Sp8 and *Sp6* regulatory functions in the limb bud ectoderm

Alejandro Castilla-Ibeas[1*#], Sandra Moreu[1*], Rocío Pérez-Gómez[1*], Laura Galán[1], Víctor Campa[1], Juan F. López-Giménez[2], Alvaro Rada-Iglesias[1] and Maria A. Ros[1#]

1) Instituto de Biomedicina y Biotecnología de Cantabria, IBBTEC (CSIC-UC-SODERCAN). Albert Einstein 22. 39011, Santander, Spain.
2) Instituto de Parasitología y Biomedicina "López-Neyra", IPBLN, CSIC. Av. del Conocimiento 17, 18016 Granada, Spain.

(*) equall contribution
(#) Corresponding authors: rosm@unican.es (M.A.R.) and alejandro.castilla @unican.es (A.C-I.)

ABSTRACT

Sp8 and *Sp6* are two closely related *Sp* genes expressed in the limb ectoderm, where they regulate proximo-distal and dorso-ventral patterning. Genetic studies in mice have shown that they act in a dose-dependent manner, with *Sp8* exerting a substantially stronger effect than *Sp6*. Here, we integrate ChIPmentation-seq and RNA-seq analyses to elucidate their genome-wide regulatory networks and mechanisms of action. Our results show that Sp8 has a dual mode of action either directly binding GC-rich motifs or indirectly binding AT-rich motifs through Dlx. In contrast, Sp6 primarily acts through AT-rich motifs, via interaction with cofactors such as Dlx5. Both factors regulate key components of the Wnt, FGF, and BMP signaling pathways, central to limb patterning, and display both cooperative and independent roles. We further show that Sp8 and Sp6 can form homo- and heterodimers, as well as interact with Dlx5, revealing a complex regulatory network. Our findings provide molecular insight into limb morphogenesis and have potential implications for understanding congenital limb malformations.

30 **INTRODUCTION**

31 The developing vertebrate limb has long been an excellent system for studying the
32 mechanisms involved in pattern formation and morphogenesis, and more recently in
33 transcriptional regulation. Limb development begins with the emergence of the limb
34 progenitor cells, which originate from the somatopleura of the lateral plate mesoderm
35 through an epithelial-to-mesenchymal transition. Subsequent limb outgrowth and
36 patterning rely on the interactions between these progenitor cells and the overlying
37 ectoderm. One of the earliest interactions induces *Fgf8* expression in the ectoderm,
38 leading to the formation of the apical ectodermal ridge (AER), a crucial signaling center
39 that supports the survival and proliferation of the limb progenitor cells (Fernandez-Teran
40 & Ros, 2008; Tickle, 2015). Simultaneously, *En1* is activated in the ventral ectoderm,
41 restricting *Wnt7a* expression to the dorsal ectoderm and thereby establishing distinct
42 dorsal-ventral (DV) gene expression domains essential for proper DV limb patterning. In
43 the dorsal ectoderm, *Wnt7a* promotes limb bud dorsalization by inducing the expression
44 of the homeobox gene *Lmx1b* in the underlying mesoderm (Fernandez-Teran & Ros,
45 2008; Tickle, 2015).

46 Elegant studies demonstrated that the activation of *Fgf8* and *En1* in the limb
47 ectoderm depends on active Wnt/βcatenin and Bmp signaling (Barrow et al., 2003; Kyung
48 Ahn et al., 2001; Soshnikova et al., 2003). The transcription factors (TF) Sp6 and Sp8 act
49 downstream of these pathways as critical mediators of *Fgf8* and *En1* induction (Haro et
50 al., 2014) but the underlying molecular mechanisms remain incompletely understood.
51 Understanding the mechanisms that lead to the induction and maintenance of the AER as
52 well as DV polarity is of maximum interest not only for advancing our understanding of
53 limb development but also for informing strategies in tissue regeneration and repair.

54 Sp6 and Sp8 belong to the Specificity protein/Krüppel-like (Sp/Klf) family found
55 across almost all metazoan species (Presnell et al., 2015; Schaeper et al., 2010; Suske,
56 1999). The Sp family is characterized by a highly conserved carboxyterminal DNA
57 binding domain consisting of three consecutive C2H2-type zinc finger (ZF) motifs and a
58 more variable N-terminal region (Suske et al., 2005). The buttonhead box and the Sp box
59 are also conserved structural domains characteristic of these proteins.

60 During limb development, Sp6 and Sp8 are exclusively expressed in the ectoderm
61 where their cooperative function is necessary for proximo-distal (PD) and DV patterning
62 (Bell et al., 2003; Haro et al., 2014; Talamillo et al., 2010; Treichel et al., 2003). The
63 analysis of *Sp8;Sp6* double mutants revealed that they act in a dose-dependent manner,

with their combined absence leading to tetra-amelia (total loss of limb formation) due to a failure in *Fgf8* activation. Partial loss of *Sp6* and *Sp8*, such as in mutants with a single functional *Sp8* allele (*Sp6*$^{-/-}$;*Sp8*$^{+/-}$), leads to split-hand/foot malformation (SHFM), characterized by irregular and immature AER formation and double dorsal digit tips (Haro et al., 2014). A reasonable conclusion of these studies, supported by some *in vitro* assays (Sahara et al., 2007), was that Sp8 and Sp6 were transcriptional activators of *Fgf8* and *En1* linking PD and DV patterning. However, the full repertoire of Sp6 and Sp8 target genes, whether shared or distinct, as well as the molecular mechanisms and genomic context orchestrating their regulation are not completely known.

Here, we combined genome-wide chromatin occupancy (ChIPmentation-seq) with transcriptomic profiling (RNA-seq) to identify shared and specific Sp8 and Sp6 direct targets in the limb ectoderm in vivo. We found that Sp8 and Sp6 act mainly as transcriptional activators via distal enhancers with many co-occupied regulatory regions, supporting a cooperative function. However, within these shared regions, Sp8 appears to be functionally more dominant. In addition, Sp8 and Sp6 also have distinct individual functions; for example, Sp8 uniquely regulates *Fgf8*, while Sp6 plays a specific role in epithelial organization.

An additional key finding of this study is the mechanistic divergence between Sp8 and Sp6 modes of action. While Sp8 directly binds the canonical GC-box motif, a hallmark of Sp transcription factors, Sp6 shows little evidence of direct DNA binding and instead primarily targets AT-rich motifs via cofactors such as Dlx5. Sp8 can also bind AT-rich motifs indirectly, showing a dual DNA-binding mechanism. The co-expression of several *Dlx* genes with *Sp8* and *Sp6* in the developing limb ectoderm, along with evidence from co-IP and EMSA assays, supports a potential functional interaction between Sp and Dlx factors. We propose that the Sp/Dlx interaction should be considered when interpreting the phenotypes of Sp and Dlx mutants and their association with congenital malformations.

RESULTS

Generation of a *Sp8*-tagged Knock-in Mouse

The genome-wide binding of Sp8 has not been previously mapped in the limb ectoderm or in any other cellular context. This is probably due to the lack of commercially available ChIP-grade antibodies for Sp8. To overcome this limitation and, thus, globally identify Sp8 binding sites in the limb ectoderm, we generated a knock-in (KI) mouse in which the

98 endogenous *Sp8* gene was tagged with three copies of the FLAG (FL) epitope using
99 homologous recombination (*Sp8FL*; Fig. 1A). This epitope was incorporated in frame at
100 the C terminus of Sp8 protein, a strategy that has been widely used in ChIP-seq studies
101 (Osterwalder et al., 2014; Vokes et al., 2008). Mice homozygous for the *Sp8FL* allele
102 were viable and fertile and displayed no obvious phenotype indicating that the tagged
103 protein was fully functional (Fig. 1B). Accordingly, the detection of the Sp8FL protein
104 in immunofluorescence assays using the αFLAG M2 antibody (Ref F1804, Sigma) was
105 similar to that of the endogenous protein detected with an αSp8 antibody (Ref 104661,
106 Santa Cruz) (Fig. S1A). The αFLAG antibody also readily detected the Sp8FL protein in
107 Western blots (Fig. S1B). Thus, the *Sp8FL* KI model provides a useful resource for future
108 studies requiring detection of Sp8.
109
110 **Global Mapping of Sp8 and Sp6 Binding Sites in the Limb Ectoderm**
111 To investigate the genome-wide occupancy of Sp8 and Sp6 in limb ectodermal cells, we
112 used ChIPmentation-seq. This method, which combines chromatin immunoprecipitation
113 (ChIP) with tagmentation, has proven highly efficient with low numbers of starting cells
114 (Schmidl et al., 2015), enabling us to reduce the initial sample size to 50 ectodermal hulls
115 of E10.5 forelimb buds (approximately $7.5x10^5$ cells). The ectodermal sheets were
116 isolated using mild trypsin digestion (see M&M). Limb buds from embryonic day (E)
117 10.5 were selected, as this stage corresponds to a fully matured AER and coincides with
118 the onset of both *Sp8* and *Sp6* null phenotypes (Bell et al., 2003; Fernandez-Teran & Ros,
119 2008; Talamillo et al., 2010; Treichel et al., 2003). The genome-wide binding of Sp8 was
120 mapped using homozygous *Sp8FL* embryos and the αFLAG. For Sp6, wild-type (WT)
121 embryos and the αSp6 antibody (Ref 21234-1-AP, Proteintech) were used. We previously
122 validated the αSp6 antibody through immunofluorescence (Fig. S1C) and confirmed its
123 specificity in Western blots by showing the absence of a band in *Sp6*-null limb buds and
124 no off-target bands from other Sp proteins (Fig. S1D).
125 Following sequencing, analysis of the ChIPmentation datasets resulted in the
126 identification of 1,263 high-confidence Sp8 binding sites (Source Data 1) and 905 high-
127 confidence Sp6 binding sites (Source Data 1). The quality and specificity of the identified
128 Sp8 and Sp6 peaks was supported by (i) their high evolutionary conservation (Fig. S1E),
129 (ii) reproducibility in an independent biological ChIPmentation replicate (Fig. S1F), (iii)
130 high correlation between the two replicates (Fig. S1G), and (iv) overlap with highly

accessible regions in the limb bud ectoderm identified by single cell ATAC-seq (Desanlis et al., 2020) (Fig. S1H).

The *de novo in-silico* analysis of sequence motifs in the Sp8 peaks using the MEME Suite (Bailey et al., 2015) revealed significant enrichment of a CG-rich/Sp motif (E-value: 3.3e-161) and an AT-rich/Sp7 motif (E-value: 1.2e-154) as the most over-represented motifs. Another C-rich motif associated with ZN transcription factors was also modestly enriched (Fig. S1I). The presence of the Sp consensus binding sequence, the GC box: 5'-CCGCCC-3', was expected as it represents the conserved DNA recognition motif shared by Sp family members (Suske, 1999). The AT-rich/SP7 motif has previously been shown to enable Sp7 to bind DNA indirectly via interaction with Dlx5 (Hojo et al., 2016), suggesting that Sp8 may also interact with Dlx. Interestingly, Sp9, another Sp family member, has also been shown to exhibit enrichment of both AT-rich motifs (Xu et al., 2018) and GC boxes (Catta-Preta et al., 2025).

In contrast, the *de novo in-silico* analysis of Sp6 peak sequences did not identify the canonical GC-box Sp family binding motif. Instead, there was significant enrichment of the AT-rich/Sp7 motif (E-value 3.9e-066), along with other motifs associated -with zinc finger transcription factors (Fig. S1I). The absence of the Sp consensus motif is consistent with findings from a previous ChIP-seq study in developing teeth (Rhodes et al., 2021), which identified a core motif (CTg/aTAATTA) closely resembling the Sp7 binding motif in bone (Hojo et al., 2016) and the motif found in the present study. These results suggest that Sp6 primarily regulates gene expression through indirect mechanisms, involving Dlx and other co-factors, at least in its major sites of expression, such as the tooth and limb.

Distinct Functional and Genomic Features of Proximal and Distal Sp6 and Sp8 Regulatory Elements

Being confident about the quality of our Sp8 and Sp6 binding profiles, we then mapped the locations of the peaks with respect to the nearest transcription start site (TSS). While both transcription factors bind proximal (<5kb from TSS) and distal (>5kb from TSS) cis-regulatory element contexts, there is a moderate preference for distal regions, 66% for Sp8 (Fig. 1C) and 73% for Sp6 (Fig. 1D) and, therefore, long-range regulatory interactions.

To explore potential functional differences based on whether binding occurred in proximal or distal regulatory regions, we performed separate *in silico* analyses of

proximal versus distal Sp8 and Sp6 peaks. Functional annotation of Sp8 peaks using GREAT (McLean et al., 2010) revealed that proximal binding was primarily associated with general biologic processes, such as “transcription regulatory region DNA binding” and “transcription factor activity” (Fig. 1E, Source Data 2). Conversely, distal binding predominantly associated with processes related to “embryonic morphogenesis”, “limb development” and “Wnt signaling pathway”. Moreover, we also examined whether Sp8 recognized different DNA sequences when binding proximal or distal regulatory sites. *de novo* motif analysis revealed that both proximal (E-value 4.3e-200) and distal (E-value 1.1e-107) Sp8-bound elements were enriched in the Sp family consensus GC-rich motif (Fig. 1E). In addition, the AT-rich/Sp7 motif was also enriched in the distal set of peaks where it was the most prominent (E-value 6.6e-204). Thus, Sp8 may use distinct regulatory mechanisms when operating at either promoters or enhancers.

A similar analysis of Sp6 peaks revealed enrichment in terms related to embryonic and epithelial morphogenesis in both proximal and distal elements (Fig. 1F). Consistent with the global peak analysis, the consensus Sp motif was absent from both proximal and distal peaks, whereas the Dlx core motif was the most enriched in distal peaks (E-value 4.5e-065). The smaller number of peaks in the proximal set limited the robustness of the analysis for this group (Fig. 1F, Source Data 1).

Together, our results suggest that Sp8 operates through two distinct functional modes. First, directly binding DNA at Sp consensus sequences, predominantly located at proximal promoters, to regulate general biological processes. Second, indirectly interacting with DNA via binding partners such as Dlx5, and potentially other Dlx family members, at distal putative enhancers, thereby controlling genes involved in specific limb developmental functions. In contrast, Sp6 appears to function primarily through indirect mechanisms involving Dlx and other cofactors, independent of genomic location.

Shared and Distinct Binding Dynamics of Sp8 and Sp6 During Limb Development

Having mapped the individual binding profiles of Sp8 and Sp6 across the genome and considering their dose-dependent cooperation during limb development (Haro et al., 2014), we analyzed the extent of overlap in their binding sites. Surprisingly, we found minimal direct overlap among the Sp8 and Sp6 peaks identified by MACS2, with only 185 peaks being shared between the two factors. This limited overlap could be attributed to differences in ChIP-seq sensitivity, potentially obscuring the detection of combinatorial binding events (Catta-Preta et al., 2025). Indeed, ChIPmentation tracks

199 revealed substantial signal overlap between Sp8 and Sp6, suggesting a more extensive
200 co-occupancy. Thus, to further characterize Sp8/Sp6 cooperative binding, we performed
201 unbiased k-means clustering of ChIPmentation signals with the merged set of peaks
202 which identified three distinct peak categories (Fig. 2A; Source Data 5): (i) peaks with
203 high signal overlap between Sp8 and Sp6 (607 peaks, 31.58%), (ii) peaks predominantly
204 bound by Sp8 (803 peaks, 41.77%), and (iii) peaks predominantly bound by Sp6 (512
205 peaks, 26.33%).

206 MEME-ChIP analysis (Fig. 2B) of regions co-occupied by Sp8 and Sp6 revealed
207 strong enrichment for the AT-rich motif (E-value 1.8e-261), with a much less significant
208 enrichment for the GC-box motif (E-value 9.1e-051). In contrast, regions bound
209 exclusively by Sp8 showed a strong preference for the GC-box motif (E-value: 6.4e-399),
210 while regions bound exclusively by Sp6 were enriched for ZN motifs, albeit with lower
211 statistical significance. These results suggest that Sp8 and Sp6 employ distinct binding
212 strategies depending on whether they act together or independently. Specifically, their
213 co-occupancy correlates with enrichment in AT-rich motifs, while Sp8 binding alone
214 correlates with GC-rich motifs, and Sp6 binding alone appears to be linked to partnerships
215 with ZN transcription factors. The presence of GC-box and AT-rich motifs was also
216 analyzed individually for every peak (Fig S2A) with FIMO algorithm (Grant et al., 2011).
217 Consistent with previous results, a significant proportion of common peaks (39%)
218 contained only AT-rich motifs, suggesting that there are genes that are regulated by Sp8/6
219 solely through AT-rich motifs.

220 Finally, we analyzed the genomic distribution of the three peak categories
221 (Fig. 2C). Co-occupied peaks and those exclusive to Sp6 were strongly biased toward
222 distal elements (79% and 70%, respectively). By contrast, Sp8-specific peaks exhibited a
223 more even distribution, with 58% in distal regions, suggesting that, in the absence of Sp6,
224 Sp8 has relatively greater engagement at promoter-proximal sites.

225

226 **Impact of Sp8 and Sp6 Loss on the Transcriptome of Limb Bud Ectodermal Cells**

227 To gain further insights into the developmental program regulated by Sp8 and Sp6 in the
228 limb ectoderm, we compared the expression profiles of the E10.5 WT and *Sp8* or *Sp6*
229 deficient limb bud ectoderms. For this, we used a robust RNA-seq protocol for low-
230 abundance RNA (see M&M for details).

231 Differential gene expression analysis identified 885 differentially expressed genes
232 (DEGs) in *Sp8*-null embryos (Fig. 2D; Source Data 3), with 511 downregulated and 374

233 upregulated, suggesting that Sp8 acts as both an activator and a repressor, with a slightly
234 stronger role as an activator. Notably, all genes previously reported to be affected by Sp8
235 loss, including *Fgf8, En1, Msx2, Rspo2,* and *Sp6* (Bell et al., 2003; Haro et al., 2014;
236 Treichel et al., 2003), were among the DEGs, validating the quality and accuracy of our
237 RNA-seq data.

238 A similar analysis in Sp6-deficient limb ectoderms identified 200 DEGs, with 114
239 downregulated and 86 upregulated (Fig. 2D; Source Data 3). This suggests Sp6 acts as
240 both an activator and a repressor, with a slightly stronger role as an activator. Consistent
241 with the limb *Sp6*-null phenotype (Talamillo et al., 2010), genes altered in the absence of
242 *Sp6* included key modulators of the Wnt/β-catenin signaling pathway, such as *Sost*, as
243 well as extracellular matrix components such as *Col4a2*, a major constituent of the
244 basement membrane.

245 Gene Ontology (GO) enrichment analysis of the DEGs using DAVID revealed
246 distinct biological processes influenced by these transcription factors. Sp8 DEGs are
247 enriched in "multicellular organism development," "embryonic limb morphogenesis,"
248 "BMP signaling pathway," and "fibroblast growth factor receptor signaling pathway",
249 indicating an important function during limb morphogenesis. Sp6 DEGs are linked to
250 "extracellular matrix organization," "canonical Wnt signaling pathway," and "cell
251 adhesion," (Fig. S2B, Source data 3). Thus, while both Sp8 and Sp6 influence key
252 biological processes in the limb ectoderm, the loss of Sp8 has a greater transcriptomic
253 impact than the loss of Sp6, consistent with their mutant phenotypes.

254

255 **Identification of Direct Sp8 and Sp6 Targets in the Limb Bud Ectoderm**

256 To distinguish direct from indirect Sp8 and Sp6 target genes and uncover the
257 transcriptional regulatory networks they govern, we intersected the list of genes assigned
258 to each binding region category (co-occupied, primarily bound by Sp8 and primarily
259 bound by Sp6) with the list of DEGs identified in *Sp8* and *Sp6* null mutants (Fig. 2E;
260 Source Data 5).

261 Among the genomic regions co-occupied by Sp8 and Sp6, we identified 24 peaks
262 linked to 12 DEGs in both *Sp8KO* and *Sp6KO* mutants (Fig. 2E and Source Data 5). This
263 set included *Cyp26a1* and *Crabp1*, regulators of retinoic acid metabolism, which showed
264 complex regulation: *Cyp26a1* was downregulated in *Sp8KO* but upregulated in *Sp6KO*,
265 while *Crabp1* was downregulated in both mutants, suggesting a complex modulation of
266 RA signaling in the limb ectoderm (Niederreither & Dollé, 2008). It also included *Sost*, a

negative modulator of Wnt signaling (Collette et al., 2010; Semënov et al., 2005) expressed in the limb bud ectoderm but not the AER (Fernandez-Guerrero et al., 2021) also implicated in BMP regulation (Winkler et al., 2003). GO analysis using DAVID confirmed enrichment for terms such as "retinoic acid binding" and "negative regulation of BMP signaling pathway" (Fig. 2E).

A subset of co-occupied regions was associated with genes differentially expressed in only one mutant, despite potential co-regulation, indicating differential dependence on Sp8 or Sp6 (Fig. 2E). Among genes co-regulated but specifically dependent on Sp8, we identified 119 genes (linked to 140 peaks), including key limb development genes (*Bambi*, *En1*, *Wnt7a*, *Hoxc9*, *Rspo2*, *Fzd1*, *Msx1, Msx2* and *Sp6*), all positively regulated by Sp8. GO analysis highlighted terms related to "embryonic limb morphogenesis" and "Wnt signaling pathway". Among genes specifically dependent on Sp6, we found 20 genes (22 peaks), including typical ectodermal genes (*Col4a2* and *Gjb5*). Corresponding GO terms included "cell-cell signaling" and "extracellular region". A substantial number of putative shared targets showed no expression changes in single mutants, suggesting mutual compensation or nonessential roles. Transcriptomic analysis of *Sp8/Sp6* double mutant limb ectoderms could clarify between these scenarios; however, their severe, early-onset limb defects limited the collection of ectoderm suitable for high-quality profiling.

Sp8 specific peaks were linked to 92 DEGs in *Sp8*-null limb buds representing regions associated with Sp8 targets not regulated by Sp6. These included key genes involved in limb development, most notably *Fgf8*, as well as *Kremen1* and *Prickle1*, all of which positively regulated by Sp8. GO analysis revealed enrichment for "fibroblast growth factor receptor signaling pathway," "positive regulation of cell division," and "positive regulation of ERK1 and ERK2 cascade" consistent with Sp8's role in regulating AER function (Fig. 2E). Conversely, Sp6 specific peaks were associated with 11 DEGs in *Sp6*-null mutants, none of which with known, relevant functions in limb development. The associated GO terms were less specific, including "negative regulation of transcription" underscoring Sp6's more limited regulatory specificity when acting independently of Sp8.

Finally, we note that some targets are associated with multiple regulatory elements across categories (Fig. S2D). 25 genes are regulated by shared and by Sp8-specific regions. For example, *Wnt5a* is linked to two distal regulatory elements co-bound by Sp8

300 and Sp6 and a third distal region bound exclusively by Sp8, suggesting complex
301 regulation (Fig S2E).
302
303 **Validation of Sp8 and Sp6 Target Genes**
304 To confirm direct Sp8 and Sp6 regulation and assess spatial expression patterns, we
305 performed in situ hybridization on direct targets selected for their relevance to limb
306 development. In addition, transgenic reporter assays were performed in selected cases to
307 assess enhancer activity.
308 *Sost* was selected from the set of Sp8 and Sp6 common target genes that require
309 the input of both TF for normal expression (Fig 3A). Sp8 and Sp6 bind to an ATAC-seq
310 accessible region within the first intron of *Sost*. In situ hybridization on tissue sections
311 revealed that *Sost* was expressed in the WT non-AER ectoderm and markedly
312 downregulated in *Sp8* and *Sp6* mutants, more severely in *Sp8*-null mutants, and
313 completely lost in double mutants. These results confirm *Sost* as bona fide direct target
314 under cooperative regulation by Sp8 and Sp6.
315 Among the extensive list of targets co-regulated by Sp8 and Sp6 but dependent
316 solely on Sp8 for normal expression, we selected *Rspo2*, *Fzd1*, *Bmp4*, *Msx1*, *En1*, *Wnt7a*,
317 and *Bhlha9*, all critical for limb development, for detailed analysis. ISH confirmed
318 specific loss of *Rspo2* in the AER and *Fzd1* expression in the ectoderm of *Sp8*-null mutant
319 with no obvious changes in *Sp6* mutants (Fig 3B-C, orange arrowheads). However, the
320 putative *Rspo2* enhancer located 217 kb upstream of the TSS, displayed only weak
321 activity in the limb ectoderm (6/15 embryos) in transgenic reporter assays (Fig 3B, green
322 arrowhead). The expression of *Bmp4* and its bona fide targets *Msx1* and *Msx2* were also
323 specifically downregulated in the ectoderm of *Sp8* mutant limb buds but not in *Sp6*
324 mutants (Fig 3D-E and S3B). The *Msx1* assigned peak has already been tested in the
325 VISTA enhancer dataset (Visel et al., 2007) and shows strong activity in the AER. Thus,
326 our results suggest that Sp8 and Sp6 modulate Wnt and Bmp signaling pathways by
327 enhancing Wnt signal reception and Bmp ligands and targets.
328 Our results also identify *En1* and *Wnt7a*, key regulators of DV limb patterning, as
329 direct targets of Sp8 and Sp6 (Fig 4A-B). *En1* expression, previously shown to be
330 downregulated in *Sp8* and double *Sp8;Sp6* mutants, is controlled by a distal downstream
331 enhancer (Fig. 4A). Transgenic reporter assays confirmed this enhancer's activity in the
332 AER with high specificity and reproducibility (3 out of 4; Fig. 4A). In *Sp8* mutants and
333 *Sp8;Sp6* double mutants, *Wnt7a* expression has been reported to expand into the ventral

334 ectoderm as a consequence of *En1* loss (Haro et al., 2014). Here we show that *Wnt7a* is
335 also directly and positively regulated by Sp8/Sp6, potentially through a distal downstream
336 enhancer. Accordingly, in addition to the expansion of its expression domain, the overall
337 expression level of *Wnt7a* is reduced in the ectoderm of *Sp8* mutants (Fig. 4B).
338 *Bhlha9* (*Fingerin*), is a basic helix-loop-helix transcription factor involved in limb
339 development that in the mouse has a phenotype similar to that of Sp6. In humans,
340 *BHLHA9* is linked to split-hand/foot malformation with long bone deficiency (SHFLD3)
341 a phenotype resembling *Sp6;Sp8* compound mutants retaining a functional copy of *Sp8*.
342 The ISH analysis (Fig 4C) confirmed a marked downregulation in Sp8 mutants.
343 Among the common targets dependent solely on Sp6 for normal expression, we
344 selected *Col4a2*. Immunofluorescence analysis for Collagen type IV confirmed a marked
345 reduction in the basement membrane in the absence of *Sp6* (Fig. 4D).
346 Among the shared target genes that do not exhibit changes in any of the mutants,
347 we examined *Pdgfa* (Fernandez-Guerrero et al., 2021). No alteration in expression was
348 observed (Fig 4E), suggesting that Sp8 and Sp6 binding may not play a regulatory role in
349 this context.
350 Among the genes exclusively regulated by Sp8 we selected *Fgf8* because of its
351 importance in limb development. The Sp8-peak assigned to *Fgf8* was in the gene
352 promoter (-326 bp) and contained several Sp1-like binding sites supporting direct
353 regulation by Sp8. Curiously, no Sp6 binding was detected in this region. In *Sp8* mutant
354 limb buds, *Fgf8* is initially activated but its expression decays concomitantly with the
355 onset of the phenotype indicating that this regulatory region is probably involved in
356 maintenance of expression (Fig. 4F) (Bell et al., 2003; Treichel et al., 2003). In contrast
357 *Fgf8* expression in *Sp6* mutants is not altered in its level of expression but in the domain
358 of expression consistent with the mild double ridge phenotype typically observed in these
359 mutants.
360 Other genes important for limb development, included *Wnt5a*, *Bambi*, *Dlx1, Dlx2*,
361 *Dlx4* and *Sp6* some of them also validated by ISH (Fig. S3B-D; Source Data 5).
362 Collectively, these findings demonstrate that Sp8 and Sp6 directly regulate essential
363 genes in the limb bud ectoderm, highlighting the cooperative as well as independent roles
364 of these transcription factors.
365

366 **Distinct DNA-Binding Modes of Sp6 and Sp8**

367 Our data show that Sp6 primarily binds to an AT-rich Dlx motif, rather than the canonical
368 GC-box. This AT-rich motif is also enriched in Sp8-bound regions, suggesting that both
369 Sp6 and Sp8 may operate through a mechanism similar to that used by Sp7 in osteoblasts.
370 To evaluate the binding affinities of Sp8 and Sp6 for AT-rich and GC-box motifs,
371 we performed an electrophoretic mobility shift assay (EMSA) using two constructs: one
372 containing the AT-rich/Dlx5 motif (Fig. 5A) (Hojo et al., 2016) and another containing
373 the GC-box motif (Fig. 5B) (Vaqué et al., 2005). Neither Sp8 nor Sp6 bound directly to
374 the AT-rich motif, suggesting that their association with this motif is likely indirect and
375 dependent on cofactors (Fig. 5A). In contrast, Sp8 (FLAG-tagged) exhibited strong and
376 specific binding to the GC-box, while Sp6 binding was barely detectable, similar to
377 previous findings for Sp7 (Fig. 5B). The Sp8-GC-box interaction was further validated
378 by the supershift in presence of the anti-FLAG antibody (Fig. 5B). These results confirm
379 the distinct modes of action for these transcription factors: whereas Sp8 functions through
380 both direct DNA binding via the GC-box and indirect interactions with other factors, Sp6
381 predominantly relies on cofactors. Protein expression levels of Sp6 and Sp8 were
382 quantified by Western blot analysis (Fig. S4A).
383

384 **Sp8 and Sp6 form Heterodimers without Affecting Sp8's Direct DNA Binding**

385 Our findings show that Sp8 and Sp6 co-occupy certain regulatory regions while binding
386 independently to others. Notably, regions bound by both Sp8 and Sp6 are highly enriched
387 in AT-rich/DLX5 motifs whereas regions exclusively bound by Sp8 are primarily linked
388 to GC-rich motifs. This suggests that Sp8 and Sp6 may have distinct DNA-binding
389 preferences, whether functioning independently or together. Considering that Sp
390 transcription factors are known to hetero dimerize, we asked whether Sp8 and Sp6 might
391 physically interact forming heterodimers that could potentially influence their DNA-
392 binding specificity and regulatory activity. To investigate this, we performed Bimolecular
393 Fluorescent Complementation (BiFC) assays to assess direct protein-protein interactions.
394 Sp8 and Sp6 were fused in-frame at their C-terminus to Yellow Fluorescent Protein (YFP)
395 or to the YFP N-terminal (YN; residues 1-172) or C-terminal (YC; residues 173-240)
396 moieties (see Materials & Methods). As expected, both Sp8 and Sp6 are localized to the
397 nucleus (Fig. S4B). BiFC assays revealed that both factors can form homo- and
398 heterodimers in the nucleus (Fig. 5C-E), with this interaction requiring the zinc finger
399 (ZF) domain (Fig. 5F). To further validate Sp8-Sp6 heterodimerization, we performed

400 co-immunoprecipitation (CoIP) assays, which confirmed their direct interaction (Fig.
401 5G).
402 Having demonstrated that Sp8 and Sp6 can heterodimerize, we next investigated
403 whether this interaction influences Sp8's ability to directly bind DNA. Using a luciferase
404 reporter assay with a construct containing four tandem SP1 sites (GC-box) (Vaqué et al.,
405 2005), we found that Sp8 alone significantly activated the luciferase reporter, whereas
406 Sp6 did not (Fig. 5H), consistent with the EMSA results. Co-expression of Sp6 with Sp8
407 did not alter Sp8's transcriptional activity, suggesting that (putative) Sp8/Sp6
408 heterodimers retain Sp8's ability to directly bind the GC-box, at least under these
409 experimental conditions (Fig. 5H). To confirm specificity, we tested a reporter with
410 mutated GC-box motifs (see M&M) and observed no transcriptional activation,
411 confirming that the effect is GC-box–dependent (Fig. S4C).
412
413 **Sp8 and Sp6 as Cofactors of the Dlx Family in Limb Ectoderm Development**
414 Despite the enrichment of the AT-rich/Dlx5 motif in Sp8- and Sp6-bound cis-
415 regulatory regions (Fig. 2C), EMSA assays demonstrated that neither factor directly binds
416 an AT-rich probe (Fig. 5A). This suggests that Sp8 and Sp6 may be recruited to some of
417 their target sites through interactions with Dlx transcription factors, a mechanism
418 previously described for Sp7 in bone development (Hojo et al., 2016). Notably, several
419 Dlx family members, including Dlx5 and Dlx6, share expression domains with Sp6 and
420 Sp8 in the limb ectoderm. Furthermore, the *Dlx5/Dlx6* double knockout embryos present
421 SHFM phenotype, resembling that of *Sp6;Sp8* compound mutants retaining a functional
422 *Sp8* allele (Haro et al., 2014; Robledo et al., 2002).
423 To investigate whether Sp8 and Sp6 physically interact with Dlx5, we performed
424 co-immunoprecipitation (CoIP) assays in HEK293 cells co-expressing epitope-tagged
425 (FLAG or Myc) of these factors. These assays confirmed that Dlx5 interacts with both
426 Sp8 and Sp6 and that this interaction requires the ZF domain of Sp proteins (Fig. 5I),
427 similar to what was previously reported for Sp7 (Hojo et al., 2016).
428 To corroborate the cooperative binding of Sp6 and Sp8 to AT-rich motifs via
429 interaction with Dlx5, we performed an additional EMSA using the probe containing the
430 AT-rich/Dlx5-binding motif. As expected, Dlx5 (Myc-tagged) alone formed a shifted
431 complex, which was supershifted upon addition of an anti-Myc antibody, confirming the
432 identity of the complex (Fig. 5J, lanes 2–3). Co-expression of Sp6 (FLAG-tagged) or Sp8
433 (FLAG-tagged) with Dlx5 resulted in a distinct shifted band, which was further

434 supershifted (in the case of Sp6) or no longer detected (in the case of Sp8) upon addition
435 of the anti-FLAG antibody. These changes indicate that both Sp6 and Sp8 are
436 incorporated into the Dlx5-DNA complex (Fig. 5J, lanes 4–7). These findings support a
437 model in which Sp6 and Sp8 do not directly bind AT-rich motifs but are recruited to these
438 regulatory elements through protein–protein interactions with Dlx cofactors. Protein
439 expression levels of Sp6, Sp8, and Dlx5 were quantified by Western blot analysis (Fig.
440 S4D).

441 The fact that Sp8 utilizes the same domain, the ZF, to recognize its cognate DNA
442 sequence and to mediate protein interactions highlights the complexity and context-
443 dependence of its regulatory function. Collectively, our results support a model in which
444 Sp8, Sp6 and Dlx5 collaboratively regulate target genes with the final functional outcome
445 depending on the relative abundance of each transcription factor.

446

447 **DISCUSSION**

448 Here, we uncover the complex regulatory networks by which Sp8 and Sp6, cooperatively
449 and independently, integrate Wnt/β-catenin, FGF, and BMP signaling, to fine-tune gene
450 expression during limb development.

451 To determine the molecular mechanisms underlying Sp8 and Sp6 transcriptional
452 activity in the limb ectoderm at a genome-wide scale, we used ChIPmentation and RNA-
453 seq. To distinguish between cooperative and specific regulatory roles, we overlapped the
454 chromatin binding profiles of Sp8 and Sp6 and classified the binding regions in three
455 categories: peaks co-occupied by both factors, and peaks specifically bound by either Sp8
456 or Sp6. To distinguish functional from non-functional binding events, we intersected
457 these peak categories with the sets of DEGs identified upon the individual deletion of Sp8
458 or Sp6. Genes associated with regulatory regions co-occupied by both transcription
459 factors and showing altered expression in either one or both single mutants were
460 considered as potential common targets. Among these, 12 genes required both factors for
461 proper expression, considered genuine common targets, while others depended solely on
462 Sp8 (119) or on Sp6 (20) for normal expression. We also identified direct target genes
463 specific to Sp8 and Sp6, each associated with regions bound exclusively by one of these
464 factors.

465 Our data also reflect the dominant transcriptional role of Sp8 compared to Sp6,
466 consistent with the more severe phenotype of *Sp8* mutants. One contributing factor to this
467 difference may be that *Sp6* is a Sp8 direct target. We previously showed that *Sp6*

468 expression is not maintained in *Sp8* mutants (Haro et al., 2014) and here we identify a
469 putative regulatory element within the *Sp6* gene through which Sp8 exerts this control.
470 As *Sp6* expression declines, *Sp8* mutants may effectively function as double mutants.
471 Additionally, Sp8's unique ability to bind GC-rich proximal regions, where Sp6 cannot
472 compensate, further explains its dominance.
473
474 **The Complexity of Sp8/Sp6 Regulatory Network in the Limb Bud Ectoderm**
475 Sp8 is a well-recognized mediator of the Wnt/βcatenin dependent induction of
476 *Fgf8* in the limb ectoderm and also in other systems such as the genital tubercle and the
477 embryonic telencephalon (Haro et al., 2014; Lin et al., 2013; Sahara et al., 2007).
478 Together with Sp5, Sp8 is also a mediator of Wnt/βcatenin dependent induction of
479 neuromesodermal progenitors (Kennedy et al., 2016). However, we did not detect
480 significant enrichment of Tcf/Lef motif in our ChIPmentation data, suggesting that in the
481 limb ectoderm Sp8/Sp6 do not operate by binding to βcatenin and rather mediate Wnt/β
482 catenin signaling by regulating the expression of key pathway components such as *Sost*,
483 *Rspo2* and *Fzd1*, which are regulated by common distal regulatory elements and
484 dependent on both (*Sost*) or solely on Sp8 (*Rspo2* and *Fzd1)*. Thus, by coordinating both
485 positive and negative components of the Wnt/β-catenin pathway Sp6/Sp8 ensure proper
486 signaling dynamics during limb development. Similarly, our ChIPmentation data showed
487 no enrichment of Smad motifs, suggesting Sp6/Sp8 modulate Bmp signaling by
488 regulating *Bmp4*, as well as its bona fide effectors *Msx1* and *Msx2*.
489 In addition to modulating key signaling pathways, Sp8 and Sp6 also directly
490 regulate crucial patterning genes in the limb ectoderm including *Fgf8* and the DV core
491 factors *En1* and *Wnt7a*. Here, we show that Sp8, but not Sp6, directly regulates *Fgf8*,
492 while both Sp8 and Sp6 contribute to the regulation of *En1* and *Wnt7a* with higher
493 contribution of Sp8.
494 *Fgf8* expression, the marker of the AER, is differentially affected in *Sp* mutants.
495 It is initiated but not maintained in *Sp8* mutants, remains unaffected in the double ridge
496 of *Sp6* mutants, and is never activated in *Sp8;Sp6* double mutants. Here, we show that
497 Sp8 directly regulates *Fgf8* by binding to GC-rich motifs within its promoter. In contrast,
498 we did not identify any *Sp6*-bound regulatory regions associated with *Fgf8*, suggesting
499 that the Sp6 input observed in double mutants is indirect, potentially mediated through its
500 role in modulating Wnt signaling in the ectoderm. In addition to promoter regulation,
501 *Fgf8* expression is governed by multiple enhancers located in a large regulatory region

502 centromeric to the gene (Hörnblad et al., 2021; Marinić et al., 2013). Duplications of this
503 genomic region have been linked to the SHFM3 syndrome (OMIM 246560), which is
504 thought to result from disrupted genomic architecture and altered AER expression (Cova
505 et al., 2023). However, our analysis found no evidence of Sp8 or Sp6 binding to this
506 complex regulatory domain.

507 We also show that *En1* and *Wnt7a* possess distal enhancer co-bound by Sp8 and
508 Sp6, but requiring only Sp8 for normal expression (Fig. 4A). The *En1* enhancer which
509 we show drives activity in the apical ectodermal ridge (AER) in mouse transgenic assays
510 and overlaps the first exon of the *En1*-associated long non-coding RNA (lncRNA) *Lmer*
511 (Ringel et al., 2024). However, deletion of the *Lmer* locus, including the Sp8/Sp6-bound
512 enhancer, did not affect *En1* expression in the limb ectoderm. Thus, despite its ability to
513 drive expression to the AER and ectoderm, this enhancer is dispensable for *En1* function,
514 likely due to redundancy with other elements such as Maenli. This suggests that the
515 downregulation of *En1* observed in *Sp8* and double mutants must depend on additional
516 factors, possibly Bmp4, which could also explain the delayed onset of this effect.

517 Our study also identifies a Sp8/Sp6 co-bound enhancer that positively regulates
518 *Wnt7a*. This may appear unexpected, as *Wnt7a* expression extends ventrally in both *Sp8*
519 and *Sp8;Sp6* mutants. However, this expansion is indirect, in parallel to the loss of *En1*.
520 When expression levels are considered, both ISH and RNA-seq reveal that *Wnt7a* is
521 reduced in *Sp8* and *Sp8;Sp6* mutants strongly suggesting that its expression is directly
522 regulated thought the Sp8/Sp6 bound enhancer. A major question that remains open is
523 how Sp8 and Sp6, initially expressed throughout the entire limb ectoderm, contributes to
524 the establishment of the defined DV expression domains.

525 These findings highlight Sp8 and Sp6's role in coordinating PD and DV patterning
526 and underscore the complexity of their cooperative and independent function.

527

528 **Sp8 Has a Dual Mode of Action**

529 Another key finding of our study is that Sp8 and Sp6 operate through distinct
530 mechanism. Motif analysis, EMSA and co-IP experiments, revealed that Sp8 uses a dual
531 strategy: it binds DNA directly at GC-rich motifs and indirectly at AT-rich regions, likely
532 via interactions with Dlx proteins. In contrast, Sp6 appears to rely primarily on indirect
533 DNA binding via co-factors, predominantly Dlx proteins. These differences challenge the
534 previously proposed functional redundancy between Sp8 and Sp6 suggested by earlier
535 genetic studies in mice (Haro et al., 2014). Furthermore, our results show that Sp8

536 preferentially binds GC-rich motifs when acting independently, whereas its cooperation
537 with Sp6 is associated with binding at AT-rich sites, a subject that merits further
538 investigation. Although we demonstrate that Sp8 and Sp6 can form heterodimers, this is
539 likely not the dominant in vivo scenario, given the limited overlap in their binding sites,
540 possibly due to interactions with distinct co-factors.

541 Another member of the Sp family, Sp7, is unable to bind the canonical Sp
542 consensus motif and instead engages DNA through interaction with Dlx5 (Hojo et al.,
543 2016). This distinct mode of action is attributed to three specific amino acid substitutions
544 in its ZF domain, which impair GC-box recognition while enhancing Dlx binding.
545 However, Sp8 and Sp6 do not share these Sp7-specific variants, indicating that these
546 residues are not required for Dlx5 interaction. Interestingly, Sp9, another Sp protein
547 expressed in the ectoderm, shows a similar binding pattern to that of Sp8, as published
548 motif enrichment analysis in ganglionic eminences indicate: AT-rich (Xu et al., 2018)
549 and GC-box (Catta-Preta et al., 2025). Our findings support the idea that interaction with
550 Dlx is a more broadly conserved feature within the Sp family, at least among members of
551 the Sp6–Sp9 clade (Fig. 6).

552

553 **Impact in Human Congenital Malformation**

554 Progressive disruption of the Sp8/Sp6 gene regulatory network in mice results in
555 a spectrum of limb malformations, including oligodactyly, ectrodactyly, and amelia,
556 mirroring human congenital disorders. In particular, the *Sp8+/-;Sp6-/-* mutant serves as a
557 robust model for human SHFM (Haro et al., 2014). Although no pathogenic variants in
558 human *SP8* or *SP6* have been identified in SHFM cases, their mechanistic interaction
559 with DLX5 suggest a potential implication in SHFM type 1, associated to *DLX5/6* and in
560 SHFM type 5, associated to *DLX1/2*. Furthermore, *Bhlha9*, a common transcriptional
561 target of Sp8 and Sp6, is linked to SHFM with long bone deficiency 3 (SHFLD3)
562 suggesting a potential role for Sp8/Sp6 in thie pathology.

563 The complexity of the Sp8/Sp6 network may rely on their ability to interact with
564 each other as well as with other TFs. Cofactor availability may influence their mechanistic
565 mode of action. In the limb ectoderm, Sp8, Sp6, and Dlx5—along with other members of
566 their respective families—likely form dynamic regulatory assemblies whose composition
567 and activity depend on context, affinity, and expression levels. Adding another layer of
568 complexity, Sp8 directly activates the transcription of several *Dlx* genes and *Sp6*,
569 establishing autoregulatory and feedforward circuits.

570 In conclusion, our study identifies Sp8 and Sp6 as key regulators of gene
571 expression in the limb ectoderm, functioning both independently and cooperatively.
572 Collectively, our results support a model in which Sp8, Sp6, and Dlx5 collaboratively
573 regulate target genes, with the final functional outcome depending on the relative
574 abundance of each transcription factor. These findings highlight the central role of Dlx–
575 Sp interactions in orchestrating limb morphogenesis and underscore the importance of
576 considering transcription factor networks, rather than individual factors, in the study of
577 congenital malformations.
578

579 **MATERIAL AND METHODS**

580 **Mouse Embryos**

581 Wild type C57BL/6 mice and the *Sp6*$^{-/-}$ (Nakamura & Ishikawa, 2008), *Sp8*^{FL} and
582 S*p8*^{CreERT2} (Treichel et al., 2003) mouse strains were used in this study. Genotyping was
583 performed using tail biopsies or embryonic membranes according to previously published
584 reports (Haro et al., 2014). Embryos of the desired embryonic day were obtained by
585 cesarean section. All animal procedures were conducted accordingly to the EU
586 regulations and 3R principles and reviewed and approved by the Bioethics Committee of
587 the University of Cantabria. All mice were maintained in a C57BL6 genetic background
588 and genotyped by PCR protocols from tail biopsies or yolk sacs samples following
589 standard protocols.
590

591 ***Sp8:3xFLAG* Knock-in Mouse Generation (*Sp8*^{FL})**

592 Three copies of the FLAG epitope (5'-
593 GACTACAAAGACCATGACGGTGATTATAAAGATCATGATATCGATTACAAG
594 GATGACGATGACAAG-3') were inserted in frame at the 3'-terminal end of the *Sp8*
595 gene with a stop codon (Fig. 1A). Mouse genomic fragments containing 5'
596 (chr12:118845063+118849869) and 3' (chr12:118849873+118852808) homology arms
597 were amplified from BAC clone by using high fidelity Taq polymerase and were
598 sequentially assembled into a targeting vector together with recombination sites and
599 selection markers (DTA and Neo cassette). Targeted ES clones were confirmed via
600 Southern Blotting and some of them were selected for blastocyst microinjection (Cyagen
601 Biosciences Inc., Santa Clara, California, US).
602

603 **Isolation of Limb Bud Ectodermal Hulls**

604 E10.5 forelimb buds were dissected in cold PBS and incubated in 0.25% trypsin
605 (SV30037.01, GE Healthcare, Logan, Utah) on ice for 20 min. After a quick rinse in 10%
606 FBS to inactivate trypsin, the limb buds were transferred to cold PBS where the ectoderm
607 was separated from the mesoderm with fine forceps.
608
609 **Transgenic Mice for LacZ activity**
610 The genomic region containing the regulatory elements of *En1* (mm10, chromosome
611 1:120,621,403-120,622,419) and *Rspo2* (mm10, chromosome 15:43,388,035-
612 43,389,518) were PCR-amplified using specific primers containing XhoI restriction
613 enzyme sites and cloned into a vector carrying the β-globin minimal promoter and the
614 LacZ coding sequence (pSK-LacZ; kindly provided by Denis Duboule lab). Transient
615 transgenic embryos were harvested at E10.5 and only those embryos containing the
616 reporter transgene were processed for detection of β-galactosidase activity according to
617 standard procedures.
618
619 **Skeletal Preparations, *in situ* Hybridization (ISH) in Paraffin Sections and**
620 **Immunofluorescence**
621 Alcian blue 8GX and Alizarin red skeletal staining was performed following standard
622 protocols, cleared by KOH treatment and stored in glycerol. *In situ* hybridization was
623 performed on paraffin sections with digoxigenin-labeled antisense riboprobes following
624 standard procedures. Immunostaining was performed in cryostat sections (14 µm) using
625 the mouse monoclonal anti FLAG M2 antibody (1:500; Ref F1804; Sigma-Aldrich, St
626 Louis, Missouri) and the anti Sp8 C18 antibody (1:500; Ref 104661, Santa Cruz
627 Biotechnology, Dallas, Texas) and then coupled with Alexa 488-conjugated anti-mouse
628 IgG secondary antibody (1:500, Invitrogen, Molecular Probes, Eugene, Oregon). Slides
629 were analyzed in a Leica Laser Scanning Confocal TCS-SP5 with a 63x 1.4 NA objective.
630
631 **Plasmids Constructions**
632 Sp8 and Sp6 tagged with FLAG or Myc epitopes at their 5' end were amplified by PCR
633 and cloned into pcDNA3 including the appropriate restriction enzyme site to be fused at
634 the 3'end to the full length YFP (1-240aa) or to its N-terminal (YN; 1-172aa) or C-
635 terminal (YC; 173-240aa) moieties. Deletion of the ZF domain of Sp8 (nt 353-486) and
636 of the ZF domain of Sp6 (nt 254-336) was constructed by PCR strategy. The following
637 clones were generated: FLAG-Sp8, FLAG-Sp6, HA-Sp6ΔZF, Sp8-YFP, Sp8-YN, Sp8-

638 YC, Myc-Sp6-YFP, Myc-Sp6-YN, Myc-Sp6-YC, Myc-Sp8ΔZF-YFP, Myc-Sp8ΔZF-
639 YN, Myc-Sp8ΔZF-YC. The Myc-Dlx5 and FLAG-Dlx5 clones were kindly provided by
640 Dr. Andrew P. McMahon (Keck School of Medicine of the University of Southern
641 California, Los Angeles, USA) (Hojo et al., 2016). The HA-Sp6ΔZF clone was
642 synthesized by NZYTech (Lisbon, Portugal).
643
644 **Bimolecular Fluorescent Complementation (BiFC)**
645 HEK293 cells were grown on microscope cover Glasses (Ø 18 mm) and transfected with
646 PEI and 1 µg total DNA. To test for homodimerization, 500ng of indicated expression
647 plasmids, each one with a complementary YFP moiety (e.g. Sp8-YN + Sp8-YC), were
648 transfected. Single transfection of the protein fused with the complete YFP or with one
649 of its moieties was used for positive and negative control. Transfected, cells were
650 incubated with Hoechst33342 and confocal images (512x512 pixels; 0.15mm pixel size)
651 were acquired sequentially on a SP5 laser-scan microscope (Leica) with a 63x 1.4 NA
652 objective. Cells were excited sequentially with 405nm 514nm laser lines and fluorescence
653 emission captured between 420-480nm (Hoechst) and 525-600nm (YFP). Images are
654 presented after digital adjustment of curve levels (gamma) to maximize signal with
655 ImageJ software. Fluorochromes and colors are as indicated in the figure legends. Images
656 are representative of at least three independent experiments.
657 HEK293 cells were grown on microscope cover Glasses (Ø 18 mm) and transfected with
658 PEI and 1 µg total DNA. To test for homodimerization, 500ng of indicated expression
659 plasmids, each one with a complementary YFP moiety (e.g. Sp8-YN + Sp8-YC), were
660 transfected. Single transfection of the protein fused with the complete YFP or with one
661 of its moieties was used for positive and negative control. Transfected, cells were
662 incubated with Hoechst33342 and confocal images (512x512 pixels; 0.15mm pixel size)
663 were acquired sequentially on a SP5 laser-scan microscope (Leica) with a 63x 1.4 NA
664 objective. Cells were excited sequentially with 405nm 514nm laser lines and fluorescence
665 emission captured between 420-480nm (Hoechst) and 525-600nm (YFP). Images are
666 presented after digital adjustment of curve levels (gamma) to maximize signal with
667 ImageJ software. Fluorochromes and colors are as indicated in the figure legends.
668
669 **Co-immunoprecipitation, SDS-PAGE, and Western Blotting**
670 HEK293 cells were co-transfected with the desired plasmids (1 µg total DNA) and lysed
671 (20 mM Hepes (pH 7.5), 10 mM EGTA pH 8, 40 mM Glycerol-phosphate, 1% NP-40,

672 2.5 mM MgCl2, 2 mM orthovanadate, 1 mM DTT, 1x PICS, 1 mM PMSF). CoIP was
673 performed with 0.5-1mg of protein, 10 µL of Dynabeads™ Protein G (Thermo Fisher
674 Scientific, 10004D) and 1µg of the desired antibody: FLAG-M2 (Sigma F1804) or MYC
675 (Cell Signaling #2272) followed by SDS-PAGE analysis of protein products. FLAG-M2
676 and MYC antibodies were used for immunoblotting and β-Actin (Santa Cruz #47778) or
677 H3 (Cell Signaling #96C10) was used as loading control when needed. Results were
678 confirmed in at least two independent experiments.
679 HEK293 cells were co-transfected with the desired plasmids (1 µg total DNA) and lysed
680 (20 mM Hepes (pH 7.5), 10 mM EGTA pH 8, 40 mM Glycerol-phosphate, 1% NP-40,
681 2.5 mM MgCl2, 2 mM orthovanadate, 1 mM DTT, 1x PICS, 1 mM PMSF). CoIP was
682 performed with 0.5-1mg of protein, 10 µL of Dynabeads™ Protein G (Thermo Fisher
683 Scientific, 10004D) and 1µg of the desired antibody: FLAG-M2 (Sigma F1804) or MYC
684 (Cell Signaling #2272) followed by SDS-PAGE analysis of protein products. FLAG-M2
685 and MYC antibodies were used for immunoblotting and β-Actin (Santa Cruz Ref 47778)
686 was used as loading control when needed. Results were confirmed in at least two
687 independent experiments.
688
689 **EMSA Assay**
690 EMSA was performed using the nuclear extracts from HEK-293 cells or HEK-293 cells
691 transfected with Sp8, Sp6 and Dlx5. DNA probes labeled with IRDye700 on 5´end were
692 synthesized by Integrated DNA Technologies, Inc. (IDT, Coralville, IA). The sequences
693 of probes are shown in (Hojo et al., 2016). Each reaction was performed by mixing 15
694 ug of the nuclear extract with the binding buffer (250 mM KCl, 100 mM HEPES-KOH
695 (pH 7.5), 0.875 mM EDTA, 25% glycerol, 1.5 µg of Poly d(I-C)). Mixtures were pre-
696 incubated at RT for 15 min. Later 0.25 mM of the Biotin-labeled probe was added, and
697 samples were incubated 15 min at RT. For supershift assay, 1 ug of anti-FLAG M2
698 antibody (Sigma-Aldrich, F1804) or anti-MYC (Cell Signaling #2272) was incubated at
699 RT for 15 min prior to binding with probes. For EMSA involving the Sp–Dlx complex
700 (Fig. 5K), all the incubations steps were extended to 30 minutes. Complexes were
701 analyzed by electrophoresis on a 6% polyacrylamide gel in 0.5x TBE buffer and
702 visualized using an Odyssey infrared imaging system (Li-Cor Bioscience). Results were
703 validated in at least two independent experiments.
704 EMSA was performed using the nuclear extracts from HEK-293 cells or HEK-293 cells
705 transfected with SP8 and SP6. DNA probes labeled with IRDye700 on 5´end were

706 synthesized by Integrated DNA Technologies, Inc. (IDT, Coralville, IA). The sequences
707 of probes are shown in (Hojo et al., 2016). The reaction was performed by mixing 15 ug
708 of the nuclear extract with the binding buffer (250 mM KCl, 100 mM HEPES-KOH (pH
709 7.5), 0.875 mM EDTA, 25% glycerol, 1.5 µg of Poly d(I-C)). Mixtures were pre-
710 incubated at RT for 15 min. Later 0.25 mM of the Biotin-labeled probe was added to each
711 condition and the resulting mixture was incubated 15 min at RT. For supershift assay, 1
712 ug of anti-FLAG M2 antibody (Sigma-Aldrich, F1804) was incubated at RT for 15 min
713 prior to binding with probes. Complexes were analyzed by electrophoresis on a 6%
714 polyacrylamide gel in Tris-Borate-EDTA buffer 0.5x. Geles were visualized using an
715 Odyssey infrared-imaging system (Li-Cor Bioscience).
716
717 **Luciferase Reporter Assay**
718 NIH 3T3 cells were plated at a density of 25,000 cells/cm2 (22Rv1) in 24-well plates and
719 the next day were transfected by Lipofectamine™ 3000 (L3000-008 Invitrogen by
720 Thermo Fisher Scientific Life Technologies, Carlsbad, California) with 50 ng/well of
721 firefly luciferase reporter SP1 sites (GC-box) (93p21-luc wt or 93p21-luc(mut 2+3+4)
722 (Vaqué et al., 2005), 15 ng of the pRL-TK Renilla luciferase control plasmid (Promega)
723 and 50ng of FLAG-Sp8 or FLAG-Sp6 expression plasmids. For dose-response
724 experiments, 12.5 ng and 25 ng of FLAG-Sp6 plasmid were also tested. After 48 hours,
725 luciferase activity was measured with the DualGlo Stop&Glo Luciferase Assay System (
726 Promega #E2920), following the manufacturer's instructions. Each experiment was
727 performed in duplicate and was repeated at least three times. Error bars represent standard
728 deviations calculated from the biological triplicates.
729 Expression profile of NIH 3T3 cells were obtained from BioGPS
730 (http://amp.pharm.mssm.edu/Harmonizome/gene_set/nih+3T3/BioGPS+Mouse+Cell+T
731 ype+and+Tissue+Gene+Expression+Profiles).
732
733 **ChIPmentation**
734 ChIPmentation protocol was based on (Schmidl et al., 2015). In order to reduce
735 background and unspecific binding, only the limb bud ectoderm was used. A total of 50
736 ectoderms from *Sp8*$^{FLAG/FLAG}$ (for Sp8) and WT (for Sp6) were used for each TF
737 ChIPmentation replicate (2 replicates). ChIPmentation sequencing data was processed by
738 a standard workflow using nf-core/chipseq/2.1.0 pipeline (Ewels et al., 2020). Because
739 of the higher signal to noise ratio, only the dataset of the first biological ChIPmentation

740 sample was used for further analysis. For proximal and distal classification, distance to
741 the nearest tss was obtained using bedtools closest function (Quinlan & Hall, 2010).
742 Pearson Correlation coefficients of the two SP8 ChIPmentation biological replicates were
743 performed with the bamCorrelate tool from deepTools (bins mode and a bin size of 10 kb
744 across the whole mouse genome) (Ramírez et al., 2014). For the generation of the
745 heatmaps, BAM files were normalized as RPGC (reads per genome coverage) and then
746 used to visualize scores associated with genomic regions using deepTools (Ramírez et
747 al., 2014). The bigwig of scATAC-seq ectoderm data was obtained from the GEO
748 accession number GSE: (Desanlis et al., 2020). Average vertebrate PhastCons score
749 profiles around the center of enhancer sequences were generated with the Conservation
750 Plot tool from the Cistrome Analysis pipeline (http://cistrome.dfci.harvard.edu/ap/root)
751 (Liu et al., 2011).
752 For the clustering analysis, Sp8 and Sp6 peaks were concatenated and merged
753 with bedtools merge function. Then, k-means ranked clustering of peaks based on Sp8
754 and Sp6 ChIPmentation signal was performed with seqMINER 1.3.4 (Ye et al., 2011),
755 targeting 12 clusters and with the seed 13,625,001. Clusters with similar profile were
756 merged to obtain 3 final clusters: Common, Sp8-, and Sp6-predominant.
757
758 **RNA-seq**
759 Ectoderms from E10.5 wild-type and homozygous *Sp8*-null (*Sp8*$^{Cre\text{-}ERT2}$ line) and Sp6-
760 null (Sp6KO line) embryos were individually collected. Experiments were performed in
761 triplicates and each replicate contained the 2 forelimb ectoderms of a single embryo. RNA
762 was extracted with RNeasy-Plus Micro Kit (Ref 74034 Qiagen GmbH, Hilden,
763 GERMANY). Pre-amplification using the Ovation RNASeq System V2 was performed.
764 Total RNA was used for first strand cDNA synthesis, using both poly(T) and random
765 primers, followed by second strand synthesis and isothermal strand-displacement
766 amplification. For library preparation, the Illumina Nextera XT DNA sample preparation
767 protocol was used, with 1 ng cDNA input. After validation (Agilent 2200 TapeStation)
768 and quantification (Invitrogen Qubit System) all six transcriptome libraries were pooled.
769 The pool was quantified using the Peqlab KAPA Library Quantification Kit and the
770 Applied Biosystems 7900HT Sequence Detection and pooled on one lane of an Illumina
771 HiSeq4000 sequencing instrument with a 2x75 bp paired-end read length.
772 To analyze the data, high-throughput next-generation sequencing analysis
773 pipeline (Wagle et al., 2015) was used. Basic read quality check was performed using

774 FastQC (Babraham Bioinformatics) and reads filtered and adaptor-trimmed with fastp
775 (Chen et al., 2018). Reads were mapped to the mouse reference assembly (GRCm38),
776 using HiSat2 (Kim et al., 2019) with the option --dta-cufflinks, which enabled direct
777 transcript assembly with StringTie (Pertea et al., 2015). Read count means, fold-change
778 (FC) and values were calculated with DEseq2 (Anders & Huber, 2010). Genes were
779 considered significant differentially expressed genes when p-adjusted value < 0.05 and
780 log2FC > 0.585 (FC > 1.5).
781 Sp6 expression was not properly calculated by the computational pipeline
782 described above due to the presence of an antisense transcript that largely overlaps with
783 Sp6 and that got assigned most of the RNA-seq reads (i.e. ENSMUSG00000087067).
784 Since according to the RNA-seq data ENSMUSG00000087067 was downregulated in
785 *Sp8KO* embryos and our RNA-seq data was not strand-specific, this indicates that Sp6 is
786 indeed downregulated in *Sp8KO* embryos. This was further confirmed by ISH and is also
787 supported by our previous findings (Haro et al., 2014). Therefore, the read values
788 measured for ENSMUSG00000087067 were assigned to *Sp6* instead.
789 Gene ontology was performed with the DAVID web-based tool
790 (davidbioinformatics.nih.gov, (Sherman et al., 2022)) with default parameters.
791
792 **Motif Analysis**
793 *De novo* motif discovery analyses of Sp8 peaks were performed using the online tool
794 MEME-ChIP (Version 5.0.1, http://meme-suite.org/tools/meme-chip) as described by
795 (Ma et al., 2014). Input sequences were centered within summit regions of recovered
796 intervals obtaining motifs that are between 6 and 20 bp wide with an E-value cut-off of
797 >0.5 for the discovery of enriched motifs.
798 The database used to identify known motifs was HOCOMOCO H12 CORE (human +
799 mouse) (Kulakovskiy et al., 2018). For the prediction of individual motif occurrences,
800 FIMO tool from the meme suite with default parameters. The AT-rich and GC-box motifs
801 obtained from MEME-ChIP with Sp8 peaks were used as input.
802
803 **Genomic Regions Enrichment of Annotations Tool (GREAT) analysis**
804 For calculating the distribution of all the Sp8 peaks with respect to the TSS of annotated
805 genes as well as for their *in silico* functional annotation analyses, we used the GREAT
806 tool (GREAT 3.0.0) (McLean et al., 2010) with the "Basal Plus Extension Association

Rule" settings. *Rspo2* was assigned to the peak chr15:43388486-43388897 given their proximity, and was subsequently validated with Lacz transgenic reporter analysis.

Acknowledgments

We are very grateful to Dr. Hojo (University of Tokyo) and Dr. McMahon (University of Southern California) for the p12xAT_luc, the p12xAT_luc and Dlx5 constructs (Hojo et al., 2016), to Susan Mackem and Juan Tena for discussions. Work in the Rada-Iglesias laboratory is supported by grants PID2021-123030NB-I00, ERC CoG PoisedLogic (862022) and Chrom_rare HORIZON-MSCA-2021-DN-01-101073334. AC-I and SM were recipients of PhD fellowships PRE2018-083421, PRE2021-098480 respectively from the Spanish Ministry of Science and Innovation. This research was supported by the Spanish Ministry of Science and Innovation PID2023-147771NB-I00 to MAR.

Author contributions: A.C-I, S.M. and R.P-G performed the experiments. V.C. contributed to experiments. A.R-I. and J.F.L-G. contributed to the design of the project and the analysis of the data. M.A.R. conceived the project, performed experiments, analyzed the data and wrote the manuscript. All the authors edited the manuscript.

Competing interests: The authors declare that they have no competing interests.

Data and materials availability: The original ChIPmentation and RNA-seq data of this paper have been deposited in GEO and can be accessed with the accession number XXX (available to reviewers upon request and will be made public upon publication).

Source Data 1-5

Source Data 1- Sp8 and Sp6 ChIPmentation peaks (Excel)
Source Data 2- In silico functional annotation of Sp8 and Sp6 peaks using GREAT (Excel)
Source Data 3- Differentially expressed genes between WT and *Sp8*- and *Sp6*- null embryos and corresponding Gene Ontology analysis (Excel)
Source Data 4- Peaks classified by category and intersections with DEGs (Excel)
Source Data 5- Gene Ontology analysis of different categories (Excel)

REFERENCES

Anders, S., & Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biology*, *11*(10), R106. https://doi.org/10.1186/gb-2010-11-10-r106

Bailey, T. L., Johnson, J., Grant, C. E., & Noble, W. S. (2015). The MEME Suite. *Nucleic Acids Research*, *43*(W1), W39–W49. https://doi.org/10.1093/nar/gkv416

Barrow, J. R., Thomas, K. R., Boussadia-Zahui, O., Moore, R., Kemler, R., Capecchi, M. R., & McMahon, A. P. (2003). Ectodermal Wnt3/β-catenin signaling is required for the establishment and maintenance of the apical ectodermal ridge. *Genes & Development*, *17*(3), 394–409. https://doi.org/10.1101/GAD.1044903

Bell, S. M., Schreiner, C. M., Waclaw, R. R., Campbell, K., Steven Potter, S., & Scott, W. J. (2003). *Sp8 is crucial for limb outgrowth and neuropore closure*. www.pnas.orgcgidoi10.1073pnas.2134310100

Catta-Preta, R., Lindtner, S., Ypsilanti, A., Seban, N., Price, J. D., Abnousi, A., Su-Feher, L., Wang, Y., Cichewicz, K., Boerma, S. A., Juric, I., Jones, I. R., Akiyama, J. A., Hu, M., Shen, Y., Visel, A., Pennacchio, L. A., Dickel, D. E., Rubenstein, J. L. R., & Nord, A. S. (2025). Combinatorial transcription factor binding encodes cis-regulatory wiring of mouse forebrain GABAergic neurogenesis. *Developmental Cell*, *60*(2), 288-304.e6. https://doi.org/10.1016/j.devcel.2024.10.004

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*(17), i884–i890. https://doi.org/10.1093/bioinformatics/bty560

Collette, N. M., Genetos, D. C., Murugesh, D., Harland, R. M., & Loots, G. G. (2010). Genetic evidence that SOST inhibits WNT signaling in the limb. *Developmental Biology*, *342*(2), 169–179. https://doi.org/10.1016/j.ydbio.2010.03.021

Cova, G., Glaser, J., Schöpflin, R., Prada-Medina, C. A., Ali, S., Franke, M., Falcone, R., Federer, M., Ponzi, E., Ficarella, R., Novara, F., Wittler, L., Timmermann, B., Gentile, M., Zuffardi, O., Spielmann, M., & Mundlos, S. (2023). Combinatorial effects on gene expression at the Lbx1/Fgf8 locus resolve split-hand/foot malformation type 3. *Nature Communications 2023 14:1*, *14*(1), 1–17. https://doi.org/10.1038/s41467-023-37057-z

Desanlis, I., Paul, R., & Kmita, M. (2020). Transcriptional trajectories in mouse limb buds reveal the transition from anterior-posterior to proximal-distal patterning at early limb bud stage. *Journal of Developmental Biology*, *8*(4), 1–16. https://doi.org/10.3390/JDB8040031

Ewels, P. A., Peltzer, A., Fillinger, S., Patel, H., Alneberg, J., Wilm, A., Garcia, M. U., Di Tommaso, P., & Nahnsen, S. (2020). The nf-core framework for community-curated bioinformatics pipelines. *Nature Biotechnology*, *38*(3), 276–278. https://doi.org/10.1038/s41587-020-0439-x

Fernandez-Guerrero, M., Zdral, S., Castilla-Ibeas, A., Lopez-Delisle, L., Duboule, D., & Ros, M. A. (2021). Time-sequenced transcriptomes of developing distal mouse limb buds: A comparative tissue layer analysis. *Developmental Dynamics*. https://doi.org/10.1002/dvdy.394

Fernandez-Teran, M., & Ros, M. A. (2008). The Apical Ectodermal Ridge: morphological aspects and signaling pathways. *The International Journal of Developmental Biology*, *52*(7), 857–871. http://www.ncbi.nlm.nih.gov/sites/entrez?Db=pubmed&DbFrom=pubmed&Cmd=Link&LinkName=pubmed_pubmed&LinkReadableName=Related

Grant, C. E., Bailey, T. L., & Noble, W. S. (2011). FIMO: Scanning for occurrences of a given motif. *Bioinformatics*, *27*(7), 1017–1018. https://doi.org/10.1093/bioinformatics/btr064

Haro, E., Delgado, I., Junco, M., Yamada, Y., Mansouri, A., Oberg, K. C., & Ros, M. A. (2014). Sp6 and Sp8 transcription factors control AER formation and dorsal-ventral patterning in limb development. *PLoS Genetics*, *10*(8), e1004468. http://www.ncbi.nlm.nih.gov/sites/entrez?Db=pubmed&DbFrom=pubmed&Cmd=Link&LinkName=pubmed_pubmed&LinkReadableName=Related

Hojo, H., Ohba, S., He, X., Lai, L. P., & McMahon, A. P. (2016). Sp7/Osterix Is Restricted to Bone-Forming Vertebrates where It Acts as a Dlx Co-factor in Osteoblast Specification. *Developmental Cell*, *37*(3), 238–253. https://doi.org/10.1016/j.devcel.2016.04.002

Hörnblad, A., Bastide, S., Langenfeld, K., Langa, F., & Spitz, F. (2021). Dissection of the Fgf8 regulatory landscape by in vivo CRISPR-editing reveals extensive intra- and inter-enhancer redundancy. *Nature Communications 2021 12:1*, *12*(1), 1–12. https://doi.org/10.1038/s41467-020-20714-y

Kennedy, M. W., Chalamalasetty, R. B., Thomas, S., Garriock, R. J., Jailwala, P., & Yamaguchi, T. P. (2016). Sp5 and Sp8 recruit β-catenin and Tcf1-Lef1 to select enhancers to activate Wnt target gene transcription. *Proceedings of the National Academy of Sciences*, *113*(13), 3545–3550. https://doi.org/10.1073/pnas.1519994113

Kim, D., Paggi, J. M., Park, C., Bennett, C., & Salzberg, S. L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology*, *37*(8), 907–915. https://doi.org/10.1038/s41587-019-0201-4

Kulakovskiy, I. V., Vorontsov, I. E., Yevshin, I. S., Sharipov, R. N., Fedorova, A. D., Rumynskiy, E. I., Medvedeva, Y. A., Magana-Mora, A., Bajic, V. B., Papatsenko, D. A., Kolpakov, F. A., & Makeev, V. J. (2018). HOCOMOCO: Towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis. *Nucleic Acids Research*, *46*(D1), D252–D259. https://doi.org/10.1093/nar/gkx1106

Kyung Ahn, Y. M. M. C. H. R. R. B. E. B. C. I., Y, M., MC, H., RR, B., EB, C., K, A., Y, M., MC, H., RR, B., & EB., C. (2001). BMPR-IA signaling is required for the formation of the apical ectodermal ridge and dorsal-ventral patterning of the limb. *Development*, *128*(22), 4449–4461.

Lin, C., Yin, Y., Bell, S. M., Veith, G. M., Chen, H., Huh, S. H., Ornitz, D. M., & Ma, L. (2013). Delineating a Conserved Genetic Cassette Promoting Outgrowth of Body Appendages. *PLoS Genetics*, *9*(1). https://doi.org/10.1371/journal.pgen.1003231

Liu, T., Ortiz, J. A., Taing, L., Meyer, C. A., Lee, B., Zhang, Y., Shin, H., Wong, S. S., Ma, J., Lei, Y., Pape, U. J., Poidinger, M., Chen, Y., Yeung, K., Brown, M., Turpaz, Y., & Liu, X. S. (2011). Cistrome: An integrative platform for transcriptional regulation studies. *Genome Biology*, *12*(8), 1–10. https://doi.org/10.1186/gb-2011-12-8-r83

Ma, W., Noble, W. S., & Bailey, T. L. (2014). Motif-based analysis of large nucleotide data sets using MEME-ChIP. *Nature Protocols*, *9*(6), 1428–1450.

Marinić, M., Aktas, T., Ruf, S., & Spitz, F. (2013). An integrated holo-enhancer unit defines tissue and gene specificity of the Fgf8 regulatory landscape. *Developmental Cell*, *24*(5), 530–542. https://doi.org/10.1016/J.DEVCEL.2013.01.025

McLean, C. Y., Bristor, D., Hiller, M., Clarke, S. L., Schaar, B. T., Lowe, C. B., Wenger, A. M., & Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. *Nature Biotechnology*, *28*(5), 495–501. https://doi.org/10.1038/nbt.1630

Nakamura, M., & Ishikawa, O. (2008). The localization of label-retaining cells in mouse nails. *Journal of Investigative Dermatology*, *128*(3), 728–730. https://doi.org/10.1038/sj.jid.5701062

Niederreither, K., & Dollé, P. (2008). Retinoic acid in development: towards an integrated view. *Nature Reviews Genetics*, *9*(7), 541–553. https://doi.org/10.1038/nrg2340

Osterwalder, M., Speziale, D., Shoukry, M., Mohan, R., Ivanek, R., Kohler, M., Beisel, C., Wen, X., Scales, S. J., Christoffels, V. M., Visel, A., Lopez-Rios, J., & Zeller, R. (2014). HAND2 targets define a network of transcriptional regulators that compartmentalize the early limb bud mesenchyme. *Developmental Cell*, *31*(3), 345–357. https://doi.org/10.1016/j.devcel.2014.09.018

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., & Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology*, *33*(3), 290–295. https://doi.org/10.1038/nbt.3122

Presnell, J. S., Schnitzler, C. E., & Browne, W. E. (2015). KLF/SP transcription factor family evolution: Expansion, diversification, and innovation in eukaryotes. *Genome Biology and Evolution*, *7*(8), 2289–2309. https://doi.org/10.1093/gbe/evv141

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*(6), 841–842. https://doi.org/10.1093/bioinformatics/btq033

Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A., & Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Research*, *42*(Web Server issue), W187-91.

Rhodes, C. S., Yoshitomi, Y., Burbelo, P. D., Freese, N. H., Nakamura, T., Chiba, Y., & Yamada, Y. (2021). Sp6/Epiprofin is a master regulator in the developing tooth. *Biochemical and Biophysical Research Communications*, *581*, 89–95. https://doi.org/10.1016/j.bbrc.2021.10.017

Ringel, A. R., Magg, A., Benetti, N., Schöpflin, R., Kühnlein, M., Stiege, A. C., Fischer, U., Wittler, L., Lorenz, S., Mundlos, S., & Allou, L. (2024). *Temporal misexpression of* En1 *during limb development causes distinct phenotypes*. https://doi.org/10.1101/2024.08.06.606766

Robledo, R. F., Rajan, L., Li, X., & Lufkin, T. (2002). The Dlx5 and Dlx6 homeobox genes are essential for craniofacial, axial, and appendicular skeletal development. *Genes & Development*, *16*(9), 1089–1101. http://www.ncbi.nlm.nih.gov/sites/entrez?Db=pubmed&DbFrom=pubmed&Cmd=Link&LinkName=pubmed_pubmed&LinkReadableName=Related

Sahara, S., Kawakami, Y., Belmonte, J. C. I., & O'Leary, D. D. M. (2007). Sp8 exhibits reciprocal induction with Fgf8 but has an opposing effect on anterior-posterior cortical area patterning. *Neural Development*, *2*(1). https://doi.org/10.1186/1749-8104-2-10

Schaeper, N. D., Prpic, N.-M., & Wimmer, E. A. (2010). A clustered set of three Sp-family genes is ancestral in the Metazoa: evidence from sequence analysis, protein domain structure, developmental expression patterns and chromosomal location. *BMC Evolutionary Biology*, *10*, 88. http://www.biomedcentral.com/1471-2148/10/88

Schmidl, C., Rendeiro, A. F., Sheffield, N. C., & Bock, C. (2015). ChIPmentation: Fast, robust, low-input ChIP-seq for histones and transcription factors. *Nature Methods*, *12*(10), 963–965. https://doi.org/10.1038/nmeth.3542

Semënov, M., Tamai, K., & He, X. (2005). SOST is a ligand for LRP5/LRP6 and a Wnt signaling inhibitor. *Journal of Biological Chemistry*, *280*(29), 26770–26775. https://doi.org/10.1074/jbc.M504308200

Sherman, B. T., Hao, M., Qiu, J., Jiao, X., Baseler, M. W., Lane, H. C., Imamichi, T., & Chang, W. (2022). DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Research*, *50*(W1), W216–W221. https://doi.org/10.1093/nar/gkac194

Soshnikova, N., Zechner, D., Huelsken, J., Mishina, Y., Behringer, R. R., Taketo, M. M., Crenshaw, E. B., & Birchmeier, W. (2003). Genetic interaction between Wnt/β-catenin and BMP receptor signaling during formation of the AER and the dorsal-ventral axis in the limb. *Genes and Development*, *17*(16), 1963–1968. https://doi.org/10.1101/gad.263003

Suske, G. (1999). The Sp-family of transcription factors. *Gene*, *238*(2), 291–300. https://doi.org/10.1016/S0378-1119(99)00357-1

Suske, G., Bruford, E., & Philipsen, S. (2005). Mammalian SP/KLF transcription factors: Bring in the family. *Genomics*, *85*(5), 551–556. https://doi.org/10.1016/J.YGENO.2005.01.005

Talamillo, A., Delgado, I., Nakamura, T., de-Vega, S., Yoshitomi, Y., Unda, F., Birchmeier, W., Yamada, Y., & Ros, M. A. (2010). Role of Epiprofin, a zinc-finger transcription factor, in limb development. *Developmental Biology*, *337*(2), 363–374. https://doi.org/10.1016/J.YDBIO.2009.11.007

Tickle, C. (2015). How the embryo makes a limb: Determination, polarity and identity. *Journal of Anatomy*, *227*(4), 418–430. https://doi.org/10.1111/joa.12361

Treichel, D., Schöck, F., Jäckle, H., Gruss, P., & Mansouri, A. (2003). mBtd is required to maintain signaling during murine limb development. *Genes and Development*, *17*(21), 2630–2635. https://doi.org/10.1101/gad.274103

Vaqué, J. P., Navascues, J., Shiio, Y., Laiho, M., Ajenjo, N., Mauleon, I., Matallanas, D., Crespo, P., & León, J. (2005). Myc antagonises Ras-mediated growth arrest in leukemia cells through the inhibition of the Ras-ERK-p21Cip1 pathway. *Journal of Biological Chemistry*, *280*(2), 1112–1122. https://doi.org/10.1074/jbc.M409503200

Visel, A., Minovitsky, S., Dubchak, I., & Pennacchio, L. A. (2007). VISTA Enhancer Browser - A database of tissue-specific human enhancers. *Nucleic Acids Research*, *35*(SUPPL. 1). https://doi.org/10.1093/nar/gkl822

Vokes, S. A., Ji, H., Wong, W. H., & McMahon, A. P. (2008). A genome-scale analysis of the cis-regulatory circuitry underlying sonic hedgehog-mediated patterning of the mammalian limb. *Genes and Development*, *22*(19), 2651–2663. https://doi.org/10.1101/gad.1693008

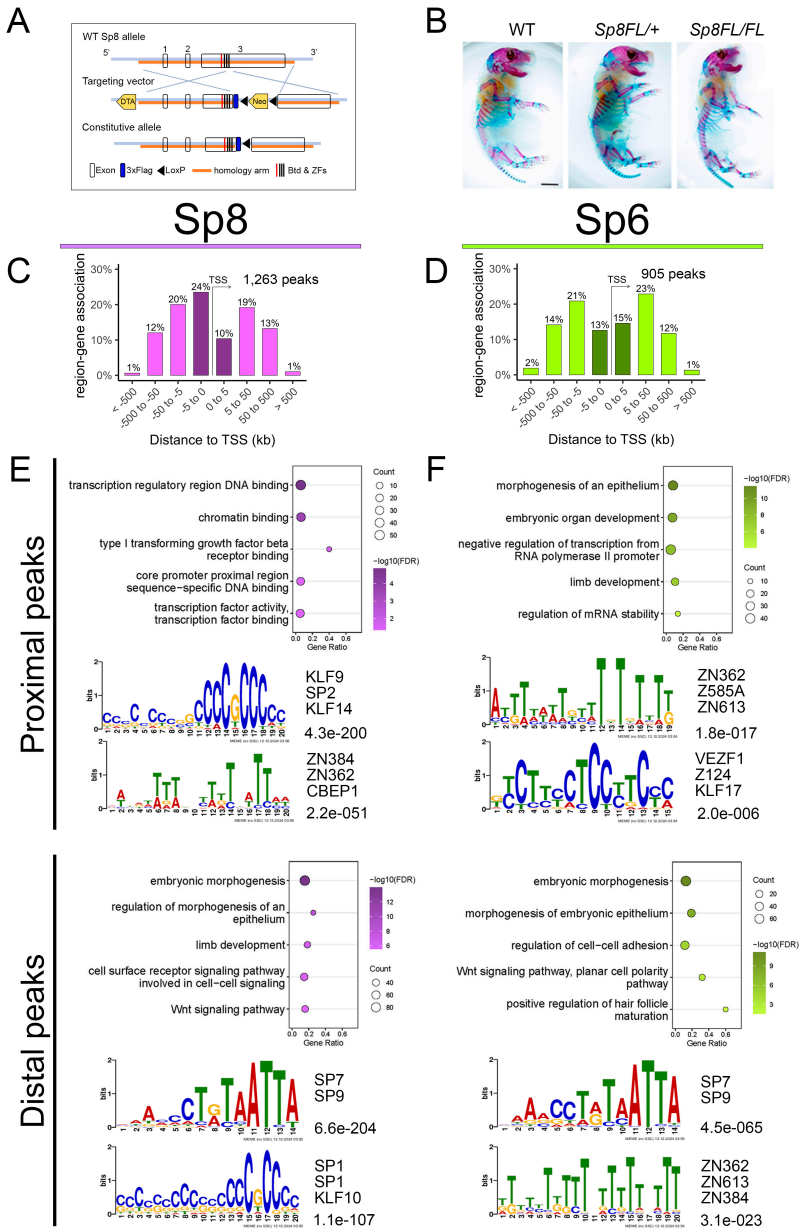
Wagle, P., Nikolić, M., & Frommolt, P. (2015). QuickNGS elevates Next-Generation Sequencing data analysis to a new level of automation. *BMC Genomics*, *16*, 487.

Winkler, D. G., Sutherland, M. K., Geoghegan, J. C., Yu, C. ; H. T., Skonier, J. E., Shpektor, D., Jonas, M., Kovacevich, B. R., Staehling-Hampton, K., Appleby, M., Brunkow, M. E., & Latham, J. A. (2003). Osteocyte control of bone formation via sclerostin, a novel BMP antagonist. *The EMBO Journal*, *22*(23), 6267–6276. https://doi.org/10.1093/emboj/cdg599

Xu, Z., Liang, Q., Song, X., Zhang, Z., Lindtner, S., Li, Z., Wen, Y., Liu, G., Guo, T., Qi, D., Wang, M., Wang, C., Li, H., You, Y., Wang, X., Chen, B., Feng, H., Rubenstein, J. L., & Yang, Z. (2018). SP8 and SP9 coordinately promote D2-type medium spiny neuron production by activating Six3 expression. *Development (Cambridge)*, *145*(14). https://doi.org/10.1242/dev.165456

1040 Ye, T., Krebs, A. R., Choukrallah, M. A., Keime, C., Plewniak, F., Davidson, I., &
1041 Tora, L. (2011). seqMINER: An integrated ChIP-seq data interpretation platform.
1042 *Nucleic Acids Research*, *39*(6). https://doi.org/10.1093/nar/gkq1287
1043
1044

1045 **FIGURE 1**



1046

1047 **Figure 1. Genome-wide analysis of Sp8 and Sp6 binding profiles and functional peak**
1048 **distribution**

1049 **(A)** Scheme showing the generation of the *Sp8-3xFLAG* (*Sp8FL*) allele by homologous
1050 recombination (homology arms in orange). Three copies of FLAG (dark blue) were introduced at the
1051 3'-terminal end of the *Sp8* gene before the STOP codon TGA (middle of third exon). Numbers 1, 2,
1052 3 indicate *Sp8* exons (white). **(B)** Skeletal preparations of WT, and Sp8FL hetero and homozygote
1053 newborn mice. **(C-D)** Distribution of Sp8 (C) and Sp6 (D) peaks relative to their distance to the
1054 nearest TSS. **(E-F)** Analysis of enrichment of GO:biological process (top) and de-novo discovered
1055 motifs (bottom) for Sp8 **(E)** and Sp6 **(F)** separately in proximal (<5kb) and distal peaks (>5kb). Top
1056 2 motifs with best matches of known motifs and e-values are shown.

FIGURE 2

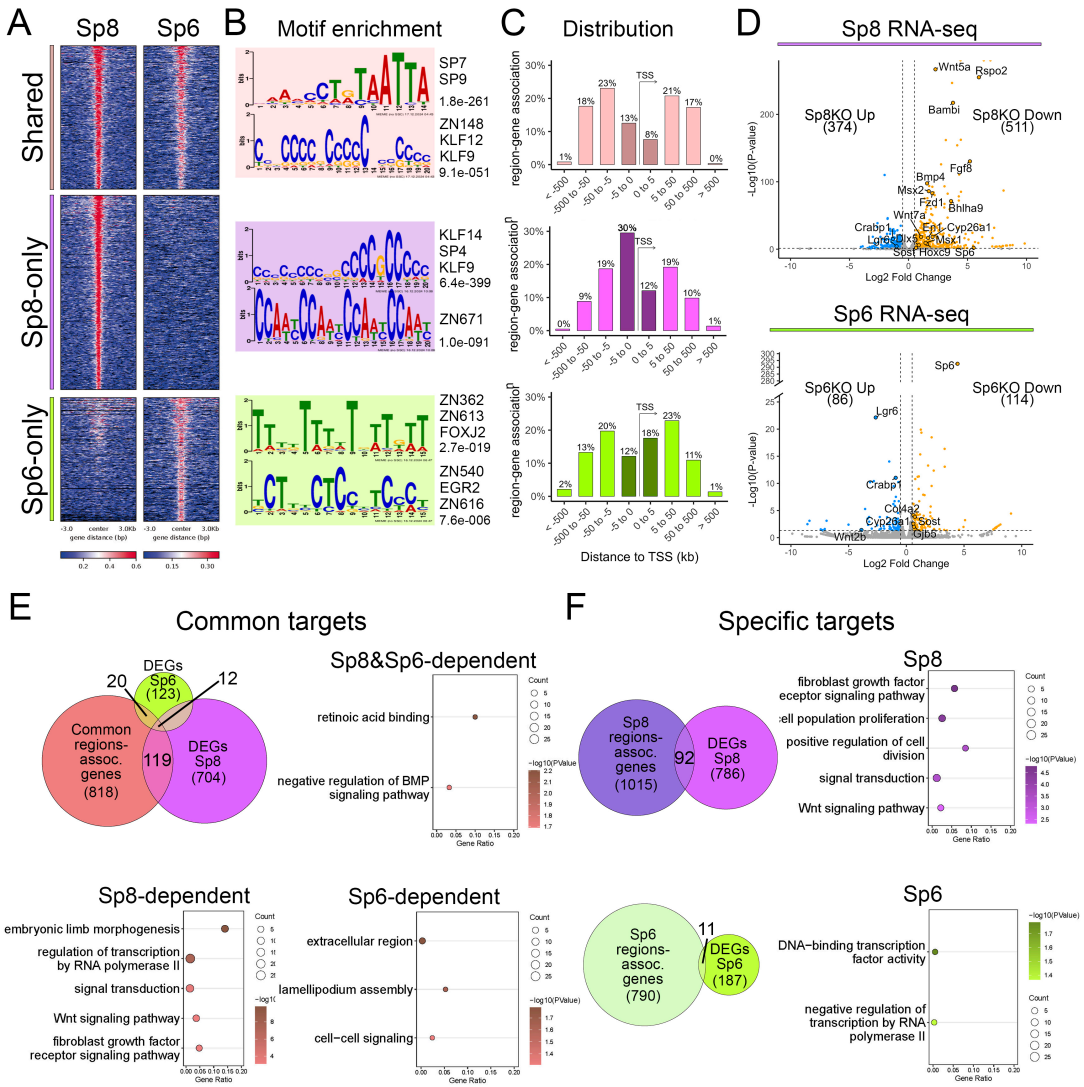
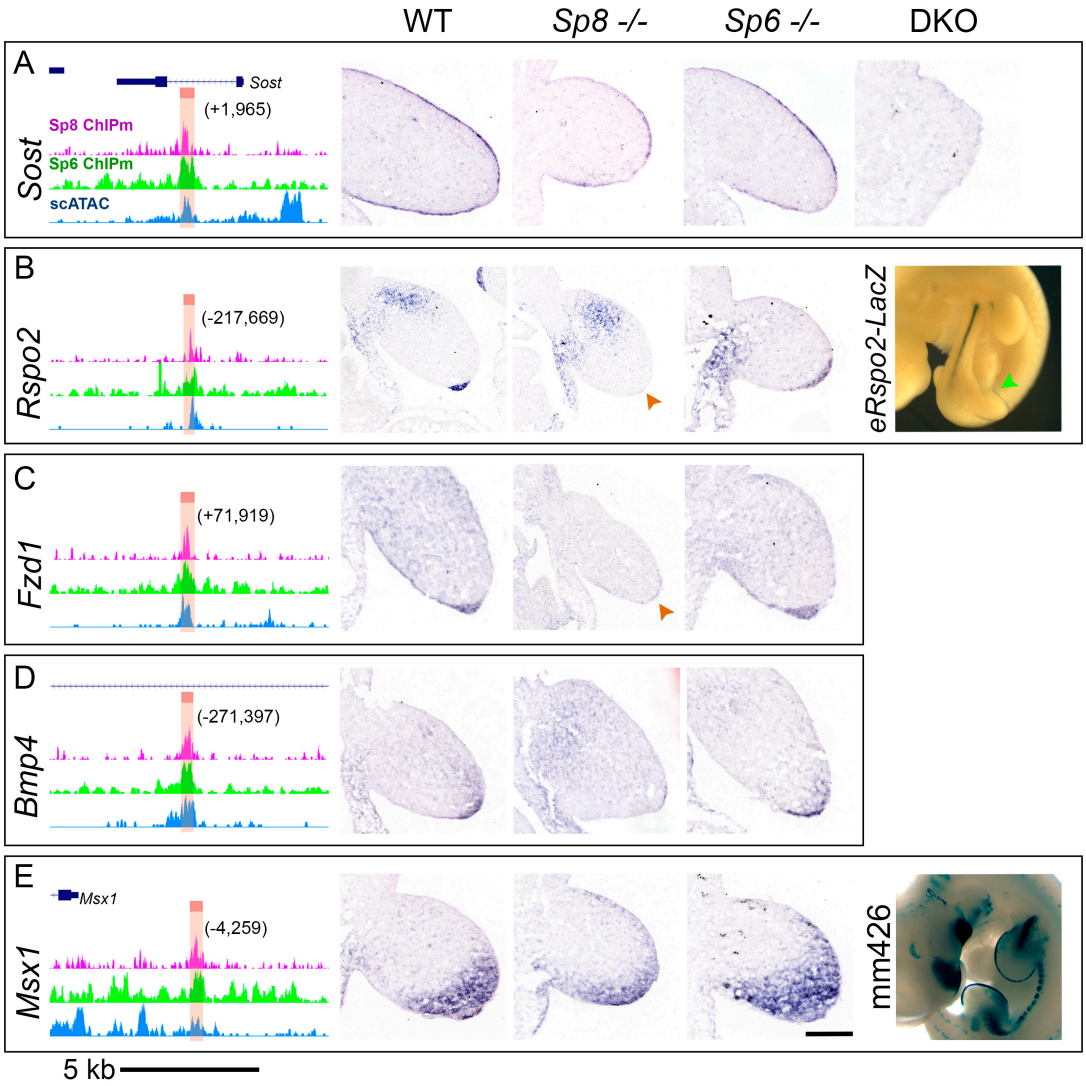


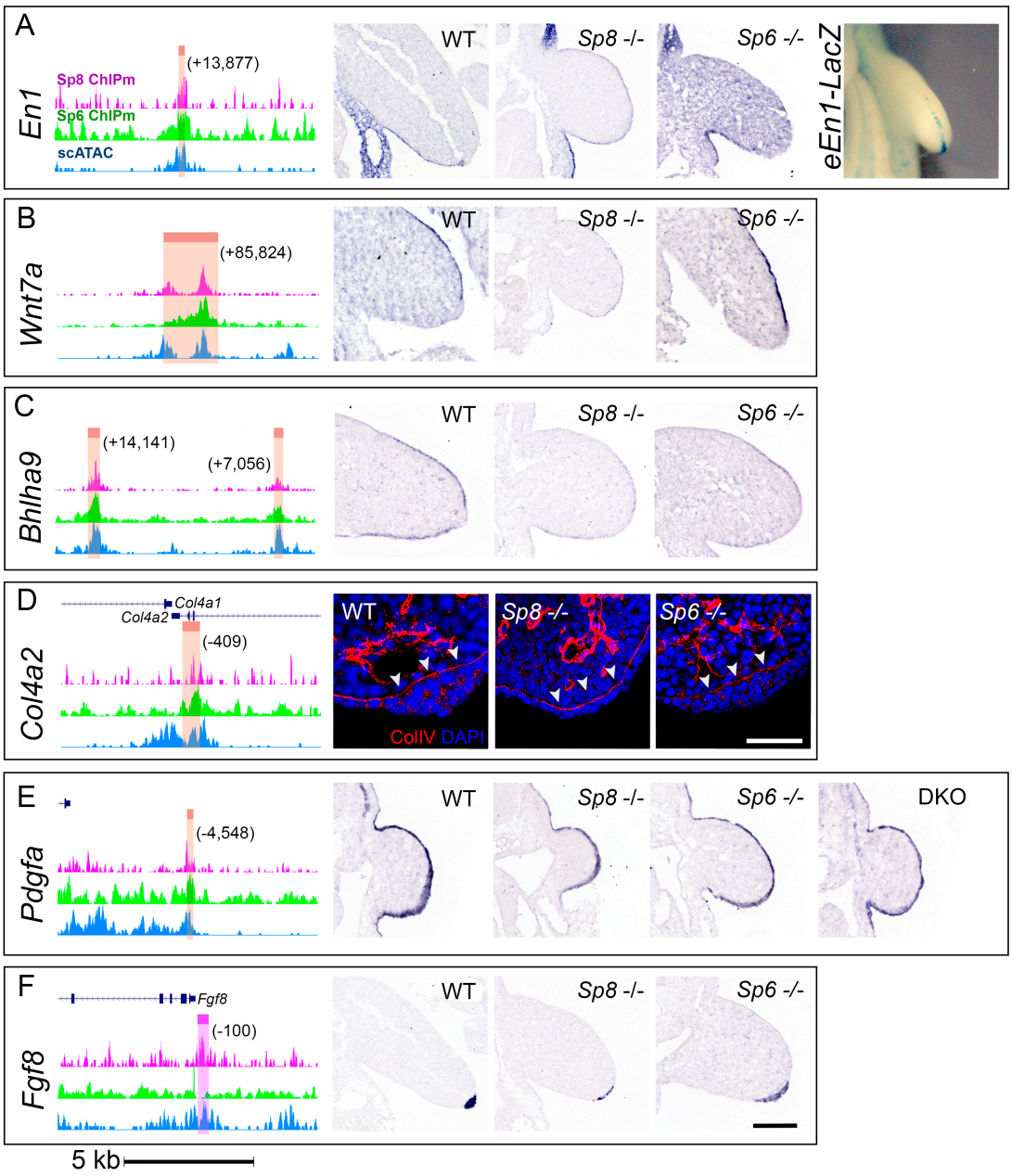
Figure 2. Genome-wide analysis of shared and specific Sp8 and Sp6 binding sites. **(A)** Heatmaps of the three peak clusters, bound by both Sp8 and Sp6 (Shared), by Sp8 only, and by Sp6 only, obtained through k-means clustering of the merged peak set. **(B)** Top 2 most enriched motifs obtained by MEME-ChIP *de novo* motif discovery tool for each cluster. Best matches of known motifs and e-values on the right side. **(C)** Distribution of peaks relative to their distance to the nearest TSS for each cluster. **(D)** Volcano plots of the differential expression analysis obtained from the comparison between WT and *Sp8* homozygous mutants (top) and between WT and *Sp6* homozygous mutants (bottom) with the number of DEGs and selected genes indicated. **(E)** Analysis of commonly bound regions. Venn plot of the intersection of the region-associated genes with Sp8 and Sp6 DEGs and plots of GO:BP terms enriched in each corresponding subcategory as indicated. **(F)** Analysis of specifically bound regions. Venn plot of the intersection of the corresponding region-associated genes with Sp8 DEGs and Sp6 DEGs and plots of GO:BP terms enriched in the corresponding category.

1070 **FIGURE 3**



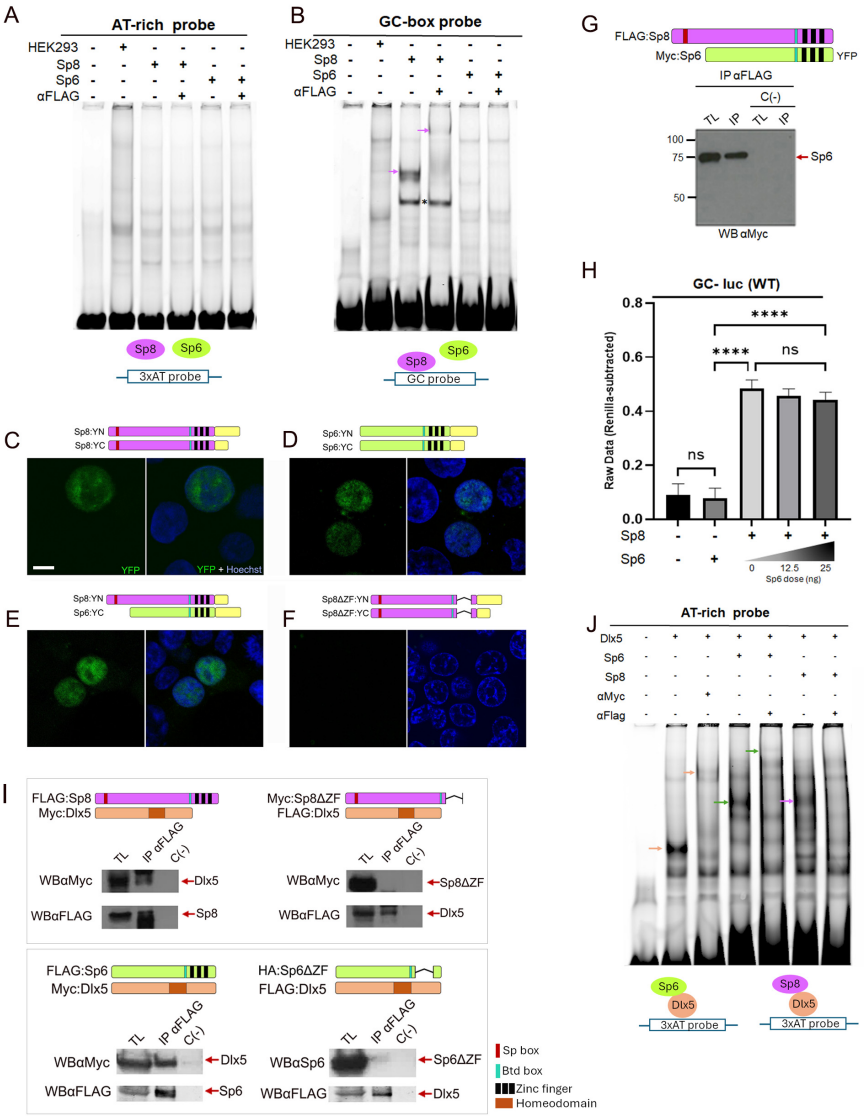
1071 **Figure 3. In situ hybridization validation of Sp8 and Sp6 target genes.**
1072 (**A-E**) For each of the indicated genes, a screenshot of the UCSC genome browser displaying Sp8
1073 (pink) and Sp6 (green) ChIPmentation tracks and the scATAC-seq accessibility in the ectoderm (blue)
1074 is shown. The Sp8/Sp6-associated regulatory modules are highlighted in orange. On the right ISH
1075 images of E10.5 longitudinal limb bud sections of the genotypes indicated at the top. LacZ reporter
1076 activity is shown for the *Rspo2* (B, right; *eRspo2-LacZ*) and the *Msx1* (E, right; mm426 from VISTA),
1077 putative enhancers. Genomic distances to each gene's transcription start site are indicated in
1078 parentheses. Scale bar in genomic tracks = 5 kb. ISH scale bar=100 µm

1079 **FIGURE 4**



1080 **Figure 4. In situ hybridization validation of Sp8 and Sp6 target genes (Continuation).**
1081 **(A-F)** Panels as in Fig. 3. LacZ reporter activity is shown in **(A)** for the *En1* enhancer. IF for Col4a2
1082 showing significant reduction in the basement membrane (indicated by white arrowheads) of *Sp6*
1083 mutants is shown in **(D)**. Scale bar in genomic tracks = 5 kb. IF scale bar=50 µm. ISH scale bar=100
1084 µm

1085 **FIGURE 5**

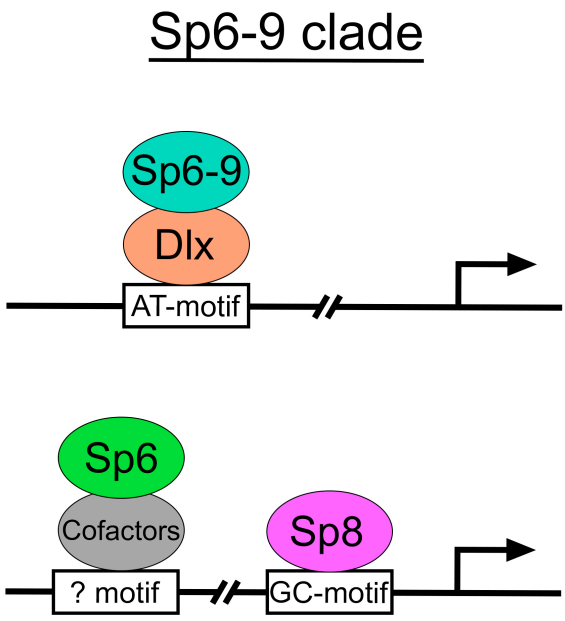


1086 **Figure 5. Characterization and functional analysis of Sp8, Sp6 and Dlx5 interactions.**
1087 **(A)** EMSA using an AT-rich probe shows no specific DNA binding by Sp8 or Sp6. **(B)** EMSA with a
1088 GC-box probe shows specific binding of Sp8, indicated by a shift and a supershift upon addition of
1089 the anti-FLAG antibody (pink arrowheads, lanes 3–4) but not of Sp6. **(C-F)** BiFC assays shows Sp8
1090 and Sp6 homo and heterodimerization; constructs used shown on top. Note the interactions require the
1091 ZF domain (F). Cells were stained with Hoechst to label nuclei. For each experiment, YFP (left) and
1092 merged YFP/Hoechst (right) channels are shown. Scale bar: 10 µm. **(G)** Co-IP confirming Sp6-Sp8
1093 heterodimerization. TL: total lysate; C−: negative control; IP: immunoprecipitation. **(H)** Luciferase
1094 reporter assay using a GC-box promoter shows that Sp8 activates transcription, while Sp6 does not.
1095 Note Sp6 does not interfere with Sp8-mediated activation. Error bars represent standard deviations
1096 calculated from biological triplicates. *P < 0.05; ****P < 0.0005. **(I)** Co-IP assays with the constructs
1097 indicated on top of each panel show that both Sp8 and Sp6 interact with Dlx5 in HEK293 cells (left
1098 panel), and that this interaction requires the ZF domain (right panel). **(J)** EMSA with AT-rich probe

shows that Myc-Dlx5 binds the probe (orange arrowhead, lane 2), with supershift upon anti-Myc addition (lane 3). Co-expression with FLAG-Sp6 or FLAG-Sp8 results in additional complexes (green and pink arrowheads, lanes 4 and 6, respectively), which are either supershifted or no longer visible upon addition of anti-FLAG antibody (lanes 5 and 7, respectively), confirming the Dlx-Sp6/Sp8 complex. The legend for the structural domains is shown in panel (I).

1104 **FIGURE 6**

1105



1106

1107 **Figure 6. Working model illustrating the mode of action of the Sp6-Sp9 clade in the limb**

1108 **ectoderm.**