Long et al. (2016a).*

When numerous active enhancers associated with the same gene(s) are found in restricted chromatin region (in linear nucleotide distance) they can be referred to super-enhancers (Hnisz et al., 2013; Whyte et al., 2013). When enhancers are spread across large linear distances in a gene desert, they can be identified as a regulatory archipelago (Bhatia et al., 2014; Montavon et al., 2011). During limb bud development, the *HoxD* locus is regulated by an archipelago of enhancers forming a global control region (GCR) whose activity depends on their TAD structure and intra-TAD interactions (Montavon et al., 2011; Spitz et al., 2003). Enhancers located at the 3′ of the landscape drive proximal limb expression while the ones located at the 5′ drive distal limb expression, reflecting the *Hox* gene collinearity.

Silencers are poorly understood modules repressing enhancer activity (Li and Arnosti, 2011; Li et al., 2014; Petrykowska et al., 2008; Stathopoulos and Levine, 2005; Vokes et al., 2008). Silencers have been described as essential to prevent ectopic gene expression (Dunipace et al., 2011; Marinic et al., 2013; Perry et al., 2011). As a result, they are essential for the fine-tuning of gene expression patterns.

Gene expression can be modulated both in quantity and in quality (spatio-temporal expression and tissue specificity) by multiple enhancers and silencers that cooperate through various possible mechanisms conferring robustness during gene expression *in vivo*.

4.2. *Gremlin1* (*Grem1*), a critical protagonist for limb bud development

4.2.1. Basic concepts: Limb bud development

The limb is an easily accessible organ whose intricate outgrowth and patterning mechanisms depend on the gene regulatory networks (GRNs) that control vertebrate organogenesis. Wide variability in morphology is observed in the vertebrate clade, emphasizing the importance of gene regulation during limb morphogenesis. Limb development is one of the foremost genetic and molecular models for developmental biologists. Limbs originate from the lateral plate mesoderm, through a localized

epithelial-to-mesenchymal transition (EMT), and growth along the three main developmental axes (Gros and Tabin, 2014; Petit et al., 2017; Zuniga, 2015, Figure 7). As a result, the limb ridge, surrounded by ectoderm arises at embryonic day 9.5 (E9.5). The limb ridge is already polarized along the antero-posterior (AP), the dorso-ventral (DV) and proximo-distal (PD) axis. The emerging limb grows in a burgeon-like shape, hence its "limb bud" appellation from E10 to E11.5 (Figure 7a), before digits condensations aggregate giving the limb a paddle-like shape at E12.0 (Drossopoulou et al., 2000; Harfe et al., 2004; Wanek et al., 1989, Figure 7b). It is during this developmental window, from the ridge- to the paddle-shaped limb, that the mesenchymal progenitors give rise to limb skeletal primordia through proliferation and differentiation, making this a critical stage for limb morphology.

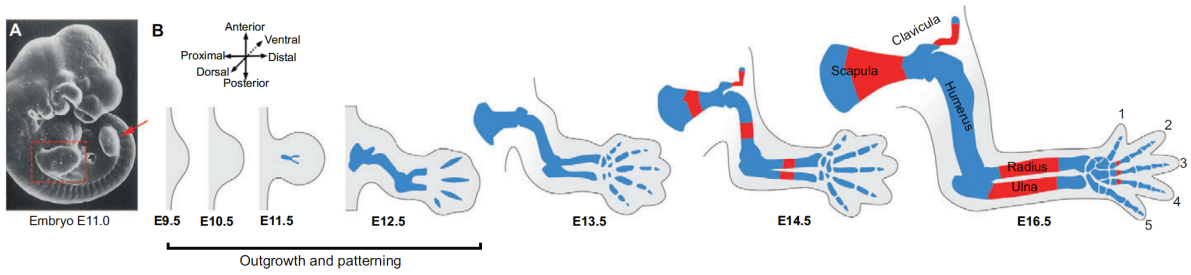


Figure 7. Schematic of limb bud development through embryogenesis. (A) Scanning electron microscopy image of a whole mouse embryo at E11.0. (B) Schematic of limb morphological evolution through embryonic development from embryonic day 9.5 to 16.5 (E9.5 to E16.5). *From Zuniga et al. (2015).*

At E9.75-E10.0, at the DV interface, the apical ectodermal ridge (AER) protrudes distally (Saunders, 1948, Figure 8). The AER represents one of the two main limb bud signaling centers. It notably produces several fibroblast growth factor (FGF) ligands (Fallon et al., 1994; Mariani et al., 2008; Niswander et al., 1993) and maintains the underlying mesenchymal progenitors in an undifferentiated state, allowing limb bud proliferative expansion and patterning (Fallon et al., 1994; Lewandoski et al., 2000; Mariani et al., 2008; Petit et al., 2017; ten Berge et al., 2008; Towers et al., 2008; Zhu et al., 2008). AER-FGFs are essential cell survival factors during early limb development (Sun et al., 2002). The morphogen *Shh* is expressed by the second main limb bud signaling center, the posterior zone of polarizing activity (ZPA, Benazet and

Zeller, 2009; Gasseling and Saunders, 1961; Zhu et al., 2008, Figure 8). In the nascent limb bud (E9.5-E9.75), the ZPA is initiated and spatially restricted as a result of pre-patterning mechanisms between the transcriptional regulators HAND2, which promotes *Shh* expression, and the GLI3R-mediated repression of *Hand2* expression (Galli et al., 2010; Osterwalder et al., 2014; Vokes et al., 2008). A *Shh* morphogen AP gradient participates in the PD axis organization through the establishment of the limb bud distal identities. Indeed, the SHH pathway has a positive effect on AER-*Fgfs* expression, leading to the clearance of retinoic acid from the distal limb bud by the AER-FGF signaling target *Cyp26b1* (Probst et al., 2011). Primarily, the SHH gradient is essential for the organization of the limb bud AP axis. The SHH pathway regulates cell survival and proliferation, ensuring an adequate number of cellular progenitors for digit formation (Zhu et al., 2008). In mice and humans, the AP axis is reflected by the digit identities (Ahn and Joyner, 2004; Drossopoulou et al., 2000; Harfe et al., 2004; Zhu et al., 2008). The most anterior digit, digit 1 (the thumb), is SHH-independent while the intermediate digits (2 to 4), and the most posterior digit, digit 5 (little finger), are dependent on the SHH gradient (Zhu et al., 2008). The digits are part of the most distal fragment of the limb, the autopod. The medial part of the limb is the zeugopod (forelimb: radius-ulna; hindlimb: tibia-fibula, Figure 7b) and the most proximal the stylopod (forelimb: humerus; hindlimb: femur, Figure 7b). AER and ZPA depend and interact with each other to coordinate limb bud outgrowth and patterning (Chiang et al., 2001; Laufer et al., 1994; Niswander et al., 1994; Tickle, 1981; Todt and Fallon, 1987). This coordination relies on key developmental genes.

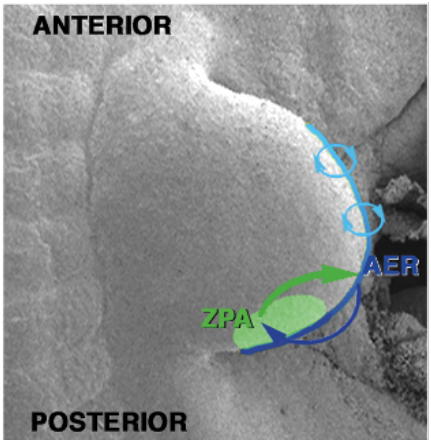


Figure 8. Schematic of the limb epithelial-mesenchymal (e-m) feedback loops. A self-regulatory feedback loop is established between the zone of polarizing activity

(ZPA, green) and the apical ectodermal ridge (AER, blue) superimposed on a mouse limb bud scanning electron micrograph. Arrows represent a simplified summary of the regulatory interactions between AER and ZPA. *From Zeller and Zuniga (2007).*

4.2.2. *Grem1* and the limb bud regulatory networks

The bone morphogenetic protein (BMP) antagonist GREM1 is a secreted protein, which directly binds to BMP ligands, blocking the activation of the BMP pathway (Figure 9a). Initially, the BMP pathway activity is required for the establishment of a functional AER, and to restrict its length (Ahn et al., 2001; Benazet and Zeller, 2013; Nissim et al., 2006; Niswander et al., 1994; Scherz et al., 2004; Zeller et al., 2009; Zuniga and Zeller, 1999, Figure 9b). High BMPs activity correlates with the transcriptional activation of *Grem1* in the limb posterior distal domain (Benazet et al., 2009; Zeller et al., 2009, Figure 9b,c).

Secreted by the *Grem1*-expressing cells in the distal posterior limb bud, GREM1 proteins prevent BMPs from binding to their cognate receptors, prompting a significant reduction of overall BMP activities (Benazet et al., 2009; Khokha et al., 2003; Michos et al., 2004; Zuniga and Zeller, 1999). In turn, low BMP activity stimulates the AER to increase FGFs production and sustains the *Shh* expression emanating from the ZPA. The BMP/GREM1 regulatory loop is called the initiator module of the limb SHH/GREM1/AER-FGF self-regulatory feedback loop (Benazet and Zeller, 2009; Zeller et al., 2009, Figure 9b). The ZPA and the AER coordinated activities support mesenchymal cell proliferation and specification (Khokha et al., 2003; Laufer et al., 1994; Niswander et al., 1994; Scherz et al., 2004; Todt and Fallon, 1987). When BMPs activity is low, *Grem1* expression is upheld by the SHH signaling pathway (Benazet et al., 2009; Khokha et al., 2003; Michos et al., 2004; Zuniga and Zeller, 1999, Figure 9b). To date, GREM1 is the only known essential active BMP antagonist at these stages. *Grem1* expression is the cornerstone of the establishment of the SHH/GREM1/AER-FGF self-regulatory feedback loop and thereby of the distal progression of limb bud development (Benazet et al., 2009, Figure 9b).

A combination of several events leads to the self-termination of this feedback loop. First, *Shh* descendant cells are refractory to *Grem1* activation, increasing the gap between the domains of *Shh* and of *Grem1* expression (Nissim et al., 2006; Scherz et al., 2004). In addition, GLI3R, from the anterior SHH-free mesenchyme, and high AER-FGF activity, inhibit *Grem1* expression (Lopez-Rios et al., 2012; Scherz et al., 2004; Verheyden and Sun, 2008; Zeller et al., 2009). As a consequence, *Grem1* expression is terminated around E12.0 and the SHH/GREM1/AER-FGF loop collapses (Figure 9b).

The tight control of the *Grem1* expression throughout limb development is reflected in its very dynamic spatio-temporal expression pattern (Figure 9c). *Grem1* is transcribed in a posterior distal domain during limb bud initiation. Its domain expends distally and anteriorly during the propagation phase before progressively fading and disappearing in the autopod after E12.0 (Figure 9c).

In the post-patterning phase, BMP activities increase, which stimulates BMP-dependent chondrogenic differentiation and inter-digital apoptosis (Barna and Niswander, 2007; Benazet et al., 2012; Ganan et al., 1996; Lopez-Rios et al., 2012; Pizette and Niswander, 2000, 2001; Yoon et al., 2005).

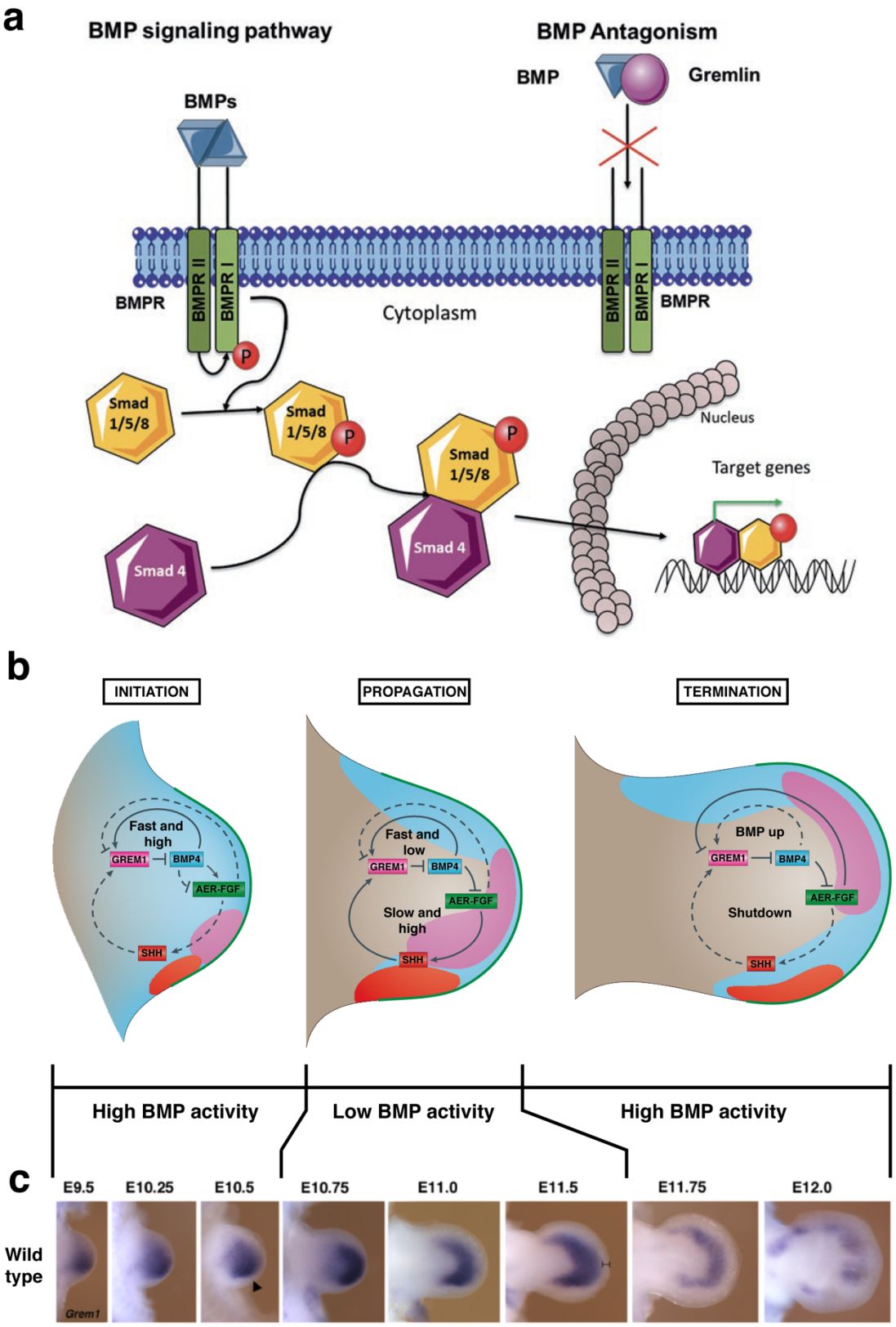


Figure 9. *Grem1* is a key factor during limb bud development. (a) A BMP ligand dimer binds its cognate receptor (BMPR) and triggers a phosphorylation cascade leading to Smad 1/5/8 pathway activation. *Gremlin1* (*Grem1*) directly binds to BMP

ligands and blocks the activation of the BMP pathway. From left to right (b) Initiation phase: High mesenchymal BMP activity (blue) results in the establishment of a functional AER (green) and activates *Grem1* transcription (pink). Propagation phase: GREM1 blocks BMP activity which permits AER-FGFs to increase ZPA *Shh* expression (red). *Shh* is then responsible for maintaining high *Grem1* expression precondition to the set-up of the SHH/GREM1/AER-FGF epithelial-mesenchymal (e-m) self-regulatory feedback loop. Termination phase: Reduction of *Grem1* transcription through an increase of AER-FGFs activity and *Grem1* refractoriness of *Shh*-expressing cells descendants prompting the regulatory feedback loop to end. (c) Dynamic *Grem1* expression throughout limb development from E9.5 to E12.0 uncovered by whole-mount *in situ* hybridization (WISH). *Adapted from (for the upper panel) Marquez-Exposito et al. (2020), (for the middle panel) Zeller et al. (2009) and (for the lower panel) Zuniga et al. (2012a)*.

4.2.3. The *Grem1 cis*-regulatory landscape

4.2.3.1. *Grem1* and mouse limb development

Until the breakthrough of Zuniga and colleagues (Zuniga et al., 2004), *limb deformity* (*ld*) phenotypes had been attributed to the LOF of the *Formin1* (*Fmn1*) gene (Figure 10a). The 24 exons of *Fmn1* span a 345 kb region in length, which is located 35 kb downstream of the *Grem1* gene 2 exons that span a much smaller 10 kb genomic region. The *Grem1* and the *Fmn1* genes are transcribed in opposite directions. Random insertional mutagenesis experiments had allowed the identification of the *ld* locus (Messing et al., 1990; Woychik et al., 1990a; Woychik et al., 1985). At the time, the *ld* locus was called the *ld* gene (Mass et al., 1990; Vogt et al., 1992; Woychik et al., 1990a; Woychik et al., 1985; Zeller et al., 1989). *Fmn1* was associated with the *ld* gene as all *ld* mutant alleles had been mapped within its sequence (Jackson-Grusby et al., 1992; Vogt et al., 1993; Vogt et al., 1992; Woychik et al., 1990b; Wynshaw-Boris et al., 1997). The *ld* mutations can result in high-frequency uni- and bilateral renal aplasia (Kleinebrecht et al., 1982; Mass et al., 1990) associated with limb long bones synostosis with autopod syndactyly and oligosyndactyly (Green et al., 1968; Kleinebrecht et al., 1982; Woychik et al., 1985; Zeller et al., 1989, Figure 10a). FMNs family proteins are involved in the regulation of actin filament and microtubule

cytoskeleton dynamics (Evangelista et al., 2002; Ishizaki et al., 2001; Manseau et al., 1996; Pring et al., 2003; Sagot et al., 2002; Wallar and Alberts, 2003). FMN1 can associate with microtubules (Zhou et al., 2006) or stimulate actin polymerization (Kobielak et al., 2004; O'Rourke et al., 2000). Nevertheless, inactivation of specific *Formin1* isoforms in the mouse was associated with partial renal agenesis phenotypes while displaying normal limb morphogenesis (Wynshaw-Boris et al., 1997). Also, not all *ld* alleles were within *Fmn1*: two spontaneous *ld* mutations, the ld^{OR} and the ld^{J} alleles affect *Grem1* open reading frame (ORF, Mass et al., 1990; Woychik et al., 1985; Wynshaw-Boris et al., 1997, Figure 10b).

In an elegant genetic study, Zuniga and colleagues excluded that *Fmn1* was involved in *ld* phenotypes. In an attempt to generate their own *ld* allele with a *lacZ* reporter, they engineered 1) the loss of *Fmn1* exon 10 ($Fmn1^{\Delta 10}$), resulting in a frameshift in the FMN1 protein C-terminal domain resulting in the degradation of FMN1 proteins, 2) mice lacking the *Fmn1* exon 10 to exon 24 ($Fmn1^{\Delta 10.24}$) region (Zuniga et al., 2004, Figure 10c). *Fmn1* was truncated in both $Fmn1^{\Delta 10}$ and $Fmn1^{\Delta 10.24}$ alleles at the level of the exon 9/10 boundary, without any aberrant splicing. The $Fmn1^{\Delta 10}$ homozygous embryos were phenotypically wild-type, despite the degradation of FMN1 isoforms, unlike the $Fmn1^{\Delta 10.24}$ ones which displayed *ld* phenotypes (Zuniga et al., 2004, Figure 10d). *Grem1* expression was maintained in $Fmn1^{\Delta 10/\Delta 10}$ but lost, exclusively in the limb bud, in the $Fmn^{\Delta 10.24/\Delta 10.24}$ embryos (Zuniga et al., 2004, Figure 10d). These mutants survived to adulthood with normal lung and kidney organogenesis.

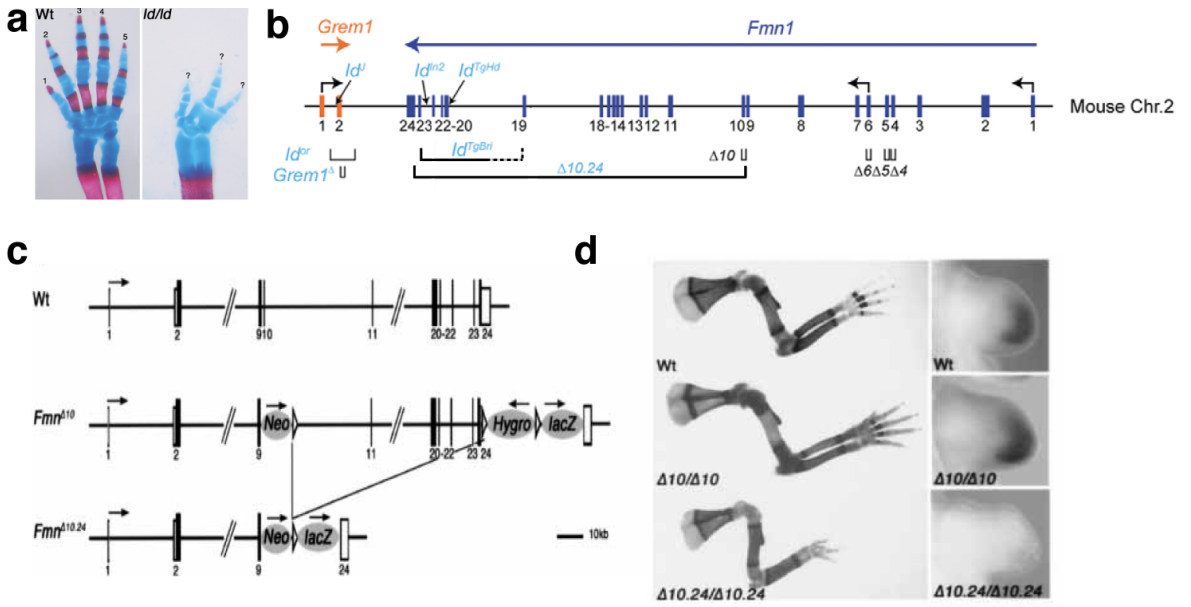


Figure 10. Mapping of limb deformity (ld) mutations within the *Grem1-Fmn1* landscape. (a) Skeletal stainings of wild-type (WT) compared to *ld/ld* mutant distal forelimbs ≥E17.0 where autopod oligosyndactyly and zeugopod synostosis is observed. (b) Schematic of the *Grem1-Fmn1* chromosomal landscape with all known mutations indicated above (*ld*^{J} point mutation, *ld*^{In2} inversions and *ld*^{TgHd}, *ld*^{TgBri} transgene insertions) or below (deletions indicated with brackets) the genomic map. *Fmn1* Δ4, Δ5, Δ6, Δ10, Δ10–24 are targeted mutations affecting *Fmn1* respective exons. *Grem1*^{Δ} is the *Grem1* null allele generated by gene targeting. Some mutations result in *ld* recessive phenotypes (blue) while others (black) do not hinder limb development. (c) Schematic of the genetically engineered *Fmn1*^{Δ10} *and Fmn1*$^{Δ10.24}$ alleles. (Neo) PGK-NeoR gene used during gene targeting; (Hygro) PGK-HygroR gene used during gene targeting; (lacZ) IRES-LacZ gene used to tag Fmn1 transcripts**.** (d) left panels: skeletal phenotypes of WT and homozygous Δ10 and Δ10–24 mutants; right panels: *Grem1* limb bud expression in WT and Δ10, Δ10–24 homozygous mice. *Adapted from (a,b) Zeller and Zuniga (2007) and (c,d) Zuniga et al. (2004).*

Finally, using homologous recombination, they inactivated the second exon of the *Grem1* gene which encodes its complete ORF. Thus, they generated the *Grem1*^{Δ} or *Grem1*^{ΔORF} allele which 1) did not complement the *Fmn1*$^{Δ10.24}$ allele, 2) phenocopied the *Fmn1*$^{Δ10.24/Δ10.24}$ mutants *ld* phenotype when homozygous (Zuniga et al., 2004). As a conclusion, they demonstrated that *Fmn1* was not implicated in *ld* phenotypes. The *ld* defects were the consequences of the loss or the alteration of the *Grem1* expression, either by the disruption of the *Grem1* ORF (e.g. *ld*^{OR} and *ld*^{J}, Figure 10a) or by the disruption of the *Grem1 cis*-regulatory landscape (e.g. *ld*^{In2}, *ld*^{TgHd}, *ld*^{TgBri}, Figure 10a), located between the exon 10 and the exon 24 of *Fmn1.* With the *Grem1*^{Δ} mice, Zuniga and colleagues phenocopied *ld* limb and kidney phenotypes that,

together with defects in lung late organogenesis, led to embryonic lethality (Zuniga et al., 2004).

4.2.3.2. *Grem1 cis*-regulatory landscape: what we know so far

After showing that the *Fmn1* exons 10 to 24 locus acts as a *cis*-regulatory region on *Grem1* transcription, Zuniga and colleagues focused on the previously described Id^{TgBri} mutation to study *Grem1* transcriptional regulation (Messing et al., 1990; Vogt et al., 1992). This allele was defined as a deletion of at least 36 kb including exons 21-22 of *Fmn1* but the boundaries of the deletion were uncertain (Vogt et al., 1992). In addition, it was not clear if further rearrangements had occurred (Vogt et al., 1992). Using a bacterial artificial chromosome (BAC, ref. Rp23:113h17, mm10 chr2:113,611,498-113,847,298) where a *lacZ* reporter gene was inserted into the *Grem1* ORF and was transcribed like *Grem1,* they revealed an essential role of a ± 70 kb *Fmn1* region located between exons 19 and 23 in the activation of the transgene (Zuniga et al., 2004, Figure 11). The BAC construct could drive reporter expression in the posterior distal part of the limb mesenchyme (Figure 11a). The deletion of the targeted region *Fmn1* exons 19 to 23 abolished the transgene reporter activity (Figure 11b). As a result, this region was renamed as the *Grem1* GCR because of its similarity with the *HoxD* GCR (Montavon et al., 2011; Spitz et al., 2003).

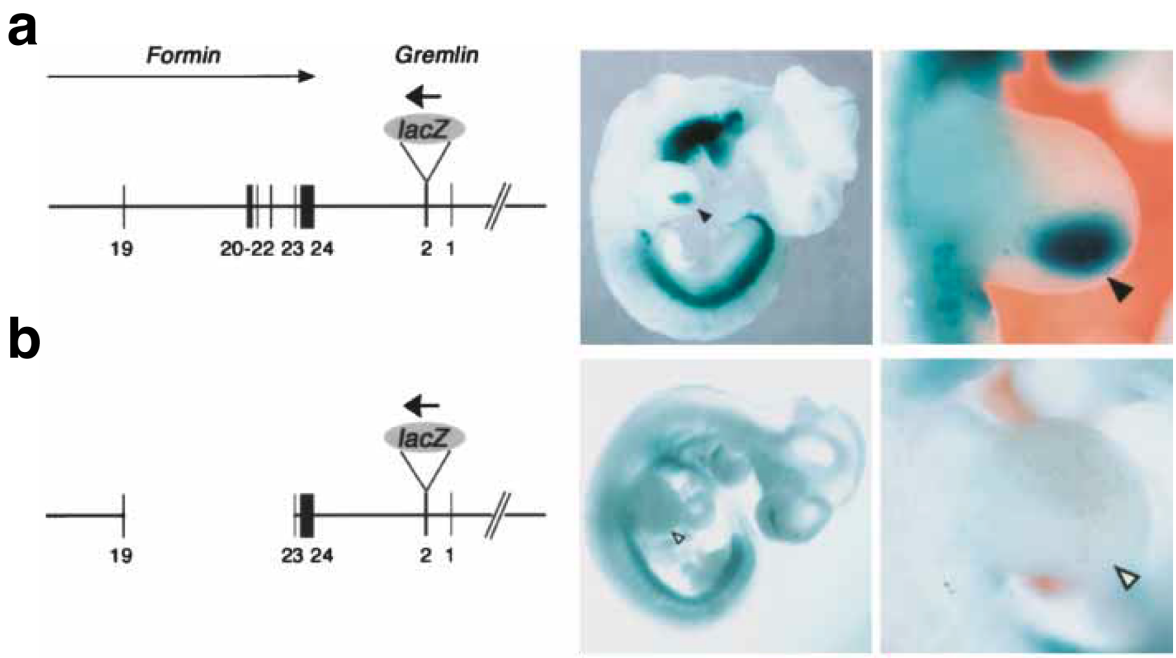


Figure 11. *Grem1* global control region (GCR) in a bacterial artificial chromosome (BAC) is essential for transgene activation. (a) The 236 kb BAC (mm10 chr2:113,611,498-113,847,298, from 25 kb upstream of *Fmn1* exon 19 to 150 kb upstream of *Grem1*) including *Grem1* GCR located between *Fmn1* exons 19 and 23 drive *lacZ* transgene inserted in-frame into *Grem1* ORF expression in the posterior distal part of the limb bud mesenchyme (black arrowhead). (b) The deletion of the GCR inactivates the BAC enhancer activity. *From Zuniga et al. (2004).*

Looking into the newly identified GCR, Zuniga and colleagues isolated the SHH pathway target GLI binding regions (GBRs, Marigo et al., 1996; Ruiz i Altaba, 1998; Vokes et al., 2008), highly conserved between human, mouse, chicken and opossum (HMCO1, HMCO2, Zuniga et al., 2012a). HMCO1 belonged to the 9.6kb *Grem1* regulatory sequence 1 (GRS1) which possessed two additional GBRs outside of HMCO1, as identified by ChIP-qPCR experiments (Zuniga et al., 2012a). Those GBRs were not as conserved as those in HMCO1. Following up with the same BAC experimental set up (Figure 12a), they showed that both GRS1 and HMCO2 were associated with *Grem1* transcriptional regulation (Figure 12b and 12c). In this BAC, GRS1 was necessary for *lacZ* transgene expression (Figure 12b), while HMCO2 removal reduced the strength of transgene expression (Figure 12c).

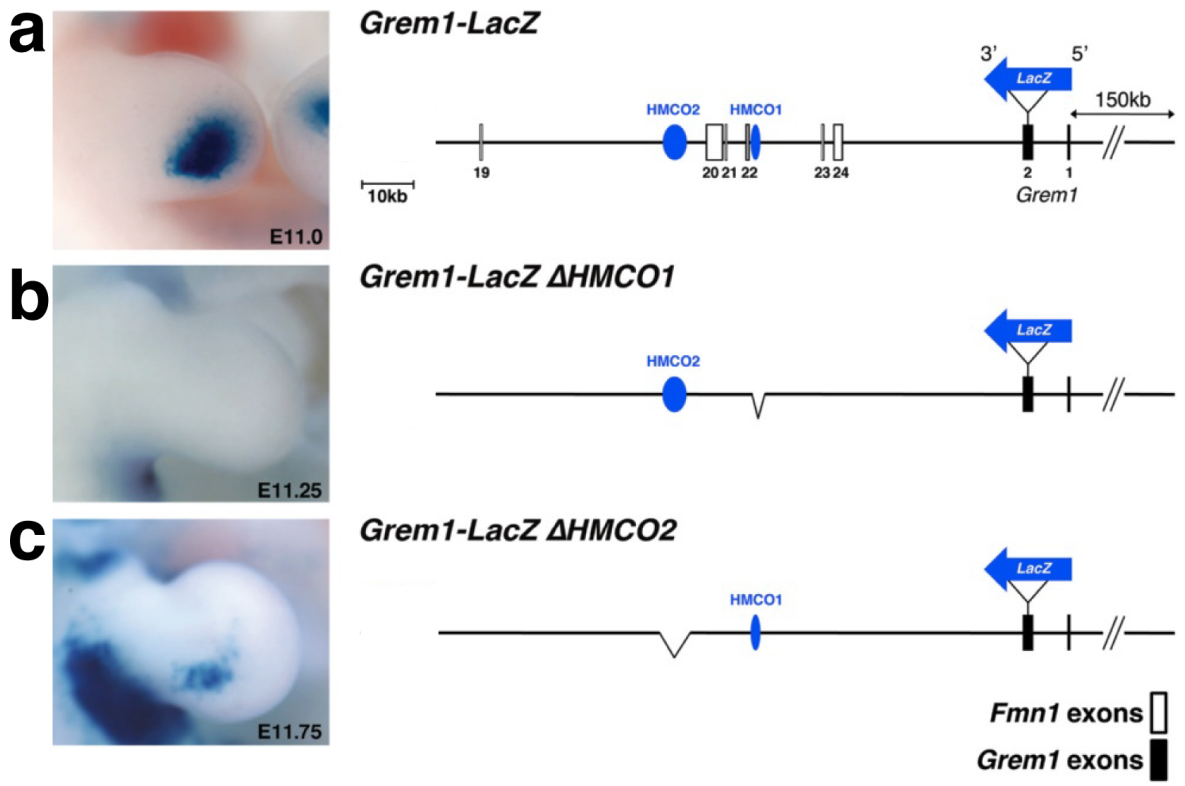


Figure 12. BAC-based strategy reveals human, mouse, chicken and opossum (HMCO1) as critical for GCR enhancer function and HMCO2 as important for the transcriptional strength of the *lacZ* transgene. (a) BAC construct with a *lacZ* cassette inserted in-frame into *Grem1* ORF drives endogenous enhancer activity in the posterior and distal limb bud mesenchyme. (b) The removal of HMCO1 from the BAC prevents *lacZ* expression. (c) The removal of HMCO2 from the BAC reduces *lacZ* expression. *From Zuniga et al. (2012a).*

One element they did not focus on during their study is the GBR4 (fourth GLI binding site after GRS1 and HMCO2) because it is not conserved like the HMCOs. The GBR4 was eventually analyzed by the group of Dr. S. Vokes as a *Grem1* enhancer in the limb posterior domain and as a *Grem1* repressor in the limb anterior domain (Li et al., 2014, Figure 13). Its enhancer or silencer activities depended on its interaction with the activator (present in the posterior limb bud mesenchyme) or repressor (present in the anterior limb bud mesenchyme) isoform of GLI3 (Ahn and Joyner, 2004; Wang et al., 2000; Wen et al., 2010). Thus, this regulatory element was renamed as GLI responsive element 1 (GRE1, Li et al., 2014; Vokes et al., 2008).

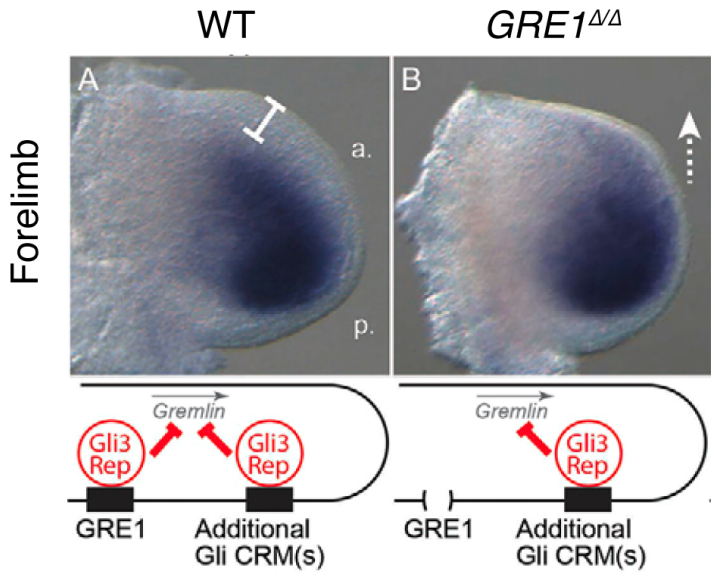


Figure 13. GRE1 interaction with GLI3 represses *Grem1* anterior expression. (A-B) Comparison of *Grem1* WISH in forelimbs of E10.5 (35-36 somites) mouse wild-type (WT) *and* GRE1 homozygous deleted embryos (*GRE1*$^{\Delta/\Delta}$). *Grem1*-free domains are indicated with white brackets. (B) Dotted white arrow shows ectopic distal-anterior *Grem1* expression in *GRE1*$^{\Delta/\Delta}$ backgrounds. Lower panel: schematic of GLI3 repression of *Grem1* might occur in WT and *GRE1*$^{\Delta/\Delta}$ backgrounds through GRE1 and/or other putative silencers. *Adapted from Li et al. (2014).*

4.2.4. *Grem1* and human limb congenital diseases

In humans, congenital limb malformations (CLMs) affect ± 1 in 500 live births (Giele et al., 2001). Cenani-Lenz (C-L) syndrome is a limb congenital malformation characterized by syndactyly of the hands and feet with complete fusion of bones and soft tissue. Individuals suffering from C-L syndrome also display shortening and fusion of zeugopod skeleton in forelimbs and hindlimbs and require reconstructive surgery (Cenani and Lenz, 1967; De Smet et al., 1992; Furniss et al., 2009; Malik, 2012). The C-L syndrome phenotypes can be associated with various systemic features such as renal defects (Bacchelli et al., 2001; Dimitrov et al., 2010). The murine *ld* model displays the most similar phenotypes to human C-L syndrome. *Grem1* transcriptional disruption in mice is responsible for the pleiotropic *ld* phenotypes. Dimitrov and colleagues in 2010 studied the case of C-L-like individuals in which the *Grem1 cis*-regulatory landscape was affected but not the *Grem1* gene itself (Dimitrov et al., 2010).

One patient showed *ld*-like phenotypes with radioulnar synostosis, syndactyly in both hands and feet plus an agenesis of the left kidney (Dimitrov et al., 2010, Figure 14a). Those phenotypes are identical to the mouse *ld* phenotypes. This patient carried a deletion of at least 246 kb, including the first 12 exons of *Fmn1.* This deletion did not encompass the *Grem1* annotated GCR (Zuniga et al., 2004) suggesting a loss of essential *Grem1* CRMS outside of the GCR and/or a structural disruption of *Grem1* *cis*-regulatory landscapes that, in humans, could cause C-L syndrome (Figure 14b).

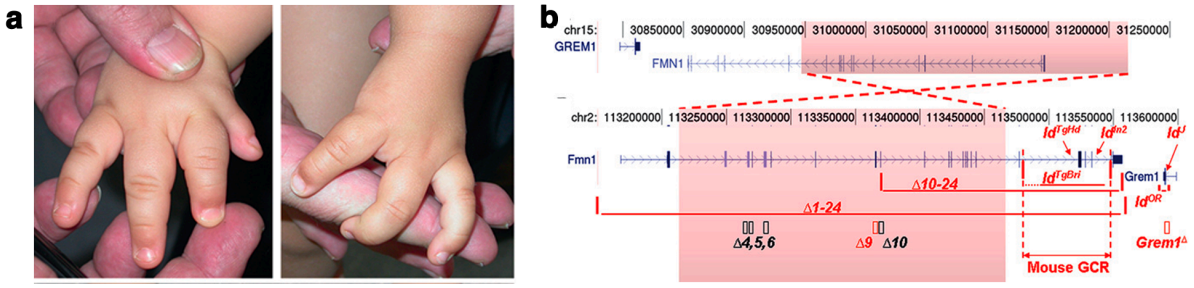


Figure 14. Cenani-Lenz-like phenotype observed in patient correlates with *Grem1* *cis*-regulatory landscape mutation. (a) young patient hands with oligosyndactyly because of (b) a 246 kb deletion (minimal size) including at least the first 12 exons of *Fmn1*. *From Dimitrov et al. (2010).*

We have discussed so far the *Grem1* LOF phenotypes. It is noteworthy that the *Grem1* gain of function (GOF) has deleterious effects. High expression of *Grem1* has been reported in different tumor types, such as in breast and colorectal cancer (Jaeger et al., 2012; Ren et al., 2019). Hereditary mixed polyposis syndrome (HMPS) characterized by autosomal dominant inheritance of various forms of colorectal polyp resulting in colorectal carcinoma (CRC) at high frequency is the consequence of a 40 kb duplication upstream of *Grem1.* Once again, this mutation does not affect the GCR.

Despite the amazing progress that have been made, much remains to be done in the understanding of *Grem1* transcriptional regulation.

5. Aims of the thesis

My Ph.D. project revolved around the understanding of the mechanisms by which the *Grem1* gene, a key developmental player of limb patterning and outgrowth, is transcriptionally regulated during mouse organogenesis.

Aim 1 – To identify all *Grem1*-associated CRMs, using chromatin profiling and transgenic reporter assays to locate possible *Grem1* CRMs.

Aim 2 – To characterize the functional role(s) of the identified CRMs in the regulation of *Grem1* expression, using a loss of enhancer function strategy based on the CRISPR/*Cas9* technology.

Aim 3 – To elucidate the chromatin organization of the *Grem1* landscape in wild-type and genetic mutant backgrounds using 4C-seq experiments.

Aim 4 – To get a more comprehensive understanding of the cellular mechanisms by which *Grem1* transcriptional defects lead to a spectrum of *ld* phenotypes; using Lyso Tracker experiments to look for abnormal apoptotic cell clusters and using fluorescence-activated cell sorting (FACS) to control the proliferation rate in mutant limb buds.

In addition to my own project, I was given the opportunity to participate in the study of the BMP signaling transducer SMAD4 functions during early limb bud development (Gamart et al., manuscript in preparation). As BMPs are direct transcriptional regulators of *Grem1*, this project complemented and enriched my own research. In addition, I was given the chance to support my colleagues working on getting insights into the underlying transcriptional mechanisms supporting vertebrate limb diversification by performing ATAC-seq experiments in pig embryonic limb bud at different stages.

6. Results

6.1. Identification of *Grem1* putative *cis*-regulatory elements

6.1.1. Characterization of the 310 kb *Grem1* TAD

Grem1 defines a functionally essential node in the SHH/GREM1/AER-FGF self-regulatory feedback loop governing limb bud outgrowth and patterning (Benazet et al., 2009). Previous studies have shown alterations of the *Grem1* chromatin landscape, leading to severe limb deformity phenotypes, reflecting major transcriptional alteration both in mouse and in human (Dimitrov et al., 2010; Jaeger et al., 2012; Khokha et al., 2003; Zuniga et al., 2004; Zuniga et al., 2012b). I thus investigated *Grem1* flanking genomic region to look for *Grem1*-associated chromatin, containing the CRMs in charge of tightly controlling *Grem1* dynamic expression in space and time during limb bud development. Hi-C maps offer the possibility to observe three-dimensional chromatin organization in order to identify regulatory blocks within which CRMs and their associated gene(s) interacted at high-frequency (Dixon et al., 2012; Nora et al., 2012; Sexton et al., 2012). Those regulatory units or TADs are organized in loop domains demarcated by divergent CTCF sites hindering inter-TAD interactions (Fudenberg et al., 2016; Rao et al., 2014). TADs are described as well-conserved throughout different cell types and species, once established they stay unchanged during embryonic development (Dixon et al., 2012).

Analysis of Hi-C experiments performed on mouse embryonic fibroblast with a 10 kb resolution by Barutcu and colleagues suggested the *Grem1* gene belonged to a 530 kb TAD at chromosomal coordinates mm10 chr2:113,298,340-113,830,145 (Barutcu et al., 2018), in which convergent CTCF sites would organize intra-TAD looping (Andrey et al., 2017, Figure 15). *Grem1* segregated together with *Scg5* and *Fmn1* genes and I observed divergent CTCF motifs 5' of *Fmn1*, separating *Grem1, Scg5* and the 3' portion of *Fmn1* from the *Fmn1* 5' region and thus generating two relatively

independent TADs (dashed lines Figure 15). This hinge, where the sub-TADs collide, was located between exon 8 and 9 of *Fmn1* (chromosomal coordinates mm10 chr2:113469703-113494902*)*, potentially isolating the *Fmn1* regulatory landscape from the *Grem1/Scg5* specific regulatory elements. By looking at the CTCF sites and according to their orientations, I presumed the regulatory region of *Grem1* was contained within a 310 kb TAD at chromosomal coordinates mm10 chr2:113,493,843-113,807,346 (Figure 15, blue dashed lines). I could not exclude that the region structurally associated with the *Fmn1* gene (chromosomal coordinates mm10 chr2:113,298,340-113,493,843, Figure 15, grey dashed lines) would be associated with *Grem1* transcriptional control as the frequency of interaction between these two loops was quite high (Figure 15, red dashed lines). These observations did not strictly exclude a role of potential CRMs located within the *Fmn1* TAD in *Grem1* transcriptional regulation.

Transcriptome profiles from mouse forelimb buds at E11.5 (and E10.5, Appendixes Figure 1) showed a high *Grem1* transcriptional level while *Fmn1* and *Scg5* were very lowly transcribed (Andrey et al., 2017, Figure 15). These observations were supported by WISH experiments performed by Victorio Palacio showing a low signal for *Fmn1* transcripts and the absence of signal for *Scg5* transcripts. In contrast, the *Grem1* signal was strong and characterized by a crescent-shape pattern at developmental stages E10.5-E11.0 (Figure 15). *Fmn1* seemed to be weakly transcribed in the mesenchymal part of the limb bud and accumulated in the AER, while *Scg5* was inactive in the limb (Iguchi et al., 1984; Pavel et al., 2007; Westphal et al., 1999, Figure 15). Zuniga and colleagues have previously demonstrated that *Fmn1* was not involved in limb development (Zuniga et al., 2004). All these observations suggested that the regulatory regions within the 530 kb TAD could be 1) active CRMs associated with *Grem1*, 2) CRMs associated with *Fmn1* or *Scg5* and weakly active or inactive in our model (Figure 15).

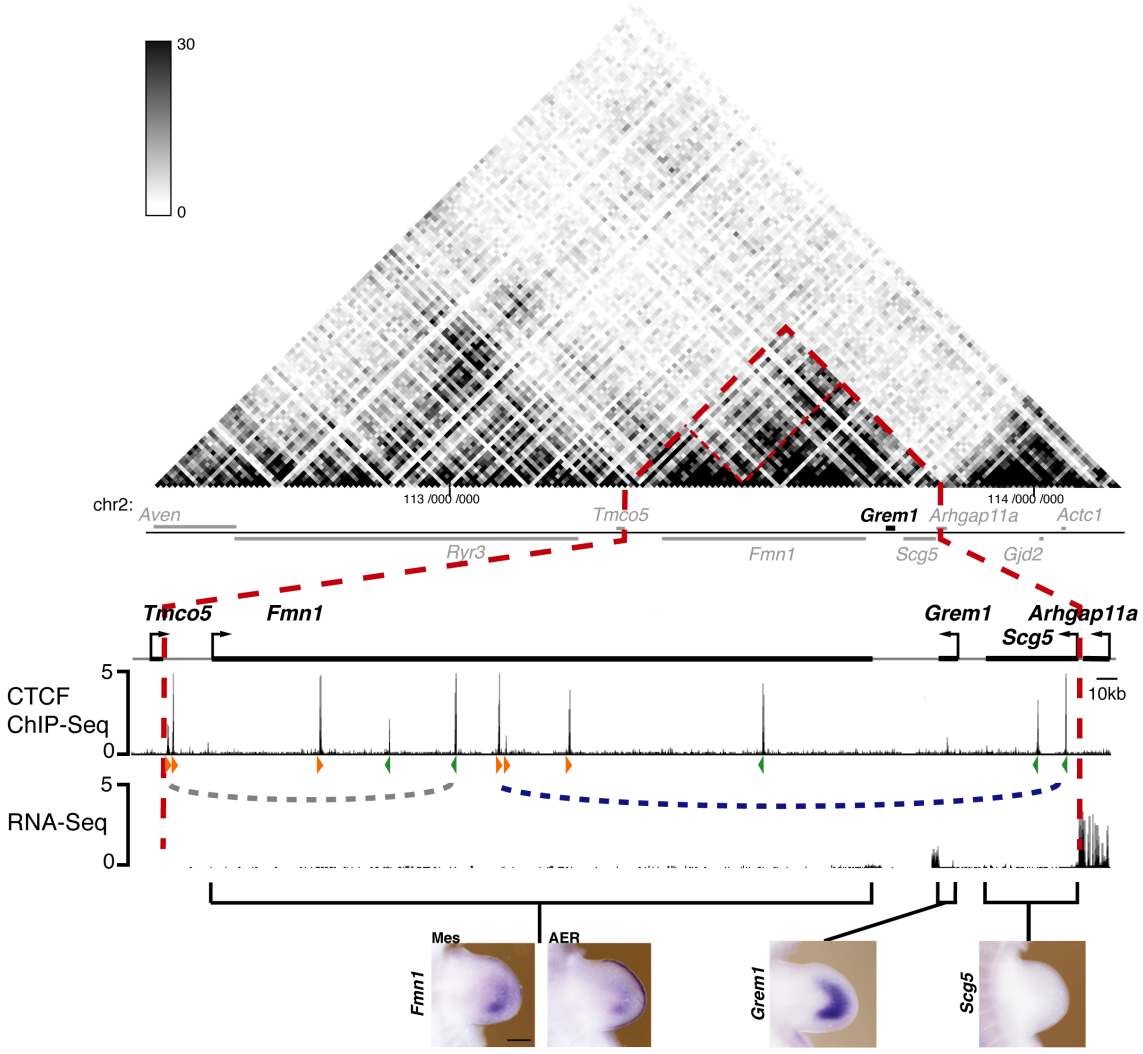


Figure 15. Identification of the *Grem1* topologically associated domain (TAD). Upper panel: Hi-C map (Barutcu et al., 2018) obtained from mouse embryonic fibroblasts with a 10kb resolution shows the distribution of TADs on either side of the *Grem1* gene in black, other genes are represented in dark grey lines, dashed lines indicate TADs. Lower panel: UCSC browser windows displaying E11.5 mouse forelimb bud CTCF ChIP-seq (Andrey et al., 2017) identify CTCF binding sites, their orientation is represented with orange and green arrowheads. Blue and grey dashed lines unveil a potential *Grem1* TAD and a potential *Fmn1* TAD respectively. E11.5 mouse forelimb bud RNA-seq (Andrey et al., 2017) showing the transcriptional activity of *Grem1* and its neighboring genes. Lower panel: WISH, performed by Victorio Palacio for *Fmn1* and *Scg5,* showing the limb expression pattern of *Grem1* TAD genes, *Fmn1* in the mesenchyme (mes) and in the AER, *Grem1* and *Scg5* that was not detected (n=3). *Scale bar: 250µm*

I identified a 530 kb TAD, that could be divided into a *Grem1*-associated TAD and a *Fmn1*-associated TAD. I predicted that the *Grem1*- *and Fmn1*-related TADs would contain CRM-like regions. Each sub-TADs being relatively independent of each other, I estimated that *Grem1* CRMs would be located within the 310 kb loop domain uncovered by Hi-C-seq experiments.

6.1.2. Identification of 14 potential CRMs in the *Grem1-Fmn1* TAD

Our CRM discovery strategy relied on well-established criteria of epigenomic profiling. CRMs are located at chromatin accessible regions. ATAC-seq experiments, performed by Dr. Rushikesh Sheth and Jens Stolte, expose open chromatin regions including CRMs but not exclusively (Buenrostro et al., 2013). In order to reveal potential *Grem1* CRMs, I aligned mouse forelimb ATAC-seq profiles at embryonic stages 10.5 and 11.5 – when *Grem1* transcription is strongest and highly dynamic – with ChIP-seq profiles of different histones marks (Andrey et al., 2017). The H3K27ac and H3K4me1 modifications are associated with active and distal enhancers while the combination of H3K27ac and H3K4me3 decoration highlights promoter sequences. As CRMs can be highly conserved through evolution, I compared the previous data sets with available mammalian conservation profiles (Figure 16). I thus identified 10 CRM-like loci (Figure 16, light blue vertical bars) within the *Grem1* loop domain (Figure 16, black bar) and could define 4 CRM-like loci (Figure 16, light blue vertical bars) in the *Fmn1* loop domain (Figure 2, grey bar). The histone “code” and CTCF ChIP-seq profiles (Andrey et al., 2017) allowed us to pinpoint promoter and boundary elements among the CRM-like regions identified by our chromatin profiling (Figure 15, Figure 16). I classified these elements from CRM1 to CRM14, the closest (in linear distances) to *Grem1* being the number 1 and the most distant the number 14. The CRM12 was decorated by the promoter mark H2K4me3 in accordance with previous studies describing the CRM12 region as an alternative *Fmn1* promoter (Jackson-Grusby et al., 1992).

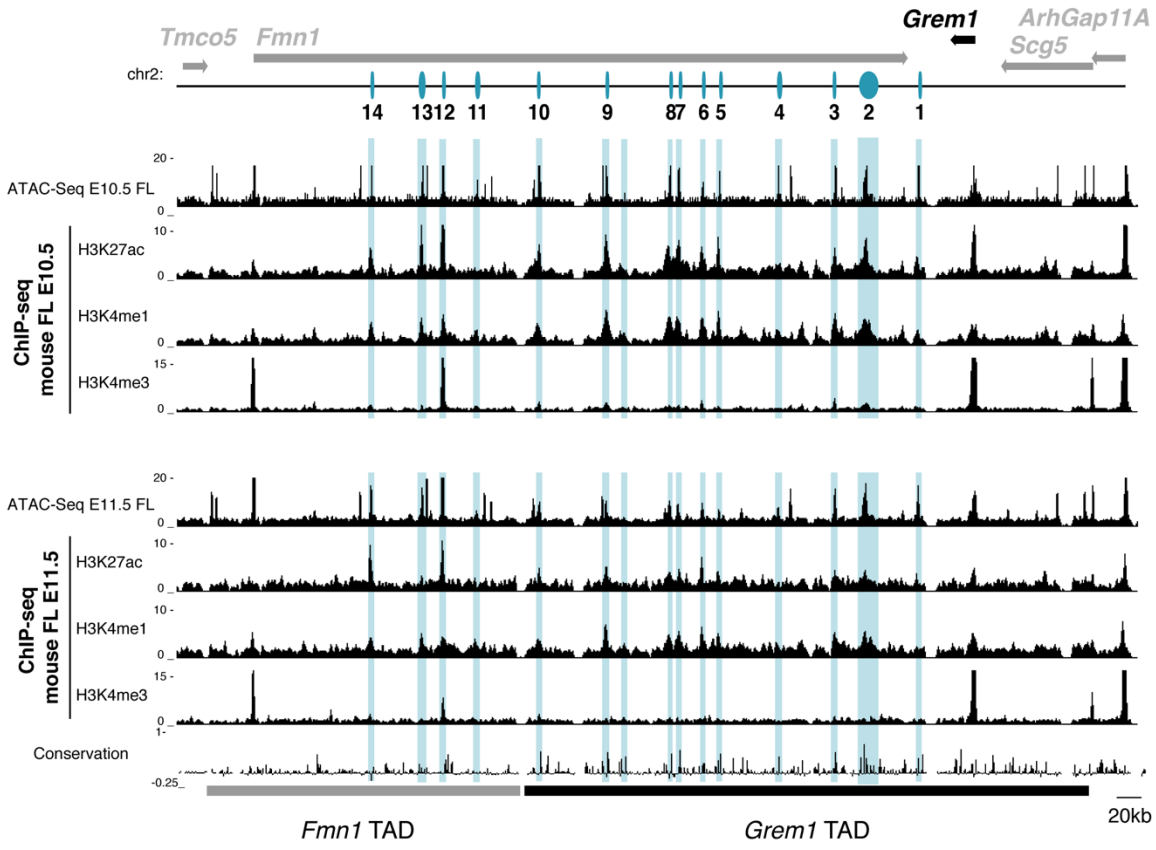


Figure 16. Identification of putative *Grem1* CRMs. Upper panel: UCSC browser windows displaying E10.5 forelimb bud ATAC-seq profile, assays by Dr. Rushikesh Sheth and Jens Stolte, ChIP-seq profiles obtained from E10.5 mouse forelimb bud tissue targeting the enhancer marks H3K27ac and H3K4me1 and the promoter signature H3K4me3 (Andrey et al., 2017). Lower panel: E11.5 forelimb bud ATAC-seq profile, ChIP-seq profiles for enhancer marks H3K27ac and H3K4me1 and the promoter signature H3K4me3 (Andrey et al., 2017). The last track shows mammalian conservation. The grey and black lines represent the *Fmn1* and the *Grem1* loop domains respectively.

Another criterion for CRM discovery is their ability to bind TFs to regulate the expression of their associated genes according to environmental cues. The signaling pathways regulating the highly dynamic transcription of *Grem1* have been extensively studied for the past 20 years. The BMP4 pathway, the SHH pathway, the WNT/ß-catenin pathway and the developmental regulators *HoxA*;*D* genes are all involved in the initiation or propagation of *Grem1* expression during limb bud development (Benazet et al., 2009; Sheth et al., 2013; Dr. Unal Ph.D. thesis; Sheth et al., manuscript in preparation). I intended to uncover possible physical interactions between the putative CRMs and the TFs relaying those signaling pathways regulatory information.

To this end, ChIP-seq experiments, in E11.5 mouse embryo forelimbs, targeting the nuclear mediators of the *Shh* pathway – GLI3 ChIP-seq (in collaboration with Dr. Kevin Peterson) – of the *Wnt* pathway – ß-catenin ChIP-seq (Sheth et al., manuscript in preparation) – of HOXA13 and HOXD13 (Sheth et al., 2016) – were analyzed (Figure 17a). BMP pathway being involved in the initiation of *Grem1* transcription (Benazet et al., 2009), I analyzed ChIP-seq data from a FLAG-tagged SMAD4 mouse line ($SMAD4^{3xF}$) at early developmental stages E9.5 and E10.5 (Figure 17b, Gamart et al., manuscript in preparation). I observed that all the CRM-like regions I identified were able to bind at least one transcription factor. Most CRM candidates were interacting with several TFs. I noted only one CRM-like locus (CRM2) bound by SMAD4 at E9.5. This interaction was lost at E10.5. This element was hypothesized to be the only CRM involved in integrating BMP pathway activity at early limb developmental stages.

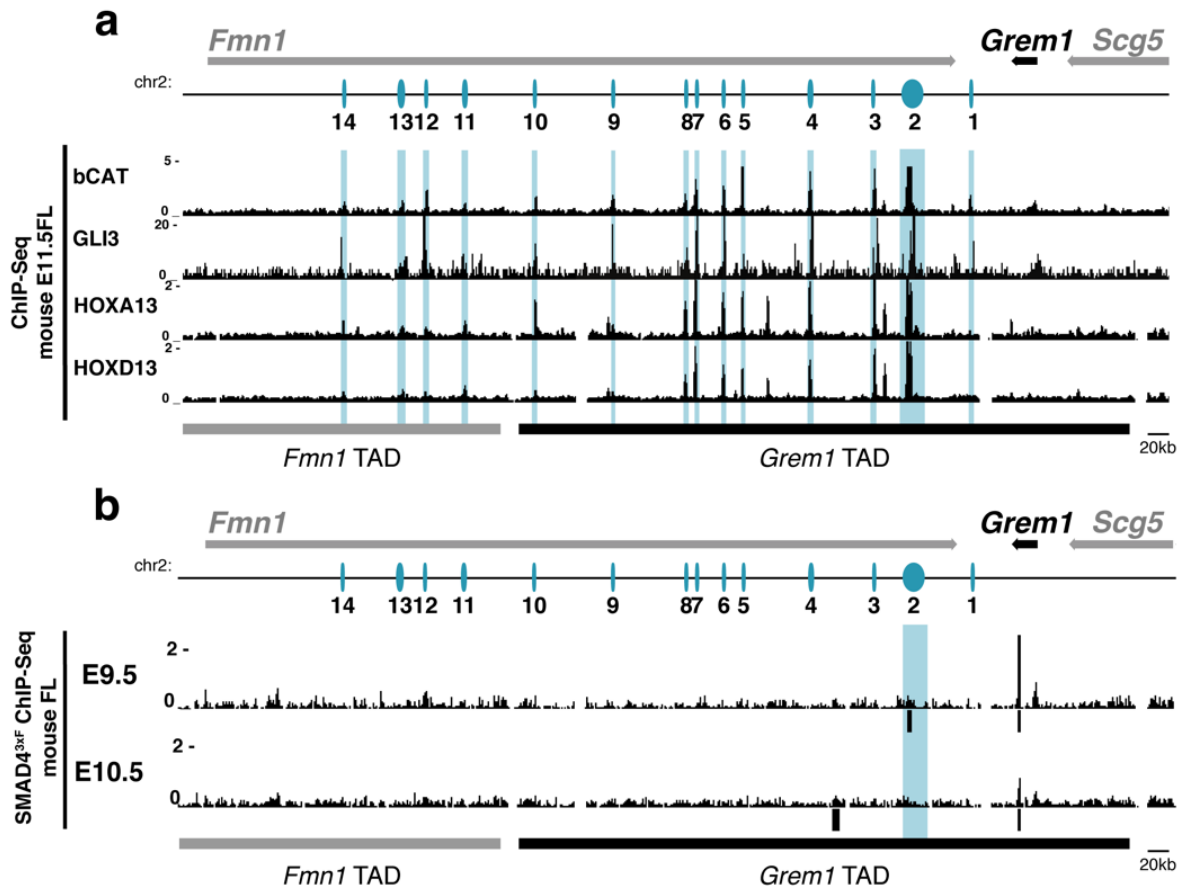


Figure 17. Identification of interactions between *Grem1* putative CRMs and *Grem1* *trans*-acting factors. (a) UCSC browser windows displaying ChIP-seq profiles in E11.5 mouse forelimb buds for the transcription factors β-catenin (Sheth et al., manuscript in preparation), GLI3 (in collaboration with Dr. Kevin Peterson),

HOXA13 and HOXD13 (Sheth et al., 2016). (b) UCSC browser windows displaying ChIP-seq profiles in E9.5 and E10.5 mouse forelimb bud targeting BMP pathway readout SMAD4^{3xF} (Gamart et al., manuscript in preparation). The grey and black lines represent the *Grem1* and the *Fmn1* loop domains respectively.

Fourteen CRM-like regions had been identified at this stage of analyses. I grouped them according to their localization within the *Grem1* TAD (CRM1 to CRM10) or the *Fmn1* (CRM11 to CRM14).

6.1.3. Analysis of *Grem1* putative CRMs enhancer activity

At this stage of the analyses, 14 putative CRM elements were identified. I intended to pinpoint which of these regions were active to drive a reporter expression in overlapping fashion with *Grem1* endogenous expression. In order to test those regions for enhancer activity, I used a *lacZ* reporter assay in transgenic mice, in collaboration with the Center for Transgenic Model (CTM, Appendixes table 1). I cloned the CRMs from mouse genomic DNA into an hsp68-*lacZ* expression vector using the Gibson Assembly® Method.

6.1.3.1. The *Grem1* TAD presents 8 limb CRMs

I screened CRM1 to CRM10 for enhancer activity. I discovered that CRM8 and CRM9, despite presenting all epigenetic characteristics of enhancers, were not able to drive a reproducible *lacZ* pattern of expression in our assay (Figure 18, Appendixes table 1). CRM8 and CRM9 were not limb bud enhancers.

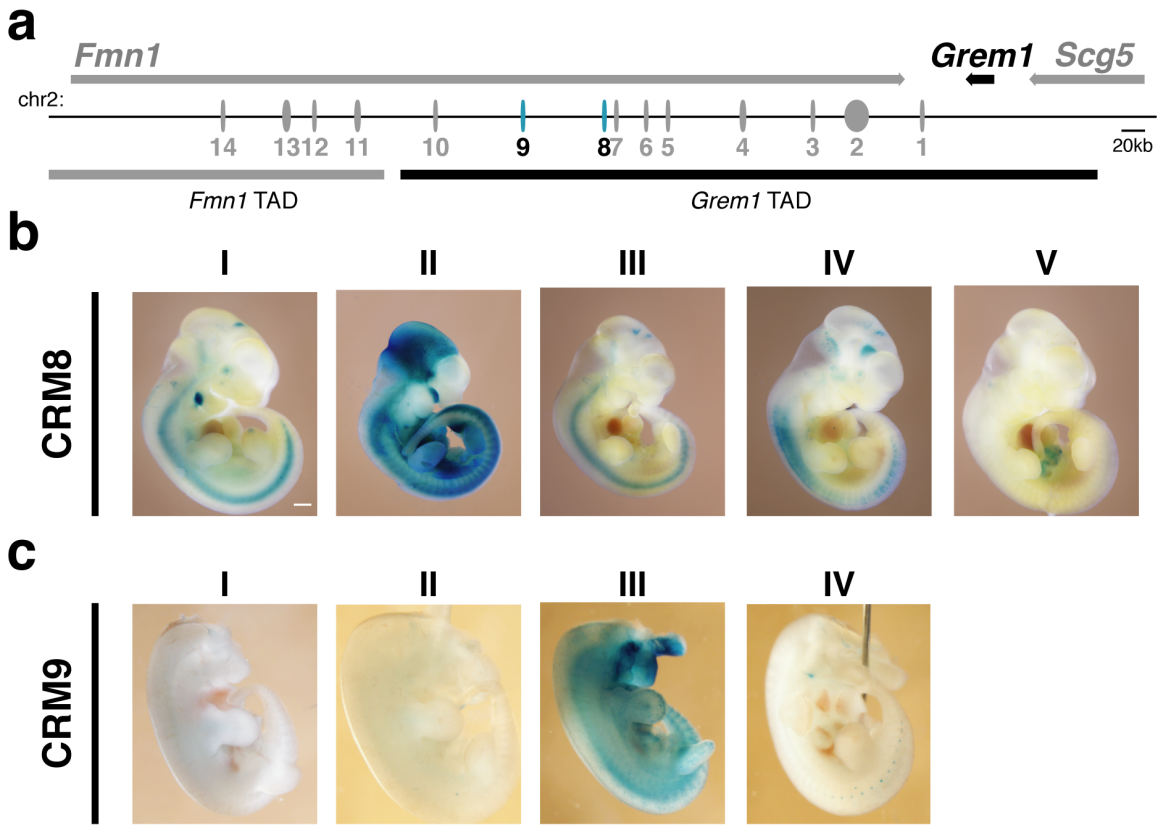


Figure 18. CRM8 and CRM9 enhancer activities in transgenic mouse embryos. (a) Schematic of the *Grem1-Fmn1* locus (b-c) CRM8 and CRM9 transgenic embryos collected at E11.5. *Scale bar: 250µm*

The remaining CRMs were all able to drive reporter expression in the limb bud (Appendixes table 1). I compared the LACZ protein spread to the pattern observed in the *Tg-GLE(tg/+)* reporter mouse line established in the laboratory (Skarnes et al., 2011, Figure 19). This mouse strain is a knock-in where the *ß-galactosidase* reporter gene, under the control of the *Grem1 cis*-regulatory region, is expressed exclusively in the limb bud.

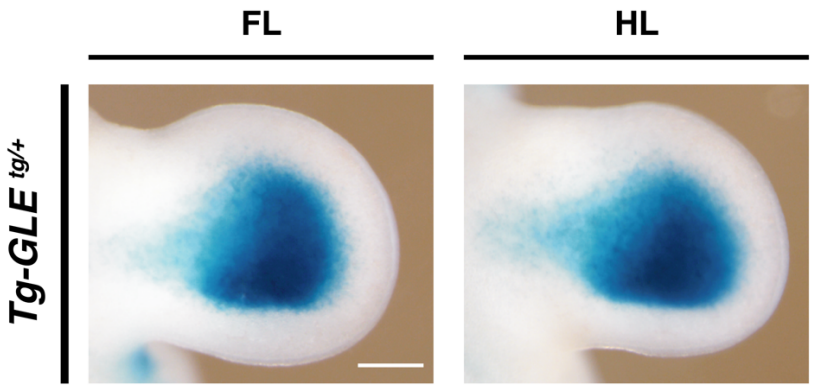


Figure 19. *LacZ* expression in the *Tg-GLE(tg/+)* transgenic mouse line forelimb and hindlimb at E11.5. The *lacZ* reporter staining overlaps with the endogenous *Grem1* expression domain in the limb buds. *Scale bar: 250µm*

These CRMs could be distinguished into two categories. The first category contained CRM1, CRM3, CRM4, CRM6, CRM7 and CRM10. This group of regulatory elements contained limb enhancers, some able to drive a reproducible reporter expression that partially overlaps with *Grem1*-expressing cells (Figure 19, Figure 20). A total of 62.5% CRM1 transgenic limb that expressed *lacZ* (n=5/8, Figure 20b) showed a reproducible signal covering the autopod entirely, including the *Grem1* expression domain as observed by WISH (Figure 15) and in the *Tg-GLE(tg/+)* embryonic limb (Figure 19). CRM3 provided a reproducible *lacZ* expression in the anterior and distal parts of the autopods in 87.5% of the transgenic limb expressers (n=7/8, Figure 20c), partially overlapping with the anterior distal *Grem1* activity. CRM4 limb expressers displayed *lacZ* expression in the posterior distal part of the limb bud for all of them (8/8, Figure 20d) and this time partially overlapping with the posterior distal *Grem1* activity. CRM6 was a limb enhancer but did not drive a reproducible *lacZ* expression as two-third of the transgenic limbs showed different activity patterns, partially overlapping with the *Grem1* expression domain, while the last embryo was stained in the AER (Figure 20e). CRM7 showed for 66% of the *lacZ*-expressing limbs a distal and anterior enhancer activity (n=4/6, Figure 20f), outside of the *Grem1* expression domain, while in forelimbs CRM10-induced expression appeared mosaic and not reproducible. CRM10-driven hindlimb expression was highly reproducible and overlapped with the *Grem1* hindlimb posterior distal domain of expression (n=3/4, Figure 20g). While forelimb expression

of the reporter gene for CRM10 constructs was quite weak, the hindlimb reporter signal suggested the CRM10 enhancer as *a Grem1* hindlimb-specific enhancer.

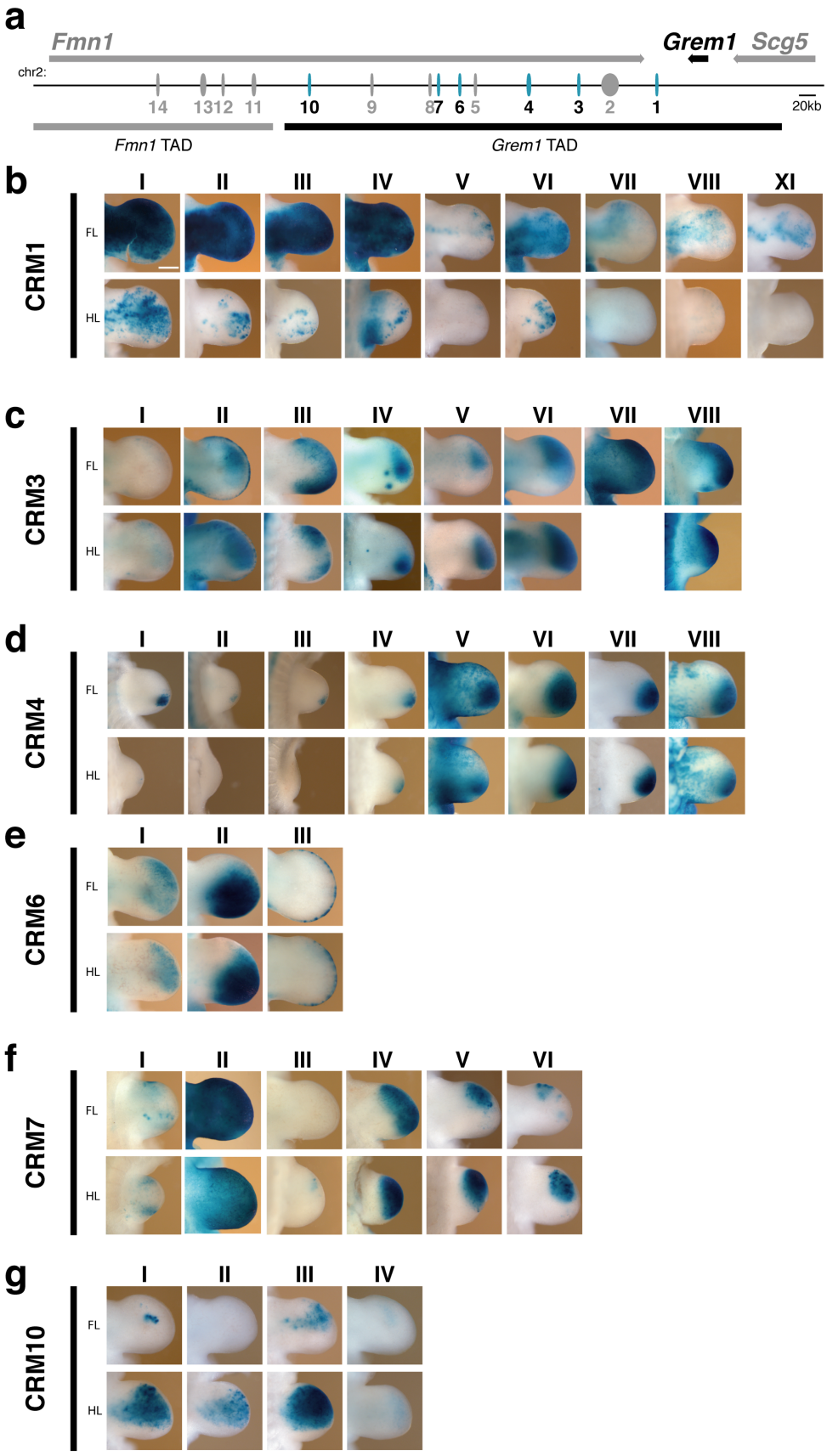


Figure 20. CRM1, CRM3, CRM4, CRM6, CRM7, CRM10 enhancer activities in transgenic mouse embryos. (a) Schematic representation of the *Grem1-Fmn1* locus (b-g) CRM1, CRM3, CRM4, CRM6, CRM7, CRM10 transgenic embryonic limb buds collected at E11.5 showing the *lacZ* reporter activity under the control of each different CRMs. *Scale bar: 250µm*

Among these active limb enhancers, CRM3 and CRM4 had already been identified as HMCO2 in our laboratory (Zuniga et al., 2012a), and by Li and colleagues as GRE1 (Li et al., 2014). While HMCO2 had been described as a region unable to drive a reproducible *lacZ* expression (Zuniga et al., 2012a), successful collaboration with the CTM revealed CRM3 as a strong limb enhancer: 87.5% of the transgenic CRM3-hsp68-*lacZ* were limb expressers (see Material and Methods section, Table 9.5.1). Together, our data suggested that all the CRMs in this category were potential modulators of *Grem1* transcription in the limb with distinct spatial activities.

According to their spatial enhancer activities, CRM2 and CRM5 did not belong to any of the heretofore described categories. *Grem1*-associated enhancers were expected to display a posterior distal enhancer activity with an anterior expansion around E11.0. These two promising loci distinguished themselves from all the previously described elements as they were able to drive reporter expression in the transgenic forelimbs similar to what was observed in the *Tg-GLE(tg/+)* mouse transgenic line (Figure 19), in accordance with our expectations. CRM2 – previously known as GRS1 (Zuniga et al., 2012a) – and CRM5 could induce a reporter expression in the limb bud mesenchyme posteriorly and distally with an anterior expansion characteristic of endogenous *Grem1*-expressing cells localization (Zuniga et al., 2012a, Figure 15, Figure 21b,c). CRM2 drove a *Grem1*-like forelimb reporter expression pattern in all limb expressers embryos, while CRM5 forelimb activity was weaker than the activity of CRM2 and detected in a *Grem1*-like domain in 50% of the forelimb expresser embryos. A total of 50% of CRM5 embryos displayed a wide hindlimb signal spread, engulfing the full hindlimb *Tg-GLE(tg/+)* pattern (Figure 19, Figure 21c). My results confirmed CRM2 as a strong limb enhancer which drove the closest endogenous *Grem1*-like expression pattern and identified CRM5 as a second enhancer with similar activity, albeit less robust. I had yet to validate their association with *Grem1* transcriptional regulation and to further validate their individual role(s).

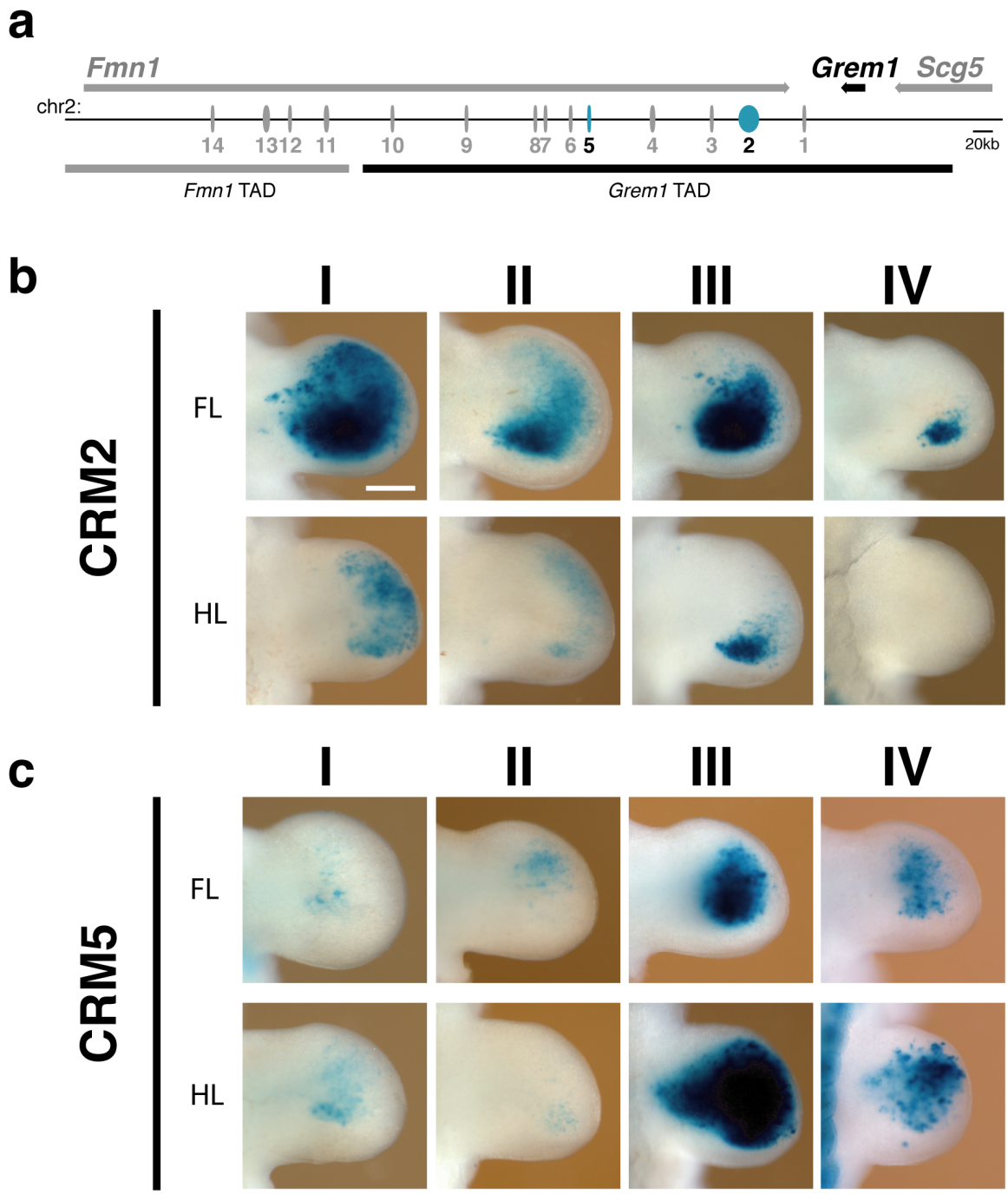


Figure 21. CRM2 and CRM5 enhancer activities in transgenic mouse embryos. (a) schematic of the *Grem1-Fmn1* locus highlighting CRM2 and CRM5 with blue ovals (b-c) CRM2 and CRM5 transgenic embryos limb bud forelimbs collected at E11.5 showing the *lacZ* reporter activity under the control of each different CRMs. *Scale bar: 250μm*

Interestingly, Zuniga and colleagues had demonstrated that a 170 kb deletion between exon 10 and exon 24 of *Fmn1* (chromosomal coordinates mm10 chr2: 113,528,475-113,709,592) was sufficient to inactivate *Grem1* expression during mouse limb bud

development, leading to severe limb deformity phenotype with a 100% penetrance (Zuniga et al., 2004). Their study suggested that outside of this region, relabeled as the *Cis* region in our study (Figure 22a, dashed lines), no enhancer was able to drive *Grem1* transcription. I was indeed able to reproduce their observations (Figure 22b). RT-qPCR and WISH analysis of $Cis^{\Delta/\Delta}$ embryos confirmed the loss of *Grem1* expression and 100% of the forelimb skeletal of E14.5 forelimb had severe *ld* phenotypes (Appendixes table 1).

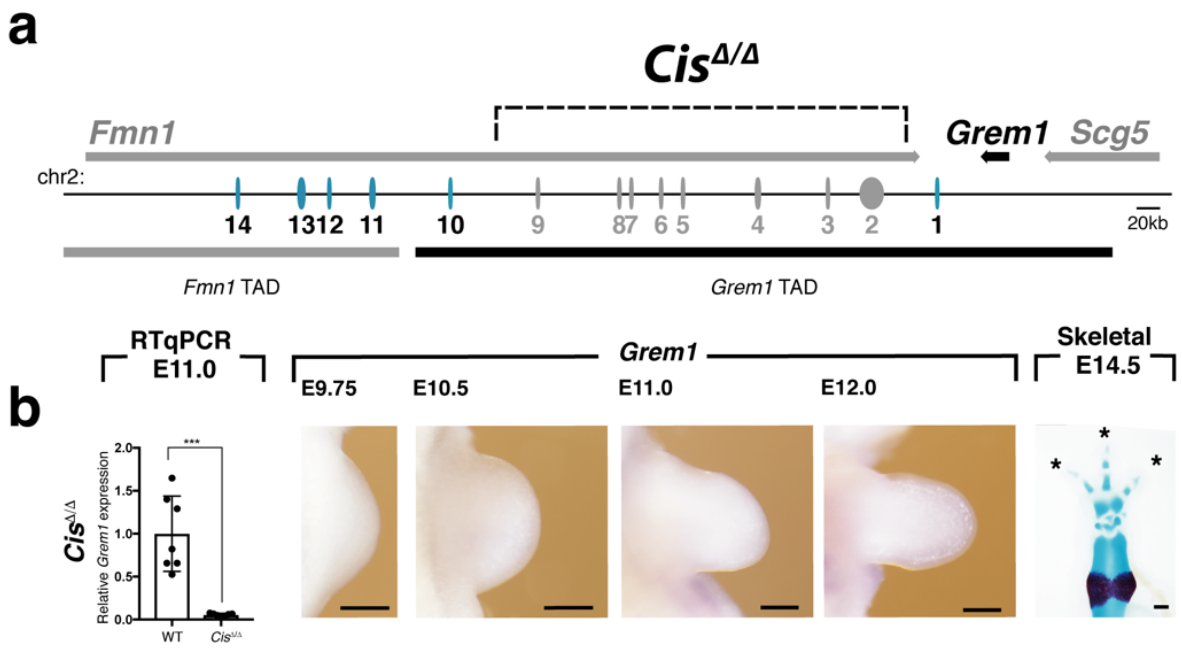


Figure 22. Effect of the $Cis^{\Delta/\Delta}$ mutation on *Grem1* transcription and limb development. (a) schematic representation of the *Cis* deletion, dashed lines highlight the *Cis* region (b) qPCR quantification of *Grem1* expression level at E11.0 (40ss to 42ss, n=7) ***P=0.0006 (n=7 littermates, two-tailed Mann-Whitney test), WISH of forelimbs (n>3) from E9.75 (26ss to 28ss), E10.5 (34ss to 36ss), E11.0 (40ss to 42ss) to E12.0 (50ss to 53ss), Skeletal stainings of E14.5 embryos. *Scale bars: 500µm (skeletal stainings), 250µm (WISH)*

The deletion of the *Cis* region did not include the *Fmn1* TAD either CRM1 or CRM10 enhancers. I had tested *Grem1* promoter activity in *lacZ* assays: it was unable to drive reporter expression in the limb on its own (Appendixes Figure 2). *Grem1* transcription needed enhancers to be activated. Nevertheless, there was no residual *Grem1* expression in this mutant (Figure 22). These observations led us to consider the remaining enhancers as non-*Grem1* associated and to remove CRM1 and CRM10 to

CRM14 from our analysis. All the CRMs able to drive reporter expression in a pattern similar to *Grem1* in the forelimb bud during mouse embryonic development were comprised within the *Cis* region.

CRM2 to CRM7 were the active limb enhancers that fulfilled all our criteria to be associated with *Grem1* expression. Based on their pattern of activity, CRM2 and CRM5 were the enhancers recapitulating the *Grem1* expression pattern the closest.

6.1.3.2. CRMs present in the *Fmn1* TAD do not seem associated to *Grem1* expression

All the CRM-like regions identified by chromatin profiling in the *Fmn1* TAD exhibited limb enhancer activity (Figure 23). The putative CRMs regions in the *Fmn1* loop domain segregated into two categories of limb enhancers. The first category grouped CRM11 and CRM13 with weak limb enhancer activity (<50% limb expression rate) and no reproducible pattern (Figure 23b-d). The second category included CRM12 and CRM14 and had a high activity rate (>50% limb expression rate, Figure 23c-e, Appendixes table 1). The LACZ protein accumulated in the AER in a reproducible fashion, 67% (n=4/6) and 91.7% (n=11/12) for CRM12 and CRM14 transgenic limb buds respectively. This pattern mimicked *Fmn1* endogenous expression in the AER (Zuniga et al., 2004, Figure 15, Appendixes table 1). Since *Grem1* is not expressed in the AER, these results suggest that 1) CRM12 and CRM14 were associated with *Fmn1* AER expression and not *Grem1* or *Fmn1* mesenchymal expression, 2) CRM11 and CRM13 were weak limb enhancers, not associated with *Grem1* in a wild-type genetic background. CRM12 was an alternative *Fmn1* promoter (Jackson-Grusby et al., 1992) with a distinctive enhancer potential.

The lack of limb mesenchymal *lacZ* activity together with the presence of activity in the AER suggest that CRM11-14 do not regulate *Grem1* expression during limb bud development.

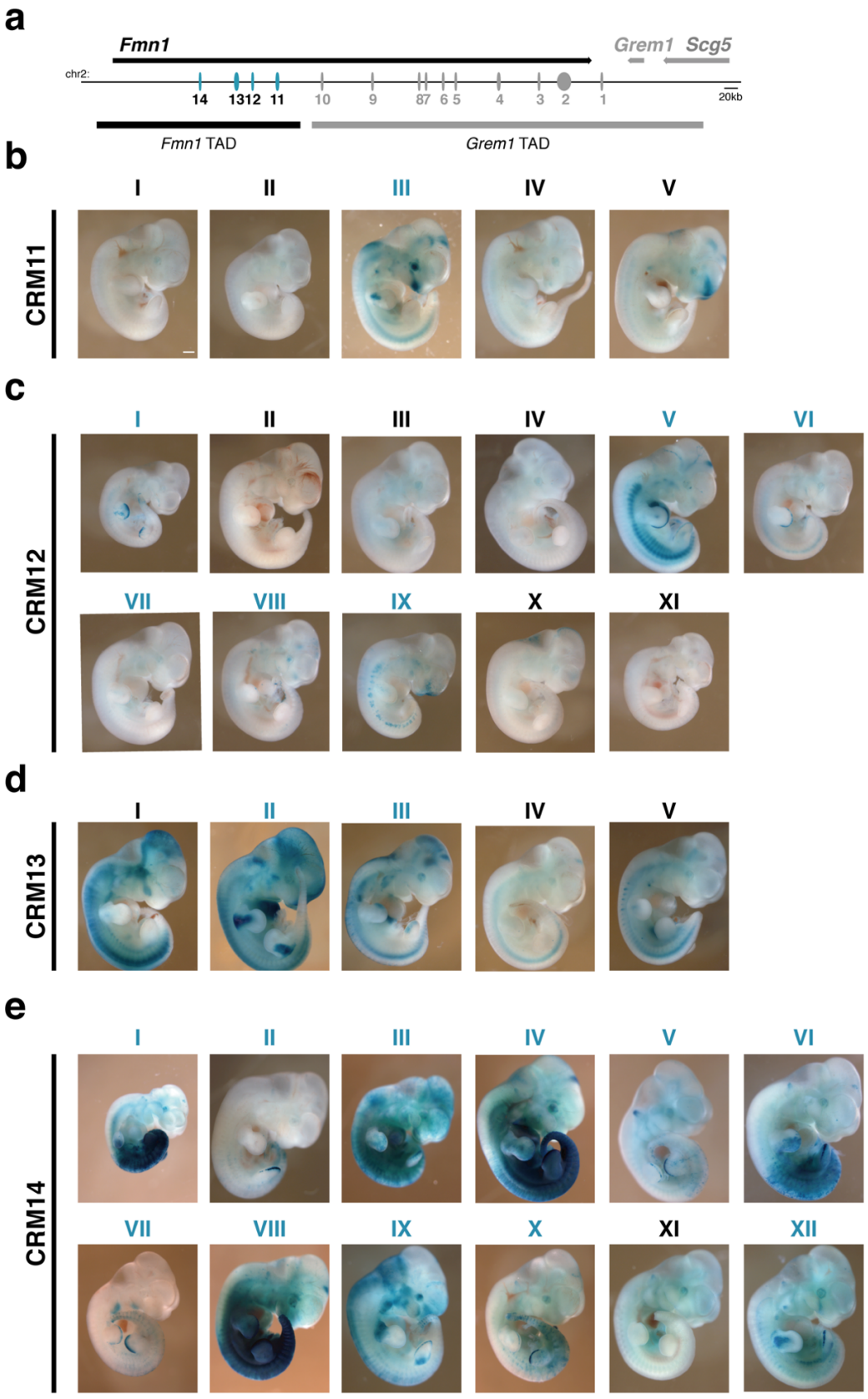


Figure 23. CRM11-14 enhancer activity in transgenic mouse embryos. (a) Schematic map of the *Grem1-Fmn1* locus (b-e) CRM11-14 transgenic embryos collected at E11.5. Transgenic embryos labeled with blue roman numerals show limb enhancer activity. *Scale bar: 500μm*

By combining TF footprints, epigenetic chromatin modifications and reporter assays, I identified and tested the enhancer properties of the CRMs of interest. Further analysis was needed to assess each CRM role(s) *in vivo*. Besides, reporter assay did not discriminate for any other *cis*-regulatory activity. CRM8 and CRM9 could potentially have regulatory properties undisclosed by this analysis.

6.2. Genetic dissection of the *Grem1* regulatory landscape

To identify *the in vivo* function of the CRMs of interest, I used CRISPR/*Cas9* technology to delete CRMs and addressed their functional relevance for the endogenous *Grem1* transcriptional regulation.

6.2.1. Disclosing Enhancer Cluster 1 (EC1) as a major regulator of *Grem1* transcriptional regulation.

Previous publication by the laboratory postulated that CRM2-4 belonged to a 70 kb GCR, necessary for the activation of *Grem1* expression in the posterior limb bud mesenchyme (Zuniga et al., 2004). I dissected this region (mm10 chr2:113,636,776-113,707,974) relabeled as enhancer cluster 1 (EC1). CRM4 was proposed to be a *Grem1* repressor in the anterior part of the limb at E10.5 (Li et al., 2014). I proceeded to analyze the deletions of CRM2 and CRM3 individually in mice, in order to clarify their respective properties within the previously described GCR (Material and Methods Table 9.5.3). RT-qPCR at E11.0 showed that the deletion of CRM2 reduced *Grem1* transcript count to ± 50% of its wild-type counterpart when *Grem1* expression was at its peak and formed a crescent shape. I performed WISH experiments to assess whether endogenous *Grem1* expression was altered by enhancer deletions. The onset of *Grem1* expression occurs at around E9.5 and is quite variable at this stage. It was therefore difficult to assess whether expression was changed in mutants (Appendixes Figure 3). At E10.5 (35-37ss), *Grem1* expression domain in $CRM2^{\Delta/\Delta}$ mutant limb buds were confined to the most distal posterior part of the limb mesenchyme (Figure 24b,c).

Previous analyses by the laboratory had quantified an approximative 50% of the *Grem1* expression levels reduction at this stage (data not shown). This observation suggested an essential role for CRM2 during the expansion phase of *Grem1*-expressing domain in the anterior distal part of the bud at E10.5. *Grem1* territory then expanded in *CRM2*$^{\Delta/\Delta}$ forelimbs and caught up with the wild-type embryos signal in a thinner crescent shape (Appendixes Figure 4, Figure 24c). At E11.0 (40-42ss), the *Grem1* crescent was narrower than what was observed in the wild-type limb buds; the proximal and anterior areas of *Grem1* domains appeared slightly truncated (Figure 24b,c). Surprisingly, E12.0 forelimb buds showed a premature termination of *Grem1* transcription. Homozygous *CRM2*$^{\Delta/\Delta}$ embryos developed normal limb skeleton despite having a strongly restricted *Grem1* domain at E10.5, a thinner domain of expression at E11.0 and an early termination at E12.0 (Figure 24b,c). Thus, *Grem1* expression at E12.0 did not seem essential for limb bud development. CRM2 enhancer appeared to be essential for proper *Grem1* transcription in terms of levels, spatial and temporal regulations. CRM3 loss did not significantly impact *Grem1* transcription levels and WISH did not reveal noticeable changes of spatial or temporal localization of *Grem1* transcript except at E11.0 where a wider distance between the *Grem1* signal and the AER region was noticed (Figure 24b-d). Limb skeletons were normal at E14.5 (Figure 24d). This suggested that CRM3 could be involved in the activation of *Grem1* in the limb bud distal domain at E11.0 but this role was not essential for forelimb development, in absence of other perturbation(s).

Individually, our single deletions did not affect the essential aspects of *Grem1* transcriptional regulation. There were neither sufficient to fully inactivate *Grem1* activity nor to hinder limb development on their own.

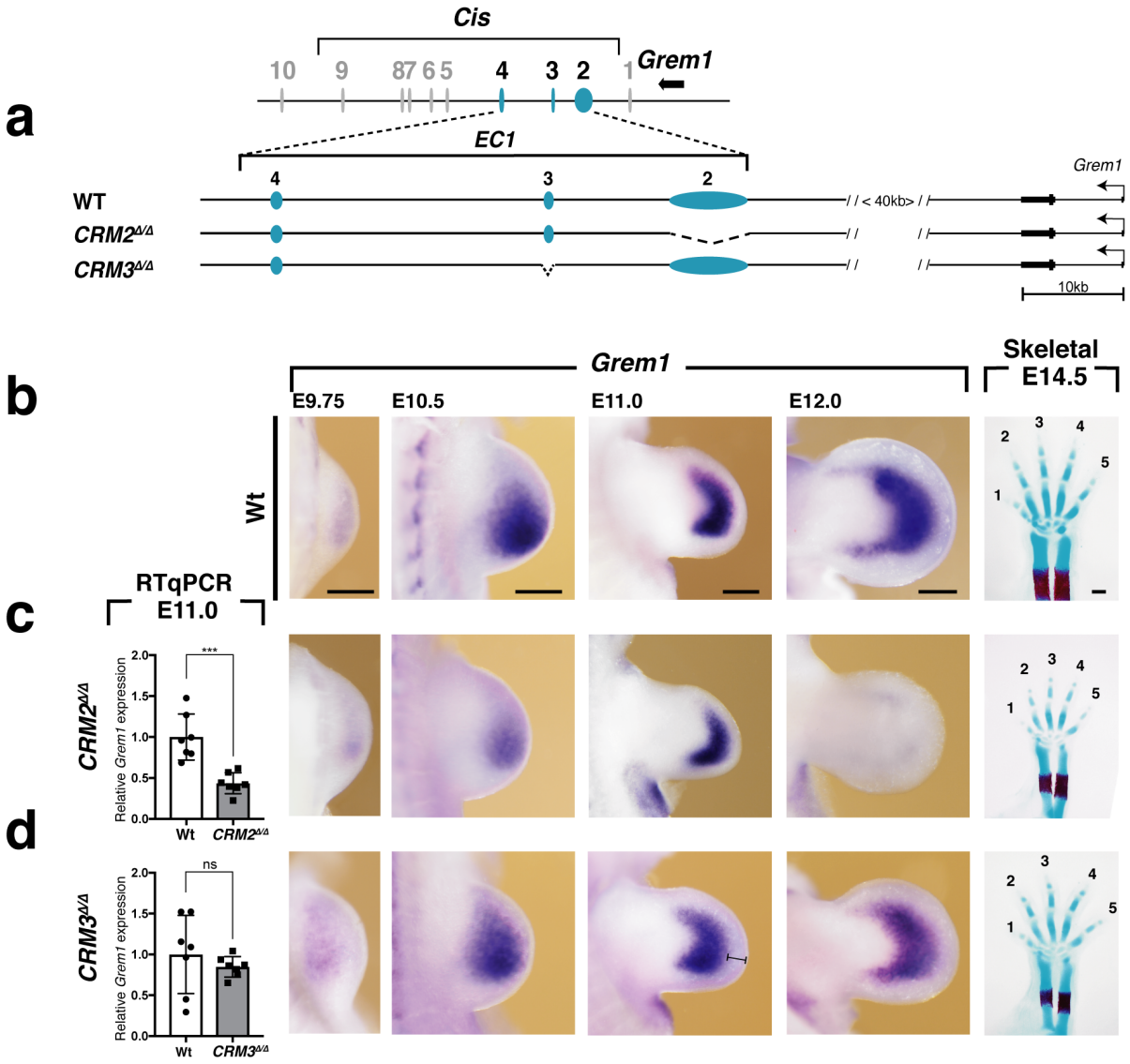


Figure 24. Genetic dissection of EC1. (a) Scheme of the *Cis* locus, active EC1 enhancers CRM2, CRM3 and CRM4 are represented in blue ovals. Dashed lines highlight deleted regions. (b-d) Comparison of *Grem1* expression levels of (b) WT, (c) *CRM2*$^{\Delta/\Delta}$ and (d) *CRM3*$^{\Delta/\Delta}$ at E11.0 ***P=0.0006 (n=7 littermates, two-tailed Mann-Whitney test). WISH of forelimbs (n≥3) from E9.75 (26ss to 28ss), E10.5 (34ss to 36ss), E11.0 (40ss to 42ss) to E12.0 (50ss to 53ss), brackets in (d) indicates the in anterior limb buds *Grem1*-free domain and of E14.5 embryos skeletal stainings. *Scale bars: 500µm (skeletal stainings), 250µm (WISH)*

To further understand the functional relevance of EC1, I next analyzed the double mutants for *CRM2CRM3* and for *CRM2CRM4* deletions (Figure 25a, Material and Methods section 9.1.2 and Table 9.5.3). I expected to see changes in *Grem1 expression* that would at least recapitulate the *CRM2*$^{\Delta/\Delta}$ phenotype (Figure 25b,c). I quantified *Grem1* expression levels by RT-qPCR at E11.0 and measured a 65%

decrease for *CRM2CRM3*$^{\Delta/\Delta}$ (Figure 25c) and a 50% decrease for *CRM2CRM4*$^{\Delta/\Delta}$ (Figure 25d) of *Grem1* transcriptional levels. The *Grem1* expression domain for these two mutants was similarly restricted distally and posteriorly (Figure 25c,d). *CRM2CRM3*$^{\Delta/\Delta}$ and *CRM2CRM4*$^{\Delta/\Delta}$ presented a much more restricted version of the crescent shape pattern of the wild-type or the *CRM2*$^{\Delta/\Delta}$ mutants (Figure 25b-d, Figure 24c). Both mutants displayed a distal and anterior loss of *Grem1* expression, suggesting the involvement of these two elements in *Grem1* expression in those domains at E11.0. Thus, combined enhancer activities of CRM3 and CRM4 at E11.0 seemed essential for *Grem1* anterior expansion at E11.0. Individually, CRM3 and CRM4 did not display a direct enhancer activity toward *Grem1* transcriptional levels. Nevertheless, they seemed to have major roles in refining the *Grem1* territory in its anterior and distal expression domains. *Grem1* transcripts from *CRM2CRM4*$^{\Delta/\Delta}$ mutants seemed, in the anterior part, to occupy a wider area than in the *CRM2CRM3*$^{\Delta/\Delta}$ homozygous mutants (white arrows Figure 25d). These results could corroborate the reported role of CRM4 as a *Grem1* limb anterior repressor at E11.0 (Li et al., 2014). The combined loss of CRM2 and CRM4 did not trigger an obvious anteriorization of its expression domain such as described by Li and colleagues at E10.5 in *CRM4*$^{\Delta/\Delta}$ mutants (Appendixes Figure 5) but reproduced the reduction of expression domain obtained by deleting CRM2 alone (Appendixes Figure 5, Figure 24c). This observation suggested 1) a critical and unique role of CRM2 for boosting the expression of *Grem1* at E10.5, 2) that CRM2 and CRM4, at least, cooperated to shape the anterior domain of *Grem1* expression at this stage but, 3) that at E11.0, different mechanisms involving CRM2, CRM3 and CRM4, at least, were set up to regulate *Grem1* expression.

Ectopic or restricted *Grem1* transcription at E10.5 did not appear to be critical for limb skeletal development, as long as *Grem1* was indeed expressed in the posterior distal part of the bud. Skeletal staining gave mild phenotypes as *CRM2CRM3*$^{\Delta/\Delta}$ and *CRM2CRM4*$^{\Delta/\Delta}$ showed variable severity of soft tissue fusion between digit 2 and 3 (Figure 25c,d, Appendixes Table 1) Our observations suggest that the E11.0 developmental stage represented a crucial time window for limb development. Those deletions were not sufficient to severely challenge limb bud development (no syndactyly was observed in forelimb skeletons) and so I considered the possibility that

the whole EC1 was required for *Grem1* transcriptional regulation. Therefore, *EC1*$^{\Delta/\Delta}$ mice were generated and analyzed.

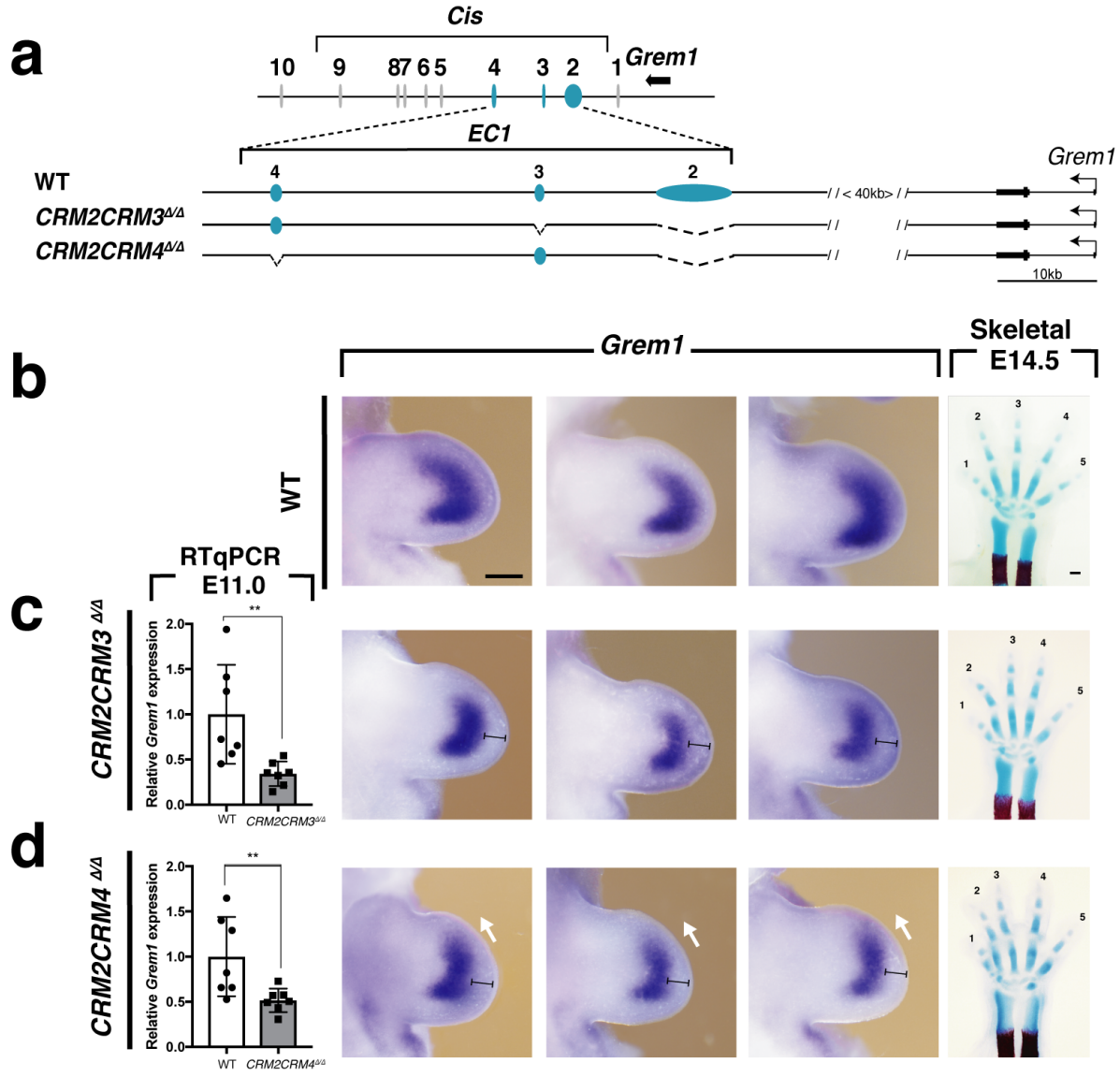


Figure 25. *Grem1* WISH at E11.0 for *CRM2CRM3*$^{\Delta/\Delta}$ and *CRM2CRM4*$^{\Delta/\Delta}$ double mutants. (a) schematic of the *Cis* locus, active EC1 enhancers CRM2, CRM3 and CRM4 are represented in blue ovals. Dashed lines highlight deleted regions. (b-d) Comparison of *Grem1* expression levels of (b) WT, (c) *CRM2CRM3*$^{\Delta/\Delta}$ and (d) *CRM2CRM4*$^{\Delta/\Delta}$ at E11.0 **P<0.01 (n=7 littermates, two-tailed Mann-Whitney test), brackets in (c-d) indicate the in anterior limb buds *Grem1*-free domain, white arrows indicate distal-anterior *Grem1* expression, WISH of forelimbs at E11.0 (40ss to 42ss) and of E14.5 embryos skeletal stainings. *Scale bars: 500µm (skeletal stainings), 250µm (WISH)*

The deletion of the EC1 locus (Figure 26a) consisted of the first 70 kb of the *Cis* region, 3' of *Grem1* (Material and Methods Table 9.5.3). CRM2, CRM3, and CRM4 were deleted simultaneously in the $EC1^{\Delta/\Delta}$ mutants (Figure 26a). The homozygous deletion of *EC1* led to a ± 50% decrease in *Grem1* expression levels (Figure 26b,c), analogous to the previously measured reductions of *Grem1* expression in $CRM2^{\Delta/\Delta}$, $CRM2CRM3^{\Delta/\Delta}$ and $CRM2CRM4^{\Delta/\Delta}$ mutant handplates. The $EC1^{\Delta/\Delta}$ homozygous mutants WISH showed a comparable distribution of *Grem1* transcript to $CRM2^{\Delta/\Delta}$ mutants at E9.75, E10.5 and E12.0 where I saw an early termination of *Grem1* expression (Figure 24c, Figure 26c). Once again, the E11.0 stage stood out as the most critical stage for *Grem1* regulation of limb development as WISH patterns greatly diverge only at this stage between $CRM2^{\Delta/\Delta}$ and $EC1^{\Delta/\Delta}$ (Figure 26b). $EC1^{\Delta/\Delta}$ combined phenotypes of both $CRM2CRM3^{\Delta/\Delta}$ and $CRM2CRM4^{\Delta/\Delta}$ in an additive manner with a narrow *Grem1* crescent shape and the slight anterior expansion described earlier at E11.0 for $CRM2CRM4^{\Delta/\Delta}$ (Figure 25d). Interestingly, I observed that the limbs adopted a trapezoidal shape at E12.0, a precursor sign of limb deformity phenotypes: the reduction of *Grem1* domain at E10.5 and the early termination of *Grem1* only combined with significant changes of *Grem1* domain at E11.0 were associated with limb developmental defects. During skeletal staining process at E14.5, I observed 100% penetrance of soft tissue fusions (digits 2 to 3, also observed in the $CRM2CRM3^{\Delta/\Delta}$ and $CRM2CRM4^{\Delta/\Delta}$ mutants, Figure 25c,d and Figure 26c) and even more compelling for our study, 20% of them had a C-L type syndactyly (Appendixes Table 1, Figure 26c). *EC1* loss caused a more severe skeletal phenotype at E14.5 than the previously described double mutants $CRM2CRM3^{\Delta/\Delta}$ or $CRM2CRM4^{\Delta/\Delta}$. By narrowing down the *Cis* region and deleting EC1, *Grem1* transcription was significantly affected and this challenged *Grem1* regulatory mechanisms beyond their compensatory capacities, resulting in mild *ld* phenotypes.

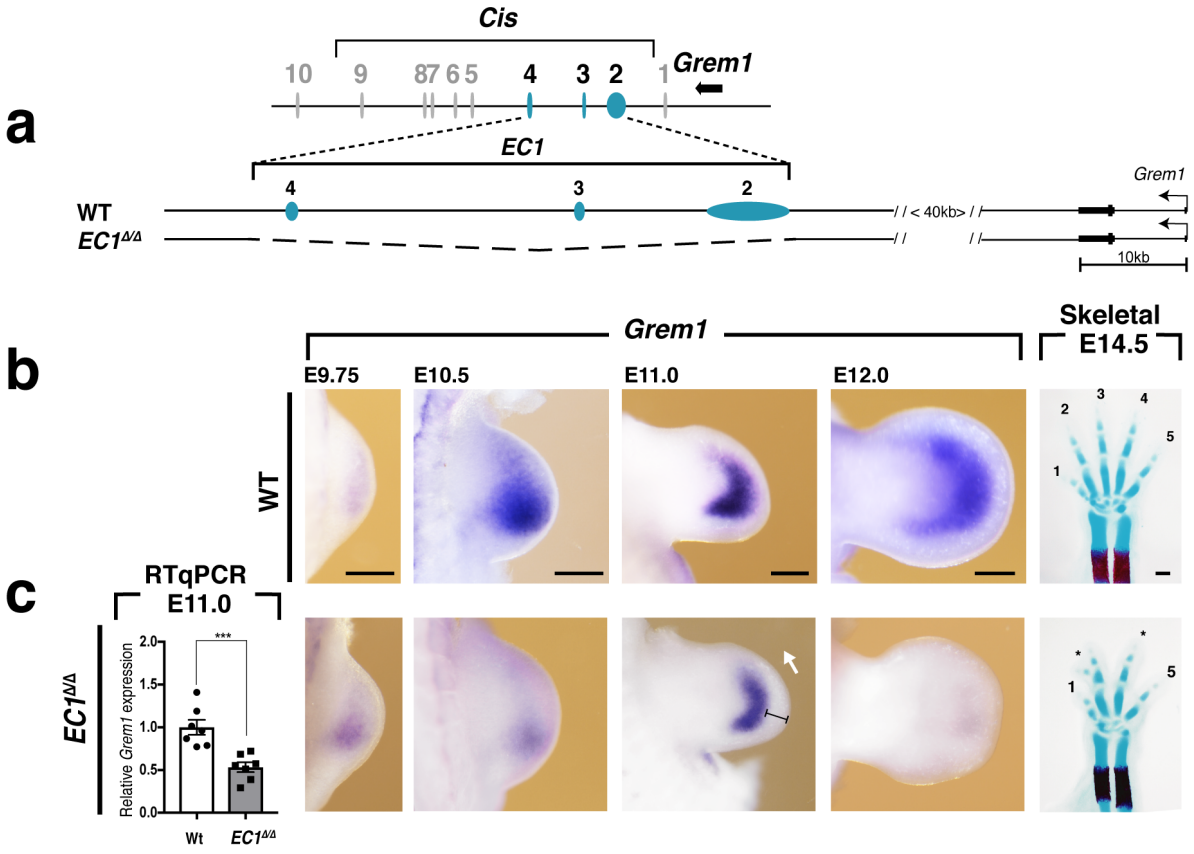


Figure 26. Analysis of *EC1* homozygous mutant. (a) Schematic of the *Cis* locus, active EC1 enhancers CRM2, CRM3 and CRM4 are represented in blue ovals. Dashed lines highlight deleted regions. (b-c) Comparison of *Grem1* expression levels of (b) WT and (c) $EC1^{\Delta/\Delta}$ at E11.0 ***P=0.0006 (n=7 littermate, two-tailed Mann-Whitney test), WISH (n≥3) from E9.75 (26ss to 28ss), E10.5 (34ss to 36ss), E11.0 (40ss to 42ss) to E12.0 (50ss to 53ss), brackets in (d) indicates the in anterior limb buds *Grem1*-free domain and of E14.5 embryos skeletal stainings. *Scale bars: 500µm (skeletal stainings), 250µm (WISH)*

These results uncover CRM2 as a preeminent *Grem1* enhancer. CRM2 is part of a system that regulated the levels and the termination of *Grem1* transcription. I observed that CRM2 worked in combination with CRM3 and CRM4 to adjust the spatial *Grem1* expression within the EC1 unit. EC1 combined the transcriptional spatial, temporal and levels enhancer properties of CRM2, CRM3 and CRM4 and was necessary for *Grem1* transcriptional regulation. Deletion of EC1 affected limb development but was not sufficient to reproduce *Grem1* LOF phenotype, suggesting that other elements were responsible for controlling *Grem1* expression outside of EC1.

6.2.2. EC2 contains all the remaining *Grem1* CRMs necessary for its expression in the limb bud

Having previously identified CRM5, CRM6 and CRM7 as active limb enhancers, I tested their requirement for *Grem1* expression. From our reporter assays, I identified CRM5 as the second most promising enhancer after CRM2. To test this hypothesis, I generated the single deletion of CRM5 and removed CRM5, CRM6 and CRM7 together as the 35 kb EC2 (Figure 27a, Material and Methods table 9.5.3). Although CRM8 was not an active limb enhancer, it was also included in EC2 as it was located only 4 kb away from CRM7.

The loss of the CRM5 enhancer led to a 30% decline in *Grem1* expression levels (Figure 27b,c). At E11.0 I detected a slight reduction in the distal region of the *Grem1* domain (Figure 27c, black brackets). However, overall the crescent shape was maintained in the *CRM5*$^{\Delta/\Delta}$ limb buds. The other stages did not differ from the wild-type profile by WISH (Figure 27c). In comparison, the deletion of EC2 resulted in a 50% reduction of *Grem1* transcriptional levels at E11.0 (Figure 27d). Nevertheless, *Grem1* spatial spread only seemed affected at E11.0 when the distal domain of its expression appeared slightly truncated in comparison to the previously described enhancer mutants with distal reduction of the *Grem1* domain (Figure 27d, black brackets). The *Grem1* spatio-temporal expression was not affected at other stages. In summary, a 50% decrease of *Grem1* and a distal and possibly anterior truncation of its domain at E11.0 upon the loss of EC2 did not affect limb bud development as no limb phenotype was observed (Figure 27d). Either those regulatory elements were not essential for *Grem1* transcription during limb development or a compensatory mechanism occurring at the level of the limb self-regulatory feedback loops in these mutants balanced the engineered deletion. EC2 on its own was an active regulatory locus with enhancer activity associated with *Grem1* transcriptional regulation. Despite a 50% drop in *Grem1* expression, no phenotype was observed. EC2 was less critical than EC1 for *Grem1* transcriptional regulation during limb bud development.

To test if other enhancers might play a compensatory role, I decided to retarget the *EC1*$^{\Delta/\Delta}$ mouse line and remove either CRM5 or EC2 in *cis*. In parallel, the *CRM2*$^{\Delta/\Delta}$

line was retargeted to generate CRM2 and CRM5 double mutants, as these two elements on their own seemed to be responsible for the most important fraction of *Grem1* transcriptional levels. At the transcriptional level, *CRM2CRM5*$^{\Delta/\Delta}$ showed additive behavior as I measured an 80% decrease in *Grem1* transcripts at E11.0 (*CRM2*$^{\Delta/\Delta}$ 50% added to CRM5 30% loss in *Grem1* transcription levels, Figure 27e). The spatial and temporal distribution of the *Grem1* signal reproduced the pattern observed in the CRM2 single deletion and these genetic alterations did not cause any limb skeleton deformity (Figure 27e). While CRM2 and CRM5 were the only enhancers able to drive a reporter expression recapitulating endogenous *Grem1* transcript localization, this double mutant still displayed a posterior *Grem1* domain at E10.5 that expanded distally and anteriorly at E11.0. In contrast, the *EC1CRM5*$^{\Delta/\Delta}$ double mutants showed novel changes in *Grem1* expression (Figure 27f). Expression levels at E11.0 were reduced by 80%, just like in the *CRM2CRM5*$^{\Delta/\Delta}$ mouse line (Figure 27e,f). Unexpectedly, at E10.5, I observed a slight ectopic expansion of the *Grem1* domain in the limb bud that was not seen in the previously described mutant limb buds (Figure 27f, black and white arrows). This early expansion preceded a significant reduction of *Grem1* territory in width at E11.0, while the signal was also detected in ectopic anterior domains (Figure 27f, black and white arrows). The *Grem1* expression levels in *EC1CRM5*$^{\Delta/\Delta}$ mutants at E11.0 were identical to the ones measured in the *CRM2CRM5*$^{\Delta/\Delta}$ forelimbs (Figure 27e,f). CRM5 had an enhancer role to amplify *Grem1* transcription at E11.0 but the anteriorization of the WISH signal in *EC1CRM5*$^{\Delta/\Delta}$ mutants suggested that CRM5 might participate to repress *Grem1* transcriptional activity in the anterior limb bud. Additionally, the normal anterior distal *Grem1* domain in *EC1CRM5*$^{\Delta/\Delta}$ limbs was completely erased. As expected from any generated deletions coupled to the removal of CRM2, I recorded an early termination of *Grem1* expression at E12.0, associated with a trapezoidal limb shape (Figure 27f). The penetrance of mice displaying syndactyly (4 digits) in *EC1CRM5*$^{\Delta/\Delta}$ background was 100% (Appendixes table 1). Those digits were organized so that the handplates followed a bilateral symmetry around the AP axis. This limb deformity phenotype was stabilized by the combined deletions of EC1 and CRM5 and tipped the scale in favor of tetradactyly. I assumed, in light of our previous skeletal analysis, that digits 2 and 3

in this mutant had fused to generate a single digit however, in absence of lineage tracing, it is difficult to prove.

At this phase of the project, the *EC1CRM5*$^{\Delta/\Delta}$ mutants had reached the threshold under which the limb bud could not recover and compensate for the introduced genetic alterations of the *Grem1* expression. Nonetheless, in the *EC1CRM5*$^{\Delta/\Delta}$ double mutants, 20% of *Grem1* expression was maintained, most likely supported by the remaining CRMs in the landscape (e.g. CRM6 and CRM7). Since these remaining 20% may be supported by multiple CRMs together, such decreases, as a result of additional enhancer deletions, might not be detectable with RT-qPCR assays. Thus, I proceeded with the analyses of *EC1EC2*$^{\Delta/\Delta}$ mutants (Figure 27g). At E10.5, I observed a barely discernible signal in the posterior distal part of bud (Figure 27g). qPCR analysis measured at E11.0 a total loss of *Grem1* transcripts (3% left versus 5% for the *Cis*$^{\Delta/\Delta}$, Figure 22b and Figure 27g). WISH could detect residual *Grem1* signal at E11.0 but in those mutants, later at E12.0, the absence of transcripts was observed in a fin-shaped limb bud, characteristic of the severe limb deformity phenotypes of *Grem1*$^{\Delta/\Delta}$ and *Cis*$^{\Delta/\Delta}$ mutant embryos. Indeed, 100% of the *EC1EC2*$^{\Delta/\Delta}$ mice had a major limb deformity phenotype with an oligosyndactyly of 3 digits. The residual *Grem1* expression in *EC1EC2*$^{\Delta/\Delta}$ at E11.0 was not sufficient to maintain normal forelimb development, not even to maintain tetradactyly, which most likely led to the breakdown of the limb self-regulatory feedback loop that controlled limb bud outgrowth and patterning. This E11.0 residual *Grem1* signal could stem from the *Grem1* proximal promoter under the control of the only CRMs left outside of EC1 and EC2: CRM1 and CRM10. This *Grem1* expression activated by the remaining CRMs in the *EC1EC2* double mutant limb buds, the intensity and/or spatial spread were not sufficient to prevent the *ld* phenotype.

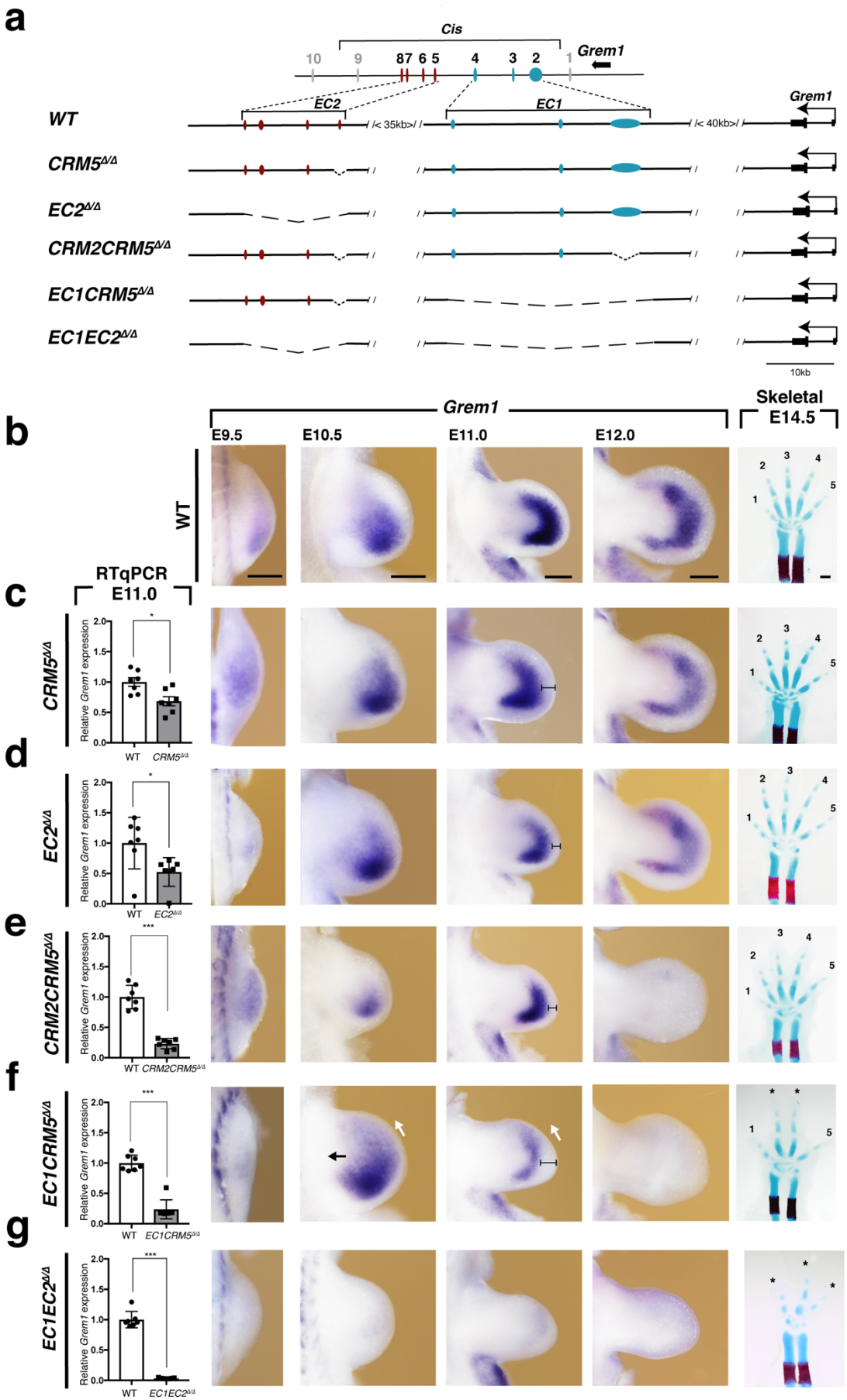

a
Cis
10
9
87 6 5
4
3
2
1
Grem1
EC2
EC1
Grem1
WT
/< 35kb>/
/< 40kb>/
CRM5Δ/Δ
EC2Δ/Δ
CRM2CRM5Δ/Δ
EC1CRM5Δ/Δ
EC1EC2Δ/Δ
10kb
b
Grem1
Skeletal
E14.5
E9.5
E10.5
E11.0
E12.0
WT
RTqPCR
E11.0
c
CRM5Δ/Δ
d
EC2Δ/Δ
e
CRM2CRM5Δ/Δ
f
EC1CRM5Δ/Δ
g
EC1EC2Δ/Δ
Relative Grem1 expression
WT
CRM5Δ/Δ
EC2Δ/Δ
CRM2CRM5Δ/Δ
EC1CRM5Δ/Δ
EC1EC2Δ/Δ

Figure 27. Genetic dissection of the *Cis* region. (a) Scheme of the *Cis* locus, active EC1 enhancers are represented in blue ovals, CRMs of EC2 are represented in maroon ovals. Dashed lines highlight deleted regions. (b-g) Comparison of *Grem1 expression* levels of (b) WT, (c) $CRM5^{\Delta/\Delta}$, (d) $EC2^{\Delta/\Delta}$, (e) $CRM2CRM5^{\Delta/\Delta}$, (f) $EC1CRM5^{\Delta/\Delta}$ and (g) $EC1EC2^{\Delta/\Delta}$ at E11.0 ***P=0.0006, *P≤0.0175 (n=7 littermates, two-tailed Mann-Whitney test). WISH of forelimbs (n≥3) from E9.75 (26ss to 28ss), E10.5 (34ss to 36ss), E11.0 (40ss to 42ss) to E12.0 (50ss to 53ss), brackets indicate the in anterior limb buds abnormal *Grem1*-free domain and arrows highlight the expansion of *Grem1* expression domains and E14.5 embryos skeletal stainings. *Scale bars: 500µm (skeletal stainings), 250µm (WISH)*

EC1 and EC2 contain all the structural and transcriptional modules necessary for *Grem1* expression and normal limb bud development. With these results I established that there was a hierarchy between those clusters. EC1 alone was sufficient to maintain *Grem1* expression in terms of time, space and intensity that are required for normal limb development in mice. The most important elements for *Grem1* spatial and temporal transcription were located within EC1. However, EC1 and EC2, in a cooperative and additive manner, controlled *Grem1* expression levels. Upon the removal of EC1, EC2 on its own was not sufficient to maintain proper limb development. The combined loss of EC1 and CRM5 preferentially resulted in tetradactyly. It appeared that the most crucial stage for the activity of *Grem1* enhancers was E11.0.

6.3. The 3D chromatin structure brings new insight into the transcriptional regulation of *Grem1* by its enhancer clusters

I next decided to address the roles of 3D chromatin architecture on the *Grem1* transcriptional regulation. Chromatin interactions are essential for many aspects of transcriptional regulation and long linear distances do not reflect close physical proximity. Using 4C-seq (Simonis et al., 2006), I was able to visualize the *Grem1* regulatory landscape interactions with its promoter as the viewpoint in wild-type and selected mutant backgrounds (Figure 28a, brackets). Following the principle of proximity ligation, 4C-seq allowed me to pinpoint the physical contacts of the *Grem1*

promoter at a kilobase resolution. This method uncovered the frequency of interactions of the *Grem1* promoter with the *Grem1* TAD was much higher than with the *Fmn1* TAD, corroborating the analysis of Hi-C seq profiles, and CTCF location and orientation (Figure 28, see section 6.1.1).

Having established a *Grem1* promoter interaction map, I could investigate whether deletions of specific CRMs or CRM clusters affected the genomic architecture of the landscape. I compared the interactions of the *Grem1* promoter in wild-type and $Cis^{\Delta/\Delta}$ mice limb buds. The results established the physical interactions with the *Grem1* promoter were drastically decreased in the full *cis*-regulatory landscape (Figure 28b). The relatively infrequent interaction between the *Grem1* promoter and the *Fmn1* TAD remained with low frequency, despite a 170 kb shorter linear distance after the *Cis* deletion. There was no recruitment of CRM outside of the *Cis* region while *Grem1* interaction with its TAD boundaries had greatly decreased. Besides, there were no regions able to contact *Grem1* promoter *de novo* (Figure 28b). Our results suggested no regulatory element outside of this region was able to enhance *Grem1* expression: all relevant interactions occurred within the *Grem1* TAD (Figure 28b).

In contrast, analysis of 4C-seq profiles for $CRM2^{\Delta/\Delta}$ mutant limb buds revealed only mild interaction changes along the TADs (Figure 28c). In particular, I identified interesting short-range changes. I observed *de novo* contacts with CRM3 and CRM4, nonexistent in the wild-type samples, in combination with the higher frequency of interactions with CRM5 (Figure 28c). Contacts with CRM6 were unchanged and contacts with CRM7 seemed to be lower but still happened at high frequency (Figure 28c). In $CRM2^{\Delta/\Delta}$ mutants the *Grem1* expression domain could be maintained by CRM3 to CRM7. No CRM was able to compensate for the spatial and temporal changes in *Grem1* expression observed with the loss of CRM2 (Figure 24c).

$EC1^{\Delta/\Delta}$ 4C-seq profiles revealed contact changes mainly within *Grem1* TAD (Figure 28d and black bar at the bottom of the figure). I could clearly distinguish a local increase in interactions frequency within the *Grem1* TAD. The EC2 cluster established more frequent contacts with the *Grem1* promoter. Those enhanced interactions could be enhancing the transcriptional levels of the targeted gene *Grem1* to compensate for the loss of EC1. However, none of the identified CRMs were able to mimic a wild-type-like *Grem1* spatial and temporal distribution in the absence of EC1. This suggested

that in the wild-type environment, EC2 could actually be responsible for less than 50% of *Grem1* transcriptional levels. Its enhancer functions could be increased through the increased frequency of contacts established in *EC1*$^{\Delta/\Delta}$ mutants. Those profiles did not show *de novo* interactions in or outside of the *Grem1* TAD in *EC1*$^{\Delta/\Delta}$ forelimbs at E11.0 (Figure 28d).

The most surprising results was the 4C-seq for *EC2*$^{\Delta/\Delta}$ (Figure 28e). I had expected high contact frequency with EC1 and in particular with CRM2, as *Grem1* spatial and temporal expression was not affected in those mutants. So far, I had attributed the decrease of the *Grem1* transcriptional levels to the loss of EC2 enhancer functions only. However, with EC2 loss, a local decrease of contacts within the *Grem1* TAD was detected; this reduction of contact frequency also affected EC1 (Figure 28e). Thus, the loss of enhancer activity measured with EC2 homozygous deletion could also reflect deficient EC1/*Grem1* promoter interactions. As I assumed CRM5 to be in charge of approximately 30% of *Grem1* transcriptional expression, I suggested the rest of EC2 enhancers activity (50-30= 20%) could rely not only on CRM6 or CRM7 but also on promoting EC1 interactions with *Grem1* locus (Figure 28e). EC2 seemed to be essential to ensure high contact frequency between *Grem1* and EC1, suggesting a structural role of this region. CRM2 and CRM3 did not interact with the promoter in those mutants.

The *Cis* region spans 170 kb from exon 10 to exon 24 of *Fmn1* and EC1 encompasses the 70 kb between exon 19 and 24 of *Fmn1.* The 30 kb EC2 is located within the remaining 100 kb region between exon 10 and exon 19 of *Fmn1*. Therefore, I asked the question if EC2 deletion and the loss of the region between exon 10 and exon 19 of *Fmn1,* called the F1Δ1019 region from now on, would have redundant or overlapping effects or if I would be seeing any noticeable differences. I engineered the *F1Δ1019* deletion mutants and some striking differences with *EC2*$^{\Delta/\Delta}$ profile were observed in the 4C-seq analysis (Figure 28e-f). The frequency of contacts with EC1 did not decrease in those mutants, on the opposite, *F1Δ1019* homozygous mutants displayed an increase in contact frequencies with EC1's CRMs (Figure 28f). The *Grem1* promoter interactions with CRM2 appeared to increase and once again, just as in the *CRM2*$^{\Delta/\Delta}$ mutants, I could distinguish that, while absent in the wild-type profiles, *de novo* contacts were initiated with CRM3 and CRM4 enhancers (Figure

28f). These mutants did not show any significant change in *Grem1* transcriptional levels and the spatial distribution of *Grem1* transcripts was similar to the wild-type limb buds at E11.0 (Figure 29b). The skeletal stainings of the *F1Δ1019* embryos forelimbs were normal (Figure 29b).

The physical proximity between the *Grem1* promoter and the *Grem1* TAD external CTCF boundaries remained strong, in all engineered mutants, despite being located from 50 kb to 275 kb away from them. Only in the *Cis*$^{\Delta/\Delta}$ mutants, those contacts appeared weaker than the wild-types (Figure 28b).

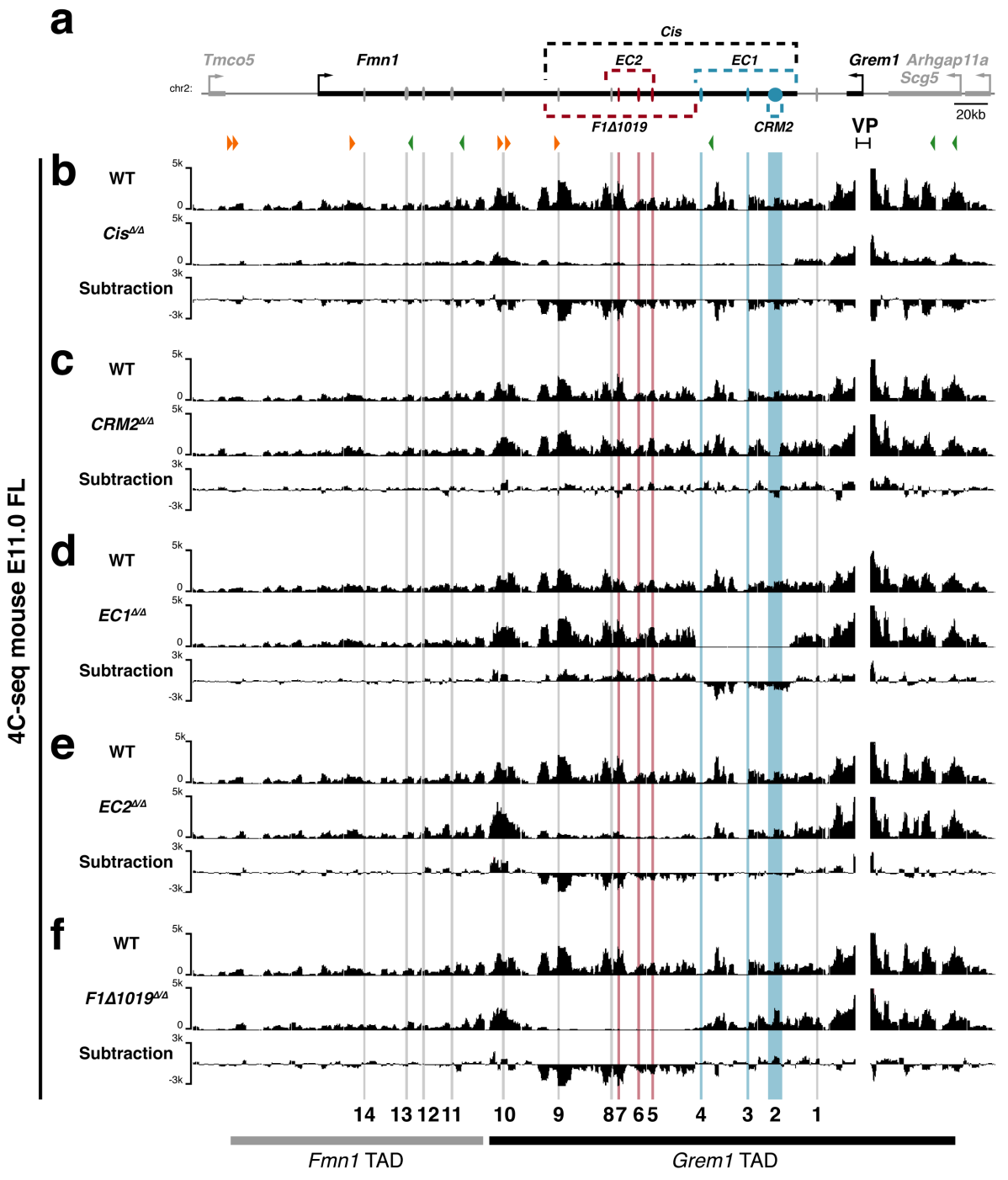


Figure 28. 4C-seq identifies various changes in chromatin configuration in wild-type and selected mutant alleles. (a) Schematic of the *Grem1-Fmn1* locus, active EC1 enhancers are represented in blue ovals, EC2 CRMs are represented in maroon ovals. Dashed lines highlight deleted regions and orange and green arrows represent CTCF binding sited and their orientation. (b-f) 4C-seq profiles of (b) *Cis*$^{\Delta/\Delta}$, (c) *CRM2*$^{\Delta/\Delta}$, (d) *EC1*$^{\Delta/\Delta}$, (e) *EC2*$^{\Delta/\Delta}$ and (f) *F1Δ1019*$^{\Delta/\Delta}$ mouse forelimb buds at E11.0 (40ss-42ss) in comparison to wild-type profiles. Vertical lines indicate the position of each CRM-like region and CRMs in the locus. Horizontal lines represent the *Fmn1* and the *Grem1* TAD. The viewpoint (VP) is highlighted with brackets.

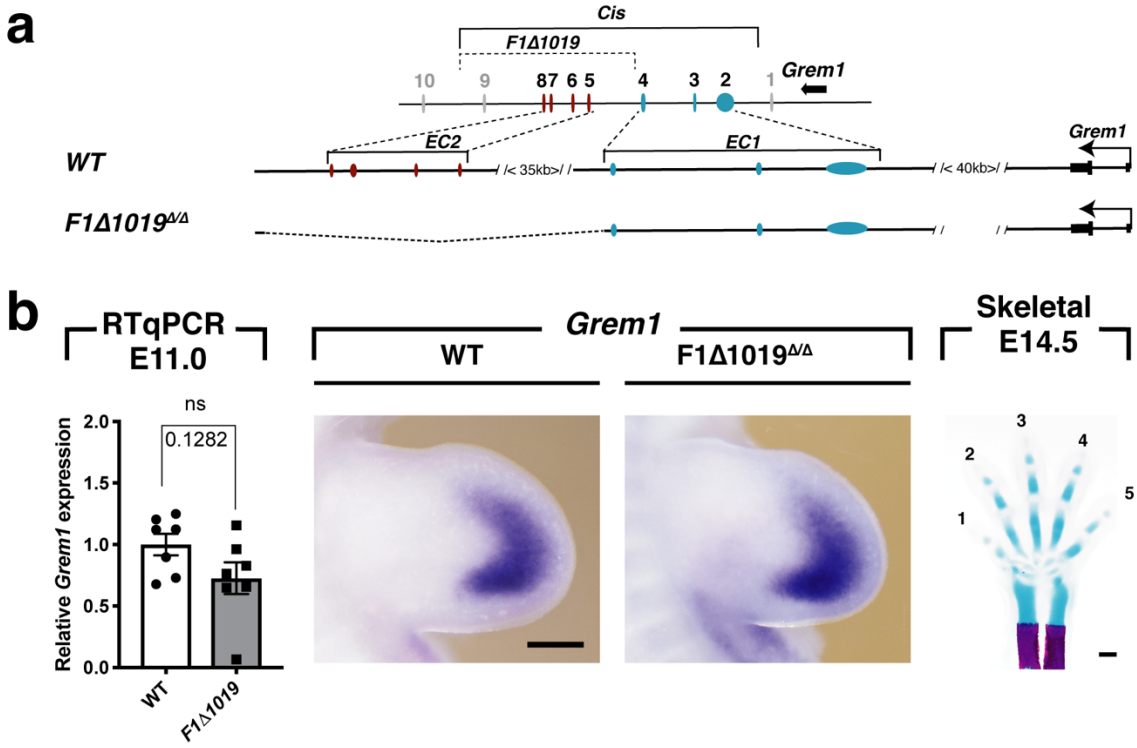


Figure 29. Genetic analysis of the *F1Δ1019* mutants shows no difference with wild-type. (a) schematic of the *Cis* locus highlighting the *FΔ1019* deletion (b) quantification of *Grem1* transcription levels at E11.0 in *F1Δ1019*$^{\Delta/\Delta}$ forelimb buds, comparison of *Grem1* expression domain at E11.0 between *F1Δ1019*$^{\Delta/\Delta}$ and wild-type forelimbs, skeletal staining of *F1Δ1019*$^{\Delta/\Delta}$ E14.5 forelimb. *Scale bars: 500µm (skeletal stainings), 250µm (WISH).*

I have identified changes in chromatin interactions with the *Grem1* promoter in selected mutants. The deletions not only caused modifications at the level of CRM-promoter contacts but also generated differences within the *Grem1* TAD, without affecting its boundaries. I observed that the promoter preferentially interacted within *Grem1* TAD CTCF boundaries in all genetic backgrounds. It appeared that transcriptional levels of *Grem1* could be associated with the presence or the absence of contacts and with subtle variations of the frequency of physical interactions between promoters and CRMs. In some cases, the increase of interactions and recruitment of CRMs to the promoter (*EC1*, *CRM2* and *F1Δ1019* mutants) were potentially parts of compensatory mechanisms. In other cases, the deletions seemed to affect the expression of the *Grem1* gene not only through the loss of CRM functions but also because of structural alteration such as a TAD-wide decrease of contacts with the promoter (*EC2* mutants) or the opposite (*EC1* mutants). Some CRMs regulated the expression of the target gene without interacting with its promoter directly (CRM3 and CRM4) but also appeared to be able to physically interact with the promoter in certain mutant contexts (*F1Δ1019*$^{\Delta/\Delta}$ and *CRM2*$^{\Delta/\Delta}$ mutants), most likely to modulate transcription. EC2 seemed to have an intrinsic structural role to maintain or enhance contacts between EC1 and the promoter.

6.4. Mechanisms underlying the phenotypic consequences of altering *Grem1 cis*-regulation

6.4.1. A glimpse into the dynamics of the limb self-regulatory feedback loop

6.4.1.1. *Bmp* and *Shh* pathways activities

Previous publications have highlighted the critical role of *Grem1* in the establishment and maintenance of the self-regulatory feedback loops controlling limb bud outgrowth and patterning (Khokha et al., 2003; Michos et al., 2004; Zuniga et al., 2004; Zuniga and Zeller, 1999). This present study described *Grem1* transcriptional regulation as critical for normal limb development. I had strong evidence that pentadactyly was in parts preserved due to 1) the spatial distribution of *Grem1* transcript and to 2) the transcriptional levels of *Grem1*. At E11.0, the *Grem1*-mediated antagonism of the BMP pathway activity in the limb mesenchyme is essential for handplate patterning and specification of digits 2 to 5 by the *Shh* pathway activity (Benazet et al., 2009). In this part of the study, I was curious about the consequences of *Grem1* transcriptional misregulation on the BMP pathway activity and how it could be correlated to the presence or absence of phenotypes. *Grem1* downregulation was supposed to result in an upregulation of the BMP pathways activity and indirectly in the repression of the SHH pathways. I proceeded to measure the possible changes in the BMP and the SHH pathway activities by measuring *Bmp4* and *Shh* transcription levels by RT-qPCR. I also quantified the expression of activity sensors *Id1* for the BMP pathway (Hollnagel et al., 1999; Korchynskyi and ten Dijke, 2002) and *Gli1* for the SHH pathway (Marigo et al., 1996; Ruiz i Altaba, 1998).

- **Mutants with little changes in *Grem1* in expression levels.**

The first tier I looked at consisted of mutants where at least 70% of *Grem1* expression remained. This category only concerned the *CRM5* enhancer deletion. In these mutants I observed a slight difference with the wild-type samples on *Grem1* WISH with a truncated distal domain at E11.0 (Figure 27c). This mouse line developed a wild-type limb. The loss of CRM5 barely affected the *Grem1* expression domain. Unsurprisingly, no significant effect was observed on the BMP or the SHH pathways (Figure 30).

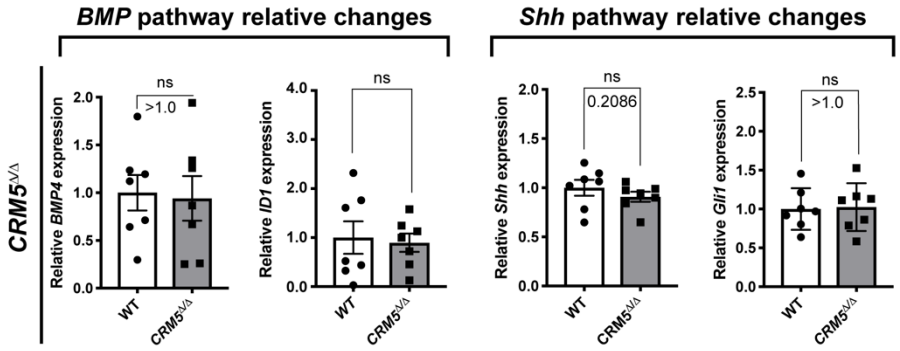


Figure 30. Rt-qPCR analysis of BMP and SHH pathway activity in *CRM5*$^{\Delta/\Delta}$ homozygous mutants at E11.0. WT and homozygous mutant handplates at E11.0 (40ss to 42ss) were dissected from littermate embryos, (n=7; mean ± SEM, two-tailed Mann-Whitney test), p-values are indicated on top.

- **Mutants with a 50% decrease in *Grem1* expression levels and no severe phenotype.**

The second tier gathered the genotypes in which approximately 50% of *Grem1* transcription was maintained. CRM2 homozygous deletion had a significant and negative effect on *Shh* transcription (only 40% of transcripts remained in comparison with wild-type littermates) but this change did not affect the *Shh* pathway activity as its downstream target *Gli1* transcription was unaffected. There was no detectible change in the BMP pathway (Figure 31a). The *CRM2CRM3*$^{\Delta/\Delta}$ forelimb buds did not showcase any significant alteration of *BMP* or *Shh* pathways, despite displaying soft-tissue fusion at later developmental stages (Figure 25c,d, Figure 31b). The *CRM2CRM4*$^{\Delta/\Delta}$ embryonic forelimbs did not present any change in the *Bmp* pathway while the *Shh* and *Gli1* levels were reduced to 30% and 80% respectively compared to wild-type. During the process of skeletal staining, these homozygous mutants revealed soft-tissue fusions comparable to the *CRM2CRM3*$^{\Delta/\Delta}$ observed handplate phenotypes (Figure 31c, data not shown). Interestingly, in *EC2*$^{\Delta/\Delta}$ forelimb buds, *Bmp4* and *Shh* expression were respectively 50% and 40% significantly lower than in their wild-type littermates, while their known targets and readouts *Id1* and *Gli1* were unchanged (Figure 31d). EC1-deficient forelimb buds displayed a two-fold increase in *Id1* transcription and no significant effects on *BMP4, Shh or Gli1* transcription levels (Figure 31e). In *EC1*$^{\Delta/\Delta}$ mutants, I observed the disruption of the self-regulatory initiator

module between GREM1 and BMPs (Benazet et al., 2009) which disturbed limb bud outgrowth and patterning to the extent of generating a spectrum of mild limb deformity phenotypes such as inter-digit soft tissue fusion and C-L type syndactyly (20% penetrance at E14.5, Appendixes table 1). SHH pathway was unaffected, which was surprising as the downregulation of BMP activity by GREM1 had been described as essential for the set-up of the SHH/GREM1/AER-FGF limb propagation loop (Benazet et al., 2009). I suggested this observed phenotype could be due to a high BMP activity-induced apoptosis (Kaltcheva et al., 2016; Zou and Niswander, 1996) independently of the SHH pathway, and will come back to this hypothesis later in section 6.4.2.1.

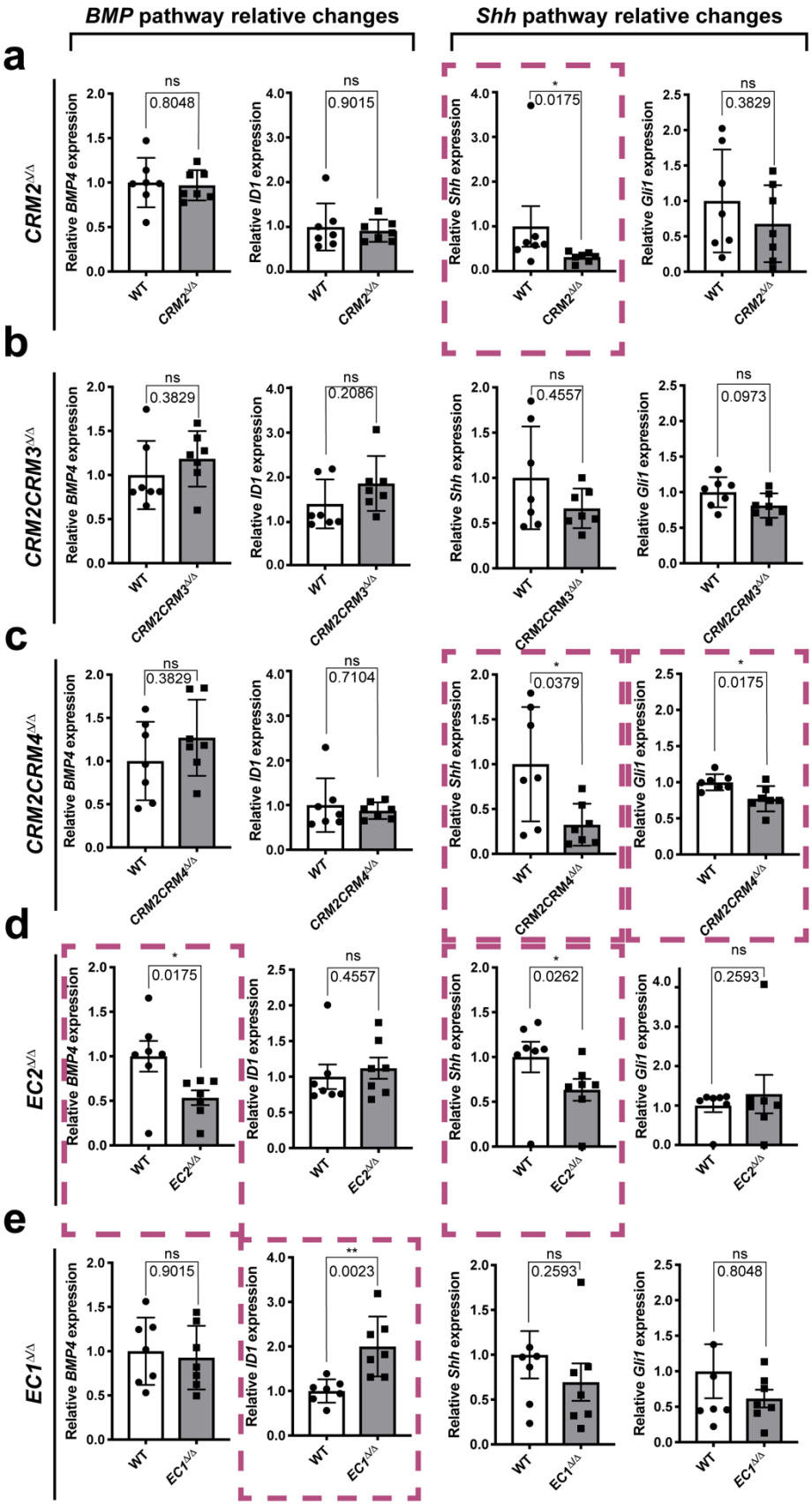


Figure 31. RT-qPCR analysis of BMP and SHH pathway activity in mutants where *Grem1* expression was quantified to approximately 50% of the wild-type levels of expression. Homozygous mutant handplates were dissected from E11.0 embryos

(40ss to 42ss) and analyzed in comparison to WT littermates at E11.0 (40ss to 42ss) for (a) *CRM2*$^{\Delta/\Delta}$, (b) *CRM2CRM3*$^{\Delta/\Delta}$, (c) *CRM2CRM4*$^{\Delta/\Delta}$, (d) *EC2*$^{\Delta/\Delta}$, (e) *EC1*$^{\Delta/\Delta}$. Dashed red boxes highlight significant changes in gene expression. (n=7; mean ± SEM, two-tailed Mann-Whitney test), p-values are indicated on top. Significant changes are highlighted with red dashed boxes.

- **Mutants with >50% decrease in *Grem1* expression levels.**

Another category consisted of *CRM2CRM5*$^{\Delta/\Delta}$ (Figure 32a) and *EC1CRM5*$^{\Delta/\Delta}$ (Figure 32b) mutants that displayed approximately an 80% reduction of *Grem1* transcription. Both mutants exhibited a significant reduction of *Bmp4* transcriptional levels (60% and 30% reductions respectively) and *Shh* (60% and 70% reductions respectively). However, in both mutants, the SHH pathway seemed to be unaffected as *Gli1* levels were normal (Figure 32). The BMP activity readout *Id1* was unaffected in *CRM2CRM5*$^{\Delta/\Delta}$ individuals (Figure 32a) while it was significantly reduced to 80% of its wild-type expression levels in *EC1CRM5*$^{\Delta/\Delta}$ handplates (Figure 32b). Besides, I had quantified a decrease of BMP pathway activity, while *Grem1* expression levels were also decreased (Figure 27f and Figure 32b) and in opposition with the two-fold increase of *Id1* transcript levels characterizing the *EC1*$^{\Delta/\Delta}$ limbs at this stage (Figure 31e). In those mutants, the significant decrease of *Grem1* expression levels and strong change in its expression pattern combined with a counterintuitive decrease of BMP activities and a normal SHH pathway activity stabilized tetradactyly (Figure 27f, Figure 32b).

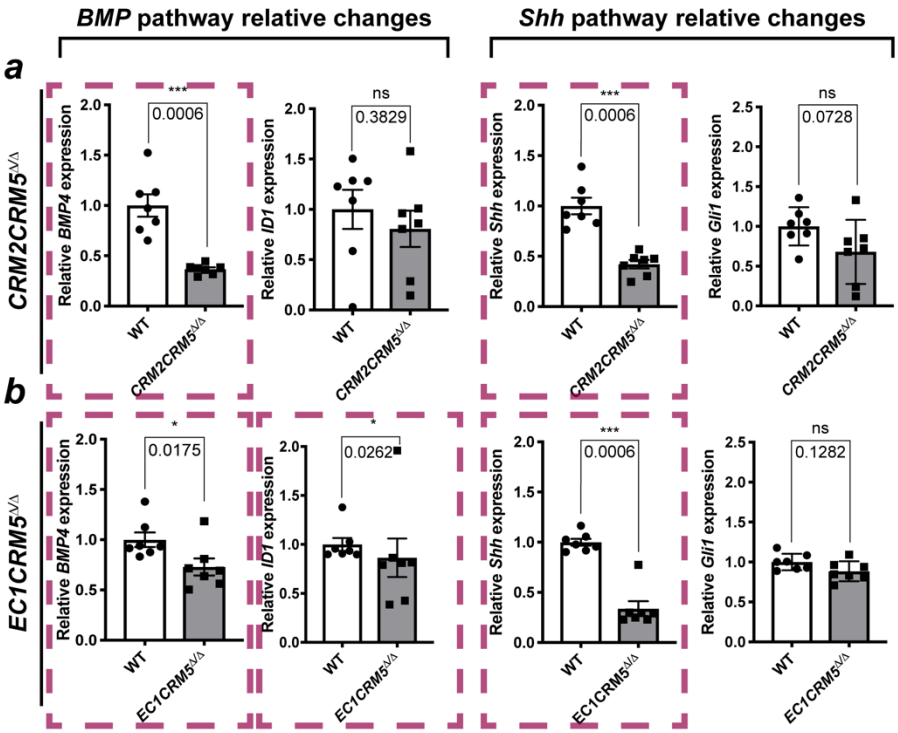


Figure 32. qPCR analysis of BMP and SHH pathway activity in mutants where *Grem1* expression was quantified to approximately 20% of the wild-type levels of expression. Homozygous mutant handplates were dissected from E11.0 embryos and analyzed in comparison with WT littermates at E11.0 (40ss to 42ss) for (a) *CRM2CRM5*$^{\Delta/\Delta}$, (b) *EC1CRM5*$^{\Delta/\Delta}$. Dashed red boxes highlight significant changes in gene expression. (n=7; mean ± SEM, two-tailed Mann-Whitney test), p-values are indicated on top.

- **Mutants with ≥95% decrease in *Grem1* expression levels and skeletal phenotype.**

The last tier I analyzed consisted of the mutants expressing very low *Grem1* levels (below 5%) at E11.0: *EC1EC2*$^{\Delta/\Delta}$ and *Cis*$^{\Delta/\Delta}$ mutants. Surprisingly, the two mutants that both suffered from a near-total loss of *Grem1* expression and similar morphological phenotypes did show differences in the expression of other limb genes. *EC1EC2*$^{\Delta/\Delta}$ handplates, despite the loss of *Grem1*, showed unchanged BMP pathway activity (Figure 33a), unlike *Cis*$^{\Delta/\Delta}$ homozygous mutants that doubled the activity of BMPs (Figure 33b), as measured with *Id1* RT-PCR. I noticed that BMP activity, for the *Cis*$^{\Delta/\Delta}$ mutants, increased in the same range as in the *EC1*$^{\Delta/\Delta}$ mutants, suggesting that at this stage the measured BMPs activity was the maximum activity one could obtain after losing GREM1 antagonism toward BMPs. Regarding the SHH pathway, both

mutants followed the same pattern with a significantly reduced activity (-50% of *Gli1* transcriptional levels for *EC1EC2*$^{\Delta/\Delta}$ and -80% for *Cis*$^{\Delta/\Delta}$).

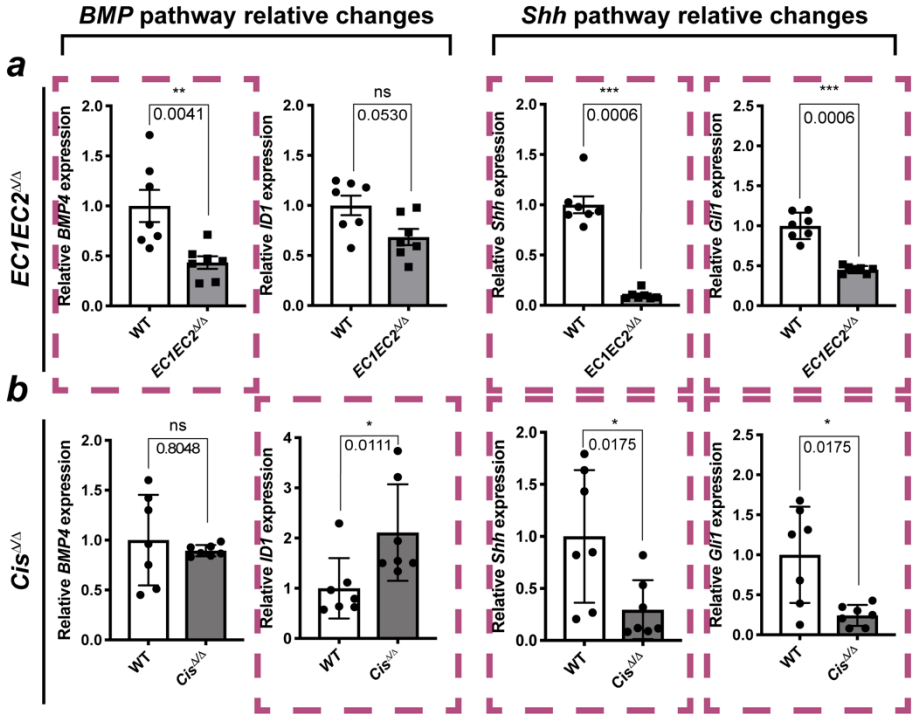


Figure 33. qPCR analysis of BMP and SHH pathway activity in mutants where *Grem1* expression was lost E11.0. Homozygous mutant handplates were dissected from E11.0 embryos and analyzed in comparison with WT littermates at E11.0 (40ss to 42ss) for (a) *EC1EC2*$^{\Delta/\Delta}$, (b) *Cis*$^{\Delta/\Delta}$. Dashed red boxes highlight significant changes in gene expression. (n=7; mean ± SEM, two-tailed Mann-Whitney test), p-values are indicated on top.

Based on the data presented here, it appeared likely that there were additional aspects to the self-regulatory feedback loop model than the ones previously established (Bastida et al., 2004; Benazet et al., 2009; Khokha et al., 2003). The *Cis*$^{\Delta/\Delta}$ forelimb buds analyses were in agreement with the published limb self-regulatory feedback loop (Benazet et al., 2009). I showed that significant fluctuations of transcription levels of *Bmp4* or *Shh* could occur without consequences on the activity of their respective downstream mediators in our *Grem1* mutants. According to the established model, *Grem1* is important for the upregulation of *Shh* during limb bud patterning (Benazet et al., 2009). The effects I observed on SHH pathway activity are coherent with the feedback loop model: lower *Grem1* expression could result in lowered SHH pathway

activity (Figure 31c, Figure 33). However, the changes with the BMP pathway were intriguing. According to the feedback loop model, *Grem1* is important for the downregulation of the BMP pathway activity during limb bud patterning (Benazet et al., 2009). When *Grem1* expression was significantly diminished, I expected an increase in BMP pathway activity. However, in the *EC1CRM5*$^{\Delta/\Delta}$ mutant embryos with only transcribe 20% of normal *Grem1* expression levels, I measured a significant handplate reduction of *Id1* expression (Figure 32b). This revealed a reduction of the entire BMP pathway activity in the limb. Especially, in the *EC1EC2*$^{\Delta/\Delta}$ mutants which suffer from a *Grem1* LOF and display a *Id* phenotype, no change was detectable in the BMP pathway activity (Figure 33a). These observations suggest that measuring transcriptional activity may be an insufficient criterion to understand changes in gene activity *per se*. Having shown that *Grem1* spatial expression varied in our mutants without differences in the overall transcriptional levels (Figures 24 to 27), I decided to investigate the *Id1* expression domain in different mutants.

Our qPCR analysis revealed that limb deformity phenotypes, induced by *Grem1* transcriptional defects, did not always result from changes in the activities of the BMP and the SHH pathways. Furthermore, fluctuations in the BMP or the SHH pathway activities did not always correlate with each other. Severe limb deformity phenotypes could arise with or without changes in BMP and/or SHH pathways. Nevertheless, all our tridactyl mutants had a significant reduction of the SHH pathway activity (independently of the BMP pathway changes). I hypothesized that change of expression domain, without any significant change in global transcription levels, could happen for the BMP pathway as well.

6.4.1.2. Changes in the *Id1* expression domain in the *Grem1* enhancer mutants

Following up on the RT-qPCR analysis, I used WISH to study possible spatial changes in BMP pathway activity. I could divide our mutants into three groups.

The first one consisted of mutants in which the *Id1* expression was unchanged (Figure 34b). Their *Grem1* expression domain followed a crescent-shaped pattern, identical or close to the normal expression pattern. Their overall *Grem1* levels were either unaffected (*CRM3*$^{\Delta/\Delta}$, *F1Δ1019*$^{\Delta/\Delta}$, Figure 24d and Figure 29b) or reduced to ±50% (*CRM2*$^{\Delta/\Delta}$, *EC2*$^{\Delta/\Delta}$, Figure 24c and Figure 27d). The mutants in this first group, *CRM2*$^{\Delta/\Delta}$, *CRM3*$^{\Delta/\Delta}$, *F1Δ1019*$^{\Delta/\Delta}$ and *EC2*$^{\Delta/\Delta}$ did not display limb deformity phenotype.

The second group exhibited ectopic expression of *Id1* in the posterior and distal part of the limb bud without any change of *Id1* transcriptional levels (Figure 34c, dashed lines). They did not present any limb skeletal deformity phenotype despite this expansion of the *Id1* expression domain. While the *Id1* expression domains in *CRM2CRM3*$^{\Delta/\Delta}$ and *CRM2CRM4*$^{\Delta/\Delta}$ mutants were overlapping, the *Id1* expression domain in *CRM2CRM5*$^{\Delta/\Delta}$ was restricted to the distal periphery of the limb bud, with a thinner mesenchymal domain of expression (Figure 34c, dashed lines).

The last group contained the mutants in which *Id1* activity was altered and its expression covered large ectopic domains along the periphery of the limb bud. *EC1*$^{\Delta/\Delta}$ and *EC1CRM5*$^{\Delta/\Delta}$ domains were nearly identical (both in localization and signal strength, Figure 34d, dashed lines), despite the fact that *EC1*$^{\Delta/\Delta}$ had a two-fold increase and *EC1CRM5*$^{\Delta/\Delta}$ a 20% decrease in *Id1* transcripts levels. The *EC1EC2*$^{\Delta/\Delta}$ and *Cis*$^{\Delta/\Delta}$ limb buds, despite their reduced size, had an *Id1* domain that was broad and ectopically covering large areas at the limb periphery (Figure 34d, dashed lines). In *EC1EC2*$^{\Delta/\Delta}$ and *Cis*$^{\Delta/\Delta}$, *Id1*-expressing domains were very similar despite their two-fold difference in *Id1* expression levels.

In summary, mutants expressing at least 50% of *Grem1* normal level in a crescent shape pattern did not show alteration in BMP pathway activity measured by *Id1* RT-qPCR as a readout (*CRM2*$^{\Delta/\Delta}$*, F1Δ1019*$^{\Delta/\Delta}$*, CRM3*$^{\Delta/\Delta}$*,* Figure 34b). The mutants that either lost the *Grem1* crescent-shaped pattern (*CRM2CRM3*$^{\Delta/\Delta}$*, CRM2CRM4*$^{\Delta/\Delta}$*, EC1*$^{\Delta/\Delta}$*, EC1CRM5*$^{\Delta/\Delta}$) or had considerably downregulated *Grem1* expression (at least an 80% downregulation for *CRM2CRM5*$^{\Delta/\Delta}$ and *EC1CRM5*$^{\Delta/\Delta}$ mutants) showed ectopic BMP pathway activity (dashed lines Figure 34b,c), similarly to what I observed in the *Grem1* LOF mutants (*Cis*$^{\Delta/\Delta}$ and *EC1EC2*$^{\Delta/\Delta}$ mutants). The mutants with simultaneous reduction of both *Grem1* levels and domains of expression at E11.0, that resulted into a change in BMP pathway activity (e.g. two-fold increase for *EC1*$^{\Delta/\Delta}$ or a 30% decrease for *EC1CRM5*$^{\Delta/\Delta}$ in *Id1* expression levels), were the ones with skeletal deformity phenotypes (syndactyly). The phenotypes caused by the downregulation of *Grem1* expression could not be interpreted through our analysis of the BMP pathway activity strength or domain.

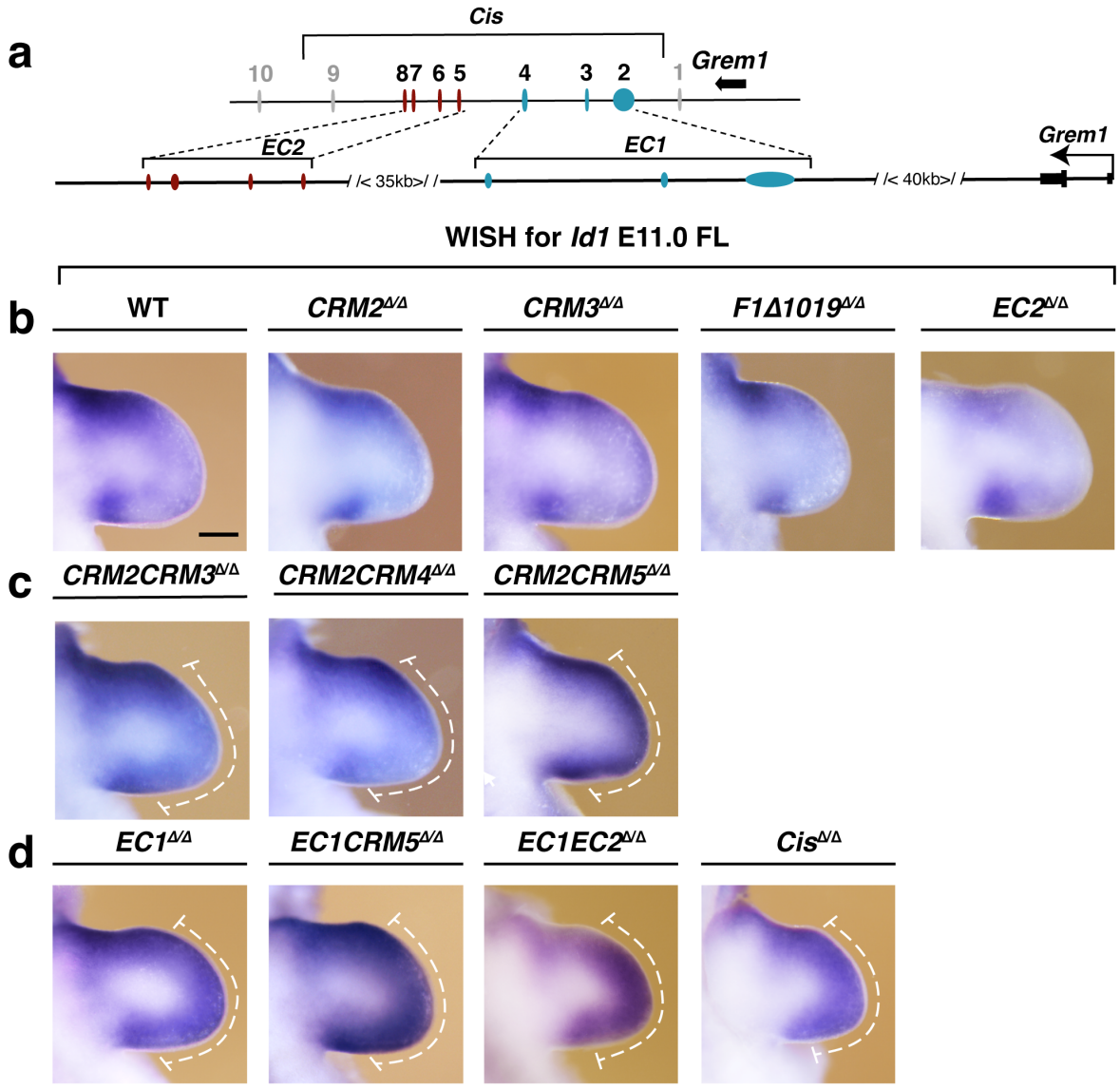


Figure 34. *Id1* was expressed in ectopic domains in the mutants that showed at least a 50% decrease of *Grem1* transcripts and reduction of the *Grem1* expression domain. (a) Schematic of the *Cis* locus, active EC1 enhancers are represented in blue ovals, EC2 CRMs are represented in maroon ovals. (b-d) WISH comparison of *Id1* expression domains of (b) WT, $CRM2^{\Delta/\Delta}$, $CRM3^{\Delta/\Delta}$, $F1\Delta1019^{\Delta/\Delta}$, $EC2^{\Delta/\Delta}$ (c) $CRM2CRM3^{\Delta/\Delta}$, $CRM2CRM4^{\Delta/\Delta}$, $CRM2CRM5^{\Delta/\Delta}$ (d) $EC1^{\Delta/\Delta}$, $EC1CRM5^{\Delta/\Delta}$, $EC1EC2^{\Delta/\Delta}$, $Cis^{\Delta/\Delta}$ in forelimbs (n≥3) from E11.0 (40ss to 42ss). Dashed lines highlight a distal expansion of the *Id1* expression domain. *Scale bars: 250μm*

A common feature in the *Grem1* enhancer mutants was the expansion of the *Id1* domain when the *Grem1* expression was reduced and restricted spatially in the limb bud mesenchyme. With RT-qPCR assays alone, I could not measure or predict those events in whole limb buds. Using WISH, I observed critical changes in the *Id1* expression domain in the distal and the posterior limb bud cell populations. This suggested that locally restricted changes in *Bmp* activity were not reflected in the global changes measured from the whole organ. These results illustrate the importance of using different methods for measuring gene activity.

6.4.2. Apoptosis or a deficit in cellular proliferation? What is the cellular defect underlying the loss of digits?

6.4.2.1. Abnormal apoptotic cell clusters correlate with loss of *Grem1* activity

The GREM1-mediated BMP antagonism in the limb bud mesenchyme has been shown to be essential for cellular survival and proliferation (Benazet et al., 2009; Michos et al., 2004; Zhu et al., 2008). The SHH signaling pathway, known to contribute to the survival and proliferation of limb mesenchymal cells, is necessary for the establishment of the future digit identities (1 to 5 along the AP axis, Zhu et al., 2008). I had identified a significant reduction of SHH pathway activity through the measurement of its target *Gli1* at this stage. I, therefore, assessed cell death in the limb buds of selected mutants with at least a 50% decrease in *Grem1* transcripts to determine if 1) I would find a correlation between *Grem1* expression levels and cell death, 2) I would uncover a connection between SHH and/or BMP pathways activities and potential cell death, 3) cell death at this stage could be an explanation for digit loss in our C-L type syndactyly phenotypes.

Using the Lyso Tracker assay (Haller et al., 1996; Zucker et al., 1999; Zucker et al., 1998) on E11.0 WT forelimb buds, I observed, in wild-types, apoptotic cells located in the proximal central part of the mesenchyme and diffused anteriorly and distally up to variable distances to the median anterior part of the limb mesenchyme. The *CRM2*$^{\Delta/\Delta}$ forelimbs were indistinguishable from the wild-type limbs (Figure 35a). The *EC2*$^{\Delta/\Delta}$ forelimb buds with an ectopic *Id1* expression displayed some variability. I detected a wild-type like apoptotic cell distribution and a pattern with slight enrichment of apoptotic cells in an anterior median locus (Figure 35a). The *EC1*$^{\Delta/\Delta}$ forelimbs behaved like wild-type forelimbs (Figure 35b). These mutants revealed that C-L type syndactyly could occur without *de novo* apoptosis during the limb bud patterning stage, even under high BMP activity conditions. The *EC1CRM5*$^{\Delta/\Delta}$ individuals, with 4 digits at E14.5 (Figure 27f), presented apoptotic patterns either identical to the *EC1*$^{\Delta/\Delta}$ or the *EC2*$^{\Delta/\Delta}$ mutants (Figure 35c). At this developmental stage, the apoptotic pattern could not be correlated to the digit loss.

In mutants displaying a near-total loss of *Grem1* expression (*Cis*$^{\Delta/\Delta}$; *EC1EC2*$^{\Delta/\Delta}$) and a decrease of SHH pathway activity, an abnormal cluster of apoptotic cells was located in the anterior distal portion of the limb (wider and more distal than the previously described in the anterior medial limb portion; Figure 35d). Those abnormal apoptotic cell clusters did not correlate with the activity of BMPs. Apoptosis in this cluster of cells, of those cell populations, at E11.0, could be a determining factor for C-L type syndactyly.

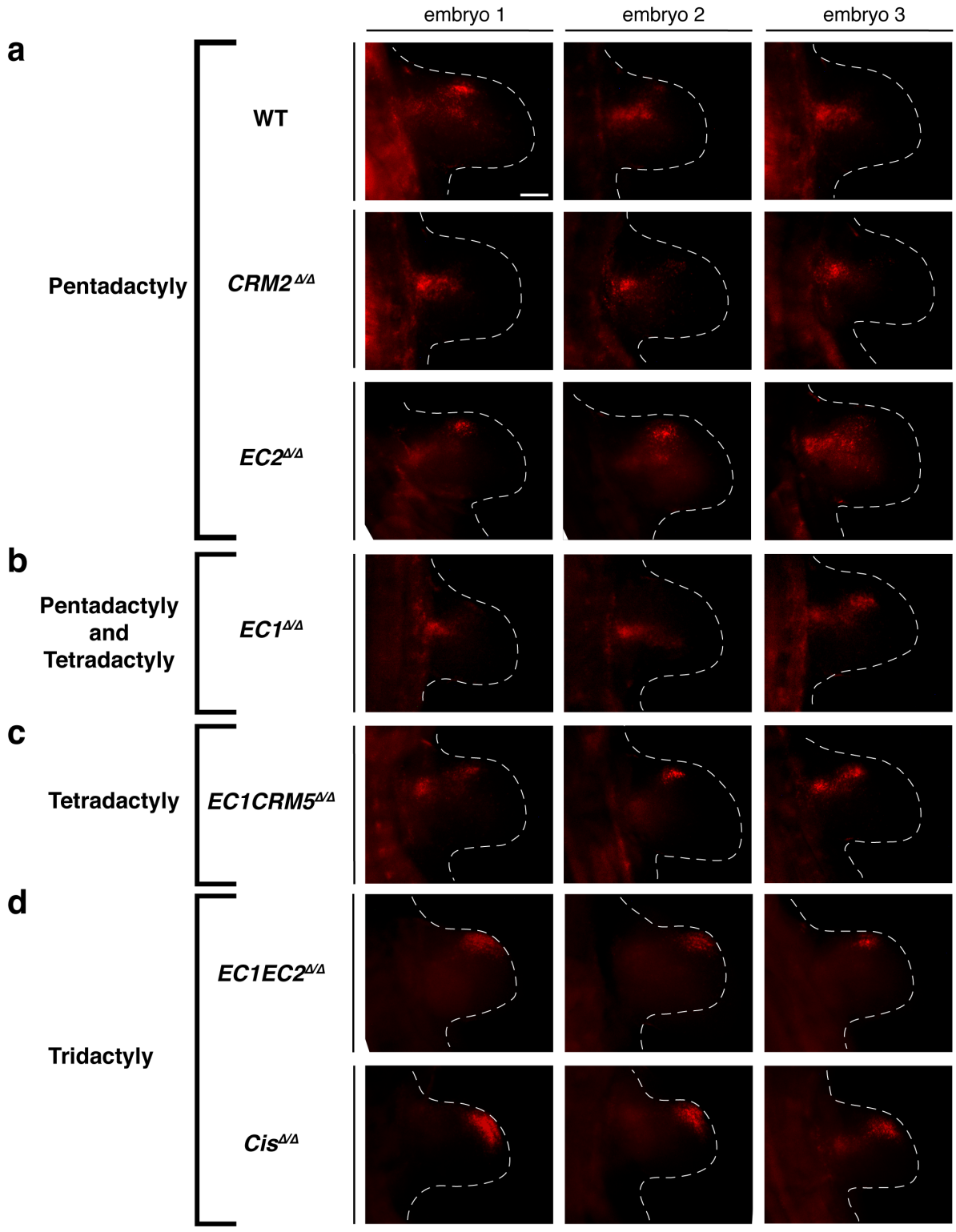


Figure 35. *De novo* apoptosis segregated tridactyl from tetradactyly and wild-type pentadactyly. Lyso Tracker assay targeting apoptotic cells in (a) pentadactyly mutants with normal BMP pathway phenotypes, (b) embryos with variable skeletal phenotypes with low penetrance of tetradactyly and high BMP activity, (c) mutants with lower BMP pathway activity and 100% of tetradactyly phenotype, (d) mutants with loss (near-loss) of *Grem1* expression and tridactyl phenotype. *Scale bars: 250μm*

The *Grem1* loss of function mutants have smaller limbs at E11.0 (Figure 9b, Figure 27g). However, at E10.5, *Grem1*$^{\Delta/\Delta}$ limb buds are indistinguishable from wild-type limb buds by shape or by apoptosis tracking (Figure 36).

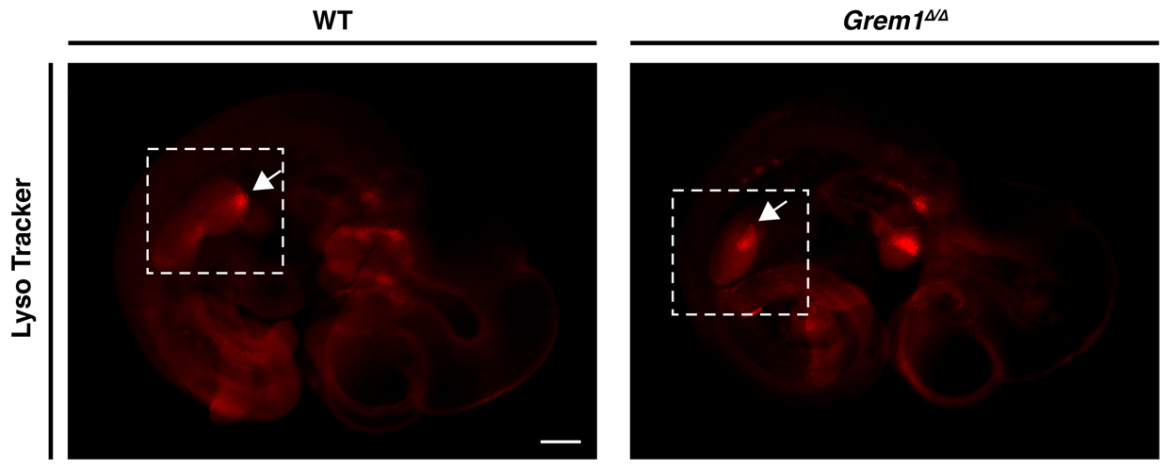


Figure 36. Lyso Tracker assay detecting programmed cell death by apoptosis in WT and *Grem1*$^{\Delta/\Delta}$ in E10.5 embryos (37ss). Dashed boxes delimitate limb buds and white arrows point toward the anterior distal apoptotic cell cluster. *Scale bar: 500µm*

The complete loss of *Grem1* expression was associated with abnormal clusters of apoptotic cells in the anterior distal limb bud mesenchymal field at E11.0, in conjunction with previous studies (Bastida et al., 2004; Michos et al., 2004; Reinhardt et al., 2019). Our observations were 1) those apoptotic clusters could arise in contexts where BMP pathway activity was unchanged in the limb (*EC1EC2*$^{\Delta/\Delta}$) and 2) high BMP activity did not always correlate with abnormal apoptosis at E11.0 (*EC1*$^{\Delta/\Delta}$). As apoptosis was normal in *EC1*$^{\Delta/\Delta}$ and *EC1CRM5*$^{\Delta/\Delta}$, their phenotypes are likely caused by other mechanisms. In particular, as these two mutants had trapezoidal limb shape at E12.0, I hypothesized that limb progenitors were either already absent at E11.0 or depleted between E11.0 and E12.0. It is highly possible that our method of detecting apoptosis was only sensitive enough to detect bigger apoptotic cell clusters. Therefore, other smaller and undetected clusters of cell death might be involved in *Grem1* mutant phenotypes.

6.4.2.2. Cell proliferation decrease: a prognosis of C-L type syndactyly at E11.0

The limb bud mesenchyme contains molecularly distinct cell populations that contribute to the osteochondrogenic lineages of the appendicular skeleton, the tendons and the connective tissue (Osterwalder et al., 2014; Reinhardt et al., 2019; Zhu et al., 2008; Zuniga, 2015). Surprisingly, little is known about the cellular heterogeneity of the limb mesenchymal progenitors (LMPs) or how they are controlled by the limb regulatory networks. During limb bud outgrowth and patterning, distal LMPs are kept in an undifferentiated and proliferative state (Karamboulas et al., 2010; Pearse et al., 2007; ten Berge et al., 2008) while the core mesenchyme undergoes a transition to chondrocyte identity under the influence of high BMP pathway activity (Akiyama et al., 2005; Barna and Niswander, 2007; Wright et al., 1995). Early termination of *Grem1* expression had been associated with premature mesenchymal condensation and chondrogenic differentiation (Bandyopadhyay 2006; Pizette and Niswander 2000; Lopez-Rios 2012). In addition, GREM1 has been directly linked to the maintenance of immature LMPs in the distal limb bud (Reinhardt et al., 2019). Together with Dr. Robert Reinhardt, we undertook cell proliferation assays using fluorescence-activated cell sorting (FACS) on mouse forelimb bud tissue at E11.0 to estimate LMPs proliferative activity by quantifying differences in the cell cycle (Figure 37). Therefore, we investigated if changes in cell proliferation could correlate with the phenotypes we observed in the *Grem1* loss of enhancer function mutants. We compared proliferation rates between wild-type and mutants with $\geq$ 50% decrease in *Grem1* expression levels at E11.0 and 1) a crescent-shaped *Grem1* domain at E11.0, no change in *Id1* expression domain and normal apoptosis (*CRM2*$^{\Delta/\Delta}$), 2) a crescent-shaped *Grem1* domain with ectopic *Id1* expression and variable apoptotic profiles (*EC2*$^{\Delta/\Delta}$), 3) reduced *Grem1* domain at E11.0, ectopic and increased *Id1* expression and no abnormal apoptosis at E11.0 (*EC1*$^{\Delta/\Delta}$) and 4) our loss of *Grem1* function mutants with abnormal apoptotic cell clusters the *Cis*$^{\Delta/\Delta}$ with high and the *EC1EC2*$^{\Delta/\Delta}$ with normal BMP pathway activity (Figure 37). After the removal of the lineage-positive cells, we established a clear correlation between the mice mutant alleles resulting in C-L type syndactyly phenotypes (*EC1*$^{\Delta/\Delta}$, *EC1EC2*$^{\Delta/\Delta}$ and *Cis*$^{\Delta/\Delta}$) and the significant

reduction in cell proliferation/cell division (mitotic or synthesis phases of the cell cycle – M phase or S phase, Figure 37a,b). The *Cis*$^{\Delta/\Delta}$ and *EC1*$^{\Delta/\Delta}$ homozygous mutant lost 40% and 30% respectively of cells in mitosis over the WT individuals (Figure 37a,b). *EC1EC2*$^{\Delta/\Delta}$ lost 25% of cells in the replication phase (Figure 37a). The measured high BMP activity changes in our RT-qPCR assay for *EC1*$^{\Delta/\Delta}$ (Figure 31e) and *Cis*$^{\Delta/\Delta}$ (Figure 33a) homozygous mutants correlated with a reduction of the fraction of progenitors occupying the M-phase while the near-total loss of *Grem1* in *EC1EC2*$^{\Delta/\Delta}$ mutant with a normal BMP activity (Figure 33b) correlated with a decrease of progenitors in S-phase.

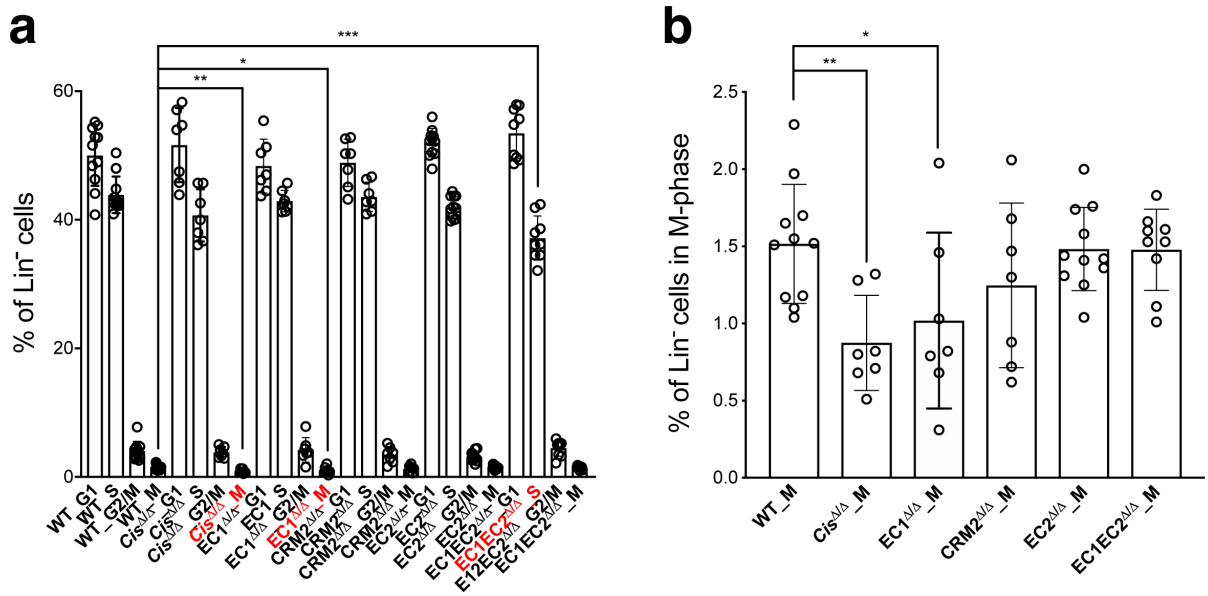


Figure 37. Forelimb bud cell cycle analysis in selected *Grem1* loss of enhancer function mutants. (a) Analysis of all cell cycle phases and (b) mitotic cells in lineage negative (Lin$^-$) cells from mouse forelimb buds at E11.0 (39-43ss), n≥7 paired forelimbs, data are mean ± SD, two-tailed Mann-Whitney test), *P≤0.1,**P≤0.01,***P≤0.001.

Cellular proliferation was gradually reduced in the enhancer mutant alleles in correlation to the phenotypic severity. Significant decreases of cellular proliferation in E11.0 forelimb buds seemed to be an immediate consequence of significant *Grem1* transcriptional alterations. Together with the elevated apoptosis levels, premature exit from the cell cycle and significantly reduced proliferation appeared to strongly correlate with limb skeletal deformity phenotypes.

7. Discussion

The findings of my Ph.D. thesis project provide insights of general importance into the *cis*-regulatory control of the *Grem1* expression dynamics during mouse limb bud development. I show that the 310 kb *Grem1* TAD (or *Grem1-Fmn1* intra-TAD) contains all the essential CRMs, with enhancer activity, that control *Grem1* expression levels and its spatio-temporal expression kinetics in limb buds. In particular, my analysis shows *Grem1* CRMs activities are dependent on their genomic arrangement within the TAD but also on *trans*-acting signaling pathways (space) and developmental stage (time). The CRM interactions are more complex than first expected, which goes beyond being simple ON/OFF switches, with respect to the *Grem1* expression. Here, I will discuss and interpret my results in the general context of the field and consider some of the major scientific questions raised.

7.1. False enhancers discovery

7.1.1. The Enhancer Chain Model

TADs have been described as conserved and discreet insulating structural chromatin interaction domains, that restrict enhancer activity toward their associated genes, and prevent inappropriate enhancer-promoter interactions (Dixon et al., 2012; Nora et al., 2012; Rao et al., 2014; Sexton et al., 2012). Together, my analysis shows that the *Grem1* TAD contains all the *cis*-regulatory enhancers participating in the *Grem1* gene transcriptional regulation during limb bud outgrowth and patterning. No elements, outside the *Grem1* TAD boundaries, are interacting with the *Grem1* promoter when the TAD boundaries are maintained (Figure 28). To identify the relevant *Grem1* CRMs, I used genomic profiling to identify the likely candidate CRMs (Figure 16, Figure 17). Interestingly, some of these regions are not enhancers, despite having the expected

chromatin state and histone modifications, TFs binding profile (Figure 17) and physical interactions with the *Grem1* promoter (Figure 15, Figure 28). These include CRM8 and CRM9, as no enhancer activity is detected in reporter assay (Figure 18), despite them displaying molecular features of enhancers. CRM8 and CRM9 are cohesin binding domains (Figure 38, Appendixes Figure 6, DeMare et al., 2013) and CRM9 is located within 1.8 kb of a CTCF binding region (Appendixes Figure 7). To date, the full spectrum of cohesin functions in gene regulation remains unclear (Dorsett, 2019; Kentepozidou et al., 2020; Merkenschlager and Nora, 2016; Rao et al., 2017). CTCF/Cohesin complexes have been described as TADs anchors, maintaining them through embryonic development (Handoko et al., 2011; Phillips and Corces, 2009; Sanyal et al., 2012) and reinforcing TAD sub-domains (100 kb to 1 Mb) around genes expressed during embryonic development (Phillips-Cremins and Corces, 2013). Cohesins were predicted to regulate gene expression through their structural role in intra-TAD interactions (<100 kb). Cohesins interact with other proteins, such as Mediator proteins, at promoter and enhancer sites to bring them in close physical proximity (Kagey et al., 2010; Phillips-Cremins and Corces, 2013). Thus, these higher chromatin structures render linear distances, on the DNA, rather irrelevant for the understanding of transcriptional regulation. Unfortunately, there is no available ChIP-seq data for Mediator in mouse limb buds before E12.5, but only data for cohesion binding to CRM8, CRM9, and CRM2 to CRM4 (Appendices Figure 6, Kagey et al., 2010). Studies by others have provided evidence that promoter/enhancer interactions are optimized when these elements are located at CTCF boundaries loci and brought in close proximity by direct and persistent contacts (Hou and Corces, 2012; Kagey et al., 2010; Tang et al., 2015). When elements are not in close proximity to CTCF binding sites, contacts are in general transient and weak as part of loop extrusion mechanisms (Fudenberg et al., 2016; Hou and Corces, 2012; Kagey et al., 2010; Tang et al., 2015). The *Grem1* promoter does not overlap with CTCF and/or cohesin binding regions but is still constitutively and robustly contacting the *Grem1* TAD CTCF boundaries. Thus, the *Grem1* promoter region is continuously associated with CTCF/cohesin binding regions, which also interact with other cohesin binding regions such as CRM8 or CRM9. Hi-C-seq analysis of human cell lines showed that the most distal enhancer at the linear genomic level coordinately controls the effects of the other

"chain enhancers", in the 3D chromatin space, that regulate the transcription of their target gene (Song et al., 2019, Figure 38). This most distal enhancer has been proposed to be the only enhancer directly interacting with the promoter. Such "first" enhancers should be enriched in CTCF and cohesin binding sites while the other enhancers in the "Enhancer Chain" (ECh) stabilize the contacts and fine-tune gene expression by physically interacting with each other and promoters through the most distal enhancer (Song et al., 2019, Figure 38a). The linear distance between enhancers in an ECh and promoter did not define their rank in the chains but rather the lowest number of intermediate contacts with other enhancers (Figure 38b). In wild-type limb buds, CRM3 and CRM4 do not directly interact with the *Grem1* promoter (Figure 28, Figure 39a). Therefore, they could regulate *Grem1* expression (Figure 25c,d) via their interactions with other *Grem1* enhancers as part on an ECh.

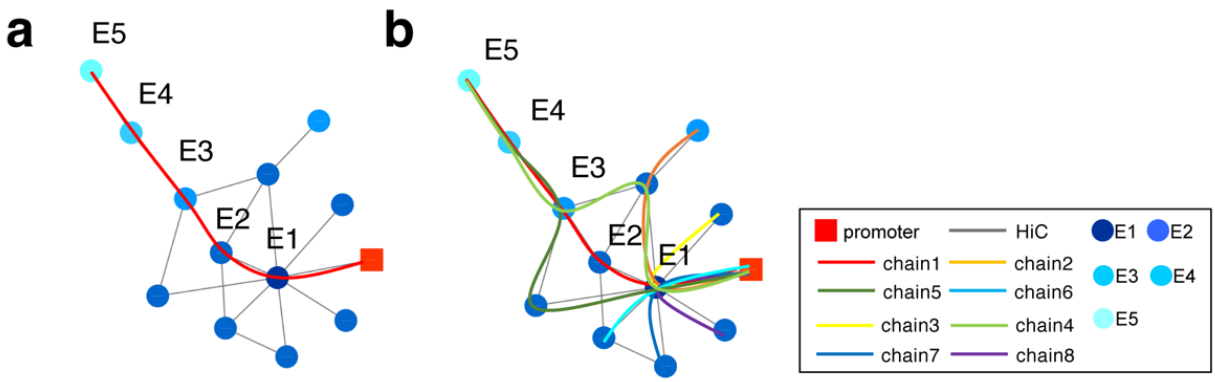


Figure 38. Examples of enhancer chains (ECh). (a) One example of regulatory element networks, (b) schematic of possible EChs, defined according to the rank of connected enhancers relative to their closest target promoter. *Adapted from Song et al. (2019)*.

Song et al. (2019) also proposed that in the absence of the first enhancer, the others would contact the promoter through a neutral genomic region. As a consequence, the transcription levels of the gene of interest would be reduced. The binding of cohesins at the CRM8 and the CRM9 loci suggests that these elements, interacting with the *Grem1* promoter at high frequency, could have intrinsic structural roles such as promoting long-range contacts with the *Grem1* promoter (Figure 28). In our model, 4C

experiments had revealed the loss of the entire EC2 region (encoding CRM5 to CRM8) results in an about 50% reduction of *Grem1* expression and widespread reduction in promoter/CRM interactions. As, within EC2, only CRM8, which wasn't a limb enhancer according to my reporter assays, was binding cohesin, this may point to a similar mechanism as proposed by Song and coworkers: CRM8 could be the first element of a *Grem1* ECh. In fact, Song et al. (2019) did not functionally assess their candidate enhancers, in contrast to my study, as they defined enhancers simply by chromatin profiling and marks as we did initially (Figure 16, Figure 17). Therefore, my analysis indicates that EC2, and by extension CRM8, could be essential for the establishment and/or maintenance of robust interactions of CRMs with the *Grem1* promoter (Figure 39a-d). In addition, 4C-seq analysis using the *Grem1* promoter as viewpoint provided deep insights into the chromatin architecture of the *Grem1* TAD in both wild-type and mutant limb buds. *De novo* and/or enhanced contacts (with CRM3 and CRM4) are observed in the *F1Δ1019* deletion, which only leaves EC1 intact, but this is not accompanied by a significant reduction in *Grem1* expression (Figure 28, Figure 29, Figure 39e). In *F1Δ1019*$^{\Delta/\Delta}$ limb buds, CRM4, which also binds cohesin, would be the most distal enhancer (Appendixes Figure 6). Therefore, CRM4 could organize a new ECh with CRM3 and CRM2 in *F1Δ1019*$^{\Delta/\Delta}$ limb buds, which could explain the maintenance of *Grem1* expression. This suggests that in the context of the *F1Δ1019* deletion, additional functions of CRM3 and CRM4 could be uncovered to compensate for the *Grem1 cis*-regulatory landscape alterations. Besides, my analysis shows that in mutant contexts, no CRM outside the critical region is recruited to the *Grem1* promoter (Figure 39f).

My study indicates that the chromatin architecture of the *Grem1* TAD functions in providing *Grem1* expression with *cis*-regulatory robustness.

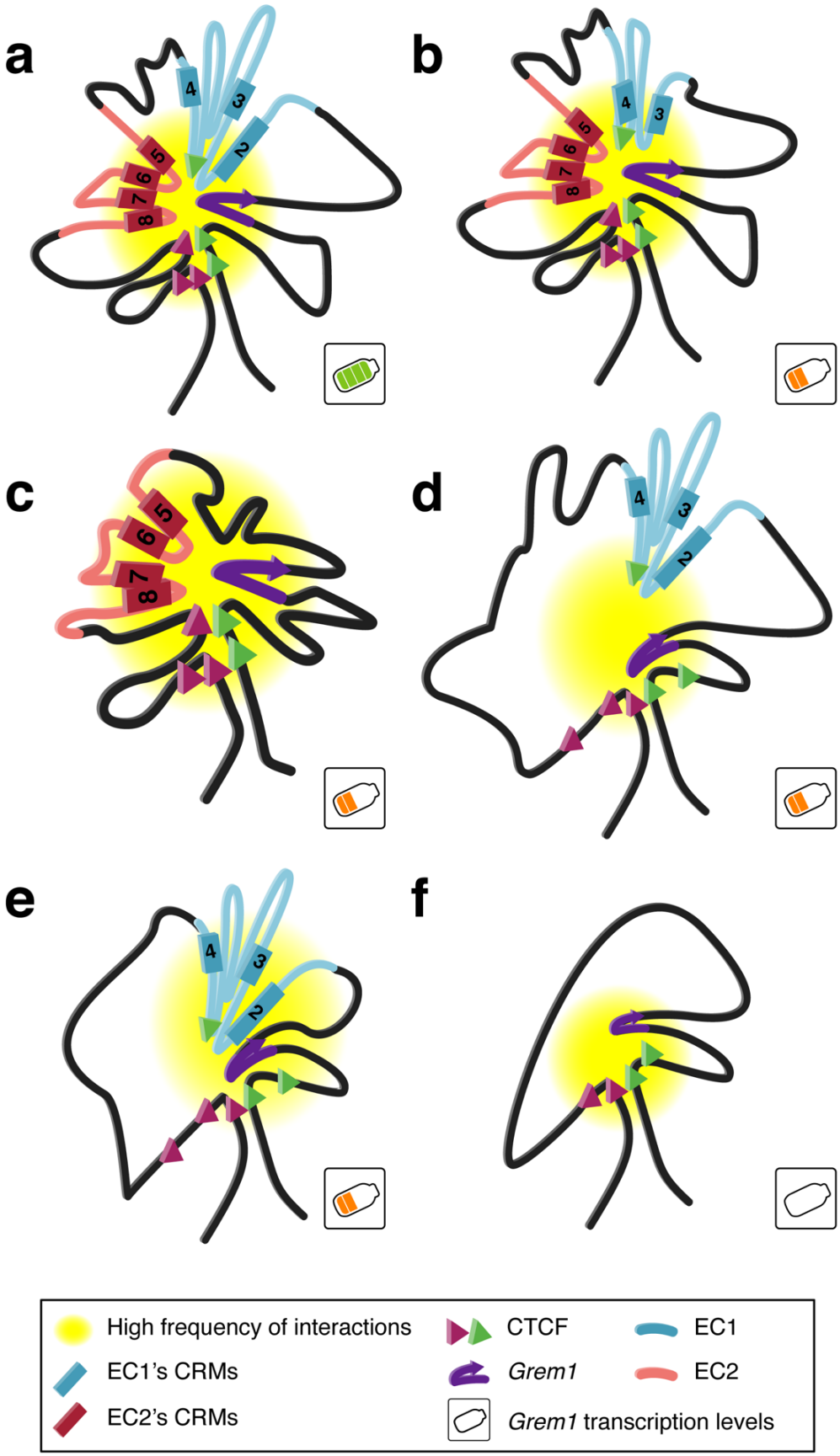


Figure 39. Schematic illustration of the *Grem1* TAD conformations in different enhancer loss-of function *Grem1* mutations. (a) WT, (b) *CRM2*$^{\Delta/\Delta}$, (c) *EC1*$^{\Delta/\Delta}$, (d)

EC2$^{\Delta/\Delta}$, (e) *FΔ1019*$^{\Delta/\Delta}$, (f) *Cis*$^{\Delta/\Delta}$ hypothesized *Grem1* intra-TAD structural organizations.

7.1.2. The Replication Domain Model

In 2014, Pope and colleagues showed that TAD boundaries overlap with a nearly one-to-one correlation with replication domain boundaries. TADs were described as “stable regulatory units of replication timing” (Pope et al., 2014, Figure 40).

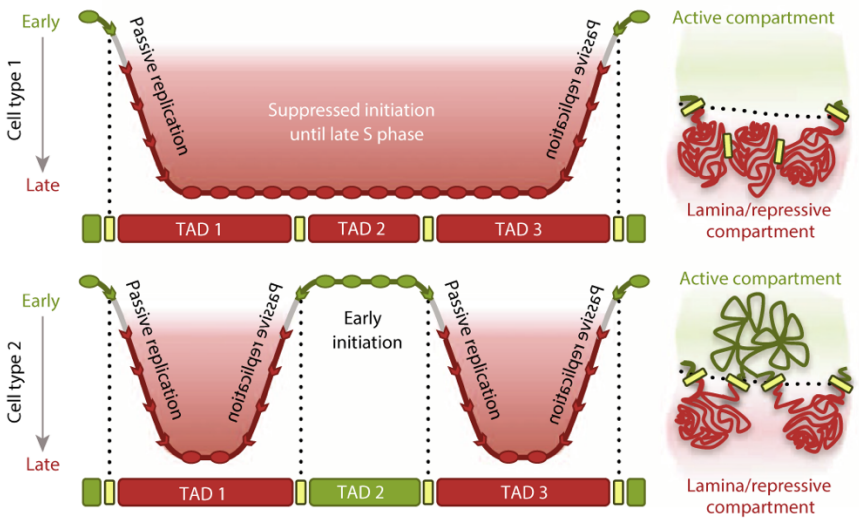


Figure 40. The replication domain model. Upper panel: Upon differentiation, TADs with uniform replication timing can switch it. Lower panel: After switching, the TAD structural boundaries, enriched in CTCF and transcription permissive marks, overlap with the boundaries of replication domains. *From Pope et al. (2014).*

Recently, loci with enhancer signatures have been associated with Early Replication Control Elements (ERCEs, Sima et al., 2019). ERCEs are open chromatin regions, enriched in enhancer histone decorations, that maintain physical contacts despite CTCF depletion (Sima et al., 2019). Their sizes vary from 3 kb to 46 kb, but the authors did not test the ECRC regions for enhancer activity. ERCEs that function in a CTCF-independent manner were described as essential for chromatin compartmentalization into A (open and transcriptionally active) and B domains (closed and transcriptionally inactive, Dixon et al., 2012; Lieberman-Aiden et al., 2009, Figure 41). It was shown

that inverting ERCEs regions creates new TAD boundaries, which demonstrated that chromatin looping is regulated independently by both CTCF and ERCEs.

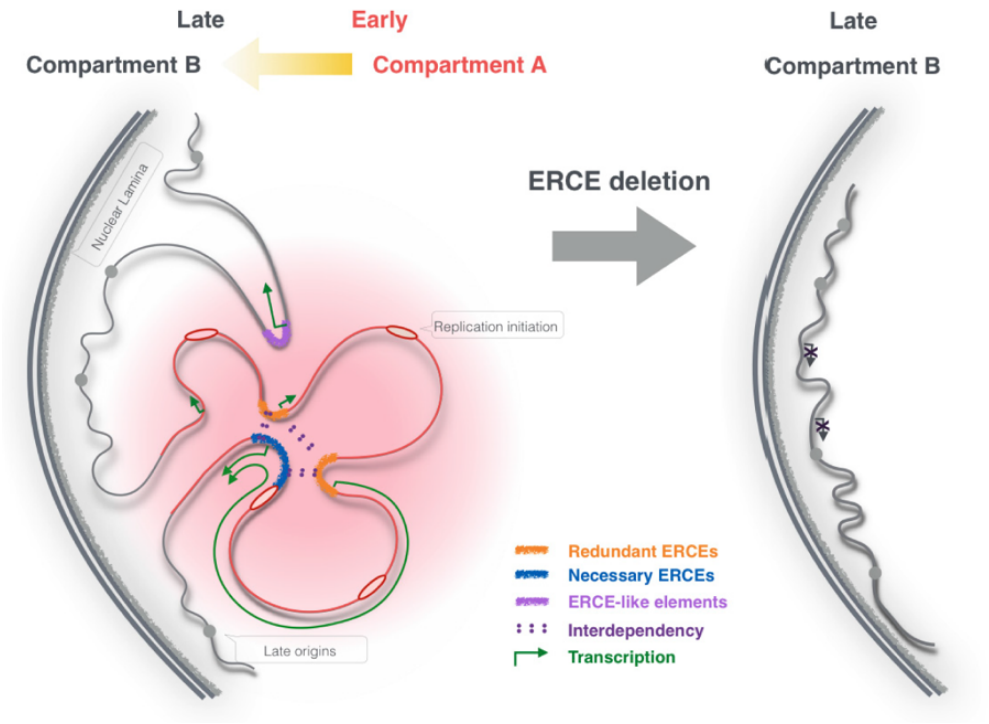


Figure 41. Model for the function of early replication control elements (ERCEs). Scheme of a co-regulation model illustrating ERCE roles in replication timing, A/B compartmentalization, transcription and chromatin looping. *From Sima et al. (2019).*

These studies showed that CRMs with enhancer-like features could also function in regulating replication during tissue development and/or homeostasis in addition to transcriptional regulation of their target gene(s).

Therefore, it is theoretically possible that some of the CRMs such as CRM9 (and possibly others), which lacks enhancer activity, could also function in structural-dependent nuclear mechanisms other than *Grem1 cis*-regulation.

7.2. The respective functions of the *Grem1* enhancers

For my research, we have generated several mouse lines harboring various CRM and/or enhancer cluster deletions in the *Grem1* critical *cis*-regulatory region. The analyses show that such of these enhancers have specific essential functions while others function in a cooperative manner.

For example, CRM2 is essential for the up-regulation of *Grem1* expression in the distal and anterior limb bud mesenchyme and for the maintenance of its expression in handplates during digit formation (Figure 24c). Benazet et al. (2009) showed that the onset and the up-regulation of *Grem1* expression depend initially on a high BMP pathway activity. Interestingly, SMAD4 ChIP-seq reveals that CRM2 is the only enhancer, in the *Grem1* landscape, enriched in SMAD4 chromatin complexes (Gamart et al., manuscript in preparation, Figure 17b), which indicates that CRM2 relays BMP signal transduction in early limb buds. In agreement, the activity of CRM2 (GRS1) remains posteriorly restricted in *Smad4*-deficient mouse limb buds (Zuniga et al., 2012a), which is corroborated by the fact that the *Smad4* deficiency reduces *Grem1* expression to a similar extent as the deletion of CRM2 in early limb buds (Figure 24c, Appendixes Figure 8).

CRM4 has been proposed to have a silencer activity, as it appears to function in the anterior restriction of the *Grem1* expression (Li et al., 2014). However, the inactivation of CRM4 in the context of the CRM2 deficiency (Appendixes Figure 5) and part of $EC1^{\Delta/\Delta}$ (Figure 26c) and $EC1CRM5^{\Delta/\Delta}$ (Figure 27f) did not result in ectopic *Grem1* expression. These results suggest 1) CRM2 is necessary for the anterior expansion of *Grem1* expression and that 2) the silencer function of CRM4 might prevent precocious anterior *Grem1* expression. In addition, $EC1CRM5^{\Delta/\Delta}$ limb buds, but not the $CRM5^{\Delta/\Delta}$ or $CRM2CRM5^{\Delta/\Delta}$ ones, display an ectopic proximal-anterior *Grem1* expression domain (Figure 27c,e). This observation implied that other CRMs than CRM2 can drive *Grem1* in ectopic domains, in the absence of EC1, when CRM5 is lost in *cis*. Thus, it appeared that in such a sensitized background ($EC1^{\Delta/\Delta}$), CRM5 is

functioning in the repression of *Grem1* expression in the anterior-proximal limb bud mesenchyme. These changes could be explained by a “CRM chain model” similar to the ECh model (Song et al., 2019), in which the CRM4 silencing activity depended on its interactions, not with the *Grem1* promoter, but with the other CRMs. This “CRM chain” could rewire to include other silencers, such as CRM5*,* in the absence of EC1, but this hypothesis would need to be tested by comparative HiC-seq analysis using the appropriate wild-type and mutant backgrounds*.* Besides, while the repressor activity of CRM4 had been associated with the GLI3 transcriptional repressor in the anterior limb bud (Li 2014), CRM5 appears not to be enriched in GLI3 chromatin complexes (Figure 17, Appendixes Figure 9). However, comparative analysis of enhancer-transcription factor interactions, in wildtype and mutant limb buds, is lacking as deletion of one CRM could potentially identify *de novo* binding sites. This genetic analysis reveals the robustness underlying the spatial regulation of the *Grem1* expression, due to the redundant activity of all the enhancers in the landscape (Figure 42). It appears that two of the CRMs (CRM4 and potentially CRM5) are not only enhancers but also spatial repressors which refine the *Grem1* expression domain.

During limb bud outgrowth, the deletion of individual CRMs had no detrimental consequences for limb skeletal development (Figure 42). This may be due to the fact that four of the enhancers in the *Grem1* TAD (CRM2 to CRM4 and CRM5) display, at least, partially overlapping spatial enhancer activities. Indeed, a recent publication showed that enhancers work in pairs of redundant elements during limb patterning, but in an additive manner to regulate the transcriptional levels (Osterwalder et al., 2018). Furthermore, *Ihh* expression, in a variety of embryonic tissues including limb buds, is maintained by interactions among nine redundant enhancers, functioning in an exclusively additive manner (Will et al., 2017). In *CRM2*-deficient limb buds, it appears that the interactions between CRM3 and CRM4 with the *Grem1* promoter could be increased (Figure 28c). Both CRM3 and CRM4 have enhancers activities in limb buds, which partially overlap in the distal limb bud mesenchyme (Figure 20c,d) and are regulated by the same TFs (Figure 17a). Their individual deletion failed to reveal major functions in the regulation of *Grem1* expression (Li et al., 2014, Figure 24d). This, together with the analysis of compound mutants lacking two or all CRMs

in EC1, reveals synergism between the CRM3/CRM4 and CRM2 enhancers (Figure 25c,d). In fact, it is likely that both primarily function as so-called shadow enhancers as their functions become apparent as CRM2 is inactivated. Shadow enhancers had been defined in invertebrates as loci providing “spatial refinement, temporal synchrony and robustness” to gene expression binding to the same TFs complexes (Hong et al., 2008). Nevertheless, CRM3/CRM4 recruitment to the *Grem1* promoter, in *F1Δ1019*-deficient limb buds, suggests their support toward CRM2 enhancer activity can operate to reinforce CRM2 action in genetically challenged backgrounds. Finally, the deletion of all enhancers encoded by EC1 results in slightly aggravated spatial changes in comparison to deleting CRM2 alone. This weakens the robustness of digit patterning system as partial fusions of anterior digits are observed in *EC1*- deficient limb buds.

My genetic analysis shows that CRM2 and CRM5 both have essential functions in regulating *Grem1* transcript levels albeit their spatial activity being rather similar (Figure 21b,c). This is due to them interacting largely with the same TF complexes, namely HOXA13, HOXD13 and ß-catenin while CRM2 also interacts with SMAD4 complexes (Figure 17). They are likely the primary enhancers of their respective ECs that have both specific and cooperative functions. Indeed, the inactivation of both enhancers reduces *Grem1* expression by 80%, but the spatio-temporal changes are similar to the single $CRM2^{\Delta/\Delta}$ deletion (Figure 24c). The observed reduction in *Grem1* transcript levels is strictly additive as it has been shown for the transcriptional regulation of other genes such as *Ihh* (Will et al., 2017), *Gli3*, *Shox2* (Osterwalder et al., 2018) or *Twist* (Hirsch et al., 2018) during mouse limb bud development. However, in spite of this drastic reduction in the absence of CRM2 and CRM5 in *cis*, the *Grem1* *cis*-regulatory landscape was able to sustain an almost normal spatial *Grem1* expression and normal digit development. Therefore, it is not unlikely that their deletion is compensated by the activity of shadow enhancers such as CRM3/CRM4 (Figure 28c).

The deletion of EC2 exposes the function of EC1 in preserving the wild-type-like spatial expression *Grem1* domain and, albeit a 50% reduction in transcript levels, limb

bud development is not significantly altered and pentadactyly is maintained. The distal margin between the AER and the *Grem1* crescent seemed broader but this is apparently not sufficient to disturb limb patterning. In addition, the deletion of EC2 shows that this enhancer cluster is not essential for the spatio-temporal regulation of the *Grem1* expression (Figure 27d; Figure 28d,e). In summary, the genetic analysis shows that EC1 and EC2 contribute in an additive manner to *Grem1* transcription level while the enhancers belonging to EC1 regulate the spatio-temporal kinetics of the *Grem1* expression (see before). In addition, EC2 has a general function in maintaining the chromatin architecture within the *Grem1* TAD, while in *EC1*-deficient limb buds, the intra-TAD contacts with the *Grem1* promoter are not reduced but elevated (Figure 28d). Furthermore, the inactivation of CRM5 in the context of the EC1 deficiency causes tetradactyly, which is paralleled by more severe spatial changes and reveals the loss of robustness of the limb bud patterning system upon deletion of both CRM5 and EC1 (Figure 27f). Previous studies showed that anteriorization of *Grem1* expression can cause preaxial polydactyly phenotypes (Capdevila et al., 1999; Lopez-Rios et al., 2012; Merino et al., 1999), however no such phenotypes were detected. The observed anterior expansion of the *Grem1* domain is, apparently, not sufficient to prevent the BMP activity from limiting digit numbers (Norrie et al., 2014).

This genetic analysis reveals both additive and cooperative interactions between individual CRMs and the two ECs. Therefore, it might be too simplistic to categorize individual enhancers/repressors and the enhancer cluster as essential primary or shadow enhancers (Cannavo et al., 2016; Degenhardt et al., 2010; Frankel et al., 2010; Montavon et al., 2011). All CRMs have specific activities and inherent properties, but their functions are likely context-dependent and the underlying compensatory mechanisms become apparent only when the genomic context is altered. During limb bud outgrowth, *Grem1* expression levels are maintained by the additive CRM activities and intra-TAD enhancer interactions (Figure 28). My analyses also provide evidence that some CRMs can adapt their functions depending on genetic background (e.g. CRM3 and CRM4) and that intra-TAD interactions and adjustments are likely part of the compensatory mechanisms that provide spatial *Grem1* expression with robustness. However, the anterior-distal and distal *Grem1* expression could be

altered by some of the single CRM deletions without leading to digit phenotypes. In contrast, *Grem1* expression in the posterior limb mesenchyme is more robust (Figure 42).

In *EC1EC2*$^{\Delta/\Delta}$ double mutant limb buds, low level residual *Grem1* expression is observed by WISH, but this results in a *Grem1* null digit phenotype. This residual expression could be driven by low level cryptic enhancer activity but none of the potential CRM located outside the EC1 and EC2 clusters displays significant enhancer activities. It might also be a possibility that in *EC1EC2*$^{\Delta/\Delta}$ double mutant limb buds, the proximal *Grem1* promoter still shows basal activity. If this was the case, this could be further quantified in cell-based luciferase assays.

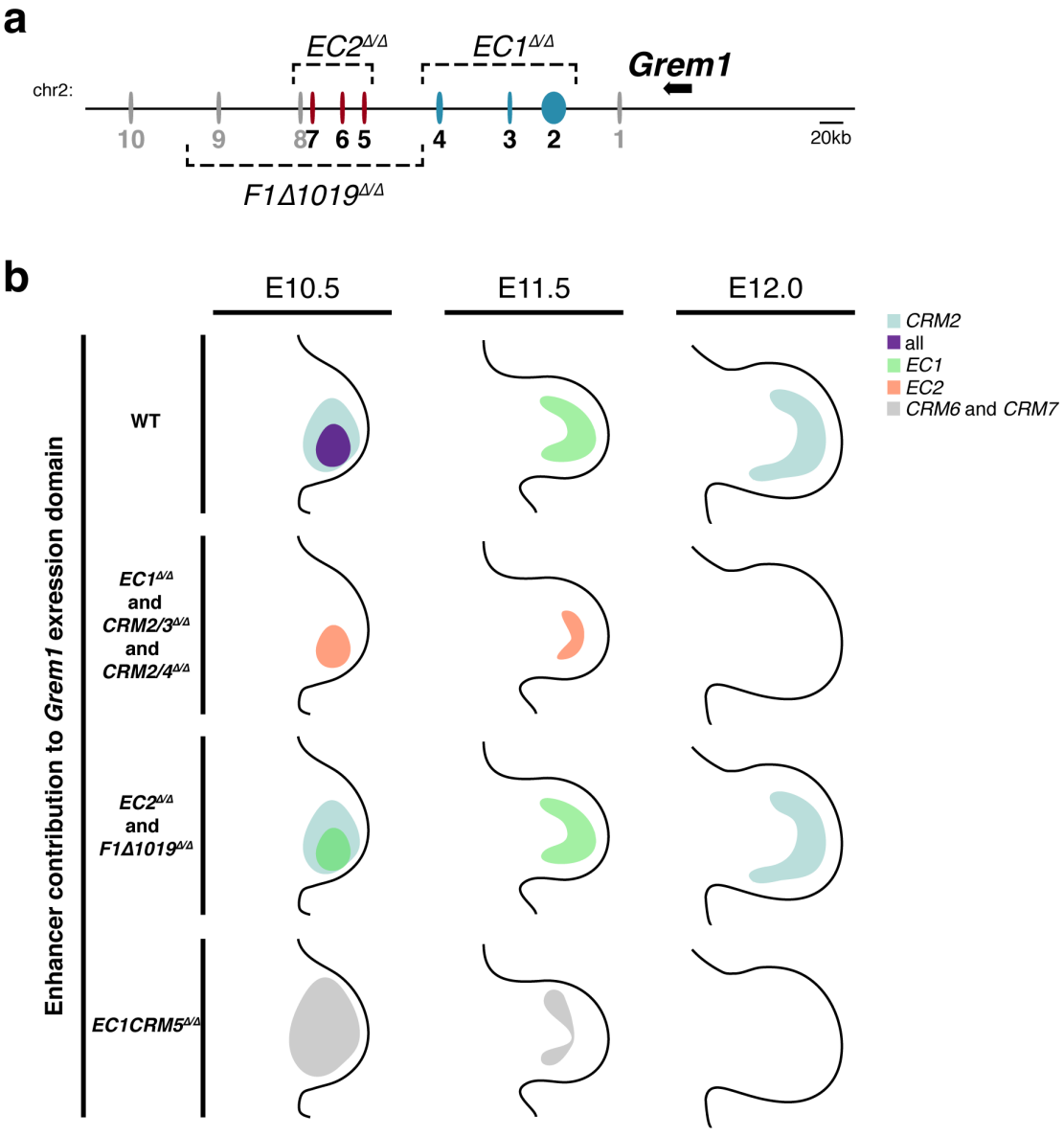


Figure 42: Schematic of CRM enhancer contributions to the spatio-temporal kinetics of *Grem1* expression. (a) schematic of *Grem1* TAD, EC1 CRMs (blue ovals) and EC2 CRMs (maroon ovals) are represented together with the other TADs CRMs (grey ovals). Dotted lines represent *EC1*, *EC2* and *F1Δ1019* deletions. (b) Schematic of the enhancer contribution to *Grem1* expression domain in different compound mutant limb buds.

Our data show how the CRM enhancers function, in an additive manner to regulate *Grem1* transcription levels and in a cooperative manner in spatial regulation of the *Grem1* expression domain in the posterior limb bud mesenchyme. Analysis of individual and compound mutants has allowed me to define the minimal requirements for apparently normal digit development, which shows that maintenance of spatial expression is most relevant (Figure 43). Disturbing the cooperativity between CRM2 to CRM4 (EC1) and CRM5 caused a significant spatial changes that disrupted the robustness of the digit patterning system, resulting in such variable fusions of anterior digits (*EC1*$^{\Delta/\Delta}$) and tetradactyly phenotypes (*EC1CRM5*$^{\Delta/\Delta}$).

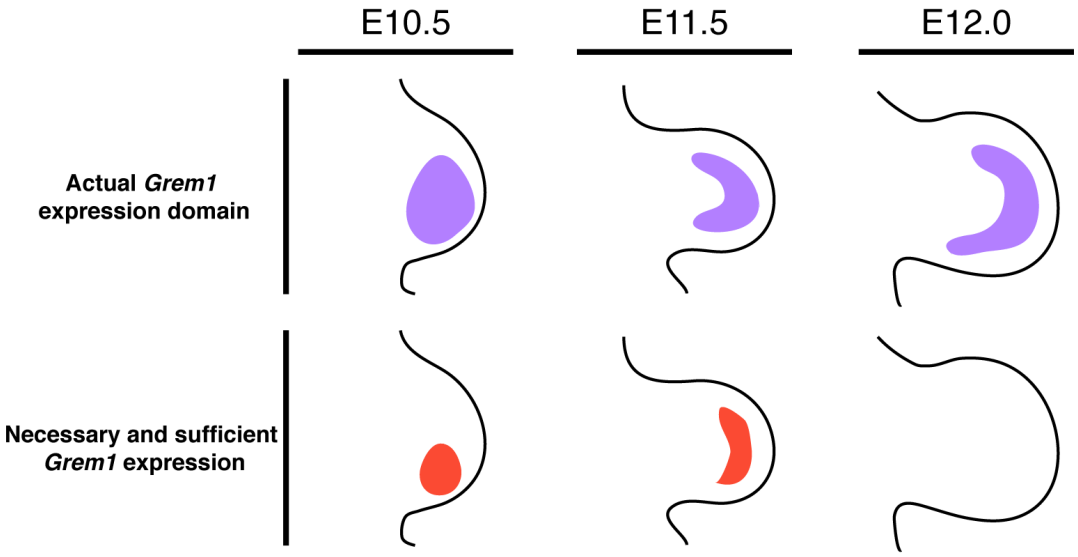


Figure 43: Schematic of the normal *Grem1* expression domains in wild-type limb buds in comparison to the minimal *Grem1* expression required for normal digit development. Upper panel: *Grem1* expression (purple) in wild-type limb buds. Lower panel: *Grem1* expression (orange) that ensures proper limb bud outgrowth and digit patterning (levels can be reduced to with 20% of wild-type *Grem1* transcripts).

The activities of individual *Grem1* CRMs are integrated in different ways. The *Grem1* transcript levels are regulated in an additive manner, while the spatial dynamics are regulated in a synergistic manner by the functionally relevant CRMs. The robustness of the limb patterning system relies more on the spatial dynamics than the overall transcript levels and on the different and partially overlapping functions between CRM2 to CRM5. These CRMs do not operate as shadow enhancers, but each of them has distinct and essential functions on their own or in interaction with the others.

7.3. Novel insights into the self-regulatory SHH/GREM1/AER-FGF feedback signaling system regulating limb bud development

In addition to analyzing the impact of the different CRM deficiencies on *Grem1* expression, I also studied the impact of these alterations on the feedback signaling system (Benazet et al., 2009; Khokha et al., 2003; Verheyden and Sun, 2008). The compound *Grem1* null mutants with digit phenotypes (*EC1EC2*$^{\Delta/\Delta}$ and *Cis*$^{\Delta/\Delta}$) show a significant reduction in SHH pathway activity (Figure 33a,b) as predicted by the feedback loop operating between SHH and GREM1 (Benazet et al., 2009). In particular, *EC1EC2*$^{\Delta/\Delta}$ limb buds showed a significant decrease in *Gli1* transcripts, reflecting the reduced SHH pathway activity, while BMP activity was not affected (Figure 33). Analysis of the allelic series of CRM mutants shows that the downstream consequences are rather complex. In some alleles, SHH pathway activity is significantly impaired without effects on the BMP pathway activity while in others, high BMP activity did not concur with reduced SHH pathway activity. These observations suggest that there are additional or alternative interactions, that regulate the interactions between the SHH and BMP pathways, in the self-regulatory limb bud signaling system (Benazet et al., 2009). It is important to remember that GREM1 is a secreted antagonist that impacts BMP ligands and not an intracellular antagonist. In certain cancer types, GREM1 has been shown to act in a TGF-β dependent manner

(Grillo et al., 2016; Lavoz et al., 2015; Marquez-Exposito et al., 2020; Mitola et al., 2010; Yin et al., 2017). At E11.0, the TGF-β signaling, together with the BMP pathway, promotes chondrogenesis (Karamboulas et al., 2010). As a consequence, this pathway, known for its cross-regulatory interactions with the *Wnt*, *Bmp* and *Shh* pathways, could play a role in the response to alteration in *Grem1* expression (Castellone and Laukkanen, 2017; Zhang et al., 2016).

In addition, my results suggest that *Id1* is not a good transcriptional readout of BMP pathway activity during limb bud outgrowth. In fact, I failed to observe strong correlations between *Id1* transcript levels, spatial changes and digit phenotypes. Furthermore, low *Grem1* expression levels did not consistently result in elevated BMP pathway activity during limb bud outgrowth (Figure 32 and Figure 33a). It is important to note that Benazet et al. (2009) showed that the GREM1/BMP initiation loop is active predominantly during the onset of limb bud development (Benazet et al., 2009). Transcriptional profiling of *Grem1*-null mutants through RNA-seq and searching for other BMP target genes among the differentially expressed genes is one way to find alternatives to *Id1*. An alternative approach to gain insight into how the progressive reduction of *Grem1* expression by CRM deletions affect the limb bud signaling system would be to genetically reduce the levels of other components (*Shh, Bmp*) in the most relevant *Grem1* enhancer mutants and assess the effects on limb bud development and digit patterning. Alternatively, additional genes and regulatory pathways might be part of the signaling system. For example, the group has been able to show that the *Wnt* pathway is required for *Grem1* normal expression during limb bud development (E. Unal Ph.D. thesis, Sheth et al., manuscript in preparation). Furthermore, the 5'HOX genes are required for *Grem1* expression (Sheth et al., 2013) and *Tbx2* is a BMP induced transcriptional regulator that inhibits *Grem1* expression (Farin et al., 2013; Nissim et al., 2006). I have generated preliminary data showing that the *Tbx2* expression domain varies. In mutants with decreased *Grem1* expression, but no limb phenotypes, *Tbx2* remains restricted, while in *Grem1* mutants with digit phenotypes, its limb bud domain is expanded. Therefore, there could be a balance between the TBX2-mediated *Grem1* repression and the spatial alterations in CRM mutant limb buds that result in digit phenotypes (Appendixes Figure 10).

It is clear that further analyses of the GRNs regulating limb bud outgrowth and patterning are necessary to understand how spatial alterations in *Grem1* expression correlate with digit phenotypes. When *Grem1* expression is reduced, rather than completely disrupted, the feedback loop could rewire, or additional gene could impact phenotypic outcomes. ChIP-seq analyses show at least two transcription factor complexes interact with the relevant CRMs and one can expect that additional TFs participate in *Grem1 cis*-regulation by interacting with specific CRMs. A deeper understanding of *Grem1* regulation in limb buds could identify additional players of the self-regulatory feedback signaling system. Analyzing total full limb buds may miss subtle events at the cellular level. This can be addressed by single-cell transcriptomics of LMPs to characterize their gene expression dynamics at cellular resolution in wild-type and select CRM mutant limb buds (Cao et al., 2019; Feregrino et al., 2019). This approach would allow the discovery of auxiliary factors/pathways (e.g. TBX2 or *Wnt* signaling components) impacting the self-regulatory feedback signaling system. In addition, ChIP-seq experiments in enhancer loss of function mutants in comparison to wild-type limb buds could help to establish alterations in TFs footprints, which would provide a better understanding of the *trans*-regulatory control of *Grem1* CRMs (Hung et al., 2019; Lee et al., 2018).

Analysis of key genes, in the self-regulatory feedback signaling system, does not fully explain the digit phenotypes observed in the allelic *Grem1* enhancer mutant series. This suggests that additional factors and regulatory interactions underlie the molecular alterations and phenotypes observed.

7.4. *Grem1* phenotypes: an explanation at the cellular level?

Prior to our research, genetic analyses had associated *Grem1* inactivation and subsequent precocious rise in BMP activity with apoptosis in the limb bud mesenchyme at E11.0 (Michos et al., 2004). In addition, at E10.5-E10.75, a small fraction (9 $\pm$ 21%) of distal-posterior limb bud progenitor cells, the SOX9-negative and

JAGGED1 (JAG1)-positive population (S9^{-}JAG1^{+}) of LMPs had been shown to undergo apoptosis in the *Grem1*$^{\Delta/\Delta}$ mutants due to precocious exposure to elevated BMP pathway activity (Reinhardt et al., 2019). This study concluded that the protection of these cells from BMP activity by the GREM1-mediate BMP antagonism is necessary to preserve the JAG1^{+} LMPs, which in turn could be essential for the specification of the skeletal condensation resulting in digit ray formation. Indeed, in all single and compound *Grem1* enhancer mutants with digit phenotypes, the expression of *Grem1* is more restricted as limb bud outgrowth progresses and terminates precociously, which indicates that the survival JAG1^{+} LMPs and their descendants could be affected, which in turn could alter the response to SHH signaling and digit identities, especially as a fraction of JAG1^{+} LMPs arises from *Shh*-expressing cells (Reinhardt et al., 2019). In two mutant alleles, *EC1*$^{\Delta/\Delta}$ and *Cis*$^{\Delta/\Delta}$ limb buds, BMP pathway activity is increased during limb bud outgrowth (Figure 31e and Figure 33b), but only in the *Cis*$^{\Delta/\Delta}$ limb buds abnormal apoptotic cell clusters are detected (Figure 35d). The same apoptotic profile is also observed in *EC1EC2*$^{\Delta/\Delta}$ limb buds, except the BMP pathway activity is not elevated (Figure 33a and Figure 35d). Furthermore, no abnormal BMP activity nor increased apoptosis is observed in *EC1CRM5*$^{\Delta/\Delta}$ limb buds, which are tetradactylous (Figure 27f and Figure 32b). As discussed before, a compensatory effect by another unidentified BMP antagonist active in limb buds cannot be excluded (Appendixes Figure 11). It is not clear how these findings fit the well-established facts that aberrant BMP activity, as a consequence of *Grem1* LOF, induces limb bud mesenchymal apoptosis (see e.g. Bastida et al., 2004; Michos et al., 2004). One additional feature of *Grem1* null mutants is a significant decrease in SHH pathway activity as measured by the *Gli1* transcriptional sensor (Figure 33). However, in CRM mutants with reduced *Grem1*, there is no clear correlation between SHH pathway activities and digit phenotypes either.

During handplate formation, the autopod shape of all enhancer mutants with digit reductions is apparent as primordia appear smaller. As I could not correlate these shape alterations with apoptosis, we assessed if the proliferation rates of CRM mutants with digit phenotypes correlated with the different autopod deformities (Boehm et al., 2010). Indeed, this analysis indicates that the limb mesenchymal cell proliferation/division rates are lowered in all enhancer mutants with digit reductions

(Figure 37a). Reduced proliferation accompanied by the apoptotic loss of limb bud mesenchymal progenitors is the likely cellular phenotype of enhancer mutants in which there is no compensatory mechanism for the reduction and/or loss of *Grem1*, resulting in digit reductions (e.g. *EC1EC2*$^{\Delta/\Delta}$ and *Cis*$^{\Delta/\Delta}$). These results show that it will be important to understand how the heterogeneity of the limb bud mesenchyme cellular population and their individual response to the different limb signaling pathways are impacted by the reduction or loss of the *Grem1* expression. For limb buds, cellular differences in *HoxD* gene cluster expression have been identified using single-cell transcriptomics, which highlights the transcriptional diversity of limb bud progenitors (Fabre et al., 2018). Considering a similar mechanism for *Grem1* expression, limb bud progenitor may regulate *Grem1* and other genes in the GRN and limb bud signaling system in a cell-type specific manner, which would add another level of complexity to transcriptional regulation.

Loss or alteration of *Grem1* expression in limb buds seems to affect the mesenchymal progenitor populations giving rise to the skeletal elements. In addition to apoptosis, I have identified changes in proliferation of LMPs. This could provide a possible explanation of the observed digit phenotypes. It is important to consider, that different limb bud progenitor populations may be affected differently depending on their transcriptional signature. Also, it cannot be excluded that the progenitors for missing digits never emerged.

7.5. Alteration of the *Grem1 cis*-regulatory landscape underlying congenital malformations and disease

In humans, TADs encoding less than three genes, such as is the case for the *Grem1-Fmn1* genomic landscape, appear more prone to mutations resulting in congenital malformations and disease (Muro et al., 2019). Deletions in the *Grem1 cis*-regulatory landscape have been associated with some rare and recessive congenital limb malformations, in both humans and mice (Dimitrov et al., 2010; Jackson-Grusby et al.,

1992; Maas et al., 1991; Zuniga et al., 2004). The *Grem1-Fmn1* locus in humans is located on chromosome 15q 13.3 (Jaeger et al., 2012) and mutations have been linked to the Cenani-Lenz (C-L) syndrome (Al-Qattan and Alkuraya, 2019). This recessive syndrome is characterized by oligosyndactyly, metacarpal synostosis, phalangeal disorganization and other variable clinical features similar to *Grem1*-deficient mice (Al-Qattan and Alkuraya, 2019; Grzeschik, 2002; Michos et al., 2004; Shi et al., 2001; Zhou et al., 2009). In particular, a 263 kb homozygous deletion in the human *Fmn1* gene and duplications upstream of *Grem1,* have also been linked to the C-L syndrome (Dimitrov et al., 2010). Such duplications can cause aberrant expression of *Grem1* in patients affected by the hereditary mixed polyposis syndrome, which predisposes them to colorectal carcinomas (predisposition to colorectal carcinoma, Jaeger et al., 2012; Jaeger et al., 2008; Venkatachalam et al., 2011). Our research provides novel insight into the mechanisms underlying these human congenital malformations and disease phenotypes. The 263 kb deletion in the *Fmn1-Grem1* landscape (Dimitrov et al., 2010) corresponds to the deletion of the region orthologous to the mouse EC2 enhancer cluster but leaves the EC1 cluster intact. In addition, this deletion also encompasses the *Fmn1* TAD and the downstream *Grem1* TAD boundary with its CTCF sites. Our 4C experiments had clearly highlighted that 1) only CRMs within *Grem1* TAD interact with its promoter at high frequency and 2) *Grem1* promoter makes the strongest contacts within its TAD boundaries (Figure 28). The boundary loss in the human loci might redirect contacts of *Grem1* CRMs to neighboring genes, which may, in turn, disrupt the *Grem1* expression and/or cause unspecific enhancers to alter *Grem1* expression during embryonic and limb bud development. It is likely that this deletion would disrupt the human *Grem1* ECh and thereby chromatin architecture. The predisposing duplications upstream of *Grem1* affected the upstream *Grem1* TAD boundary and CTCF sites (Jaeger et al., 2012). This 40 kb duplication results in a *Grem1* overexpression in colonic crypts and predisposes to colorectal cancer. This overexpression of *Grem1* points to 1) *de novo* enhancer activity upstream of the *Grem1* gene and 2) ectopic *Grem1* expression and GREM1 secretion in cells, promoting an abnormal survival/proliferation. *Grem1* TAD boundaries prevent such an excessive *Grem1* transcription by limiting *Grem1* promoter upstream contacts to the CRMs within its TAD. Therefore, this type of genetic alteration results in TADopathies

(Matharu and Ahituv, 2015). Dimitrov et al. (2010) have also identified a 1.7 Mb duplication, including the entire *Fmn1-Grem1* landscape that results in a C-L-like syndrome, but without a predisposition to colon carcinomas.

The analyses of our data bring new perspectives in the understanding of human congenital limb malformations and *Grem1*-related cancer.

8. Conclusions and Outlook

In the past ten years, tremendous advances have uncovered the large scope of the mechanisms associated with transcriptional regulation. When starting my Ph.D. project, I aimed to gain a molecular understanding of the transcriptional regulation of the *Grem1* gene during mouse limb bud development. I used a variety of genome-wide molecular and genetic approaches to identify and study the functions of the *Grem1*-associated CRMs during limb bud development.

My research established that all functionally relevant CRMs for *Grem1* regulation during limb bud development are located within a critical region in the *Grem1* TAD. The functionally relevant CRMs with enhancer activity are organized in two enhancer clusters, EC1 and EC2. These CRMs interact with the *Grem1* promoter to control *Grem1* transcript levels in an additive fashion and the spatial expression dynamic in a cooperative manner. For one of the CRMs, an essential function in regulating the temporal kinetics during handplate development was identified. By analyzing individual and compound mutant mice, it becomes apparent that CRMs have both specific and context-dependent functions that depend on 1) genetic alterations and 2) the activity of the other CRMs in the landscape. Furthermore, intra-TAD interactions are altered by genomic modifications and likely contribute to compensatory efforts by increasing or decreasing existing contacts and also by the *de novo* recruitment of regulatory regions. As a consequence, individual CRMs may appear dispensable when removed individually, but essential and cooperative functions are uncovered in compound mutants that result in larger alterations of the *Grem1* TAD structures. These interactions confer robustness to the *cis*-regulation of the *Grem1* expression. This robustness does not entirely rely on redundancy among CRMs but synergies from coordinated CRM interactions, that provide in particular the spatio-temporal *Grem1* expression with robustness. The progressive loss of spatial robustness in *Grem1* expression eventually results in increasingly more severe digit reductions and loss (ranging from mild soft-tissue fusion to tridactyly). In addition to mesenchymal apoptosis during limb bud outgrowth, alterations of cell proliferation correlating with

the severity of digit defects were identified. Through this study, I gave evidence that the existing limb self-regulatory system might be incomplete and does not apply to all *Grem1* mutants. My Ph.D. thesis research has provided deep molecular insights into the *cis*-regulatory robustness underling the spatial dynamics of the *Grem1* expression. Furthermore, I shed light on the functional redundancy, synergy and hierarchy of ECs in the regulation of a single gene during embryonic tissue development. Another novel finding of my studies is how CRMs cooperate in different ways to maintain expression levels and the spatial domain of *Grem1*. The functional analysis of the *Grem1* CRMs is directly relevant to a better understanding of human C-L syndromes and *Grem1*-related tumors. Another fascinating field that will profit from my studies is the evolution and diversification of the deeply conserved CRMs that regulate *Grem1* expression in tetrapods. For example, I have designed a compound mutant mouse strain lacking several CRMs (*EC1CRM5*$^{\Delta/\Delta}$) which results in spatial changes in *Grem1* expression, a trapezoidal limb bud shape and tetradactyly that is highly reminiscent to what is observed for limb buds of artiodactyl species such as pig embryos (Jonas Malkmus, personal communication). We have initiated studies to test the possible association between the limb diversity from different vertebrates and the evolution of the *Grem1* *cis*-regulatory landscape.

In summary, my Ph.D. project has uncovered the complex *cis*-regulatory mechanisms that confers robustness to the dynamic regulation of *Grem1* expression, which is key to normal mouse limb bud development.

9. Materials and Methods

9.1. Mouse husbandry and embryo analysis

9.1.1. Ethics statements

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

9.1.2. Mouse strains

For this study, the following mouse strains were used:

Experimental Models: Organisms/Strains	Reference	Scientist designer of the deletion
Mouse : *Tg-GLE(tg/+)*	EUCOMM project ID: 23892	European Conditional Mouse Mutagenesis program
Mouse: *ΔCRM2**	this study	N. Riesen/ A. Zuniga
Mouse: *ΔCRM2CRM3**	this study	N. Riesen
Mouse: *ΔCRM2CRM4**	this study	N. Riesen
Mouse: *ΔCRM2CRM5**	this study	L. Ramos Martins/ N. Riesen
Mouse: *ΔCRM5**	this study	L. Ramos Martins
*Mouse: ΔEC1**	this study	N. Riesen

Mouse: *ΔEC1CRM5**	this study	L. Ramos Martins/ N. Riesen
Mouse*: ΔEC1EC2**	this study	L. Ramos Martins/ N. Riesen
Mouse: *ΔEC2**	this study	L. Ramos Martins
Mouse: *F1Δ1019**	this study	L. Ramos Martins
Mouse: *F1Δ1024*	*Zuniga et al., 2004*	A. Zuniga
Mouse*: Grem1* null	*Michos et al., 2004*	A. Zuniga
Mouse: Smad4^{3xF}	*Gamart et al., unpublished*	F. Laurent/ A. Zuniga

*: strains generated in the laboratory (see Table 9.5.3).

All mouse alleles produced during this study were generated either in the laboratory by aggregation chimera (Aline Baur, Dr. Julie Gamart, Laurène Ramos Martins, Nathalie Riesen, Dr. Robert Reinhardt, PD.Dr. Aimée Zuniga) or at the Center for Transgenic Model (CTM) of the University of Basel by pronuclear injection of the CRISPR/*Cas9* protein and guides. Double mutant mouse lines were generated by re-targeting ESCs carrying one of the two deleted alleles (namely EC1 and CRM2) for the other allele (namely EC2, CRM3, CRM4 and CRM5). All mice lines were maintained in a Swiss Albino background. All embryos and mice were genotyped by PCR for relevant alleles using primers listed in the table (Table 4).

9.1.3. Embryo collection and staging

All embryos, regardless of their species, were isolated in ice-cold PBS (140mM NaCl, 2.7mM KCl, 25mM Tris-HCl pH 7.5) except stated otherwise. The staging was performed by counting somite number. During this study, collection of forelimb buds was restricted to the handplate for ChIP and 4C assays.

9.1.4. Digoxigenin-labelled RNA probe preparation

To generate riboprobes for RNA in situ hybridization, the corresponding cDNA clone, obtained from the laboratory stock and previously described, were linearized using the required restriction enzyme and transcribed using the either T3, T7 or SP6 RNA polymerases (Roche or Fermentas) and Digoxigenin labeling mix (Roche 11277073910). Plasmid DNA was removed by digesting the reaction with 8 units of TURBO DNase (Invitrogen AM2238) for 45 minutes at 37°C. DIG-labelled RNA probes were purified in 0.015% linear polyacrylamide (LPA, Gaillard and Strauss, 1990) and 0.5M LiCl and precipitated with ethanol (EtOH) twice before dissolving in TE buffer (10 mM TrisHCl, pH 8.0,100 μM EDTA). For the WISH screen performed for Gamart et al. (unpublished), all cDNAs of interest were generated by PCR amplification using primers that include a SP6 binding site before DNase treatment as mentioned before (Gamart et al., unpublished). The DIG-UTP riboprobes were purified using the mini Quick Spin RNA columns (Quiagen).

The riboprobes were pre-heated at 85°C for 5 min and equilibrated in prehybridization buffer (50% deionized formamide, 5xSSC (0.75M NaCl, 0.075M sodium citrate dihydrate, pH adjusted to 4.5 with 1N HCl), 2% Blocking Reagent (Roche), 0.1% Tween-20; 0.5% CHAPS (Sigma C3023), 50 mg.ml^{-1} yeast RNA (Sigma R8759), 5mM EDTA, 50 mg.ml^{-1} heparin (Sigma H5515)) at 70°C. Probes in prehybridization buffer were stored at -20°C and re-used several times.

9.1.5. Whole-mount in situ hybridization (WISH)

Embryos were fixed in 4% paraformaldehyde (PFA) in PBS at 4°C overnight. The next day, they were rinsed twice in PBS-0.1% Tween-20 (PBT) and dehydrated in a gradient series of 25%, 50%, 75% and PBT:methanol (MeOH:PBT) before storage in 100% methanol at -20°C. Except if stated otherwise, all washes were performed at room temperature (RT) for 5 minutes. On the first day, embryos were rehydrated in reverse gradient series of 75%, 50%, 25% MeOH:PBT and washed twice in PBT before splitting them in half along the AP axis. Embryos were bleached in 6% hydrogen

peroxide (H_2O_2):PBT for 15 min and washed 3 times in PBT. Then, the embryos were treated for 15 min with 10μg.ml^{-1} proteinase K in PBT for mesenchymal expressed genes, 4 min 5μg.ml^{-1} proteinase K in PBT for AER expressed genes . To inactivate the proteinase K, embryos were washed with fresh 2mg.ml^{-1} Glycine in PBT. After 2 washes in PBT, embryos were post-fixed in fresh 0.2% glutaraldehyde, 0.1% Tween-20 in 4% PFA for 20 min, and rinsed in PBT twice. Then, they were equilibrated in 2ml of prewarmed prehybridization buffer at 70°C for at least 1hr. The prehybridization buffer was then replaced by 1ml of fresh prewarmed prehybridization buffer containing 10μl.ml^{-1} of digoxigenin-labelled RNA riboprobe (see above) and incubated overnight at 70°C.

On the second day, the riboprobe-containing prehybridization buffer was recovered, stored at -20°C and substituted by pre-warmed prehybridization buffer, followed by serial washes in 75%, 50% and 25% prehybridization buffer: 2xSSC (0.3M NaCl, 0.03M sodium citrate pH 4.5) at 70°C. Embryos were then washed twice in 2xSSC, 0.1% CHAPS for 30 min at 70°C on a rotating wheel. The excess of RNA probe was eliminated with a 20μg.ml^{-1} RNase A (Roche 10109169001) in 2xSSC,0.1% CHAPS treatment for 45 min at 37°C. This step was followed by 2 times for 10 min washes at RT in maleic acid solution (100mM maleic acid disodium, 150mM NaCl pH 7.5) before incubations in maleic acid for 30 min 2 times at 70°C. Next, embryos were washed 3 times in fresh PBST (140mM NaCl, 2.7mM KCl, 25mM Tris-HCl pH 7.5, 1% Tween-20). The samples were blocked in 10% lamb serum in PBST for at least 1hr after which the solution was replaced with a solution containing anti-digoxigenin- Alkaline Phosphatase Fab Fragments (Roche 11093274910) diluted 1:2000 in 1% lamb serum in PBST. Embryos were incubated overnight (O/N) at 4°C with gentle rocking.

The next day, embryos were washed 5 times in 0.1% BSA:PBST for 45 min and 2 times for 30 min in PBST. Embryos were then incubated 3 times for 10 min in NTMT (100mM NaCl, 100mM Tris-HCl pH 9.5, 50mM MgCl2, 1% Tween-20), and transferred into 1ml of prewarmed BM purple Alkaline Phosphatase substrate (Roche) at RT. The *in situs* were developed in the dark and checked roughly every 30 min. The development was stopped at an appropriate timepoint (before the background was too strong) by washing the stained embryos 3 times in PBT (10 min) and twice in PBS. Pictures were taken using a Leica MZ FLII stereomicroscope and a Leica Application

Suite V3 software or a Nikon SMZ25 stereomicroscope and a Nikon application NIS Elements Basic Research. Embryos were stored in 30% azide (in PBS) at (4°C).

9.1.6. Generation of *lacZ* reporter transgenic embryos

9.1.6.1. Constructs preparation for injection

CRMs analyzed in this study were amplified by PCR from genomic DNA extracted from mouse embryonic tissue (for primers see table 2). Amplified CRMs were then cloned into a hsp68-*lacZ* reporter vector (Pennacchio et al., 2006) using the Gibson Assembly ® Method (Gibson et al., 2009). The ligated products were dialyzed against milliQ water on MF nitrocellulose membrane filters (0.22µm pore size; Millipore Ref. VSWP01300). Transformation was performed into competent bacteria (XL1-Blue or DH5α) in Gene Pulser 0.2cm cuvettes with a Biorad MicroPulser. The colonies were plated on ampicillin supplemented LB-agar plates (prepared by Z. Boudebaba and S. Rabéhi). Plasmids were recovered using the NucleoBond Xtra Midi kit (Machery Nagel, ref. 740410.10). Positive bacterial colonies were) and digested by appropriate restriction enzyme to assess transgene insertion. Reporter constructs were linearized using the appropriate restriction enzyme and purified by phenol chloroform extraction in phase lock gel™ (VWR™).

9.1.6.2. Constructs pronuclear injection

Transgenic embryos were generated by pronuclear injection at the CTM. Pawel Pelczar and Heide Oller (CTM, Basel) injected the linearized reporter constructs in at least 150 C57BL/6 zygotes. The surviving embryo were transferred in pseudo pregnant females and then were collected at E11.0. Founder embryos were stained for LACZ activity as described below (9.1.6.3).

9.1.6.3. Whole-mount LACZ staining of mouse embryos

Embryos were fixed in 1% formaldehyde, 0.2% glutaraldehyde, 0.02% NP40 (Igepal CA-630, Cas. No: 9002-93-1 MERK), 0.01% sodium deoxycholate in PBS for precisely 20 min at 4°C. Then, they were washed 3 times 5 min in PBS at RT. The revelation step was done in the dark in 1mg.ml^{-1} X-Gal in dimethyl formamide, 0.25mM $K_3Fe(CN_6)$, 0.25mM $K_4Fe(CN_6)$, 0.01% NP40, 0.4mM $MgCl_2$. The reaction was regularly monitored and was stopped by washing 3 times in PBS at RT.

9.1.7. Skeletal preparation

Performed together by Aline Baur and Victorio Palacio. E14.5 embryos biopsies were taken for genotyping before fixation in EtOH 95% O/N. The following day, embryos were transcfered to acetone 100% O/N. The samples were then incubated 5 hrs to 6 hrs in alcian blue staining solution (15mg alcian blue 8GX (Sigma) in an 80:20 mix of EtOH:glacial acetic acid). Embryos were then rinsed for approximately 1hr and 2hr in an 80:20 mixture of EtOH:glacial acetic acid before an O/N incubation in EtOH 95%. The next day, embryos were cleared in 1% potassium hydroxide (KOH) from 45 min to 1 hr before counterstaining them with alizarin red (25mg.L^{-1} Alizarin red (Sigma) in 1% KOH) for 4 hrs to 5 hrs. Embryos were then cleared in 20:80 1% KOH: Glycerol for 1hr. The samples were washed through progressively higher ratios of 70%EtOH:Glycerol:water (1:2:7, 2:2:6, 3:3:4, 4:4:2, 5:5:0) O/N. Pictures were taken using a Leica MZ FLII stereomicroscope.

9.1.8. Cell death detection with Lyso Tracker assay

Uteri of pregnant females were dissected in prewarmed (37°C) HBSS buffer (Gibco, ref.55021C). Embryos were transferred into a glass vial with 4ml prewarmed (37°C) HBSS buffer and Lyso Tracker Red (5μL.ml^{-1}) dye that accumulates in the lysosome under acidic conditions during apoptosis, when lysosomes engulf apoptotic bodies, for 45 min at 37°C. Embryos were extensively washed 5 times 30 min with (37°C) HBSS buffer and fixed overnight in 4% PFA at 4°C. The next day, samples were dehydrated in a MeOH:PBT series. The embryos were cleared (to avoid diffraction of the fluorescent signal in Benzyl Benzoate/Benzyl Alcohol 2:1, BBBA) 30 min in 50% methanol/50% BB/BA and then left to sit in 100% BB/BA. The signal was detected by epifluorescence (Leica MZ FLIII fluorescence stereomicroscope).

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

9.1.9.1. Mouse embryonic fibroblasts culture

Mouse embryonic fibroblasts (MEF) were cultured at 37°C, 5% CO_2 in MEF medium (DMEM 4.5g.L^{-1} Glucose/3.5g.L^{-1} $NaHCA_3$ (Gibco), 10% Fetal Bovine Serum, 1μg.ml^{-1} Penicillin Streptomycin (Sigma), 2mM L-Glutamine (Sigma)). Medium was changed every 2 days and the MEF split every 2 to 3 days depending on their growth rate. Before splitting, cells were rinsed with 37°C PBS and incubated in Trypsin/EDTA (Sigma T-3924), for 5 min at 37°C 5% CO_2 as follow:

- 3ml for 10 cm plates
- 1ml for 6 well plates
- 100-150μL for 96 well plates

Cells were dissociated by gentle pipetting 3 times before stopping the trypsinization with MEF medium:

- 7ml for 10 cm plates
- 4ml for 6 well plates
- 100-150μL for 96 well plates

Then, the cells were resuspended in MEF medium and transferred to 15ml Falcon tubes and centrifuged for 5 min at 1200 rpm. After supernatant removal, cells were resuspended in MEF medium and plated at varying dilutions (1:3 to 1:4 depending on their density). MEFs were passaged a maximum twice after thawing.

9.1.9.2. MEF mitomycin C treatment

Confluent MEF plates were treated with 10μg.ml^{-1} of Mitomycin C (Sigma, M 0503) in MEF medium 2 hrs at 37°C, 5% CO2. The plates were rinsed 3 times with PBS before fresh medium was added. Growth-arrested MEF were maintained maximally 1 week after treatment ant their medium changed every two days.

9.1.9.3. Mouse Embryonic Stem Cells (ESCs) culture

ESCs from the G4 line were cultured on growth-arrested MEF at 37°C, 5% CO2 in ESC medium (DMEM (Gibco 41966029), 15% filtered FCS (PanBiotech, P30.3302), 100U Penicillin, 0.1mg.ml^{-1} Streptomycin (Sigma P-0781), 200mM L-Glutamine (Sigma G-7513), 0.1mM βMercapto-Ethanol (Gibco 31350-010), 107U.ml^{-1} EsGRO LIF (Gibco 13275-029), 1XNon-Essential Amino Acids (Gibco 11140-035), 100mM Sodium Pyruvate (Gibco 11360-39)). Medium was changed every day and ESCs split every 2 days to 3 days during deletion positive clone recovery period. Before splitting, cells were rinsed with DPBS and incubated in Trypsin/EDTA,15 min at 37°C 5% CO2:

- 3ml for 10 cm plates
- 1ml for 6 well plates
- 100-150μL for 96 well plates

Colonies were dissociated by gentle pipetting 3 times. Trypsinization was stopped with MEF medium:

- 7ml for 10 cm plates
- 4ml for 6 well plates

- 100-150μL for 96 well plates

The cells were transferred to 15ml Falcon tubes and centrifuged for 5 min, 1200rpm. After supernatant removal, cells were resuspended and plated onto Mitomycin treated MEFs in ES medium with a 1:3 dilution.

9.1.9.4. ES cell targeting

The CRISPR sgRNA (see Table 3) were designed using the CRISPOR program (see Table 5) and selected based on predicted specificity and efficiency scores. For transfection, G4 ES cells (ESCs) were plated in a 6cm dish coated with a monolayer of MEFs in ES medium -DMEM (Gibco 41966029), 15% filtered FCS (PanBiotech P30.3302), 100U Penicillin, 0.1mg.ml^{-1} Streptomycin (Sigma P- 0781), 200mM L-Glutamine (Sigma G-7513), 0.1mM βMercapto-Ethanol (Gibco 31350-010), 107U.ml^{-1} EsGRO LIF (Gibco 13275-029), 1X Non-Essential Amino Acids (Gibco 11140-035), 100mM Sodium Pyruvate (Gibco 11360-39). The next morning, the medium of each plate was replaced by 1.75ml ES medium without antibiotics. In the late evening, the cells were transfected using FuGENE kit (Promega TM-328, (Kraft et al., 2015)). Puromycin-resistant plasmids (4μg, px459-variants) were diluted in OptiMEM-Glutamax medium in a final volume of 125μL. This DNA-OptiMEM mixture was combined with 125μL of FuGENE master mix (100μL of OptiMEM-Glutamax supplemented by 25μL of FuGENE reagent added directly to the liquid to prevent adhesion to plastic) and incubated 15 min at RT. This final mix of 250μL was added in a dropwise manner to the ESC. Precisely, 12 hrs later, the transformation was stopped by changing the medium. 24 hrs later, cells were split into 6cm dishes (2 wells with 1:3 dilution, 1 well with 1:6 dilution, 2 wells with 1:12 dilution) coated with mitomycin-treated DR4 puromycin-resistant feeders. Cells positives for the lipofection were selected using puromycin (1.25μg.ml^{-1}, Sigma-P8833) for 48 hrs. The selection was stopped by replacing the puromycin medium by ESC medium. PCR was used to screen for the correct CRISPR deletion. The surviving clones were picked and expanded. The selected clones were used to generate transgenic mice by aggregation chimera.

9.1.9.5. Superovulation of oocyte donors and priming of pseudo-pregnant females

Female Swiss Albino mice 13 weeks old (approximately 10-20 females) were injected intraperitoneally with 25G needle equipped syringes filled with 5 I.U PMSG (Pregnant Mare Serum Gonadotropin – Pregnyl from Organon). Sybsequently, 48 hrs after the first injections, all females were injected with 5 IU hCG (human chorionic gonadotropin Folligon from Intervet). PMSG and hCG were dissolved in sterile PBS to 50 $IU.ml^{-1}$ and aliquots of 1ml were stored at -20°C. Matings with Swiss Albino males were set immediately after the second injection.

The day after the first injection, 2 months old females were primed by adding bedding from a fertile Swiss Albino male in their cages. Two days later, these females were mated with vasectomized Swiss Albino males Only females presenting a clear plug on the day after the mating were selected as recipients of for the aggregated embryos.

9.1.9.6. ESCs preparation for aggregation

Five days later, ESCs from a confluent 6cm dish were trypsinized and dissociated by pipetting up and down in ES medium. Cells were pre-plated on gelatin-coated dishes for 45 min 37°C, 5% CO2, to remove an excess of MEF. ESCs were collected by centrifugation and counted. After centrifugation (1200rpm, 5 min), ESCs were re-suspended to $5x10^{6}cells.ml^{-1}$ in filtered (0.20μm) aggregation medium (18ml DMEM $4.5g.L^{-1}$ glucose (Gibco ref.41966029), 66mg Calcium lactate (Sigma, ref.21185) and 10% FCS). Drops of 50μL ESCs were made on a 10cm Petri dish, then covered by mineral oil (Sigma, ref.M5310). Drops were incubated 10 min before aggregation.

9.1.9.7. Embryo collection for aggregation

Performed by Aline Baur, Dr. Julie Gamart, Dr. Robert Reinhardt, Nathalie Riesen and PD.Dr. Aimée Zuniga. All the solutions were pre-warmed at 37°C. Embryos were manipulated using a mouth pipette. Except if stated otherwise, all embryo

manipulation occurred in M2 medium. On day 5, embryos (E2.5, morula stage) were flushed out of uteri by injecting M2 medium through the infundibulum. They were pooled in a 3cm dish in M2 Medium. At this stage, a selection was performed to keep only live embryos with a nicely formed zona pellucida and 8 to 16 blastomeres. These embryos were then treated with Tyrode's acid solution (Sigma T1788) for a few seconds to remove the zona pellucida. Then they were rinsed 3 times in M2 medium.

Composition of the M2 Medium:
6.4mg.ml^{-1} NaCl (Merck, ref.1064041000), 350μg.ml^{-1} NaHCO3 (Merck ref.1063290500), 36μg.ml^{-1} Na-Pyruvate (Merck, ref.1066190050), 50μg.ml^{-1} Streptomycin. Sulf. (Sigma, ref.56501) 160μg.ml^{-1} KH2PO4 (Merck 1048731000), 465μg.ml^{-1} Ca-Lactate.3H20 (Sigma, ref.44388), 356μg.ml^{-1} KCl (Merck 1049361000), 294μg.ml^{-1} $MgSO_4$.7H20 (Merck, ref.1058860500), 1mg.ml^{-1} Glucose (Sigma, ref.G8270), 621μg.ml^{-1} HEPES (Sigma 54457), 75μg.ml^{-1} K-PenG (Sigma, ref.P7794) and 4mg.ml^{-1} BSA (Sigma A3311) in Aqua ad inject (Braun/Aichele Medico, ref.530108).

9.1.9.8. ES cell-embryo aggregation and preparation for transfer

Performed by Aline Baur, Dr. Julie Gamart, Dr. Robert Reinhardt, Nathalie Riesen and PD. Dr. Aimée Zuniga. This section of the material and method was written by Dr. Julie Gamart. 10 embryos were gently placed in a drop of ESCs; without touching the monolayer of ESCs and incubated for 2 hrs at 37°C 5% CO2. During this time, Petri dishes with 50μL drops of KSOM (Millipore MR^{-1}06-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 ESCs. Ten aggregated embryos were placed in a drop of KSOM, well separated from each other and incubated overnight at 37°C in 5% CO2. Around 20 hrs after aggregation, embryos were mostly blastocysts and were ready for transfer. They were washed once in prewarmed M2 medium and

kept in M2 at 37°C until the transfer. Embryo transfer was performed by Julie Gamart or Aline Baur.

9.1.9.9. Zygote CRISPR/*Cas9* targeting

The CTM of the University of Basel performed microinjection of the CRISPR/*Cas9* protein and the single guides RNA (see Table 3) in at least 150 C57BL/6 WT or mutant (*CRM2*$^{\Delta/+}$, *EC1*$^{\Delta/+}$ zygotes). The surviving embryos were transferred into pseudo pregnant females. Founder littermates were screened using PCR and the mice positive for the deletion were sequenced to determine the exact coordinates of the deletion. The founders carrying the desired deletion were outbred in Swiss Albino background for 2 generations before further analysis.

9.2. Molecular biology

9.2.1. RNA extraction and cDNA synthesis

RNA was extracted from embryonic limb buds (40-42ss) dissected in ice-cold PBS. After PBS removal, samples were covered with 100μl of RNAlater® (R0901, Sigma-Aldricht) and stored at -20°C until extraction. Total RNA was extracted using the RNeasy Micro Kit (Qiagen). The PrimeScript^{TM} 1ST strand cDNA synthesis kit (TAKARA Bio Inc.) was used to synthetize cDNA. Total RNA (<5μM) was incubated with 50μM Oligo dT Primers and 10mM dNTP mixture to a total volume of 12μl of annealing reaction, for 5 min at 65°C and immediately cooled on ice. The 12μL of annealed product were retrotranscribed in presence of 200U of PrimeScriptRTase (200U.μL^{-1}), 20U of RNase Inhibitor (40U.μL^{-1}), 1X PrimeScript Buffer to a total volume of 20μL at 50°C for 1h. The enzyme was inactivated at 75°C for 15 min and immediately cooled on ice. cDNAs were stored at -20°C before amplification by RT-qPCR.

9.2.2. Real Time-quantitative PCR (RT-qPCR)

RT-qPCR reactions were performed with the Bio-Rad CFX96 Real-Time PCR system using iQ SYBR Green Supermix (Bio-Rad). PCR reactions were set in 20μL containing 0.3μM each primer (see Table 5), 50% SYBR green and either 0.5μL cDNA or 1μL ChIP/Input samples (for ChIP-qPCR). The program for qPCR was: 95°C 3 min 95°C 10sec 13 cycles 60°C 1 min 95°C 10sec 65°C 5sec 95°C 5sec (Benazet et al., 2009). For gene expression analysis, the ribosomal protein L19 (RPL19) transcript was used as normalizer. Relative Cq values of the target transcripts were normalized to the Cq values of RPL19, and normalized fold expression level ($2^{-\Delta\Delta Cq}$) are shown as mean ± SEM. (for primers see Tables 6 and 7)

9.2.3. Chromatin Immunoprecipitation and sequencing (ChIP-seq)

9.2.3.1. SMAD4^{3xFlag} ChIP-seq

The ChIP experiments targeting the Smad4$^{3xF/3xF}$ were performed by Dr. Julie Gamart (Gamart et al., manuscript in preparation) who taught me this protocol. This protocol is adapted from Visel et al., 2009a for tissue collection and crosslinked chromatin preparation, on Kim et al., 2007 and Vokes et al., 2008 for subsequent steps, and had been modified and optimized by Dr. Marco Osterwalder, Dr. Javier Lopez-Rios and Dr. Julie Gamart.

Embryos were collected in ice-cold PBS, and dissected forelimbs from embryos at E9.75 (28-31ss), E10.5 (33-35ss) of the same genotype were pooled into 2ml Eppendorf tubes. Dissected tissues were washed twice with ice-cold DPBS w/$Ca^{2+}Mg^{2+}$ (Gibco) and transferred to a glass douncer (Tissue Grind Tube Size 2ml, Kimble-Chase) on ice. Tissue disaggregation was achieved by applying 25 strokes with pestle A (Tissue Grind Pestle LC 2ml, Kimble-Chase) followed by 25 strokes with

pestle B (Tissue Grind Pestle SC 2ml, Kimble-Chase). The nuclei solution was transferred back to a 2ml Eppendorf tube, and the douncer was rinsed with 300μL of cold DPBS w/$Ca^{2+}Mg^{2+}$ to recover the maximum amount of each sample. After centrifugation (3000rpm, 3 min, 4°C), supernatants were discarded and nuclei resuspended in 1.5ml DPBS w/$Ca^{2+}Mg^{2+}$ containing 150μL 11Xcrosslinking buffer (0.1M NaCl, 1mM EDTA, 0.5mM EGTA, 50mM Hepes pH 8.0; supplemented with 11% formaldehyde just before use) at RT The solutions were incubated 5 min at RT with gentle shaking on a horizontal platform. Crosslinking was interrupted with 75μl of 2.5M Glycine followed by a 5 min incubation at RT without shaking. The samples were then centrifuged (3000rpm, 3 min, 4°C), supernatants removed, and pellets resuspended in 1.5ml ice-cold DPBS w/$Ca^{2+}Mg^{2+}$. After centrifuging once again (3000rpm, 3 min,4°C), the supernatants were discarded and the nuclear pellets snap-frozen in liquid nitrogen (N_2) and stored at -80°C. 3 samples of crosslinked nuclei from similar amounts of dissected forelimbs were processed for ChIP in parallel per experiments. Before performing the actual ChIP, antibodies coupled to magnetic beads were prepared. For one ChIP sample, 20μl of Dynabeads® Protein G (Invitrogen) were used and transferred to a 2ml Eppendorf tube containing 1ml freshly prepared ice-cold BSA (5mg.ml^{-1} in DPBS w/$Ca^{2+}Mg^{2+}$). Beads were washed with 1ml cold BSA/PBS on a magnetic rack 6 times. Then they were resuspended in 2.5x the original volume of beads (50μl per sample) with BSA:PBS and transferred to a 2ml screw cap microtube (Starstedt). 2μg of anti-FLAG M2 antibody were used for each ChIP sample and incubated overnight on a rotating wheel (4°C).The following day, the Dynabeads-antibodies complexes were transferred back to a 2ml Eppendorf tube and washed 6 times with 1ml cold BSA:PBS before being resuspended in the original volume of suspended beads (20μl per sample) and kept on ice. In the meantime, snap-frozen crosslinked nuclear pellets were thawed on ice and resuspended in 6ml cold Lysis buffer (50mM Hepes pH 7.5, 140mM NaCl, 1mMEDTA pH 8.0, 10% Glycerol, 0.5% NP-40, 0.25% Triton X-100 (SIGMA-ALDRICHT), 1X Complete mini protease inhibitor cocktail (Roche)). Lysates were incubated for 10 min (4°C) with gentle rocking, followed by centrifugation (10 min, 2500rpm, 4°C). After discarding the supernatants, the pellets were resuspended in 2ml 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) and supplemented with 3ml of Protein Extraction Buffer at RT. The samples were incubated for 10 min at RT on a rocking platform and the nuclei pelleted by centrifuging (10 min, 2500rpm, 4°C). After discarding the supernatant, the nuclei were resuspended in 1ml ice-cold Chromatin Extraction Buffer (1mM EDTA pH 8.0,1460.5mM EGTA pH 8.0, 10mM Tris-HCl pH 8.0, 0.1% Na-Deoxycholate, 0.5% N-lauroylsarcosine, 3xComplete mini protease inhibitor cocktail). Sonication was proceeded in Covaris AFA Tubes Fiber & Cap 12x12mm (part number 520081, Covaris) with the S220 Covaris Ultra-sonicator. Each sample was sonicated for 15 min using the 5% duty cycle, 140 watts peak Incident power and 200 cycles per Burst. The sheared chromatin was then transferred to 1.5ml Eppendorf tubes, insoluble debris were pelleted by centrifugation (10 min, 13000rpm, 4°C). The supernatants were transferred to new 2ml screw cap microtubes, and the volume was topped up to 1.060ml with Chromatin Extraction Buffer. At this stage, 2x 30*µ*l aliquots were taken and separated to new Eppendorf tubes: one for input control and the other to assess the quality of the sonicated chromatin on an agarose gel. Aliquots were kept overnight (4°C) and treated in parallel with the ChIP sample. 300*µ*l of ChIP cocktail Mix (130*µ*l Triton X-100 10%, 3*µ*l Na-Deoxycholate 10%, 26*µ*l Complete protease inhibitor solution (from a 50x tablet dissolved in 1ml mQ H2O),131*µ*l TE buffer (100mM Tris HCl pH 7.4, 10mM EDTA pH 8.0) and 10*µ*l mQ H2O were added to each 1ml ChIP sample to adjust to RIPA buffer conditions (see below). Twenty microliters of the freshly washed Dynabead-antibody complexes were added to each ChIP samples and incubated overnight on a rotating wheel (4°C).The following day, the Dynabeads were transferred to 1.5ml Eppendorf tubes containing 1ml of freshly prepared ice-cold RIPA buffer (50mM Hepes pH 8.0, 1mM EDTA pH 8.0, 1% NP-40, 0.7% Na-Deoxycholate, 0.5M LiCl, 1XComplete mini protease inhibitor cocktail), and washed 6 times with 1ml cold RIPA buffer on a magnetic rack. The beads were rinsed with 1ml TE-plus (100mM Tris HCl pH 8.0, 10mMEDTA pH 8.0, 50mM NaCl, 1X Complete mini protease inhibitor cocktail) and centrifuged (3 min, 1000rpm, 4°C). All residual liquid was removed from the beads using the magnetic rack, and the beads were resuspended in 100*µ*l freshly prepared Elution Buffer (1mM EDTA, 1% SDS, 10mM Tris-HCl pH 8.0). The crosslinked protein-DNA complexes were eluted from the Dynabead-antibody complexes at 65°C for 15 min on heat blocks shaking at 1300rpm. Beads were then

centrifuged for 1 min at 1300rpm and retained in the tube using a magnetic rack, while 100μl of supernatant containing the protein-DNA complexes were transferred to a PCR microtube. In parallel, the 30μl input aliquot and the 30μl chromatin aliquot were transferred to PCR microtubes containing 120μl of Elution Buffer. Reverse crosslinking was performed for all three samples by an overnight incubation at 65°C (≥12 hrs).The samples were then transferred to 1.5ml Eppendorf tubes and treated with RNase A (from bovine pancreas, Sigma) in TE (100mM Tris HCl pH 8.0, 10mMEDTA pH 8.0) at a final concentration of 0.2μg.μl^{-1}. After 1hr incubation at 37°C, Proteinase K (0.2μg.μl^{-1}) was added and proteins were digested for 2 hrs at 55°C. Finally, DNA was purified from using the QIAquick Gel Extraction Kit (Qiagen). Elution was achieved in using twice 30μl Buffer EB, and the eluted ChIP and input samples were aliquoted and stored at -20°C before use for qPCR. Half of the purified chromatin control samples were loaded on an 1.5% agarose gel to estimate the range of DNA fragments sizes (between 150 bp to ~1.5 kb with a maximal intensity at ~300bp) and evaluate sonication quality. Libraries were made by the sequencing facility and Sequencing was done on an Illumina 2000 sequencer in a 50-cycles paired-end configuration (Genomic core facility, EMBL, Heidelberg). (for Antibody see Table 11)

9.2.3.2. ChIP-seq general protocol

For GLI3, ChIP-seq was performed on tissue obtained from homozygous mice carrying a 3XFLAG epitope knocked-in to the endogenous *Gli3* locus (Lopez-Rios et al., 2014; Lorberbaum et al., 2016). Briefly, stage-specific embryonic limb buds were collected in ice cold PBS and pooled prior to crosslinking in 1% formaldehyde/PBS for 20 minutes at room temp. Following tissue cross-linking samples were processed for ChIP as previously described (Peterson et al. 2012). DNA was purified using QiaQuick columns (Qiagen) and quantitated using a Qubit (ThermoFisher). Sequencing libraries were prepared according to manufacturer instructions (Takara). All samples were sequenced on NextSeq instrument (Illumina) to generate 1X75 bp reads. The resulting reads were mapped to genome assembly GRCm38 (mm10) using bwa (Li and Durbin, 2009). Peak calls were made relative to input controls using macs2 (Zhang et al. 2008)

callpeak function and default parameters with the following options: --Call-summits -B --trackline. (Paragraph written by Dr. Kevin Peterson)

All other ChIP-Seq and ChIP-qPCR experiments (see table 9.5.7 for primers) were performed by and Dr. Rushikesh Sheth and Jens Stolte. This protocol for embryonic tissue is adapted from Sheth et al., 2016 and I participated to the establishment of this protocol in the laboratory. In the morning, magnetic beads were coupled with the corresponding antibody. For one ChIP sample, 20μl of Dynabeads® Protein A and 20μl of Dynabeads® Protein G and were transferred to a 1.5ml Eppendorf tube containing 1ml freshly prepared ice-cold ChIP DIL buffer (Triton 1.0%, Tris pH8 10mM, NaCl 150mM , EDTA pH8 2mM) and nutated for 2 min. The tubes were briefly centrifuged before resting for 1 min on magnetic racks and washed as described earlier two more times. The appropriate antibodies (1 to 10μg) were resuspended in 1ml of ChIP DIL buffer and incubated 3 hrs to 6 hrs with the beads on the rotating platform (4°C).

Embryos were collected in ice cold PBS and dissected forelimbs from somites-matched embryos of the same genotype were pooled into 1.5ml low-bind Eppendorf's. The pellet was disaggregated in 1ml of buffer A (Triton 0.25%, Tris pH8 10mM, EDTA pH8 10mM, EGTA pH8 0.5mM) with 13 strokes of pestle A (Tissue Grind Pestle LC 2ml, Kimble-Chase), transferred to 1.5ml Eppendorf followed by 10 min incubation on ice. The samples were transferred to 1.5ml Low-Bind Eppendorf's and centrifugated for 3 min at 4000rpm (4°C). After discarding the supernatant, the pellets were resuspended in 1ml of buffer B (NaCl 200mM, Tris pH8 10mM, EDTA pH8 1mM, EGTA pH8 0.5M) and incubated for 30 min on ice while inverting the tubes every 10 min. After a 3 min centrifugation 4000rpm (4°C), pellets were resuspended in 300μL of buffer CD (SDS 0.5%, Triton 0.5%, Sodium deoxycholate 0.5%, NaCl 140mM, Tris pH8 10mM, EDTA pH8 1M, EGTA pH8 0.5mM). DNA was sonicated with the Bioruptor® Plus sonication device (Diagenode) in Covaris vials. Each sample was sonicated through 17 cycle of 30s ON and 30s OFF. The sheared chromatin was then centrifuged for 15 min at 12500rpm (4°C), 20μL of supernatant was kept as input stored at -20°C and the recovered supernatant ≈280μL was transferred in a new tube. The complexes antibodies-beads were retrieved, quickly centrifuged, set on the

magnetic racks where the supernatant was delicately aspired and replaced by 500μL of ChIP DIL buffer. The 280μL of supernatant was added to the beads and incubated O/N (4°C) on the rotating platform for crosslinking.

The next day, the bead-DNA complexes were washed with 1ml of IB buffer (Triton 1.0%, SDS 0.1%,NaCl 150mM, Tris pH8 20mM, EDTA pH8 2mM) and incubated for 2 min on the rotating platform, transferred to a new LoBind tube, followed by a 2 min successive washes on the rotating platform with IC buffer (Triton 1.0%, SDS 0.1%, NaCl 500mM, Tris pH8 20mM, EDTA pH8 2mM), ID buffer (Triton 1.0%, SDS 0.1%, NaCl 500mM, Tris pH8 20mM, EDTA pH8 2mM) and ITEN buffer (NaCl 50mM, Tris pH8 10mM, EDTA pH8 1mM) before transferring the sample to a new LoBind tube. After a 1 min centrifugation 2000rpm RT, 2/3 of the supernatant was discarded. The tubes were put on magnetic racks and the rest of the supernatant was discarded. Reverse cross-linking was performed using 120μL of buffer E (SDS 1.0%, Tris pH8 10mM, EDTA pH8 10mM) for the beads-DNA complexes and 200μL of buffer E for the inputs with a 6 hrs incubation at 65°C. Input and DNA were treated with 50ng of RNAse (SIGMA-ALDRICHT) for 15 min RT followed by a 1hr treatment with 100ng of Proteinase K (SIGMA-ALDRICHT) at 55°C. Finally, DNA and 50μL of input were purified from all samples using the QIAquick Gel Extraction Kit (Qiagen). Elution was achieved in two steps using 50μl Buffer EB, and the eluted ChIP and input samples were aliquoted and stored at -20°C before use. To assess the quality of the sonication, half of the purified chromatin control samples were loaded on an 1.5% agarose gel to estimate the range of DNA fragments sizes (between 150bp to ~1.5 kb with a maximal intensity at ~300 bp, data not shown). Libraries were made with the Nextera DNA Library Preparation Kit (FC-121-1030 , Illumina). Sequencing was done on an Illumina 2000 sequencer in a 50-cycles paired-end configuration (Genomic core facility, EMBL, Heidelberg). (for Antibody see Table 11)

9.2.4. ATAC-seq

ATAC-seq was performed by following the protocol of Buenrostro et al., 2013, for 2 biological replicates performed by Dr. Rushikesh Sheth and Jens Stolte. One replicate consisted of 75 000 cells for each transposition. After two washes in ice-cold DPBS

w/$Ca^{2+}Mg^{2+}$, dissected tissues were transferred to a glass douncer (Tissue Grind Tube Size 2ml, Kimble- Chase) on ice. Then, 20 stokes with pestle A (Tissue Grind Pestle LC 2ml, Kimble-Chase) followed by 20 stokes with the pestle B (Tissue Grind Pestle SC 2ml, Kimble-Chase) were applied to disaggregate the cells clusters. The volume corresponding to 75000 cells was centrifuged (5 min at 500g 4°C) and supernatants gently removed before rinsing with 100μL ice-cold DPBS w/$Ca^{2+}Mg^{2+}$ followed by a centrifugation (5 min at 500g 4°C). DPBS was discarded and the samples were lysed in 50μL Lysis buffer (10mM Tris-HCl pH 7.5, 10mM NaCl, 3mM MgCl2 and 0.1% NP40) by gently pipetting up-and-down and centrifuging (10 min, 2500rpm 4°C). The nuclei pellets, freed of supernatant, were transposed for 30 min at 37°C in a 25μL reaction mix (25μL TD 2x reaction buffer), 2.5μL TDE1 (Nextera Tn5 Transposase (Illumina 15028212)) and 22.5μL nuclease-free H20. Samples were immediately purified using a Qiagen minElute PCR Purification Kit. The transposed DNA was eluted in 13μL EB buffer (Tris-HCl pH 8.5, 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 (sequence: 5'-AAT GAT ACG GCG ACC ACC GAG ATC TAC ACT CGT CGG CAG CGT CA GAT GTG-3') and 25μM barcoded PCR Primer 2.1 (sequence: 5'-CAA GCA GAA GAC GGC ATA CGA GA**T TCG CCT TA**G TCT CGT GGG CTC GGA GAT GT-3') or 2.2 (sequence: 5'-CAA GCA GAA GAC GGC ATA CGA GA**T CTA GTA CG**G TCT CGT GGG CTC GGA GAT GT-3') from the kit for each replicates in a total volume of 50μL.

PCR program details:

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

The amplified library was purified with QIAquick PCR purification kit (Qiagen) and eluted in 30μL EB buffer. Sequencing was done using an Illumina NextSeq 500 system (SE75) by the Genomics Facility Basel - ETH Zürich.

9.2.5. Circular Chromosome Conformation Capture and sequencing, 4C-seq (Matelot and Noordermeer, 2016)

9.2.5.1. Forelimb buds collection

One biological replicate consists in 10 pairs of mouse forelimb buds at 40-42ss (2-4x10^6 cells) collected in cold DPBS (PCB with w/$Ca^{2+}Mg^{2+}$ Gibco Ref.D1283) supplemented by 10% Fetal Bovine Serum (FBS, filter sterilized through 0.45μm filter membrane) in 2ml Eppendorf tubes. After a 1 min centrifugation 300g at 4°C, the medium was replaced with 500μL of fresh 10% FBS/DPBS and the pellet was disaggregated using 3μL of collagenase D (ROCHE, 300μg.ml^{-1}) at 37°C 750rpm and through mechanical disruption by pipetting up and down 20 times every 3 min. Single cell suspensions were obtained by passing the tissue through pluriStrainer mini (70μm, Pluriselect) in a Falcon 15ml and rinsing the pluriStrainer with 11.4ml of RT 10%FBS/DPBS allowing us to maximize the recovered amount of cells. Protein-DNA complexes were cross-linked by a 10 min incubation with 650μL of formaldehyde 37% (SIGMA-ALDRICHT) tumbling. Immediately after cross-linking, samples were transferred on ice and the formaldehyde reactivity was quenched using 0.1M cold glycine. After an 8 min centrifugation at 400g (4°C), the supernatants were discarded and cells were lysed in 5ml lysis buffer (Tris-HCl pH7.5 50mM, NaCl 150mM, EDTA 50mM, NP40 0.5%, Trition 1%, complete proteinase inhibitor 1X) for 10 min on ice. The lysis buffer was gently discarded after a 5 min 400g (4°C) centrifugation and the pellet was resuspended in the left-over medium and transferred in a 1.5ml Eppendorf. Lastly, the cells were pelleted by a 1 min centrifugation at 800g (4°C), the supernatant was discarded and the samples were frozen in N_2 and stored at -80°C.

9.2.5.2. 4C-seq

The 4C protocol was adapted from Noordermeer et al., 2011. The nuclei pellets were washed with 450μL 1.2x DpnII Buffer (NEB). A 1 min 800g centrifugation at 4°C was performed to replace the supernatant by 500μL of 1.2x DpnII Buffer 0.3% SDS. The tubes were incubated first at 55°C for 10 min at 600rpm then at 37°C for 50 min at 600rpm. SDS was quenched by adding 50μL of Triton 20% to the samples before incubating at 37°C for 1hr 600 rpm. A 5μL aliquot was treated with 100μg of Proteinase K in 10mM Tris-HCl pH 7.5 in a final volume of 100μL for 1hr at 65°C and stored at -20°C as undigested controls. In parallel, the sample were digested with 400U of DpnII for 6 hrs at 37°C 600rpm. The reaction was spiked with 400U of DpnII O/N, 37°C 600rpm.

The next day, a 5μL aliquot was treated with 100μg of Proteinase K in 10mM Tris-HCl pH 7.5 in a final volume of 100μL for 1hr at 65°C and run on a 1.5% agarose gel with the undigested controls to assess the digestion. When the digestion was successful, the restriction enzyme was inactivated for 25 min at 65°C with 80μL of SDS 10%. The sample were transferred to a 15ml tube and incubated 1hr at 37°C 400rpm with 375μL of 20% Triton in 1.15x ligation buffer (ATP 10mM, Dithiothreitol 50mM (DTT, MERK), $MgCl_2$ 50mM, Tris-Hcl pH7.5 0.7M). The samples were cooled down at RT before the addition of 100U of T4 ligase (Promega). After a 4 hrs incubation at 16°C, the reaction was pursued for 30 min at RT. De-crosslinking was processed O/N at 65°C in presence of 30μg of Proteinase K.

The next morning, the samples were treated with RNase for 45 min at 37°C. Purification of the samples was performed by adding 7ml of phenol:chlorophorm:isoamyl alcohol (SIGMA ALDRICHT) pre-warmed at RT and shaking the tubes until they became white. The DNA was retrieved after a 15 min centrifugation at 2200g (RT) and transferred to a 50ml Falcon tube (SIGMA-ALDRICHT). In order to facilitate DNA precipitation, 7ml of nuclease free water and 1.5ml of 2M Sodium acetate pH5.6 was gently added to the samples. 35ml of pure ethanol was added to the samples and gently mixed before a 2 hrs incubation at -20°C. After a 45 min centrifugation at 4500g (4°C), DNA pellets were cleaned with

10ml of 70% ethanol followed by a 15 min centrifugation 4500g (4°C). The pellets were dried and then resuspended in 150μL of 10mM Tris-HCl pH 7.5. After measuring the DNA concentration, 200ng per sample were loaded on a 1.5% agarose gel to control the ligation. The DNA was diluted to 100ng/μL in the appropriate 1X digestion buffer and 1U/μg of secondary enzyme NlaIII was added to the solution and incubated O/N at 37°C 600rpm.

On the fourth day, the restriction enzyme was inactivated by a 20 min incubation at 65°C. The digested DNA was mixed 1:1 with RT Phenol:Chlorophorm:Isoamyl alcohol and transferred to a phase lock gel™. After a 5 min centrifugation 13000 (RT) the upper phase was transferred to a new 1.5ml Eppendorf tube and gently mixed with 1:10 (depending on the NlaII digestion final volume) 2M Sodium acetate pH5.6 and 2:1 100% ethanol. After a 30 min incubation at -80°C, the DNA was pelleted thanks to a 20 min centrifugation at 13000 rpm (4°) and cleaned with 1ml of 70% ethanol. The supernatant was discarded after a 5 min centrifugation 13000 rpm (4°C) and the pellet was dried 1 min (RT) and resuspended in 100μL of clean water. A 5μL aliquot was run on a 1.5% agarose gel to control the digestion. After transferring the DNA to a new 50ml tube in large volume of water (≈12.46ml), the ligation was performed with ligation buffer in a 1X final concentration with 200U of T4 DNA ligase HC (Promega) for 4 hrs at 16°C and then for 30 min at RT. The ligation products were purified with 14ml of Phenol:Chlorophorm:Isoamyl alcohol followed by a 10 min 3000g (RT) centrifugation. The supernatant was transferred to a new 50ml tube and mixed with 14ml of water and 5μg.μL (≈112μL) of glycogen (ROCHE) and 1:10 (approximately 1.5ml) 2M Sodium acetate pH5.6 to facilitate DNA precipitation. The samples were divided into two 50ml tubes incubates O/N with 2 volumes of 100% ethanol at -20°C.

On the fifth day, the DNA was pelleted by centrifuging for 45 min at 4500rpm (4°C), cleaned with 10ml of 70% ethanol followed by a 15 min centrifugation 4500rpm (4°C) and left to dry for 10 min (RT). The purified DNA pellet was dissolved in 200μL of 10mM Tris-HCl pH7.5 and left to sit at least for 30 min at 37°C. The samples were purified with the QIAquick Nucleotide removal kit (QIAgen). Libraries were constructed using the highest concentration of 4C template that was still linear on a 1.5% agarose after PCR with the index free view point primers designed by PD Dr. Aimée Zuniga (C1: 5'-CTT CCG TAG ATG CTG GCC GAT-3' and C6: 5'AAG CAC CAG GAC CGA

GTT TG-3') using the Expand Long Template PCR system (Roche) polymerase. (see Table 8 for primer sequences)

PCR program details:	94°C 2 min	
	94°C 15sec	30 cycles
	55°C 1 min	
	68°C 3 min	
	68°C 7 min	
	4°C ∞	

The pooled PCR products were purified using a PCR purification kit (Qiagen) such as no more than 10μg was loaded on a column. After purification, 2ng of DNA was loaded on a fragment analyzer (ThermoFisher) and the samples with the best profiles (no primer contamination) were sequenced using an Illumina Next Seq 500 system or Mi-seq single end sequencing with 75 cycles by the Genomics Facility Basel - ETH Zürich.

9.3. Bioinformatics Analysis

Bioinformatic analysis for SMAD4 ChIP-seq was performed by Iros Barozzi (Lawrence Berkeley National Laboratory, Berkeley, CA, USA). The ß-catenin ChIP-seq and the ATAC-seq experiments were analyzed by the DBM Bioinformatics Core Facility member Florian Geier. This part of the material and methods section was partly written by Iros Barozzi for Gamart et al., manusctipt in preparation. Bioinformatic analysis for 4C-seq were performed by Dr. Shalu Janhwar and the 4C section of the material and methods was written by her. The GLI ChIP-seq were performed and analyzed by Dr. Kevin Peterson.

9.3.1. 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 subsignificant 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 pvalue (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 kb 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.

9.3.2. 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. Performed by the DBM Bioinformatics Core Facility member Florian Geier.

9.3.3. 4C-seq analysis

4C-seq libraries of WT and different *Grem1* mutant samples were sequenced with Illumina technology according to standard protocols, generating reads of comparable lengths (76, 81 or 85bp). To achieve an overall high quality of the raw sequencing data, firstly reads from different samples were trimmed to 76bp and subsequently, those reads that did not match with the primer sequence were discarded. After clipping the primer sequence, cleaned reads were aligned onto the mouse reference genome mm10 using Bowtie v 2.2.9. To get the valid restriction fragments, the mouse reference genome was in-silico digested using the two restriction enzymes. Restriction fragments that did not contain a cutting site of the second restriction enzyme or are smaller than 20bp were filtered out (Klein et al. 2015), finally yielding valid restriction fragments that were used for the quantification of 4C-seq profiles. Read counts were then computed in each valid fragment and subsequently visualized using the UCSC genome browser. To visualize the data, bedGraph formatted files containing read counts of either each fragment or in a specified window of fragments were generated. 4C-seq contacts were analyzed in the mouse region chr2:113326224-113894862. As suggested by Lupiáñez, Darío G., et al. (Lupiáñez, Darío G., et al. cell 2015), the viewpoint, adjacent undigested fragments and fragments 10 kb up and downstream were excluded during the procedure. Finally, a range of 5 informative fragments was used to normalize the data per million reads (RPM) over a sliding window using custom scripts and the continuous-valued data was displayed in the figures as well as onto the UCSC genome browser tracks. The differences track was calculated by

subtracting fragment reads between different samples (Klein et al., 2015; Lupianez et al., 2015)

9.4. Cell proliferation analysis by fluorescence-activated cell sorting (FACS)

The flow cell analysis was performed by Dr. Robert Reinhardt and this part is adapted from Reinhardt et al., 2019.

Forelimb buds were dissected and pools of 4 limbs per replicates at E11.0 (39ss to 43ss) were dissociated using collagenase D. Following the exclusion of lineage-positive cells (ectodermal cells, erythrocytes, endothelial cells, hematopoietic cells, macrophages, dendritic cells, myeloid lineage, see Table 9.5.11) all cells were fixed in 70% ethanol at -20°C O/N. Phospho-histone H3 antibodies (BD Biosciences, clone HTA28 Alexa Fluor 647) were used to detect mitotic cells. Cells were also incubated with 50μl.ml^{-1} propidium iodide (Sigma) and 50μl.ml^{-1} RNase A (Sigma) to measure their DNA content. FACS was used to determine the fractions of cells in the different phases of the cell cycle and the fraction of phospho-histone H3-positive mitotic cells among the lineage-negative (Lin^{-}) limb bud mesenchymal cells. FACS was used to determine the fractions of cells in the different phases of the cell cycle and the fraction of phospho-histone H3-positive mitotic cells among the Lin^{-} limb bud mesenchymal cells. FACS analysis was carried out using a BD FACS Aria III machine. After exclusion of lineage-positive cells, the numbers of cells in different phases of the cell cycle and mitotic cells were determined and fractions calculated using the FlowJo 10.5.3 software.

9.5. Tables

9.5.1. Table 1: Summary of transgenic collections for CRM screening

CONSTRUCTS	Embryo collected	% Transgenic embryos	Expresser embryos	Limb expresser embryos
CRM1	76	22.4% (n=17/76)	9 (52.9%)	9/9 (100%)
CRM2	59	6.8% (n=4/59)	4 (100%)	4/4 (100%)
CRM3	73	19% (n=14/73)	11 (78%)	8/11 (73%)
CRM4	113	15.9% (n=18/113)	15 (83%)	10/15 (70%)
CRM5	15	40% (n=6/15)	4 (67%)	3/4 (75%)
CRM6	17	17.65% (n=3/17)	3 (100%)	2/3 (66.7%)
CRM7	47	12.7% (n=6/47)	6 (100%)	6/6
CRM8	26	15.4% (n=4/26)	3 (75%)	0/3
CRM9	26	19.2% (n=5/26)	1 (20%)	0/1
CRM10	42	11.9% (n=5/42)	4 (80%)	4/4 (100%)
CRM11	46	10.8% (n=5/46)	3 (60%)	1/5 (33%)
CRM12	36	30.6% (n=11/36)	6 (45.5%)	6/12 (50%)
CRM13	42	12% (n=5/42)	12% (n=5/42)	2/5 (40%)
CRM14	74	21.7% (n=5/24)	12 (75%)	11/12 (91.7%)

CTCF-V	67	7.5% (n=5/67)	5 (100%)	0
CRM2_delAB	118	5.9% (n=7/118)	5 (71.4%)	0
Exon9	51	9.8% (n=5/51)	4 (100%)	0
***Grem1* promoter _GRE1**	89	12.4% (n=11/89)	0	0
***Grem1*_promoter_2.1kb**	66	7.6% (n=5/66)	2 (40%)	1/2 (50%)
***Grem1ORF*_700bp**	65	20% (n=13/65)	8 (61.5%)	2/8 (25%)
Grem1PromlacZ	52	9.6% (n=5/52)	60%(n=3/5)	0
HMCO2_700bp	75	13% (n=10/75)	6 (60%)	3/6 (50%)
HMCO3	49	16.3% (n=8/49)	6 (75%)	1/6 (16.7%)

9.5.2. Table 2: Reporter assay primers

Target	Coordinates (mm10)	Size (bp)		Sequence (5' to 3')
CRM1	chr2:113,724,491-113,724,715	225	F	CACTAAAGGGAACAAAAGCTGGTACCGCATGTTTTTGGCCTTCATT
			R	TTTGGATGTTCCTGGAGCTCGGTACCTGAATCCCGCTATTGGTAGG
CRM2	chr2:113,689,752-113,693,024	3262	F	CACTAAAGGGAACAAAAGCTGGTACCTCAGATCTCATCTTTAATGTAGAAACC
			R	CACTGTTGTTTCAAAACTGCAGAAGGTAGGGGGAAAC
	chr2:113,692,996-113,696,244	3249	F	CTACCTTCTGCAGTTTTGAAACAACAGTGGGATATTTTG
			R	GTTAAAACAAACAACATTTACCTTTGGGTCTTTCTCTTTG
	chr2:113,696,216-113,699,421	3206	F	ACCCAAAGGTAAATGTTGTTTGTTTTAACCCCACTTC

			R	TCACACAGGTAAGATGGGCCGGGC CCAGTAGCTTTCGCA TCATGATG
CRM3	chr2:113,674,074-113,675,273	1246	F	CACTAAAGGGAACAAAAGCTGGTAC CCCGCGTAGTTTCCATTCAAC
			R	TTTGGATGTTCCTGGAGCTCGGTAC CGCAACTGCCTTGAAGGTGAT
CRM4	chr2:113,640,376-113,641,758	1389	F	CACTAAAGGGAACAAAAGCTGGTAC CTTCCTCAGAGCAGCTGAGTGT
			R	TTTGGATGTTCCTGGAGCTCGGTAC CAGCGCTGAGAATTATAGGCTGA
CRM5	chr2:113,605,542-113,606,069	528	F	CACTAAAGGGAACAAAAGCTGGTAC CGAAGTCCCTCGAACACCAAG
			R	TTTGGATGTTCCTGGAGCTCGGTAC CCTTTTCCAGCAAGGTGTGGT
CRM6	chr2:113,595,444-113,596,148	705	F	CACTAAAGGGAACAAAAGCTGGTAC CAGGTGTTGAGCCACTACTGC
			R	TTTGGATGTTCCTGGAGCTCGGTAC CTCATCTTTTGTGAGGGGAAAA
CRM7	chr2:113,581,140-113,581,955	1029	F	CACTAAAGGGAACAAAAGCTGGTAC CGGTGCTCGAGAGCAGAAGTT
			R	TTTGGATGTTCCTGGAGCTCGGTAC CATACCAGGTGTGGTGGGTGT
CRM8	chr2:113,575,889-113,576,624	736	F	CACTAAAGGGAACAAAAGCTGGTAC CGGGATGAAAGAACGCCTGTA
			R	TTTGGATGTTCCTGGAGCTCGGTAC CTGTCACTCTCCTGTCCTCCA
CRM9	chr2:113,537,701-113,538,435	735	F	CACTAAAGGGAACAAAAGCTGGTAC CTCATTCACCGTTGTCCATGT
			R	TTTGGATGTTCCTGGAGCTCGGTAC CTGCATCTTCTGACGCTTGAC
CRM10	chr2:113,497,608-113,498,439	832	F	CACTAAAGGGAACAAAAGCTGGTA CCATGGTGTGGAACAAGGCTTC
			R	TTTGGATGTTCCTGGAGCTCGGTAC CTGCATCTTCTGACGCTTGAC
CRM11	chr2:113,460,271-113,461,676	1406	F	GGAACAAAAGCTGGTACCACCTAAG CCCGTGGGTAC
			R	CCTGGAGCTCGGTACCGACAGACA GACAAGAAAACACTATG
CRM12	chr2:113,439,964-113,441,240	1277	F	GGAACAAAAGCTGGTACCTCTTTCT CTGTCCCTTCC

			R	CCTGGAGCTCGGTACCGCTCTCTCA GTTTATCCGC
CRM13	chr2:113,427,755-113,429,093	1339	F	GGAACAAAAGCTGGTACCTTTATGA GGGTCTGTCTCCTC
			R	CCTGGAGCTCGGTACCAGGGACTG AGGCACAGAG
CRM14	chr2:113,397,272-113,398,270	999	F	GGAACAAAAGCTGGTACCAAGGAAA TCTGTGGTTGGATAATG
			R	CCTGGAGCTCGGTACCAAGGGCAC ATGATTGTACAG
CTCF-V	chr2:113,648,020-113,648,733	714	F	CACTAAAGGGAACAAAAGCTGGTAC CGTGATGGCCCCAGTTTCTAA
			R	TTTGGATGTTCC TGGAGCTCGGTA CCTGAGTGGAGTTGACCAGAGC
Grem1 ORF	chr2:113,749,490-113,750,190	701	F	CACTAAAGGGAACAAAAGCTGGTAC CAAGCAACTGCTGGTTCTTCTG
			R	TTTGGATGTTCCTGGAGCTCGGTA CCTTGTCTGTGTCCCCCTCTCT
RYR3	chr2:113215912-113217517	1606	F	AAGGGAACAAAGCTGGTACCTTAAA AAGGACAGTGCAA G
			R	TTTGGATGTTCCTGGAGCTCGGTAC GGTACCACATGCACCGTGCAA
Grem1 Promoter	chr2:113,757,063-113,759,318	2256	F	AACTGAGCAGCCGGGACCGGTGGC AAACACTAAGCTGAAG
			R	TCATCGCGAGCCATGACCGGTAGAG GGTATCCTTTTGCAG

9.5.3. Table 3: CRISPR/*Cas9* deletions coordinates and guides

Deletion	Coordinates (mm10)		sgRNa sequence	Size (kb)
ΔCRM2	chr2:113,689,603-113,699,473	5'	AGCGGCAGTTCGGCTTCCGG	9.9
		3'	TCTCATACGATCCAGGAGAA	
ΔCRM3	chr2:113,672,786-113,681,077	5'	CACCGCCTGTGATCCATCGAATGCC	8.6
		3'	AAACGGCATTCGATGGATCACAGGC	

ΔCRM4	chr2:113,640,518-113,641,588	5'	CACCGCAGCCTATAATTCTCAGCGC	1.3
		3'	AAACGCGCTGAGAATTATAGGCTGC	
ΔCRM5	chr2:113,603,760-113,606,572	5'	CTTGTAATGCTAGGACGGCC	2.8
		3'	AGTCAAGGACACACCTGTA	
ΔEC1	chr2:113,636,776-113,707,974	5'	CACCGTGGCTTACCAGACTAGCGGT	71.2
		3'	CACCGAATGGTCTGATCGCCA	
ΔEC2	chr2:113,572,310-113,606,574	5'	TTCAGCTGCATTGCGTCCTA	34
		3'	AGTCAAGGACACACGCTGTA	
F1Δ1019	chr2:113,528,475-113,636,771	5'	CACCGGTCTTTCAGGAACTCCCCGG	108
		3'	CCAGAAAGTCCTTGAGGGGC	

9.5.4. Table 4: Genotyping Primers

Locus	Forward	Reverse	Size (bp)	Allele
Cis	GGGCGGATCTCAAACTC TCCTC	CTGCATTCTAGTTGTGGTTTGTCC	350	Δ
	GAAAGACTGCTGCAGAAGGAAGC	TCC AAA GAA GGC ACA GGG GAC TT	200	WT
CRM12	CTCCCATATGCTCACCGGTTT	GGAAAATCCCTGCCCAAGAGA	908	Δ
	CTCCCATATGCTCACCGGTTT	GGTGGGAGTGGAGTTTGACC	429	WT
CRM2	AACAAACAGTGCAATTCTGAAGAG	AGGAAGCCAGTGATCTCAAATATC	605	Δ
	GACTGGGATACATGTGATGGTAAA	AGGAAGCCAGTGATCTCAAATATC	932	WT
CRM3	CCTTTACATGCACACTCACACA	TGCTTTGGATGTCTACAAGTGG	2.4kb	Δ
	AGCAGTGTGTCCCCCTAATGAGC	GAGCCTGGAGCTTACTGAGCAAC	389	WT

CRM4	AATTGAAGGAACAAAAG GCTCA	GGCTGACAGTAGTTTGC TGTTG	205	Δ
	GGCCTTCGAAACCATGA TGC	GGGAGGTGCTGGAATTA GGG	513	WT
EC1	GGGACAAGTCACAGAT CTTTTTG	TCCTTCATGTCTCGTTTT GTTTT	439	Δ
	TCCAGTTAAATGCAAAA AGGAAA	CTCTTCCTTCATCTCTCC TAGCC	593	WT
EC2	AGAAGCACTTGGCATGT GATG	GGAAATTCAGAGCCATC CAA	750	Δ
	AGAAGCACTTGGCATGT GATG	TGAAAGAGTCGGCAACT GTG	500	WT
F1Δ1019	TCCTCAGCAGTCACCTC CTG	AGGTCATCTCAGTCCAG GGG	569	Δ
	TCCTCAGCAGTCACCTC CTG	GAG GGA GGG AGG GAA GGA AT	609	WT
Grem1	ATGAATCGCACCGCATA CACTG	TCCAAGTCGATGGATATG CAACG	600	Δ
	CCCAGGAGGCACATGG AACACAA	GGCACATGGCTGAATAT CGACGG	700	WT
Tg-GLE	AATTCCTTACTGGCTCC TGACAC	CCAGAGAAATTAAAGACC TACCTGG	542	*lacZ*
	CAACGGGTTCTTCTGTT AGTCC	CCAGAGAAATTAAAGACC TACCTGG	570	WT

9.5.5. Table 5: Software and algorithms table

Software / Algorithms	Provider	website
Adobe Illustrator CS6	Adobe System	https://www.adobe.com/products/illustrator.html
Adobe photoshop CS6	Adobe System	https://www.adobe.com/products/photoshop.html
CRISPOR	Maximilian Haeussle et al., 2016	http://crispor.tefor.net/

Galaxy	CyVerse, Advanced Computing Center Texas	https://usegalaxy.org/
GraphPad Prism 7	GraphPad Software, Inc	https://www.graphpad.com/scientific-software/prism/
HOMER 4.8.1	Heinz et al., 2010	N/A
Leica Application Suite V3 software	Leica Microsystems	https://www.leica-microsystems.com/products/microscope-software/details/product/leica-application-suite/
liftOver	Hinrichs et al., 2006	N/A
NIS Elements Basic Research	Nikon	https://www.microscope.healthcare.nikon.com/products/software/nis-elements/nis-elements-basic-research
UCSC	Kent et al., 2002	http://genome.ucsc.edu/cite.html
3D Genome Browser	Yue Lab	http://promoter.bx.psu.edu/hi-c/view.php
Universal ProbeLibrary Assay Design Center	ROCHE	https://lifescience.roche.com/en_ch/brands/universal-probe-library.html#assay-design-center

9.5.6. Table 6: RT-qPCR Primers table

Target	Forward	Reverse	Size (bp)
Gli1	CAAGTGCACGTTTGAAG	CAACCTTCTTGCTCACACATGTAAG	76
Rpl19	ACCCTGGCCCGACGG	TACCCTTTCCTCTTCCCTATGCC	53
BMP4	AGCCGAGCCAACACTGTGA	GTTCTCCAGATGTTCTTCGTG	51
Grem1	CCCACGGAAGTGACAGAATGA	AAGCAACGCTCCCACAGTGTA	53
HMBS	TCTAGATGGCTCAGATAGCATGCA	TGGACCATCTTCTTGCTGAACA	76
ID1	GCGAGATCAGTGCCTTGG	CTCCTGAAGGGCTGGAGT	111

Shh	GAGGTGCAAAGACAAGTTAAATGC	GGCCACTGGTTCATCACAGAG	57
SMOC1	GATAAGGTCATCTCACTGCCTGA	AAGCTGCCAAGGCTACCAC	74

9.5.7. Table 7: ChIP-qPCR Primers table

Primer Name	Coordinates (mm10)	Forward	Reverse	% Eff.	Size (bp)
CRM1	chr2:113724513-113724657	ACCACCCCAGTCGCCTATAT	TGTCCCCTCCCCTTTGATCT	88.3	145
CRM12	chr2:113440509-113440601	TCCAGATGAAACCTGCCAGC	GAGTGCCCACAACTCTGAGA	98.1	93
CRM14	chr2:113397688-113397758	TCCCAAGGTTTCAAAGCCAGA	GCAGGGCCATGAAACTTGAA	88.0	71
CRM2 A Bcat	chr2:113692121-113692000	TAATAGTCACAGGGCTTCGGG	TGTGGCCATCCTGAATACTGAG	97.0	80
CRM2 A GLI	chr2:113532291-113532421	GAAGCTTGGAGTGGGTGAGG	TGAAGGCATGCTGTCCACTT	84.3	131
CRM2 B	chr2:113693479-113693578	TGG CAG ACA GTA CAG ACA AA	TGTGTGAAGTGTGGCCTGAG	104.8	100
CRM2 D	chr2:113693471-113693578	CCTTGCTGTGGCAGACAGTA	TGTGTGAAGTGTGGCCTGAG	86.9	108
CRM2 Neg	chr2:113690722-113690834	CCCCTTGAAAACACACACACA	TGGCCTGGGGATTTCTCTTTC	100.0	113
CRM3	chr2:113674690-113674805	GAATAAAGTTAGTGACCC GCCA	GCAAAGCCCAGATTAATTCCCC	99.5	116
CRM4	chr2:113640898-113640993	CCTGTGCAGCCATTTGTGTG	GCATCATGGTTTCGAAGGCC	102.3	96
CRM5	chr2:113605801-113605890	TAAGGCCTTGGTGTGAACCC	ACCTCAAAGCCTCATCAAATCTT	95.4	90
CRM6	chr2:113595746-113595859	GGGAAGGGAAATACACAGGCA	GCCAAGCTACTTTATGACCCTG	91.0	114
CRM7	chr2:113581322-113581412	GGCCCGGCTAATTTCAAACAC	TCCTTCCCAGATACGTTCGGA	97.0	91
CRM7	chr2:113581688-113581768	CCCTTCCTGGTAATGGGCAA	AAGTCTTTCTTCCCGGCTGG	95.7	81

CRM8	chr2:113576238-113576348	ACCCAATAAAGTTCCCGGCT	TAGTTGTTCCCGCCTCACAC	95.3	111
CRM9	chr2:113538009-113538078	TGCATGGCTGTAAACACGAG	GGGGTCTCAGGTTTCTTTTGC	86.1	70
CTCF V	chr2:113648274-113648362	CAGTGACAGAGGAGACGTTCA	GTGACGTTGTGGATTTTGGCA	103.0	89
Fmn1	chr2:113573195-113573335	AGCATGTTGTCTTTCCTCCCA	TGAAGGGCGGTGTAAGGAAG	93.6	141
Grem1 int GLI3	chr2:113758014-113758094	AGCCCTGTACCCAACTCTCA	GGAGCTGGCATCTCTGTGAG	97.6	81
Grem1 exon2	chr2:113749207-113749297	TTCCGAGTCTAATGCAGCCC	TGTGTTCCATGTGCC CCTG	87.4	91
HMCO3	chr2:113618667-113618997	AGCATCCCATTGTACAGCGA	GGGCCACATTGTCAATTCGC	98.8	131
KP*-Ptch	chr13:63562838-63562988	CCCCTCCCATCAAACTTGCT	GCACAAACTGGAACGCATGT	85.5	151
KP*-Ptch1-p7	chr13:63719677-63719827	TATCTTGCAAATTGGAGGCTTTCG	ACATCTGATAAGCAGCCCCAGAAC	80.4	150
KP*-Ptch1-p9	chr13:63980451-63980649	GGTCCTGCCTTGATTGCTAA	GGGAGTGGAGATGACTGGGA	89.9	199
Neg LRM	chr2:113695093-113695179	TGATCTAATGCGAACTTGTGGT	TGTATGGATGTTCAGAA TGCCT	80.2	87

*: *designed by dr. Kevin Peterson*

9.5.8. Table 8: 4C-seq primers table

Name	Sequence
Grem1_C1_TUA	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCTTCCGTAGATGCTGGCCGAT
Grem1_C6_index12	CAAGCAGAAGACGGCATACGAGAT**TACAAG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTAAGCACCAGGACCGAGTTTG
Grem1_C6_index2	CAAGCAGAAGACGGCATACGAGAT**ACATCG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTAAGCACCAGGACCGAGTTTG
Grem1_C6_index4	CAAGCAGAAGACGGCATACGAGAT**TGGTCA**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTAAGCACCAGGACCGAGTTTG
Grem1_C6_index5	CAAGCAGAAGACGGCATACGAGAT**CACTGT**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTAAGCACCAGGACCGAGTTTG

Grem1_C6_index6	CAAGCAGAAGACGGCATACGAGAT**ATTGGC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTAAGCACCAGGACCGAGTTTG
Grem1_C6_index7	CAAGCAGAAGACGGCATACGAGAT**GATCTG**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTAAGCACCAGGACCGAGTTTG
Grem1_C6_index9	CAAGCAGAAGACGGCATACGAGAT**CTGATC**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTAAGCACCAGGACCGAGTTTG

9.5.9. Table 9: 4C-seq qPCR primers

Primer Name	Coordinates (mm10)	Forward	Reverse	% Eff.	Size (bp)
4CqPCR	chr2:113724513-113724657	TTGGACCAATTAGTGCATCTGT	AACCTAGCCTCTCAGGTCTCG	111	64

9.5.10. Table 10: WISH probes

Gene	Forward	Reverse	Size (bp)
ID1	CTCAGGAGACAGCTCAGTGTACC	CTAACTCAAGCCAGACCCCCTAC	71
Grem1	ATGAATCGCACCGCATACACTG	TCCAAGTCGATGGATATGCAACG	500

9.5.11. Table 11: Antibodies

Antibody	Catalogue number	Manufacturer
Monoclonal ANTI-FLAG® AN	F1804	Sigma
Anti-Digoxigenin-AP, Fab fragments	11093274910	Sigma
Beta catenin polyclonal antibody	71-2700	Thermofisher
Biotin Ep-CAM	clone G8.8 ref.118204	BioLegend
Biotin TER-119	clone TER-119 ref.116204	BioLegend
Biotin CD31	clone 390 ref.102404	BioLegend
Biotin CD45	clone 30-F11 ref.103104	BioLegend
Biotin CD11b	clone M1/70 ref.101204	BioLegend
Biotin Gr-1	clone RB6-8C5 ref.108403	BioLegend
APC/Cy7 Streptavidin	405208	BioLegend

10. Acknowledgement

First, I would like to express my gratitude to my Ph.D. supervisors Prof. Dr. Rolf Zeller and PD. Dr. Aimée Zuniga for giving me the opportunity to tackle a very challenging project, allowing me to grow and mature into a research scientist. Special thanks to my first supervisor and supporter PD. Dr. Aimée Zuniga who has trusted me and encouraged me to develop my own ideas throughout this project.

I thank my righteous Co-Referee, Prof. Dr. Markus Affolter, for his wise guidance and warm support during my PAC meetings. I express my deepest gratitude to Prof. Dr. Guillaume Andrey, who honored me by being part of my thesis committee. Prof. Andrey has been an expert from whom I got brilliant advice and unlimited support throughout my Ph.D. project. I would like to thank Dr. Patrick Tschopp whom I have the privilege of counting among my thesis committee members. During different local and international seminars where I could present him my work, we had constructive discussions that pushed me to think about my results through different perspectives, helping my project to move forward. Last but not least, I am deeply thankful to the esteemed Prof. Dr. Walter Salzburger who accepted the Chair position of my Ph.D. committee.

I would like to thank Dr. Julie Gamart as my whole project started from a side project she was leading. She has been a mentor and a great advisor as Ph.D. senior and then a Postdoctoral fellow in the laboratory. She took the time to teach me how to perform complex ChIP-seq experiments and I will never forget our nightly experiments (and dance contests) in the laboratory. She has also been a friend who introduced me to the warmth of “les gens du Nord, du vrai Nord!”. A last word for this generous and so valuable person that always knew how to make me smile: Perroquet.

I would like to thank Dr. Robert Reinhardt who has been a mentor for me since my arrival in the laboratory and who corrected my thesis manuscript. He has guided me throughout my project and kept me on the right path. I sincerely value everything I

have learned from him and hope I can do the same thing for my juniors. He believed in me, even when I didn't myself, and is major contributor to my achievements. I will never be able to thank him enough for everything. I feel privileged to have met such a generous and precious friend (and cat-sitter!).

I would like to thank the members of the Developmental Genetics laboratory, past and present. First, thanks to my junior Ph.D. students, "my boys", Jonas Malkmus, Victorio Palacio and Research assistant Tommy "Oberhölzerly" Oberholzer, who greatly helped to fasten the completion of the WISH and skeletal analyses. Thanks to Dr. Rushikesh Sheth and research assistant Jens Stolte who provided astute technical inputs and also teamed up to performed all the ChIP-seq experiments that are exclusive to this study. Research assistants Nathalie Riesen, who trained me with great enthusiasm during the year we overlapped in the laboratory, Dario Speziale who introduced me to the joy of bacterial work. Dr. Javier Lopez-Rios and Dr. Gretel Nusspaumer who provided invaluable scientific inputs during all laboratory seminars but also took time to provide sharp recommendations while not being involved in my project. Many many many thanks to Aušra Girdžiušaitė, my dear fellow Ph.D. student, who has been such an important ally in every aspect of my Ph.D. adventure. Thanks to my valuable friend, Dr. Virginie Tissière, for all the lengthy scientific (or not) discussions and for teaching me how to tame ESCs culturing like a pro and double-checking the material and method part of my thesis manuscript. Special thanks to Dr. Shalu Jhanwar who has performed all the talented bioinformatic analyses of my 4C-seq experiments. I would like to thank Dr. Fabiana Gullotta and Dr. Maëva Luxey who spent a short time in the laboratory but have always been available for constructive discussions and spreading their contagious good mood.

Many thanks to our collaborators from the CTM. Dr. Pawel Pelczar, Head of the CTM, who always took time to discuss my project with me and also provided moral support and scientific suggestions regarding the design of the early steps of my Ph.D. research. Together with his team members, notably Heide Oller. Aline Baur, Alain Brülhart and Andreas Siegrist, as they provided excellent services that allowed me to proceed with this challenging project.

I thank Christian Beisel, Elodie Burcklen, Katja Eschbach and Ina Nissen from the Genomics Facility, for sharing their expertise and consistent good work. Many thanks to Angelika Offinger (Head of the Animal Facility MS28), Emilia Stanislawa Terszowska and all the competent animal facility workers for taking care of our mice colonies.

Thanks to Zoubida Boudebaba and Salέha Rabehi who always provided us autoclaved solutions, glassware, bright smiles and sweet snacks out of pure kindness. I am thankful to Prof. Zeller past and present personal assistants - Barbara Widmer, Sonia Calzascia, Debora Neyer, Cansu Sahinbay and Antoinette Walser-van Gelderen- for their outstanding administrative support and unlimited helpfulness.

Thanks to Prof. Dr. Verdon Taylor who always assisted to all Monday's Taylor/Zeller seminars and proposed bright and well-considered thoughts on my Ph.D. project. His team members have also provided great technical and scientific support during my Ph.D. and I would like to name some of them that have been of huge help for my project: Dr. Anna Engler who advised my for the cloning of big DNA fragments and Dr. Tanzila Mukthar who took time to read and offer clever comments on the result part of my thesis manuscript, Dr. Chiara Rolando who double-checked my qPCR analyses in unbiased ways, Dr. Zahra Ehsaei who taught me how to take care of the cell culture room, and Dr. Elena Parmigiani who offered valuable input regarding my project in general.

I would like to thank my family, in particular my mother, Elsa, and my father, Manuel, for our weekly phone calls, for cheering me up, for giving me strength and for their continuous encouragement. I thank my friends, for sticking up for me, for being such kind souls and dependable peoples for so many years. Special thanks to Fanny Vazzoler, I am so lucky I had such a strong supporter throughout my thesis.

"Alone we can do so little; together we can do so much." – Helen Keller

11. Bibliography

Ahituv, N., Zhu, Y., Visel, A., Holt, A., Afzal, V., Pennacchio, L.A., and Rubin, E.M. (2007). Deletion of ultraconserved elements yields viable mice. PLoS Biol *5*, e234.

Ahn, K., Mishina, Y., Hanks, M.C., Behringer, R.R., and Crenshaw, E.B., 3rd (2001). BMPR-IA signaling is required for the formation of the apical ectodermal ridge and dorsal-ventral patterning of the limb. Development *128*, 4449-4461.

Ahn, S., and Joyner, A.L. (2004). Dynamic changes in the response of cells to positive hedgehog signaling during mouse limb patterning. Cell *118*, 505-516.

Akiyama, H., Kim, J.E., Nakashima, K., Balmes, G., Iwai, N., Deng, J.M., Zhang, Z., Martin, J.F., Behringer, R.R., Nakamura, T.*, et al.* (2005). Osteo-chondroprogenitor cells are derived from Sox9 expressing precursors. Proc Natl Acad Sci U S A *102*, 14665-14670.

Al-Qattan, M.M., and Alkuraya, F.S. (2019). Cenani-Lenz syndrome and other related syndactyly disorders due to variants in LRP4, GREM1/FMN1, and APC: Insight into the pathogenesis and the relationship to polyposis through the WNT and BMP antagonistic pathways. Am J Med Genet A *179*, 266-279.

Alexander, J.M., Guan, J., Li, B., Maliskova, L., Song, M., Shen, Y., Huang, B., Lomvardas, S., and Weiner, O.D. (2019). Live-cell imaging reveals enhancer-dependent Sox2 transcription in the absence of enhancer proximity. Elife *8*.

Andrey, G., Schopflin, R., Jerkovic, 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.

Bacchelli, C., Goodman, F.R., Scambler, P.J., and Winter, R.M. (2001). Cenani-Lenz syndrome with renal hypoplasia is not linked to FORMIN or GREMLIN. Clin Genet *59*, 203-205.

Banerji, J., Rusconi, S., and Schaffner, W. (1981). Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. Cell *27*, 299-308.

Barna, M., and Niswander, L. (2007). Visualization of cartilage formation: insight into cellular properties of skeletal progenitors and chondrodysplasia syndromes. Dev Cell *12*, 931-941.

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

Bartman, C.R., Hsu, S.C., Hsiung, C.C., Raj, A., and Blobel, G.A. (2016). Enhancer Regulation of Transcriptional Bursting Parameters Revealed by Forced Chromatin Looping. Mol Cell *62*, 237-247.

Barutcu, A.R., Maass, P.G., Lewandowski, J.P., Weiner, C.L., and Rinn, J.L. (2018). A TAD boundary is preserved upon deletion of the CTCF-rich Firre locus. Nat Commun *9*, 1444.

Bastida, M.F., Delgado, M.D., Wang, B., Fallon, J.F., Fernandez-Teran, M., and Ros, M.A. (2004). Levels of Gli3 repressor correlate with Bmp4 expression and apoptosis during limb development. Dev Dyn *231*, 148-160.

Bell, A.C., West, A.G., and Felsenfeld, G. (1999). The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. Cell *98*, 387-396.

Benabdallah, N.S., Williamson, I., Illingworth, R.S., Kane, L., Boyle, S., Sengupta, D., Grimes, G.R., Therizols, P., and Bickmore, W.A. (2019). Decreased Enhancer-Promoter Proximity Accompanying Enhancer Activation. Mol Cell *76*, 473-484 e477.

Benazet, J.D., Bischofberger, M., Tiecke, E., Goncalves, 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.

Benazet, 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.

Benazet, J.D., and Zeller, R. (2009). Vertebrate limb development: moving from classical morphogen gradients to an integrated 4-dimensional patterning system. Cold Spring Harb Perspect Biol *1*, a001339.

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.

Bender, M.A., Bulger, M., Close, J., and Groudine, M. (2000). Beta-globin gene switching and DNase I sensitivity of the endogenous beta-globin locus in mice do not require the locus control region. Mol Cell *5*, 387-393.

Berlivet, S., Paquette, D., Dumouchel, A., Langlais, D., Dostie, J., and Kmita, M. (2013). Clustering of tissue-specific sub-TADs accompanies the regulation of HoxA genes in developing limbs. PLoS Genet *9*, e1004018.

Bernstein, E., Duncan, E.M., Masui, O., Gil, J., Heard, E., and Allis, C.D. (2006). Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. Mol Cell Biol *26*, 2560-2569.

Bhatia, S., Monahan, J., Ravi, V., Gautier, P., Murdoch, E., Brenner, S., van Heyningen, V., Venkatesh, B., and Kleinjan, D.A. (2014). A survey of ancient conserved non-coding elements in the PAX6 locus reveals a landscape of interdigitated cis-regulatory archipelagos. Dev Biol *387*, 214-228.

Bhatia, V., Cao, Y., Ko, T.C., and Falzon, M. (2016). Parathyroid Hormone-Related Protein Interacts With the Transforming Growth Factor-beta/Bone Morphogenetic Protein-2/Gremlin Signaling Pathway to Regulate Proinflammatory and Profibrotic Mediators in Pancreatic Acinar and Stellate Cells. Pancreas *45*, 659-670.

Boehm, B., Westerberg, H., Lesnicar-Pucko, G., Raja, S., Rautschka, M., Cotterell, J., Swoger, J., and Sharpe, J. (2010). The role of spatially controlled cell proliferation in limb bud morphogenesis. PLoS Biol *8*, e1000420.

Boffelli, D., Nobrega, M.A., and Rubin, E.M. (2004). Comparative genomics at the vertebrate extremes. Nat Rev Genet *5*, 456-465.

Bolt, C.C., and Duboule, D. (2020). The regulatory landscapes of developmental genes. Development *147*.

Bonn, S., Zinzen, R.P., Girardot, C., Gustafson, E.H., Perez-Gonzalez, A., Delhomme, N., Ghavi-Helm, Y., Wilczynski, B., Riddell, A., and Furlong, E.E. (2012). Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. Nat Genet *44*, 148-156.

Bothma, J.P., Garcia, H.G., Ng, S., Perry, M.W., Gregor, T., and Levine, M. (2015). Enhancer additivity and non-additivity are determined by enhancer strength in the Drosophila embryo. Elife *4*.

Boyle, A.P., Davis, S., Shulha, H.P., Meltzer, P., Margulies, E.H., Weng, Z., Furey, T.S., and Crawford, G.E. (2008). High-resolution mapping and characterization of open chromatin across the genome. Cell *132*, 311-322.

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.

Bulger, M., and Groudine, M. (1999). Looping versus linking: toward a model for long-distance gene activation. Genes Dev *13*, 2465-2477.

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

Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D.M., Hill, A.J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F.J.*, et al.* (2019). The single-cell transcriptional landscape of mammalian organogenesis. Nature *566*, 496-502.

Capdevila, J., Tsukui, T., Rodriquez Esteban, C., Zappavigna, V., and Izpisua Belmonte, J.C. (1999). Control of vertebrate limb outgrowth by the proximal factor Meis2 and distal antagonism of BMPs by Gremlin. Mol Cell *4*, 839-849.

Carter, D., Chakalova, L., Osborne, C.S., Dai, Y.F., and Fraser, P. (2002). Long-range chromatin regulatory interactions in vivo. Nat Genet *32*, 623-626.

Castellone, M.D., and Laukkanen, M.O. (2017). TGF-beta1, WNT, and SHH signaling in tumor progression and in fibrotic diseases. Front Biosci (Schol Ed) *9*, 31-45.

Cattoni, D.I., Cardozo Gizzi, A.M., Georgieva, M., Di Stefano, M., Valeri, A., Chamousset, D., Houbron, C., Dejardin, S., Fiche, J.B., Gonzalez, I.*, et al.* (2017). Single-cell absolute contact probability detection reveals chromosomes are organized by multiple low-frequency yet specific interactions. Nat Commun *8*, 1753.

Cenani, A., and Lenz, W. (1967). [Total syndactylia and total radioulnar synostosis in 2 brothers. A contribution on the genetics of syndactylia]. Z Kinderheilkd *101*, 181-190.

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.

Consortium, E.P., Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigo, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T.*, et al.* (2007).

Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature *447*, 799-816.

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

De Smet, L., Winnepenninckx, B., Fryns, J.P., and Fabry, G. (1992). Cenani-Lenz type of syndactyly: a complex type of syndactyly with multiple synostoses. Genet Couns *3*, 145-147.

Degenhardt, K.R., Milewski, R.C., Padmanabhan, A., Miller, M., Singh, M.K., Lang, D., Engleka, K.A., Wu, M., Li, J., Zhou, D.*, et al.* (2010). Distinct enhancers at the Pax3 locus can function redundantly to regulate neural tube and neural crest expressions. Dev Biol *339*, 519-527.

Dekker, J., and Mirny, L. (2016). The 3D Genome as Moderator of Chromosomal Communication. Cell *164*, 1110-1121.

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. Science *295*, 1306-1311.

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.

Despang, A., Schopflin, R., Franke, M., Ali, S., Jerkovic, I., Paliou, C., Chan, W.L., Timmermann, B., Wittler, L., Vingron, M.*, et al.* (2019). Functional dissection of the Sox9-Kcnj2 locus identifies nonessential and instructive roles of TAD architecture. Nat Genet *51*, 1263-1271.

Dickel, D.E., Ypsilanti, A.R., Pla, R., Zhu, Y., Barozzi, I., Mannion, B.J., Khin, Y.S., Fukuda-Yuzawa, Y., Plajzer-Frick, I., Pickle, C.S.*, et al.* (2018). Ultraconserved Enhancers Are Required for Normal Development. Cell *172*, 491-499 e415.

Dimitrov, B.I., Voet, T., De Smet, L., Vermeesch, J.R., Devriendt, K., Fryns, J.P., and Debeer, P. (2010). Genomic rearrangements of the GREM1-FMN1 locus cause oligosyndactyly, radio-ulnar synostosis, hearing loss, renal defects syndrome and Cenani--Lenz-like non-syndromic oligosyndactyly. J Med Genet *47*, 569-574.

Dixon, J.R., Gorkin, D.U., and Ren, B. (2016). Chromatin Domains: The Unit of Chromosome Organization. Mol Cell *62*, 668-680.

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.

Dorsett, D. (2019). The Many Roles of Cohesin in Drosophila Gene Transcription. Trends Genet *35*, 542-551.

Drossopoulou, G., Lewis, K.E., Sanz-Ezquerro, J.J., Nikbakht, N., McMahon, A.P., Hofmann, C., and Tickle, C. (2000). A model for anteroposterior patterning of the vertebrate limb based on sequential long- and short-range Shh signalling and Bmp signalling. Development *127*, 1337-1348.

Dunipace, L., Ozdemir, A., and Stathopoulos, A. (2011). Complex interactions between cis-regulatory modules in native conformation are critical for Drosophila snail expression. Development *138*, 4075-4084.

El-Sherif, E., and Levine, M. (2016). Shadow Enhancers Mediate Dynamic Shifts of Gap Gene Expression in the Drosophila Embryo. Curr Biol *26*, 1164-1169.

Epner, E., Reik, A., Cimbora, D., Telling, A., Bender, M.A., Fiering, S., Enver, T., Martin, D.I., Kennedy, M., Keller, G.*, et al.* (1998). The beta-globin LCR is not necessary for an open chromatin structure or developmentally regulated transcription of the native mouse beta-globin locus. Mol Cell *2*, 447-455.

Evangelista, M., Pruyne, D., Amberg, D.C., Boone, C., and Bretscher, A. (2002). Formins direct Arp2/3-independent actin filament assembly to polarize cell growth in yeast. Nat Cell Biol *4*, 260-269.

Fabre, P.J., Leleu, M., Mascrez, B., Lo Giudice, Q., Cobb, J., and Duboule, D. (2018). Heterogeneous combinatorial expression of Hoxd genes in single cells during limb development. BMC Biol *16*, 101.

Fallon, J.F., Lopez, A., Ros, M.A., Savage, M.P., Olwin, B.B., and Simandl, B.K. (1994). FGF-2: apical ectodermal ridge growth signal for chick limb development. Science *264*, 104-107.

Farin, H.F., Ludtke, T.H., Schmidt, M.K., Placzko, S., Schuster-Gossler, K., Petry, M., Christoffels, V.M., and Kispert, A. (2013). Tbx2 terminates shh/fgf signaling in the

developing mouse limb bud by direct repression of gremlin1. PLoS Genet *9*, e1003467.

Felsenfeld, G. (1996). Chromatin unfolds. Cell *86*, 13-19.

Felsenfeld, G., and Groudine, M. (2003). Controlling the double helix. Nature *421*, 448-453.

Feregrino, C., Sacher, F., Parnas, O., and Tschopp, P. (2019). A single-cell transcriptomic atlas of the developing chicken limb. BMC Genomics *20*, 401.

Finn, E.H., Pegoraro, G., Brandao, H.B., Valton, A.L., Oomen, M.E., Dekker, J., Mirny, L., and Misteli, T. (2019). Extensive Heterogeneity and Intrinsic Variation in Spatial Genome Organization. Cell *176*, 1502-1515 e1510.

Flyamer, I.M., Gassler, J., Imakaev, M., Brandao, H.B., Ulianov, S.V., Abdennur, N., Razin, S.V., Mirny, L.A., and Tachibana-Konwalski, K. (2017). Single-nucleus Hi-C reveals unique chromatin reorganization at oocyte-to-zygote transition. Nature *544*, 110-114.

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.

Fromm, M., and Berg, P. (1983). Simian virus 40 early- and late-region promoter functions are enhanced by the 72-base-pair repeat inserted at distant locations and inverted orientations. Mol Cell Biol *3*, 991-999.

Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., and Mirny, L.A. (2016). Formation of Chromosomal Domains by Loop Extrusion. Cell Rep *15*, 2038-2049.

Fulco, C.P., Nasser, J., Jones, T.R., Munson, G., Bergman, D.T., Subramanian, V., Grossman, S.R., Anyoha, R., Doughty, B.R., Patwardhan, T.A.*, et al.* (2019). Activity-by-contact model of enhancer-promoter regulation from thousands of CRISPR perturbations. Nat Genet *51*, 1664-1669.

Furlong, E.E.M., and Levine, M. (2018). Developmental enhancers and chromosome topology. Science *361*, 1341-1345.

Furniss, D., Kan, S.H., Taylor, I.B., Johnson, D., Critchley, P.S., Giele, H.P., and Wilkie, A.O. (2009). Genetic screening of 202 individuals with congenital limb malformations and requiring reconstructive surgery. J Med Genet *46*, 730-735.

Gaillard, C., and Strauss, F. (1990). Ethanol precipitation of DNA with linear polyacrylamide as carrier. Nucleic Acids Res *18*, 378.

Galli, A., Robay, D., Osterwalder, M., Bao, X., Benazet, 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.

Ganan, Y., Macias, D., Duterque-Coquillaud, M., Ros, M.A., and Hurle, J.M. (1996). Role of TGF beta s and BMPs as signals controlling the position of the digits and the areas of interdigital cell death in the developing chick limb autopod. Development *122*, 2349-2357.

Gasseling, M.T., and Saunders, J.W., Jr. (1961). Effects of the apical ectodermal ridge on growth of the versene-stripped chick limb bud. Dev Biol *3*, 1-25.

Gaszner, M., and Felsenfeld, G. (2006). Insulators: exploiting transcriptional and epigenetic mechanisms. Nat Rev Genet *7*, 703-713.

Gerstein, M.B., Lu, Z.J., Van Nostrand, E.L., Cheng, C., Arshinoff, B.I., Liu, T., Yip, K.Y., Robilotto, R., Rechtsteiner, A., Ikegami, K.*, et al.* (2010). Integrative analysis of the Caenorhabditis elegans genome by the modENCODE project. Science *330*, 1775-1787.

Ghavi-Helm, Y., Jankowski, A., Meiers, S., Viales, R.R., Korbel, J.O., and Furlong, E.E.M. (2019). Highly rearranged chromosomes reveal uncoupling between genome topology and gene expression. Nat Genet *51*, 1272-1282.

Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., 3rd, and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods *6*, 343-345.

Giele, H., Giele, C., Bower, C., and Allison, M. (2001). The incidence and epidemiology of congenital upper limb anomalies: a total population study. J Hand Surg Am *26*, 628-634.

Gillies, S.D., Morrison, S.L., Oi, V.T., and Tonegawa, S. (1983). A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. Cell *33*, 717-728.

Gomez-Marin, C., Tena, J.J., Acemel, R.D., Lopez-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 U S A *112*, 7542-7547.

Gonen, N., Quinn, A., O'Neill, H.C., Koopman, P., and Lovell-Badge, R. (2017). Normal Levels of Sox9 Expression in the Developing Mouse Testis Depend on the TES/TESCO Enhancer, but This Does Not Act Alone. PLoS Genet *13*, e1006520.

Green, W.T., Wyatt, G.M., and Anderson, M. (1968). Orthoroentgenography as a method of measuring the bones of the lower extremities. Clin Orthop Relat Res *61*, 10-15.

Grillo, E., Ravelli, C., Corsini, M., Ballmer-Hofer, K., Zammataro, L., Oreste, P., Zoppetti, G., Tobia, C., Ronca, R., Presta, M.*, et al.* (2016). Monomeric gremlin is a novel vascular endothelial growth factor receptor-2 antagonist. Oncotarget *7*, 35353-35368.

Gros, J., and Tabin, C.J. (2014). Vertebrate limb bud formation is initiated by localized epithelial-to-mesenchymal transition. Science *343*, 1253-1256.

Gross, D.S., and Garrard, W.T. (1988). Nuclease hypersensitive sites in chromatin. Annu Rev Biochem *57*, 159-197.

Grosveld, F., van Assendelft, G.B., Greaves, D.R., and Kollias, G. (1987). Position-independent, high-level expression of the human beta-globin gene in transgenic mice. Cell *51*, 975-985.

Grzeschik, K.H. (2002). Human limb malformations; an approach to the molecular basis of development. Int J Dev Biol *46*, 983-991.

Haarhuis, J.H.I., van der Weide, R.H., Blomen, V.A., Yanez-Cuna, J.O., Amendola, M., van Ruiten, M.S., Krijger, P.H.L., Teunissen, H., Medema, R.H., van Steensel, B.*, et al.* (2017). The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. Cell *169*, 693-707 e614.

Haller, T., Dietl, P., Deetjen, P., and Volkl, H. (1996). The lysosomal compartment as intracellular calcium store in MDCK cells: a possible involvement in InsP3-mediated Ca2+ release. Cell Calcium *19*, 157-165.

Handoko, L., Xu, H., Li, G., Ngan, C.Y., Chew, E., Schnapp, M., Lee, C.W., Ye, C., Ping, J.L., Mulawadi, F.*, et al.* (2011). CTCF-mediated functional chromatin interactome in pluripotent cells. Nat Genet *43*, 630-638.

Hardison, R.C., and Taylor, J. (2012). Genomic approaches towards finding cis-regulatory modules in animals. Nat Rev Genet *13*, 469-483.

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

Hark, A.T., Schoenherr, C.J., Katz, D.J., Ingram, R.S., Levorse, J.M., and Tilghman, S.M. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. Nature *405*, 486-489.

Harmston, N., Baresic, A., and Lenhard, B. (2013). The mystery of extreme non-coding conservation. Philos Trans R Soc Lond B Biol Sci *368*, 20130021.

Harmston, N., Ing-Simmons, E., Tan, G., Perry, M., Merkenschlager, M., and Lenhard, B. (2017). Topologically associating domains are ancient features that coincide with Metazoan clusters of extreme noncoding conservation. Nat Commun *8*, 441.

Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A*., et al.* (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet *39*, 311-318.

Heinz, S., Benner, C., Spann, N., Bertolino, E., 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.

Hirsch, N., Eshel, R., Bar Yaacov, R., Shahar, T., Shmulevich, F., Dahan, I., Levaot, N., Kaplan, T., Lupianez, D.G., and Birnbaum, R.Y. (2018). Unraveling the transcriptional regulation of TWIST1 in limb development. PLoS Genet *14*, e1007738.

Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-Andre, V., Sigova, A.A., Hoke, H.A., and Young, R.A. (2013). Super-enhancers in the control of cell identity and disease. Cell *155*, 934-947.

Hollnagel, A., Oehlmann, V., Heymer, J., Ruther, 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.

Hong, J.W., Hendrix, D.A., and Levine, M.S. (2008). Shadow enhancers as a source of evolutionary novelty. Science *321*, 1314.

Hou, C., and Corces, V.G. (2012). Throwing transcription for a loop: expression of the genome in the 3D nucleus. Chromosoma *121*, 107-116.

Huang, S., Li, X., Yusufzai, T.M., Qiu, Y., and Felsenfeld, G. (2007). USF1 recruits histone modification complexes and is critical for maintenance of a chromatin barrier. Mol Cell Biol *27*, 7991-8002.

Hung, S., Saiakhova, A., Faber, Z.J., Bartels, C.F., Neu, D., Bayles, I., Ojo, E., Hong, E.S., Pontius, W.D., Morton, A.R.*, et al.* (2019). Mismatch repair-signature mutations activate gene enhancers across human colorectal cancer epigenomes. Elife *8*.

Iampietro, C., Gummalla, M., Mutero, A., Karch, F., and Maeda, R.K. (2010). Initiator elements function to determine the activity state of BX-C enhancers. PLoS Genet *6*, e1001260.

Iguchi, H., Chan, J.S., Seidah, N.G., and Chretien, M. (1984). Tissue distribution and molecular forms of a novel pituitary protein in the rat. Neuroendocrinology *39*, 453-458.

Ishizaki, T., Morishima, Y., Okamoto, M., Furuyashiki, T., Kato, T., and Narumiya, S. (2001). Coordination of microtubules and the actin cytoskeleton by the Rho effector mDia1. Nat Cell Biol *3*, 8-14.

Istrail, S., and Davidson, E.H. (2005). Logic functions of the genomic cis-regulatory code. Proc Natl Acad Sci U S A *102*, 4954-4959.

Jackson-Grusby, L., Kuo, A., and Leder, P. (1992). A variant limb deformity transcript expressed in the embryonic mouse limb defines a novel formin. Genes Dev *6*, 29-37.

Jaeger, E., Leedham, S., Lewis, A., Segditsas, S., Becker, M., Cuadrado, P.R., Davis, H., Kaur, K., Heinimann, K., Howarth, K.*, et al.* (2012). Hereditary mixed polyposis syndrome is caused by a 40-kb upstream duplication that leads to increased and ectopic expression of the BMP antagonist GREM1. Nat Genet *44*, 699-703.

Jaeger, E., Webb, E., Howarth, K., Carvajal-Carmona, L., Rowan, A., Broderick, P., Walther, A., Spain, S., Pittman, A., Kemp, Z.*, et al.* (2008). Common genetic variants at the CRAC1 (HMPS) locus on chromosome 15q13.3 influence colorectal cancer risk. Nat Genet *40*, 26-28.

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.

Jing, H., Vakoc, C.R., Ying, L., Mandat, S., Wang, H., Zheng, X., and Blobel, G.A. (2008). Exchange of GATA factors mediates transitions in looped chromatin organization at a developmentally regulated gene locus. Mol Cell *29*, 232-242.

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S.*, et al.* (2010). Mediator and cohesin connect gene expression and chromatin architecture. Nature *467*, 430-435.

Kaltcheva, M.M., Anderson, M.J., Harfe, B.D., and Lewandoski, M. (2016). BMPs are direct triggers of interdigital programmed cell death. Dev Biol *411*, 266-276.

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

Kentepozidou, E., Aitken, S.J., Feig, C., Stefflova, K., Ibarra-Soria, X., Odom, D.T., Roller, M., and Flicek, P. (2020). Clustered CTCF binding is an evolutionary mechanism to maintain topologically associating domains. Genome Biol *21*, 5.

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, T.H., Barrera, L.O., and Ren, B. (2007). ChIP-chip for genome-wide analysis of protein binding in mammalian cells. Curr Protoc Mol Biol *Chapter 21*, Unit 21 13.

Klein, F.A., Pakozdi, T., Anders, S., Ghavi-Helm, Y., Furlong, E.E., and Huber, W. (2015). FourCSeq: analysis of 4C sequencing data. Bioinformatics *31*, 3085-3091.

Kleinebrecht, J., Selow, J., and Winkler, W. (1982). The mouse mutant limb-deformity (ld). Anat Anz *152*, 313-324.

Kobielak, A., Pasolli, H.A., and Fuchs, E. (2004). Mammalian formin-1 participates in adherens junctions and polymerization of linear actin cables. Nat Cell Biol *6*, 21-30.

Korchynskyi, O., and ten Dijke, P. (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.

Kraft, K., Geuer, S., Will, A.J., Chan, W.L., Paliou, C., Borschiwer, M., Harabula, I., Wittler, L., Franke, M., Ibrahim, D.M.*, et al.* (2015). Deletions, Inversions, Duplications: Engineering of Structural Variants using CRISPR/Cas in Mice. Cell Rep *10*, 833-839.

Kragesteen, B.K., Brancati, F., Digilio, M.C., Mundlos, S., and Spielmann, M. (2019). H2AFY promoter deletion causes PITX1 endoactivation and Liebenberg syndrome. J Med Genet *56*, 246-251.

Kragesteen, B.K., Spielmann, M., Paliou, C., Heinrich, V., Schopflin, R., Esposito, A., Annunziatella, C., Bianco, S., Chiariello, A.M., Jerkovic, I.*, et al.* (2018). Dynamic 3D chromatin architecture contributes to enhancer specificity and limb morphogenesis. Nat Genet *50*, 1463-1473.

Krebs, A.R., Imanci, D., Hoerner, L., Gaidatzis, D., Burger, L., and Schubeler, D. (2017). Genome-wide Single-Molecule Footprinting Reveals High RNA Polymerase II Turnover at Paused Promoters. Mol Cell *67*, 411-422 e414.

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

Kvon, E.Z. (2015). Using transgenic reporter assays to functionally characterize enhancers in animals. Genomics *106*, 185-192.

Lam, D.D., de Souza, F.S., Nasif, S., Yamashita, M., Lopez-Leal, R., Otero-Corchon, V., Meece, K., Sampath, H., Mercer, A.J., Wardlaw, S.L.*, et al.* (2015). Partially redundant enhancers cooperatively maintain Mammalian pomc expression above a critical functional threshold. PLoS Genet *11*, e1004935.

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.

Laufer, E., Nelson, C.E., Johnson, R.L., Morgan, B.A., and Tabin, C. (1994). Sonic hedgehog and Fgf-4 act through a signaling cascade and feedback loop to integrate growth and patterning of the developing limb bud. Cell *79*, 993-1003.

Lavoz, C., Alique, M., Rodrigues-Diez, R., Pato, J., Keri, G., Mezzano, S., Egido, J., and Ruiz-Ortega, M. (2015). Gremlin regulates renal inflammation via the vascular endothelial growth factor receptor 2 pathway. J Pathol *236*, 407-420.

Leddin, M., Perrod, C., Hoogenkamp, M., Ghani, S., Assi, S., Heinz, S., Wilson, N.K., Follows, G., Schonheit, J., Vockentanz, L.*, et al.* (2011). Two distinct auto-regulatory loops operate at the PU.1 locus in B cells and myeloid cells. Blood *117*, 2827-2838.

Lee, H.K., Willi, M., Shin, H.Y., Liu, C., and Hennighausen, L. (2018). Progressing super-enhancer landscape during mammary differentiation controls tissue-specific gene regulation. Nucleic Acids Res *46*, 10796-10809.

Lettice, L.A., Heaney, S.J., 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.

Levine, M. (2010). Transcriptional enhancers in animal development and evolution. Curr Biol *20*, R754-763.

Lewandoski, M., Sun, X., and Martin, G.R. (2000). Fgf8 signalling from the AER is essential for normal limb development. Nat Genet *26*, 460-463.

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

Li, L.M., and Arnosti, D.N. (2011). Long- and short-range transcriptional repressors induce distinct chromatin states on repressed genes. Curr Biol *21*, 406-412.

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., Haarhuis, J.H.I., Sedeno Cacciatore, A., Oldenkamp, R., van Ruiten, M.S., Willems, L., Teunissen, H., Muir, K.W., de Wit, E., Rowland, B.D.*, et al.* (2020). The structural basis for cohesin-CTCF-anchored loops. Nature.

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O.*, et al.* (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science *326*, 289-293.

Long, H.K., Prescott, S.L., and Wysocka, J. (2016). Ever-Changing Landscapes: Transcriptional Enhancers in Development and Evolution. Cell *167*, 1170-1187.

Lopez-Rios, J., Duchesne, A., Speziale, D., Andrey, G., Peterson, K.A., Germann, P., Unal, 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.

Lopez-Rios, J., Speziale, D., Robay, D., Scotti, M., Osterwalder, M., Nusspaumer, G., Galli, A., Hollander, 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.

Lorberbaum, D.S., Ramos, A.I., Peterson, K.A., Carpenter, B.S., Parker, D.S., De, S., Hillers, L.E., Blake, V.M., Nishi, Y., McFarlane, M.R.*, et al.* (2016). An ancient yet flexible cis-regulatory architecture allows localized Hedgehog tuning by patched/Ptch1. Elife *5*.

Lu, D., Myers, A.R., George, N.P., and Keck, J.L. (2011). Mechanism of Exonuclease I stimulation by the single-stranded DNA-binding protein. Nucleic Acids Res *39*, 6536-6545.

Lupianez, D.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J.M., Laxova, R.*, et al.* (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. Cell *161*, 1012-1025.

Maas, R.L., Jepeal, L.I., Elfering, S.L., Holcombe, R.F., Morton, C.C., Eddy, R.L., Byers, M.G., Shows, T.B., and Leder, P. (1991). A human gene homologous to the formin gene residing at the murine limb deformity locus: chromosomal location and RFLPs. Am J Hum Genet *48*, 687-695.

Maeda, R.K., and Karch, F. (2011). Gene expression in time and space: additive vs hierarchical organization of cis-regulatory regions. Curr Opin Genet Dev *21*, 187-193.

Maekawa, T., Imamoto, F., Merlino, G.T., Pastan, I., and Ishii, S. (1989). Cooperative function of two separate enhancers of the human epidermal growth factor receptor proto-oncogene. J Biol Chem *264*, 5488-5494.

Majumder, P., and Boss, J.M. (2010). CTCF controls expression and chromatin architecture of the human major histocompatibility complex class II locus. Mol Cell Biol *30*, 4211-4223.

Malik, S. (2012). Syndactyly: phenotypes, genetics and current classification. Eur J Hum Genet *20*, 817-824.

Manseau, L., Calley, J., and Phan, H. (1996). Profilin is required for posterior patterning of the Drosophila oocyte. Development *122*, 2109-2116.

Mariani, F.V., Ahn, C.P., and Martin, G.R. (2008). Genetic evidence that FGFs have an instructive role in limb proximal-distal patterning. Nature *453*, 401-405.

Marigo, V., Johnson, R.L., Vortkamp, A., and Tabin, C.J. (1996). Sonic hedgehog differentially regulates expression of GLI and GLI3 during limb development. Dev Biol *180*, 273-283.

Marinic, M., Aktas, T., Ruf, S., and Spitz, F. (2013). An integrated holo-enhancer unit defines tissue and gene specificity of the Fgf8 regulatory landscape. Dev Cell *24*, 530-542.

Marquez-Exposito, L., Cantero-Navarro, E., R, R.R.-D., Orejudo, M., Tejera-Munoz, A., Tejedor, L., Rayego-Mateos, S., Randez-Carbayo, J., Santos-Sanchez, L., Mezzano, S.*, et al.* (2020). Molecular Regulation of Notch Signaling by Gremlin. Adv Exp Med Biol *1227*, 81-94.

Martowicz, M.L., Grass, J.A., Boyer, M.E., Guend, H., and Bresnick, E.H. (2005). Dynamic GATA factor interplay at a multicomponent regulatory region of the GATA-2 locus. J Biol Chem *280*, 1724-1732.

Mass, R.L., Zeller, R., Woychik, R.P., Vogt, T.F., and Leder, P. (1990). Disruption of formin-encoding transcripts in two mutant limb deformity alleles. Nature *346*, 853-855.

Matelot, M., and Noordermeer, D. (2016). Determination of High-Resolution 3D Chromatin Organization Using Circular Chromosome Conformation Capture (4C-seq). Methods Mol Biol *1480*, 223-241.

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.

Merino, R., Rodriguez-Leon, J., Macias, D., Ganan, Y., Economides, A.N., and Hurle, J.M. (1999). The BMP antagonist Gremlin regulates outgrowth, chondrogenesis and programmed cell death in the developing limb. Development *126*, 5515-5522.

Merkenschlager, M., and Nora, E.P. (2016). CTCF and Cohesin in Genome Folding and Transcriptional Gene Regulation. Annu Rev Genomics Hum Genet *17*, 17-43.

Messing, A., Behringer, R.R., Slapak, J.R., Lemke, G., Palmiter, R.D., and Brinster, R.L. (1990). Insertional mutation at the Id locus (again!) in a line of transgenic mice. Mouse Genome *87*, 107.

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.

Mihaly, J., Barges, S., Sipos, L., Maeda, R., Cleard, F., Hogga, I., Bender, W., Gyurkovics, H., and Karch, F. (2006). Dissecting the regulatory landscape of the Abd-B gene of the bithorax complex. Development *133*, 2983-2993.

Mir, M., Bickmore, W., Furlong, E.E.M., and Narlikar, G. (2019). Chromatin topology, condensates and gene regulation: shifting paradigms or just a phase? Development *146*.

Mitola, S., Ravelli, C., Moroni, E., Salvi, V., Leali, D., Ballmer-Hofer, K., Zammataro, L., and Presta, M. (2010). Gremlin is a novel agonist of the major proangiogenic receptor VEGFR2. Blood *116*, 3677-3680.

Montavon, T., Soshnikova, N., Mascrez, B., Joye, E., Thevenet, L., Splinter, E., de Laat, W., Spitz, F., and Duboule, D. (2011). A regulatory archipelago controls Hox genes transcription in digits. Cell *147*, 1132-1145.

Moon, A.M., and Ley, T.J. (1990). Conservation of the primary structure, organization, and function of the human and mouse beta-globin locus-activating regions. Proc Natl Acad Sci U S A *87*, 7693-7697.

Muro, E.M., Ibn-Salem, J., and Andrade-Navarro, M.A. (2019). The distributions of protein coding genes within chromatin domains in relation to human disease. Epigenetics Chromatin *12*, 72.

Nagano, T., Lubling, Y., Stevens, T.J., Schoenfelder, S., Yaffe, E., Dean, W., Laue, E.D., Tanay, A., and Fraser, P. (2013). Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. Nature *502*, 59-64.

Navratilova, P., Fredman, D., Hawkins, T.A., Turner, K., Lenhard, B., and Becker, T.S. (2009). Systematic human/zebrafish comparative identification of cis-regulatory activity around vertebrate developmental transcription factor genes. Dev Biol *327*, 526-540.

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

Nichols, M.H., and Corces, V.G. (2018). A tethered-inchworm model of SMC DNA translocation. Nat Struct Mol Biol *25*, 906-910.

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.

Niswander, L., Jeffrey, S., Martin, G.R., and Tickle, C. (1994). A positive feedback loop coordinates growth and patterning in the vertebrate limb. Nature *371*, 609-612.

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

Nora, E.P., Goloborodko, A., Valton, A.L., Gibcus, J.H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L.A., and Bruneau, B.G. (2017). Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. Cell *169*, 930-944 e922.

Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J.*, et al.* (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature *485*, 381-385.

Norrie, J.L., Lewandowski, J.P., Bouldin, C.M., Amarnath, S., Li, Q., Vokes, M.S., Ehrlich, L.I.R., Harfe, B.D., and Vokes, S.A. (2014). Dynamics of BMP signaling in limb bud mesenchyme and polydactyly. Dev Biol *393*, 270-281.

O'Rourke, D.A., Liu, Z.X., Sellin, L., Spokes, K., Zeller, R., and Cantley, L.G. (2000). Hepatocyte growth factor induces MAPK-dependent formin IV translocation in renal epithelial cells. J Am Soc Nephrol *11*, 2212-2221.

Ong, C.T., and Corces, V.G. (2014). CTCF: an architectural protein bridging genome topology and function. Nat Rev Genet *15*, 234-246.

Osterwalder, M., Barozzi, I., Tissieres, V., Fukuda-Yuzawa, Y., Mannion, B.J., Afzal, S.Y., Lee, E.A., Zhu, Y., Plajzer-Frick, I., Pickle, C.S.*, et al.* (2018). Enhancer redundancy provides phenotypic robustness in mammalian development. Nature *554*, 239-243.

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.

Paliou, C., Guckelberger, P., Schopflin, R., Heinrich, V., Esposito, A., Chiariello, A.M., Bianco, S., Annunziatella, C., Helmuth, J., Haas, S.*, et al.* (2019). Preformed

chromatin topology assists transcriptional robustness of Shh during limb development. Proc Natl Acad Sci U S A *116*, 12390-12399.

Pavel, E., Zhao, W., Powell, K.A., Weinstein, M., and Kirschner, L.S. (2007). Analysis of a new allele of limb deformity (ld) reveals tissue- and age-specific transcriptional effects of the Ld Global Control Region. Int J Dev Biol *51*, 273-281.

Pearse, R.V., 2nd, Scherz, P.J., Campbell, J.K., and Tabin, C.J. (2007). A cellular lineage analysis of the chick limb bud. Dev Biol *310*, 388-400.

Pennacchio, L.A., Ahituv, N., Moses, A.M., Prabhakar, S., Nobrega, M.A., Shoukry, M., Minovitsky, S., Dubchak, I., Holt, A., Lewis, K.D.*, et al.* (2006). In vivo enhancer analysis of human conserved non-coding sequences. Nature *444*, 499-502.

Perry, M.W., Boettiger, A.N., and Levine, M. (2011). Multiple enhancers ensure precision of gap gene-expression patterns in the Drosophila embryo. Proc Natl Acad Sci U S A *108*, 13570-13575.

Petit, F., Sears, K.E., and Ahituv, N. (2017). Limb development: a paradigm of gene regulation. Nat Rev Genet *18*, 245-258.

Petrykowska, H.M., Vockley, C.M., and Elnitski, L. (2008). Detection and characterization of silencers and enhancer-blockers in the greater CFTR locus. Genome Res *18*, 1238-1246.

Phillips, J.E., and Corces, V.G. (2009). CTCF: master weaver of the genome. Cell *137*, 1194-1211.

Phillips-Cremins, J.E., and Corces, V.G. (2013). Chromatin insulators: linking genome organization to cellular function. Mol Cell *50*, 461-474.

Phillips-Cremins, J.E., Sauria, M.E., Sanyal, A., Gerasimova, T.I., Lajoie, B.R., Bell, J.S., Ong, C.T., Hookway, T.A., Guo, C., Sun, Y.*, et al.* (2013). Architectural protein subclasses shape 3D organization of genomes during lineage commitment. Cell *153*, 1281-1295.

Pizette, S., and Niswander, L. (2000). BMPs are required at two steps of limb chondrogenesis: formation of prechondrogenic condensations and their differentiation into chondrocytes. Dev Biol *219*, 237-249.

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

Pope, B.D., Ryba, T., Dileep, V., Yue, F., Wu, W., Denas, O., Vera, D.L., Wang, Y., Hansen, R.S., Canfield, T.K.*, et al.* (2014). Topologically associating domains are stable units of replication-timing regulation. Nature *515*, 402-405.

Pring, M., Evangelista, M., Boone, C., Yang, C., and Zigmond, S.H. (2003). Mechanism of formin-induced nucleation of actin filaments. Biochemistry *42*, 486-496.

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. Development *138*, 1913-1923.

Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S.A., Flynn, R.A., and Wysocka, J. (2011). A unique chromatin signature uncovers early developmental enhancers in humans. Nature *470*, 279-283.

Rao, S.S., 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.

Rao, S.S.P., Huang, S.C., Glenn St Hilaire, B., Engreitz, J.M., Perez, E.M., Kieffer-Kwon, K.R., Sanborn, A.L., Johnstone, S.E., Bascom, G.D., Bochkov, I.D.*, et al.* (2017). Cohesin Loss Eliminates All Loop Domains. Cell *171*, 305-320 e324.

Reik, A., Telling, A., Zitnik, G., Cimbora, D., Epner, E., and Groudine, M. (1998). The locus control region is necessary for gene expression in the human beta-globin locus but not the maintenance of an open chromatin structure in erythroid cells. Mol Cell Biol *18*, 5992-6000.

Reinhardt, R., Gullotta, F., Nusspaumer, G., Unal, E., Ivanek, R., Zuniga, A., and Zeller, R. (2019). Molecular signatures identify immature mesenchymal progenitors in early mouse limb buds that respond differentially to morphogen signaling. Development *146*.

Ren, J., Smid, M., Iaria, J., Salvatori, D.C.F., van Dam, H., Zhu, H.J., Martens, J.W.M., and Ten Dijke, P. (2019). Cancer-associated fibroblast-derived Gremlin 1 promotes breast cancer progression. Breast Cancer Res *21*, 109.

Ritter, D.I., Li, Q., Kostka, D., Pollard, K.S., Guo, S., and Chuang, J.H. (2010). The importance of being cis: evolution of orthologous fish and mammalian enhancer activity. Mol Biol Evol *27*, 2322-2332.

Rodriguez-Carballo, E., Lopez-Delisle, L., Yakushiji-Kaminatsui, N., Ullate-Agote, A., and Duboule, D. (2019). Impact of genome architecture on the functional activation and repression of Hox regulatory landscapes. BMC Biol *17*, 55.

Ruiz i Altaba, A. (1998). Combinatorial Gli gene function in floor plate and neuronal inductions by Sonic hedgehog. Development *125*, 2203-2212.

Rusche, L.N., Kirchmaier, A.L., and Rine, J. (2003). The establishment, inheritance, and function of silenced chromatin in Saccharomyces cerevisiae. Annu Rev Biochem *72*, 481-516.

Sagai, T., Hosoya, M., Mizushina, Y., Tamura, M., and Shiroishi, T. (2005). Elimination of a long-range cis-regulatory module causes complete loss of limb-specific Shh expression and truncation of the mouse limb. Development *132*, 797-803.

Sagot, I., Klee, S.K., and Pellman, D. (2002). Yeast formins regulate cell polarity by controlling the assembly of actin cables. Nat Cell Biol *4*, 42-50.

Sandelin, A., Bailey, P., Bruce, S., Engstrom, P.G., Klos, J.M., Wasserman, W.W., Ericson, J., and Lenhard, B. (2004). Arrays of ultraconserved non-coding regions span the loci of key developmental genes in vertebrate genomes. BMC Genomics *5*, 99.

Sanyal, A., Lajoie, B.R., Jain, G., and Dekker, J. (2012). The long-range interaction landscape of gene promoters. Nature *489*, 109-113.

Saunders, J.W., Jr. (1948). The proximo-distal sequence of origin of the parts of the chick wing and the role of the ectoderm. J Exp Zool *108*, 363-403.

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.

Schoenfelder, S., and Fraser, P. (2019). Long-range enhancer-promoter contacts in gene expression control. Nat Rev Genet *20*, 437-455.

Schubeler, D., Groudine, M., and Bender, M.A. (2001). The murine beta-globin locus control region regulates the rate of transcription but not the hyperacetylation of histones at the active genes. Proc Natl Acad Sci U S A *98*, 11432-11437.

Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B., and Cavalli, G. (2007). Genome regulation by polycomb and trithorax proteins. Cell *128*, 735-745.

Schwarzer, W., Abdennur, N., Goloborodko, A., Pekowska, A., Fudenberg, G., Loe-Mie, Y., Fonseca, N.A., Huber, W., Haering, C.H., Mirny, L.*, et al.* (2017). Two independent modes of chromatin organization revealed by cohesin removal. Nature *551*, 51-56.

Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., and Cavalli, G. (2012). Three-dimensional folding and functional organization principles of the Drosophila genome. Cell *148*, 458-472.

Shen, Y., Yue, F., McCleary, D.F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenkov, V.V.*, et al.* (2012). A map of the cis-regulatory sequences in the mouse genome. Nature *488*, 116-120.

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.

Sheth, R., Gregoire, D., Dumouchel, A., Scotti, M., Pham, J.M., Nemec, S., Bastida, M.F., Ros, M.A., and Kmita, M. (2013). Decoupling the function of Hox and Shh in developing limb reveals multiple inputs of Hox genes on limb growth. Development *140*, 2130-2138.

Shi, W., Zhao, J., Anderson, K.D., and Warburton, D. (2001). Gremlin negatively modulates BMP-4 induction of embryonic mouse lung branching morphogenesis. Am J Physiol Lung Cell Mol Physiol *280*, L1030-1039.

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

Sima, J., Chakraborty, A., Dileep, V., Michalski, M., Klein, K.N., Holcomb, N.P., Turner, J.L., Paulsen, M.T., Rivera-Mulia, J.C., Trevilla-Garcia, C.*, et al.* (2019). Identifying cis Elements for Spatiotemporal Control of Mammalian DNA Replication. Cell *176*, 816-830 e818.

Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B., and de Laat, W. (2006). Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nat Genet *38*, 1348-1354.

Skarnes, W.C., Rosen, B., West, A.P., Koutsourakis, M., Bushell, W., Iyer, V., Mujica, A.O., Thomas, M., Harrow, J., Cox, T.*, et al.* (2011). A conditional knockout resource for the genome-wide study of mouse gene function. Nature *474*, 337-342.

Sofueva, S., Yaffe, E., Chan, W.C., Georgopoulou, D., Vietri Rudan, M., Mira-Bontenbal, H., Pollard, S.M., Schroth, G.P., Tanay, A., and Hadjur, S. (2013). Cohesin-mediated interactions organize chromosomal domain architecture. EMBO J *32*, 3119-3129.

Song, W., Sharan, R., and Ovcharenko, I. (2019). The first enhancer in an enhancer chain safeguards subsequent enhancer-promoter contacts from a distance. Genome Biol *20*, 197.

Spitz, F., Gonzalez, F., and Duboule, D. (2003). A global control region defines a chromosomal regulatory landscape containing the HoxD cluster. Cell *113*, 405-417.

Stalder, J., Larsen, A., Engel, J.D., Dolan, M., Groudine, M., and Weintraub, H. (1980). Tissue-specific DNA cleavages in the globin chromatin domain introduced by DNAase I. Cell *20*, 451-460.

Stam, M., Tark-Dame, M., and Fransz, P. (2019). 3D genome organization: a role for phase separation and loop extrusion? Curr Opin Plant Biol *48*, 36-46.

Stanojevic, D., Small, S., and Levine, M. (1991). Regulation of a segmentation stripe by overlapping activators and repressors in the Drosophila embryo. Science *254*, 1385-1387.

Stathopoulos, A., and Levine, M. (2005). Localized repressors delineate the neurogenic ectoderm in the early Drosophila embryo. Dev Biol *280*, 482-493.

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.

Stine, Z.E., McGaughey, D.M., Bessling, S.L., Li, S., and McCallion, A.S. (2011). Steroid hormone modulation of RET through two estrogen responsive enhancers in breast cancer. Hum Mol Genet *20*, 3746-3756.

Sun, X., Mariani, F.V., and Martin, G.R. (2002). Functions of FGF signalling from the apical ectodermal ridge in limb development. Nature *418*, 501-508.

Symmons, O., Pan, L., Remeseiro, S., Aktas, T., Klein, F., Huber, W., and Spitz, F. (2016). The Shh Topological Domain Facilitates the Action of Remote Enhancers by Reducing the Effects of Genomic Distances. Dev Cell *39*, 529-543.

Symmons, O., Uslu, V.V., Tsujimura, T., Ruf, S., Nassari, S., Schwarzer, W., Ettwiller, L., and Spitz, F. (2014). Functional and topological characteristics of mammalian regulatory domains. Genome Res *24*, 390-400.

Tang, Z., Luo, O.J., Li, X., Zheng, M., Zhu, J.J., Szalaj, P., Trzaskoma, P., Magalska, A., Wlodarczyk, J., Ruszczycki, B.*, et al.* (2015). CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription. Cell *163*, 1611-1627.

ten Berge, D., Brugmann, S.A., Helms, J.A., and Nusse, R. (2008). Wnt and FGF signals interact to coordinate growth with cell fate specification during limb development. Development *135*, 3247-3257.

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

Tickle, C. (1981). The number of polarizing region cells required to specify additional digits in the developing chick wing. Nature *289*, 295-298.

Todt, W.L., and Fallon, J.F. (1987). Posterior apical ectodermal ridge removal in the chick wing bud triggers a series of events resulting in defective anterior pattern formation. Development *101*, 501-515.

Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F., and de Laat, W. (2002). Looping and interaction between hypersensitive sites in the active beta-globin locus. Mol Cell *10*, 1453-1465.

Towers, M., Mahood, R., Yin, Y., and Tickle, C. (2008). Integration of growth and specification in chick wing digit-patterning. Nature *452*, 882-886.

Vakoc, C.R., Letting, D.L., Gheldof, N., Sawado, T., Bender, M.A., Groudine, M., Weiss, M.J., Dekker, J., and Blobel, G.A. (2005). Proximity among distant regulatory elements at the beta-globin locus requires GATA-1 and FOG-1. Mol Cell *17*, 453-462.

Valenzuela, L., Dhillon, N., Dubey, R.N., Gartenberg, M.R., and Kamakaka, R.T. (2008). Long-range communication between the silencers of HMR. Mol Cell Biol *28*, 1924-1935.

Valenzuela, L., and Kamakaka, R.T. (2006). Chromatin insulators. Annu Rev Genet *40*, 107-138.

Venkatachalam, R., Verwiel, E.T., Kamping, E.J., Hoenselaar, E., Gorgens, H., Schackert, H.K., van Krieken, J.H., Ligtenberg, M.J., Hoogerbrugge, N., van Kessel, A.G.*, et al.* (2011). Identification of candidate predisposing copy number variants in familial and early-onset colorectal cancer patients. Int J Cancer *129*, 1635-1642.

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

Visel, A., Blow, M.J., Li, Z., Zhang, T., Akiyama, J.A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F.*, et al.* (2009a). ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature *457*, 854-858.

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

Vogt, T.F., Jackson-Grusby, L., Rush, J., and Leder, P. (1993). Formins: phosphoprotein isoforms encoded by the mouse limb deformity locus. Proc Natl Acad Sci U S A *90*, 5554-5558.

Vogt, T.F., Jackson-Grusby, L., Wynshaw-Boris, A.J., Chan, D.C., and Leder, P. (1992). The same genomic region is disrupted in two transgene-induced limb deformity alleles. Mamm Genome *3*, 431-437.

Vokes, S.A., Ji, H., Wong, W.H., and McMahon, A.P. (2008). A genome-scale analysis of the cis-regulatory circuitry underlying sonic hedgehog-mediated patterning of the mammalian limb. Genes Dev *22*, 2651-2663.

Wallace, J.A., and Felsenfeld, G. (2007). We gather together: insulators and genome organization. Curr Opin Genet Dev *17*, 400-407.

Wallar, B.J., and Alberts, A.S. (2003). The formins: active scaffolds that remodel the cytoskeleton. Trends Cell Biol *13*, 435-446.

Wanek, N., Muneoka, K., Holler-Dinsmore, G., Burton, R., and Bryant, S.V. (1989). A staging system for mouse limb development. J Exp Zool *249*, 41-49.

Wang, B., Fallon, J.F., and Beachy, P.A. (2000). Hedgehog-regulated processing of Gli3 produces an anterior/posterior repressor gradient in the developing vertebrate limb. Cell *100*, 423-434.

Wen, X., Lai, C.K., Evangelista, M., Hongo, J.A., de Sauvage, F.J., and Scales, S.J. (2010). Kinetics of hedgehog-dependent full-length Gli3 accumulation in primary cilia and subsequent degradation. Mol Cell Biol *30*, 1910-1922.

Wendt, K.S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S., Nagae, G., Ishihara, K., Mishiro, T.*, et al.* (2008). Cohesin mediates transcriptional insulation by CCCTC-binding factor. Nature *451*, 796-801.

Westphal, C.H., Muller, L., Zhou, A., Zhu, X., Bonner-Weir, S., Schambelan, M., Steiner, D.F., Lindberg, I., and Leder, P. (1999). The neuroendocrine protein 7B2 is required for peptide hormone processing in vivo and provides a novel mechanism for pituitary Cushing's disease. Cell *96*, 689-700.

Whyte, W.A., Orlando, D.A., Hnisz, D., Abraham, B.J., Lin, C.Y., Kagey, M.H., Rahl, P.B., Lee, T.I., and Young, R.A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell *153*, 307-319.

Will, A.J., Cova, G., Osterwalder, M., Chan, W.L., Wittler, L., Brieske, N., Heinrich, V., de Villartay, J.P., Vingron, M., Klopocki, E.*, et al.* (2017). Composition and dosage of a multipartite enhancer cluster control developmental expression of Ihh (Indian hedgehog). Nat Genet *49*, 1539-1545.

Williamson, I., Kane, L., Devenney, P.S., Flyamer, I.M., Anderson, E., Kilanowski, F., Hill, R.E., Bickmore, W.A., and Lettice, L.A. (2019). Developmentally regulated Shh expression is robust to TAD perturbations. Development *146*.

Woychik, R.P., Generoso, W.M., Russell, L.B., Cain, K.T., Cacheiro, N.L., Bultman, S.J., Selby, P.B., Dickinson, M.E., Hogan, B.L., and Rutledge, J.C. (1990a). Molecular and genetic characterization of a radiation-induced structural rearrangement in mouse chromosome 2 causing mutations at the limb deformity and agouti loci. Proc Natl Acad Sci U S A *87*, 2588-2592.

Woychik, R.P., Maas, R.L., Zeller, R., Vogt, T.F., and Leder, P. (1990b). 'Formins': proteins deduced from the alternative transcripts of the limb deformity gene. Nature *346*, 850-853.

Woychik, R.P., Stewart, T.A., Davis, L.G., D'Eustachio, P., and Leder, P. (1985). An inherited limb deformity created by insertional mutagenesis in a transgenic mouse. Nature *318*, 36-40.

Wright, E., Hargrave, M.R., Christiansen, J., Cooper, L., Kun, J., Evans, T., Gangadharan, U., Greenfield, A., and Koopman, P. (1995). The Sry-related gene Sox9 is expressed during chondrogenesis in mouse embryos. Nat Genet *9*, 15-20.

Wutz, G., Varnai, C., Nagasaka, K., Cisneros, D.A., Stocsits, R.R., Tang, W., Schoenfelder, S., Jessberger, G., Muhar, M., Hossain, M.J.*, et al.* (2017). Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. EMBO J *36*, 3573-3599.

Wynshaw-Boris, A., Ryan, G., Deng, C.X., Chan, D.C., Jackson-Grusby, L., Larson, D., Dunmore, J.H., and Leder, P. (1997). The role of a single formin isoform in the limb and renal phenotypes of limb deformity. Mol Med *3*, 372-384.

Yin, M., Tissari, M., Tamminen, J., Ylivinkka, I., Ronty, M., von Nandelstadh, P., Lehti, K., Hyytiainen, M., Myllarniemi, M., and Koli, K. (2017). Gremlin-1 is a key regulator of the invasive cell phenotype in mesothelioma. Oncotarget *8*, 98280-98297.

Yoon, B.S., Ovchinnikov, D.A., Yoshii, I., Mishina, Y., Behringer, R.R., and Lyons, K.M. (2005). Bmpr1a and Bmpr1b have overlapping functions and are essential for chondrogenesis in vivo. Proc Natl Acad Sci U S A *102*, 5062-5067.

Zang, C., Schones, D.E., Zeng, C., Cui, K., Zhao, K., and Peng, W. (2009). A clustering approach for identification of enriched domains from histone modification ChIP-Seq data. Bioinformatics *25*, 1952-1958.

Zeller, R., Jackson-Grusby, L., and Leder, P. (1989). The limb deformity gene is required for apical ectodermal ridge differentiation and anteroposterior limb pattern formation. Genes Dev *3*, 1481-1492.

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

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

Zentner, G.E., Tesar, P.J., and Scacheri, P.C. (2011). Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. Genome Res *21*, 1273-1283.

Zhang, J., Tian, X.J., and Xing, J. (2016). Signal Transduction Pathways of EMT Induced by TGF-beta, SHH, and WNT and Their Crosstalks. J Clin Med *5*.

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.

Zhou, F., Leder, P., and Martin, S.S. (2006). Formin-1 protein associates with microtubules through a peptide domain encoded by exon-2. Exp Cell Res *312*, 1119-1126.

Zhou, F., Leder, P., Zuniga, A., and Dettenhofer, M. (2009). Formin1 disruption confers oligodactylism and alters Bmp signaling. Hum Mol Genet *18*, 2472-2482.

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.

Zou, H., and Niswander, L. (1996). Requirement for BMP signaling in interdigital apoptosis and scale formation. Science *272*, 738-741.

Zucker, R.M., Hunter, E.S., 3rd, and Rogers, J.M. (1999). Apoptosis and morphology in mouse embryos by confocal laser scanning microscopy. Methods *18*, 473-480.

Zucker, R.M., Hunter, S., and Rogers, J.M. (1998). Confocal laser scanning microscopy of apoptosis in organogenesis-stage mouse embryos. Cytometry *33*, 348-354.

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

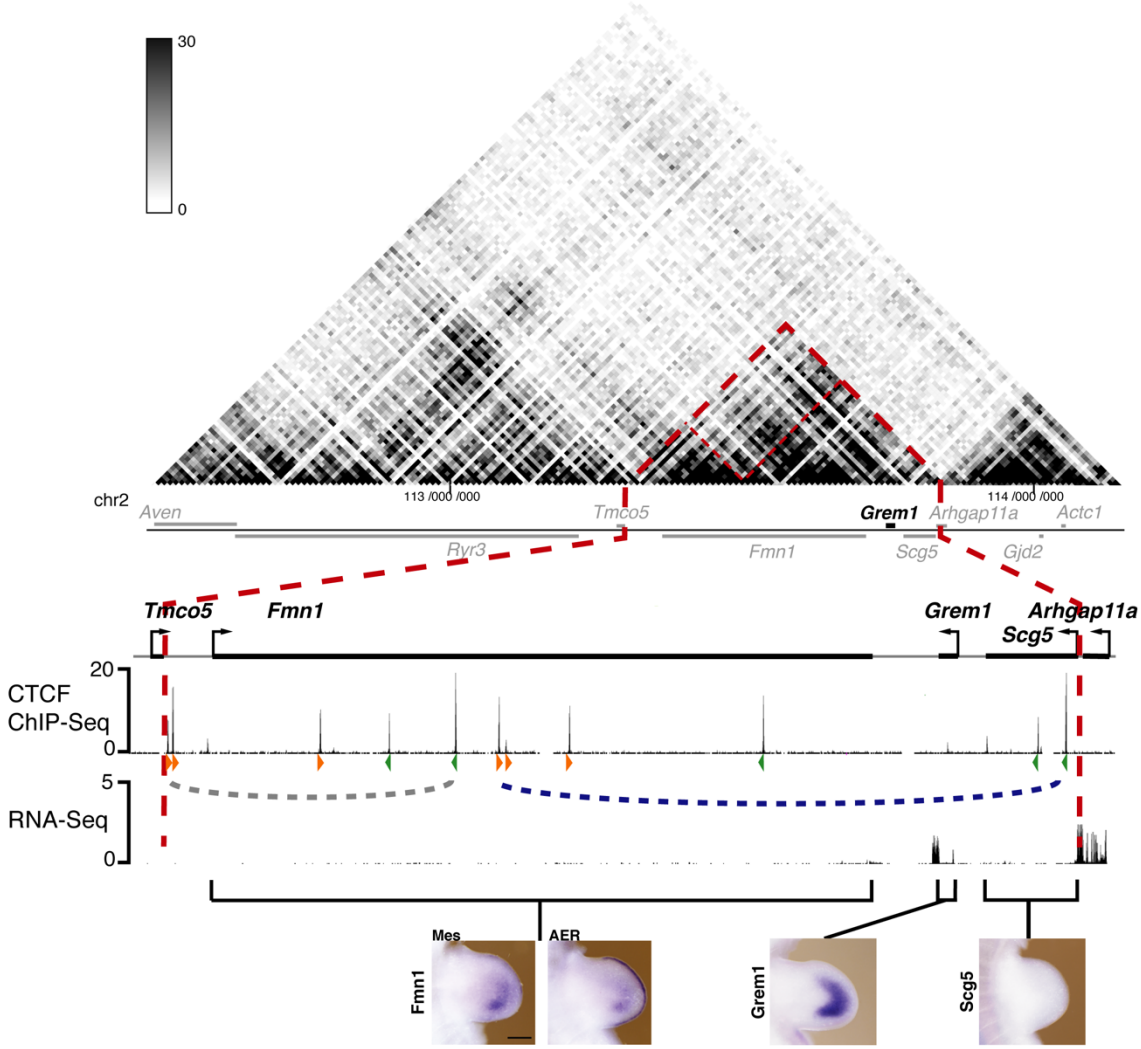
Zuniga, A., Laurent, F., Lopez-Rios, J., Klasen, C., Matt, N., and Zeller, R. (2012a). Conserved cis-regulatory regions in a large genomic landscape control SHH and BMP-regulated Gremlin1 expression in mouse limb buds. BMC Dev Biol *12*, 23.

Zuniga, A., Michos, O., Spitz, F., Haramis, A.P., 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.

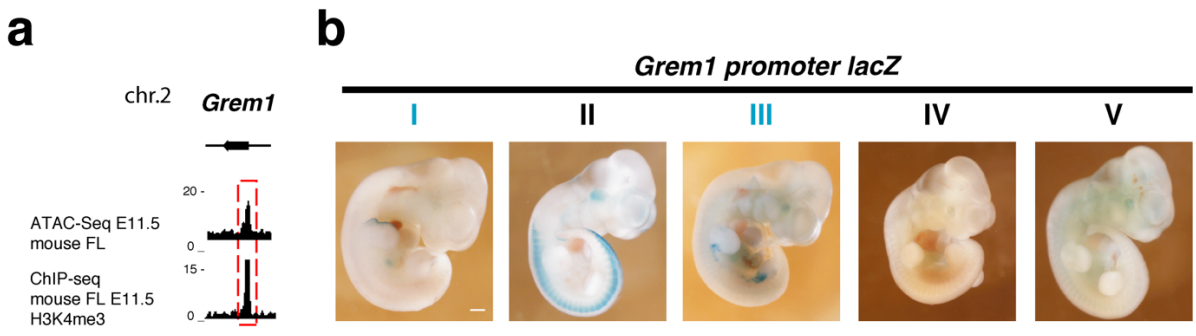
Zuniga, A., and Zeller, R. (1999). Gli3 (Xt) and formin (ld) participate in the positioning of the polarising region and control of posterior limb-bud identity. Development *126*, 13-21.

Zuniga, A., Zeller, R., and Probst, S. (2012b). The molecular basis of human congenital limb malformations. Wiley Interdiscip Rev Dev Biol *1*, 803-822.

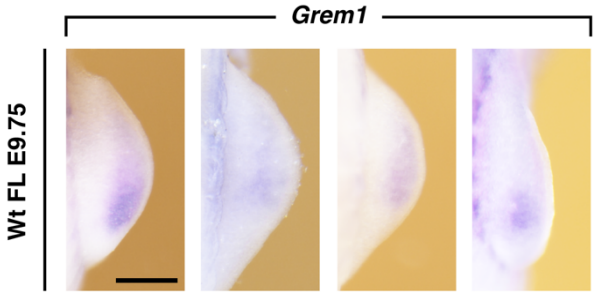
12. Appendixes



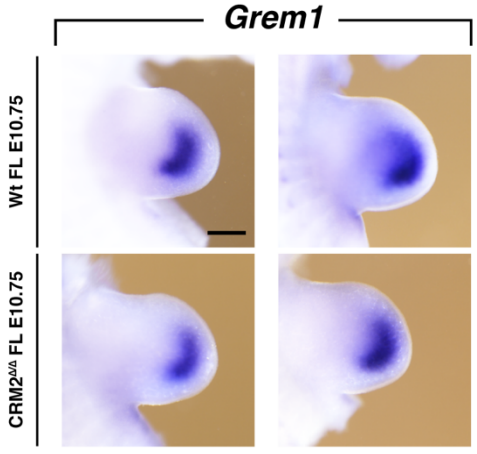
Appendixes Figure 1. Identification *Grem1* topologically-associated domain. Upper panel: Hi-C map (Barutcu et al., 2018) obtained from mouse embryonic fibroblasts with a 10kb resolution show the distribution of TADs on either side of *Grem1* gene in black, other genes are represented in dark grey lines, dashed lines indicate TADs, E10.5 mouse forelimb bud CTCF ChIP-Seq (Andrey et al., 2017) identify CTCF binding sites, their orientation are represented with orange and green arrowheads. Blue and grey dashed lines unveil a potential *Grem1* TAD and a potential *Fmn1* TAD respectively. E10.5 mouse forelimb bud RNA-Seq (Andrey et al., 2017) shows the transcriptional activity of *Grem1* neighboring genes. Lower panel: WISH performed by Victorio Palacio and Laurène Ramos Martins showing the limb expression of *Grem1* TAD genes (n=3). *Scale bar: 250µm*



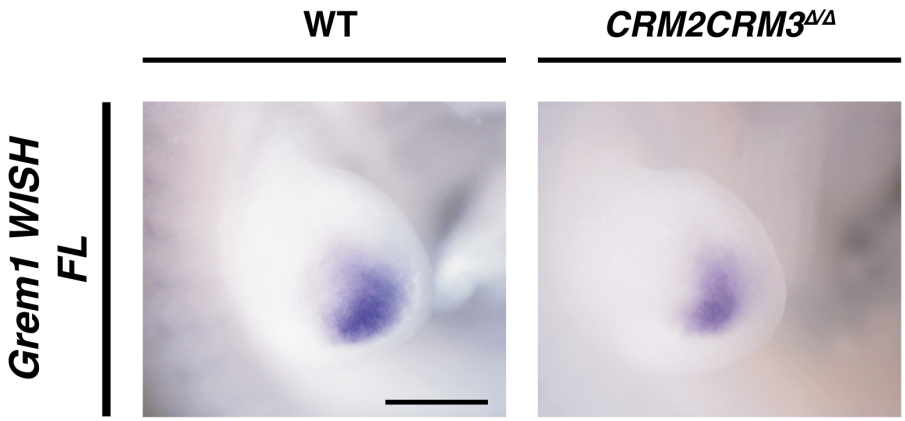
Appendixes Figure 2. *LacZ* expression in the *Grem1-promoter lacZ* transgenic embryos forelimb at E11.5. (a): E11.5 forelimb bud ATAC-seq profile, ChIP-seq profiles obtain from E11.5 mouse forelimb bud tissue targeting the promoter histone mark H3K4me3 (Sheth et al., 2016) on a 60kb region centered on *Grem1*, the red dashed box highlight *Grem1* promoter. (b) *Grem1*_promoter_*lacZ* plasmid transgenic embryos collected at E11.5, blue numbers show limb expressers. *Scale bar: 500µm*



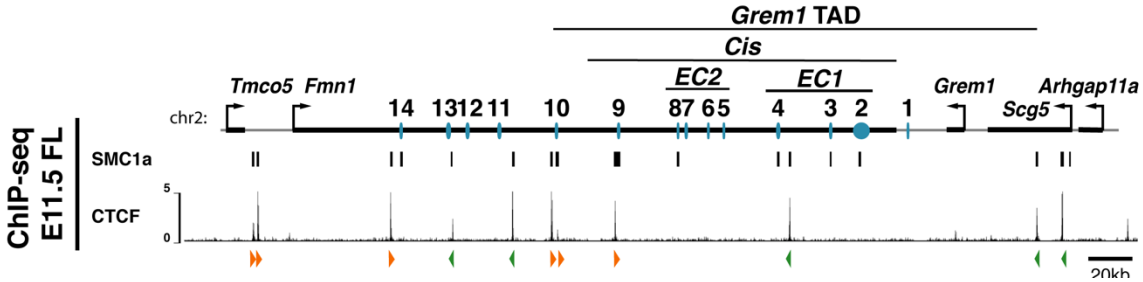
Appendixes Figure 3. *Grem1* WISH at early developmental stage E9.75 (26-28ss) stage is quite variable. *Grem1* WISH on 4 forelimbs buds at 26-28ss. *Scale bar: 250µm*



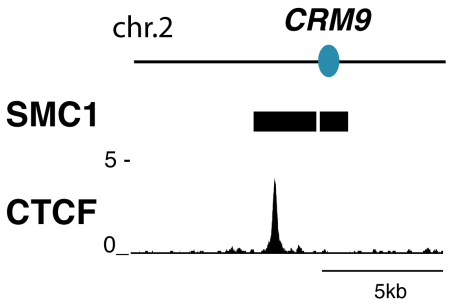
Appendixes Figure 4. *Grem1* WISH at developmental stage E10.75 (39ss) for *CRM2*$^{\Delta/\Delta}$ mutants. *Grem1* WISH on 2 wild-type forelimbs buds at 26-28ss in comparison with two *CRM2*$^{\Delta/\Delta}$ forelimb buds at the same stage. *Scale bar: 250µm*



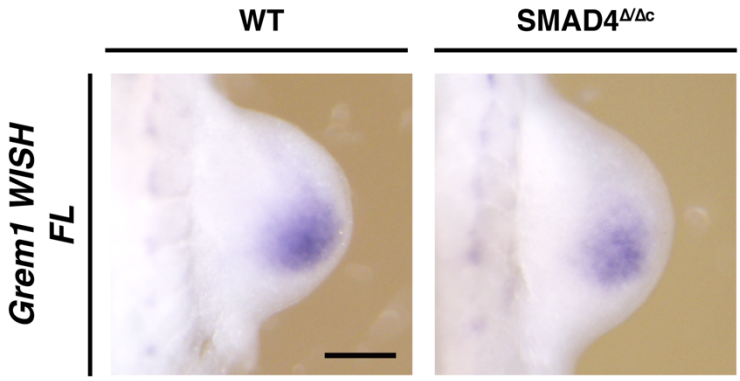
Appendixes Figure 5. *Grem1* WISH on E10.5 WT and *CRM2CRM3*$^{\Delta/\Delta}$ mutants forelimb buds *Grem1* WISH on wild-type forelimbs buds at 35-36ss in comparison with *CRM2CRM3*$^{\Delta/\Delta}$ forelimbs at the same stage *Scale bar: 250µm*



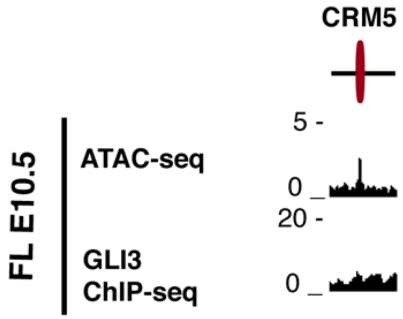
Appendixes Figure 6. Identification of cohesin binding site in the *Grem1/Fmn1* landscape. Peak calling from ChIP-seq experiments targeting cohesin subunit SMC1a (Demare et al., 2013) compared to CTCF binding sites at E11.5 in mouse FL (Andrey et al., 2017). Orange and green arrowheads represent CTCF motifs orientation.



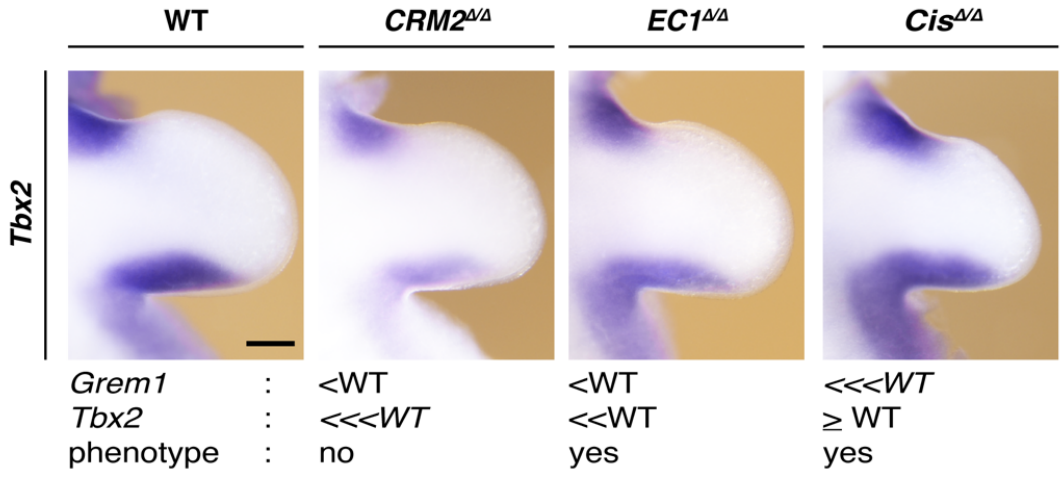
Appendixes Figure 7. ChIP-seq revealing cohesin binding at *CRM9* in mouse forelimb at E11.5 SMC1 (Demare et al., 2013) is a subunit of cohesin, ChIP-seq peak calling of SMC1 reveal cohesin binding site. CTCF ChIP-seq Andrey et al., 2017) on the lower track shows no overlapping with *CRM9* locus. (coordinates mm10 chr2:113,529,824-113,542,826)



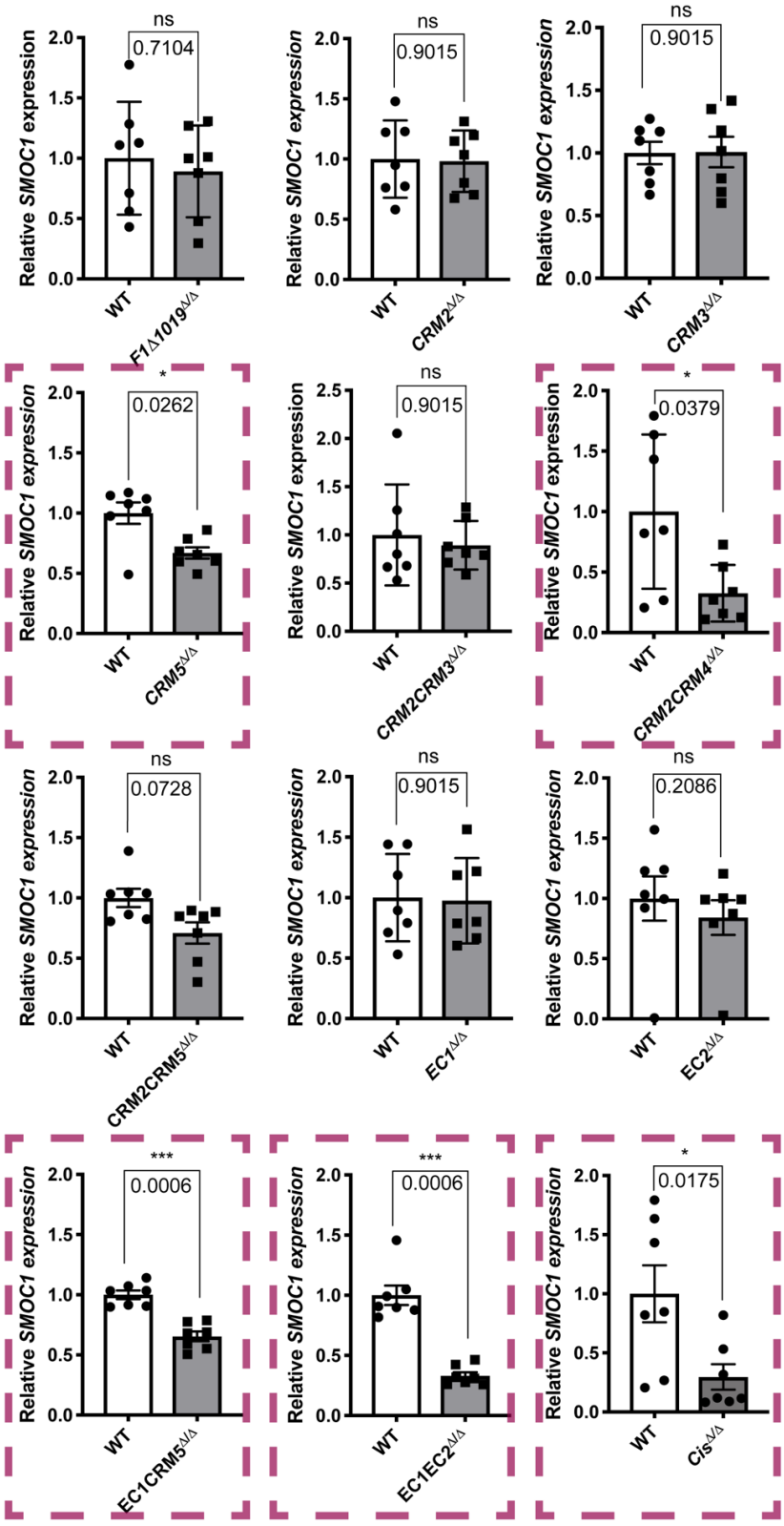
Appendixes Figure 8. *Grem1* WISH at E10.5 in WT and SMAD4 depleted forelimb (Gamart et al., manuscript in preparation). Comparison of *Grem1* expression domain in WT versus SMAD4 depleted forelimb bud at E10.5 (n=3). *Scale bar: 250µm*



Appendixes Figure 9. *CRM5* at E10.5 did not bind GLI proteins. Alignment of GLI3 ChIP-seq (in collaboration with Dr. Kevin Peterson) with the *CRM5* locus in wild-type forelimb buds at E10.5.



Appendixes Figure 10. *Tbx2* WISH at E11 in WT, *CRM2*, *EC1* and *Cis* homozygous forelimb buds. From left to right: *Tbx2* territory is broad in the anterior and posterior proximal wild-type limb bud at E11.0, in comparison, its territory is reduced in *CRM2*$^{\Delta/\Delta}$ and *EC1*$^{\Delta/\Delta}$ *Scale bar: 250µm*



Appendixes Figure 11. RT-qPCR analyses of the BMP pathway antagonist *Smoc1* transcription in our homozygous mutant's forelimb buds at E11.0. WT and homozygous mutant handpates at E11.0 were dissected from littermates embryos. Red dotted boxes highlight mutants with significant changes (n=7; mean ± SEM, two-tailed Mann-Whitney test), p-values are indicated on top.

Genotype	Sample size	Forelimb phenotype
Cis$^{\Delta/\Delta}$	9	100% oligosyndactyly (3 digits)
CRM2CRM3$^{\Delta/\Delta}$	4	100% soft-tissue syndactyly
CRM2CRM4$^{\Delta/\Delta}$	3	100% soft-tissue syndactyly
CRM2CRM5$^{\Delta/\Delta}$	3	no phenotype
CRM5$^{\Delta/\Delta}$	10	no phenotype
EC1$^{\Delta/\Delta}$	12	20% C-L type syndactyly (4 digits)- 80% soft-tissue syndactyly (digits 2-3)
EC1CRM5$^{\Delta/\Delta}$	10	100% (4 digits)
EC1EC2$^{\Delta/\Delta}$	7	100% oligosyndactyly (3 digits)
EC2$^{\Delta/\Delta}$	9	no phenotype
Grem1$^{\Delta/\Delta}$	2	100% oligosyndactyly (3 digits)

Appendixes Table 1. Skeletal phenotype scoring. List of each homozygous mutants generated in this study and their associated limb phenotype