CONTEXT-DEPENDENT ENHANCER FUNCTION REVEALED BY TARGETED INTER-TAD RELOCATION

Christopher Chase Bolt[1,2,*], Lucille Lopez-Delisle[1], Aurélie Hintermann[2], Bénédicte Mascrez[2], Antonella Rauseo[3], Guillaume Andrey[3] and Denis Duboule[1,2,4,*]

[1] School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland

[2] Department of Genetics and Evolution, Faculty of Sciences, University of Geneva, 30 quai Ernest Ansermet, 1211, Geneva 4, Switzerland

[3] Department of Medical Genetics, Faculty of Medicine, University of Geneva, Rue Michel Servet 1, 1211, Geneva 4, Switzerland

[4] Collège de France, 11 Place Marcelin Berthelot, 75231 Paris, France

Keywords: TADs, *Hox* genes, regulatory landscapes, enhancers, limb development

*Authors for correspondence:

denis.duboule@epfl.ch

christopher.bolt@unige.ch

ABSTRACT

The expression of genes with a key function during development is frequently controlled by large regulatory landscapes containing multiple enhancer elements. These landscapes often match Topologically Associating Domains (TADs) and sometimes integrate range of similar enhancers, thus leading to TADs having a global regulatory specificity. To assess the relative functional importance of enhancer sequences *versus* the regulatory domain they are included in, we set out to transfer one particular enhancer sequence from its native domain into a TAD with a closely related, yet different functional specificity. We used *Hoxd* genes and their biphasic regulation during limb development as a paradigm, since they are first activated in proximal limb cells by enhancers located in one TAD, which is then silenced at the time when the neighboring TAD starts to activate its enhancers in distal limb cells. We introduced a strong distal limb enhancer into the 'proximal limb TAD' and found that its new context strongly suppresses its distal specificity, even though it continues to be bound by HOX13 transcription factors, which normally are responsible for this activity. Using local genetic alterations and chromatin conformation measurements, we see that the enhancer is capable of interacting with target genes, with a pattern comparable to its adoptive neighborhood of enhancers. Its activity in distal limb cells can be rescued only when a large portion of the surrounding environment is removed. These results indicate that, at least in some cases, the functioning of enhancer elements is subordinated to the local chromatin context, which can exert a dominant control over its activity.

INTRODUCTION

Genes with important functions during vertebrate development are frequently multifunctional, as illustrated by their pleiotropic loss-of-function phenotypes. This multi-functionality generally results from the accumulation of tissue-specific enhancer sequences (Banerji et al., 1981) around the gene. At highly pleiotropic loci, the collection of enhancers can be distributed across large genomic intervals, which are referred to as regulatory landscapes (Bolt and Duboule, 2020; Spitz et al., 2003). Such landscapes often correspond in linear size to Topologically Associating Domains (TADs)(Dixon et al., 2012; Nora et al., 2012; Sexton et al., 2012), defined as chromatin domains wherein the probability of DNA-to-DNA interactions is higher than with neighboring domains. These collections of enhancers likely reflect an evolutionary process whereby the emergence (or modification) of morphologies was accompanied by novel regulatory sequences with the ability to be bound by tissue specific factors, and hence alter the transcription of nearby target genes (Carroll et al., 2008).

This model, however, does not explicitly account for the potential influence of local genomic context to modulate interactions between promoters and nascent enhancers, thus adding another layer of complexity to our understanding of developmental gene regulation. For instance, many developmental gene loci are initially decorated with the chromatin mark H3K27me3 deposited by the Polycomb complex PRC2, before their activation (Bernstein et al., 2005), possibly as a way to maintain silencing or limit transcription to very low levels. These chromatin modifications disappear along with gene activation and are re-deposited once the gene and its enhancers are progressively decommissioned, thus preventing inappropriate transcription that may cause severe developmental problems (see Schuettengruber et al., 2017). This indicates that enhancer or promoter function can be affected by their local chromatin environments that may amplify both repressive or activating cues.

In the case of large regulatory landscapes containing many enhancers with similar tissue specificity, the question arises as to how their activation and decommissioning are coordinated and enforced over very large genomic intervals to overcome any contradictory inputs. Two scenarios can be considered in this context. The first possibility is that each enhancer sequence within the landscape independently responds to information delivered solely by the factors bound to it in a sequence-specific manner. Such a mechanism would act entirely in *trans* and its coordination would be ensured by the binding of multiple similar positive or negative factors. In the second scenario, a higher level of regulation is imposed by the general environment of the landscape, for example corresponding to a TAD-wide positive or negative regulation, which would tend to dominate individual enhancer-specific controls in order to minimize dangerous misexpression events.

80 To address this issue, we used the development of the vertebrate limb as a paradigm and, in
81 particular, the essential function of *Hox* genes in patterning and producing the main pieces of
82 tetrapod appendages (Zakany and Duboule, 2007). The emergence of the tetrapod limb
83 structure was a major evolutionary change that facilitated animal migration onto land. A key
84 step in this process was the acquisition of a fully developed distal part (hands and feet),
85 articulating with the more ancestral proximal parts of the limbs (e.g. the arm and the forearm).
86 During development, the distal piece requires the specific function of several key genes,
87 amongst which are *Hoxa13* and *Hoxd13*. While the absence of either gene function induces a
88 moderate phenotype in hands and feet (Dolle et al., 1993; Fromental-Ramain et al., 1996), the
89 double loss-of-function condition leads to the agenesis of these structures (Fromental-Ramain
90 et al., 1996) thus suggesting a critical function for these two genes both during the development
91 of distal limbs and its evolutionary emergence (see Woltering and Duboule, 2010). In this
92 context, the regulatory mechanisms at work to control expression of both genes in most distal
93 limb cells were carefully analyzed either for *Hoxa13* (Berlivet et al., 2013; Gentile et al., 2019),
94 or for *Hoxd13* (Montavon et al., 2011; Spitz et al., 2003), and found to be quite distinct from
95 one another.

96 In the case of *Hoxd* genes, the evolutionary co-option of *Hoxd13* into both the distal limbs and
97 the external genitals was accompanied by the emergence of a novel TAD (C-DOM) containing
98 multiple enhancers specific to either tissue or common to both (Acemel et al., 2016; Amândio
99 et al., 2020; Montavon et al., 2011; Woltering et al., 2014). This TAD is inactive until late in
100 development, when limb cells with a distal fate start to appear. In contrast, the expression of
101 other *Hoxd* genes (*Hoxd9*, *Hoxd10* and *Hoxd11*), which are all critical for the formation of
102 more proximal parts of the limb (Davis et al., 1995), are controlled by enhancers acting earlier
103 and located in the adjacent TAD (T-DOM)(Andrey et al., 2013). The *HoxD* cluster thus lies in
104 between these two TADs and contains several CTCF binding sites that create a boundary
105 between the two adjacent chromatin domains, insulating *Hoxd13* and *Hoxd12* from the rest of
106 the gene cluster (Rodriguez-Carballo et al., 2017). In the early stages of limb bud formation,
107 *Hoxd9*, *Hoxd10,* and *Hoxd11* are activated by the regulatory landscape located telomeric to the
108 cluster (T-DOM) leading to the patterning and growth of long bones of the arms and legs.
109 Subsequently, in a small portion of cells at the posterior and distal end of the early limb bud,
110 enhancers located in the centromeric regulatory landscape (C-DOM) are switched on, driving
111 expression of 5'-located *Hoxd* genes in the nascent hands and feet (Andrey et al., 2013).

112 The operations of these two TADs are mutually exclusive and as soon as C-DOM starts to
113 upregulate expression of *Hoxd13* in distal cells, enhancers within T-DOM are decommissioned
114 and a large part of this chromatin domain becomes decorated by H3K27me3 marks (Andrey et

115 al., 2013). This switch in TAD implementation involves the HOX13 proteins themselves since
116 their production in response to C-DOM enhancers participate in the repression of T-DOM
117 enhancers, likely through direct binding (Beccari et al., 2016; Sheth et al., 2016). In parallel,
118 HOX13 proteins positively regulate C-DOM enhancers, thus re-enforcing the functional switch
119 between the two TADs. In the absence of both HOXA13 and HOXD13, C-DOM is never
120 activated whereas T-DOM enhancers continue to function due to the absence of both
121 decommissioning and H3K27me3 coverage (Beccari et al., 2016). Therefore, in this context,
122 HOXD13 shows properties of both a transcriptional activator and a transcriptional repressor at
123 different times and in different chromatin environments.

124 At the time when *Hoxd13* becomes activated by distal limb (digit) enhancers, it is critical that
125 all proximal limb (forearm) enhancers are rapidly switched off, to prevent that the latter may
126 act on the former, a situation that was shown to be detrimental to limb morphology (Bolt et al.,
127 2021). This bimodal regulatory situation thus provides a paradigm to assay for the presence of
128 a 'context-dependent' repression of enhancer activity. Accordingly, we set out to introduce into
129 the proximal limb-specific T-DOM domain, the strongest distal enhancer normally working
130 within C-DOM to activate *Hoxd13* in digit cells, asking whether this distal limb regulatory
131 sequence would still be able to exert its function when relocated into a TAD where limb
132 enhancers are being repressed in distal cells. We report that this enhancer element, which is
133 functionally very penetrant when introduced at various random positions by non-targeted
134 transgenesis, loses most of its distal limb activity when recombined within T-DOM, even
135 though it continues to recruit the HOX13 factors, which are essential for its function in distal
136 limb cells. We further show that part of the distal limb activity is restored when large portions
137 of T-DOM are deleted *in-cis*. We conclude that the function of this enhancer is inhibited by an
138 *in cis* mechanism acting at the level of an entire chromatin domain, suggesting the existence,
139 in this particular case, of a level of regulation higher than that of the enhancer sequences
140 themselves.

141

142 **RESULTS**

143 **A distal-limb specific enhancer**

144 To evaluate how a tissue specific enhancer would behave when relocated into a
145 different chromatin and regulatory context, we searched for a candidate enhancer element that
146 would display the strongest possible specificity for developing distal limb cells at the *HoxD*
147 locus. We performed ATAC-seq in wild type E12.5 distal limb cells and compared the signals

148 with a previously reported H3K27ac ChIP-Seq dataset (Rodriguez-Carballo et al., 2017).
149 Because distal limb enhancers located at *Hox* loci were shown to colocalize with the binding
150 of HOX13 proteins as assayed in ChIP-seq experiments (Beccari et al., 2016; Sheth et al.,
151 2016), we used a CUT&RUN approach for both HOXA13 and HOXD13 (together referred to
152 as HOX13) to more precisely delineate enhancer elements controlled, at least in part, by these
153 transcription factors.

154 We identified a small element within a region previously described as Island II, one of
155 the islands of the C-DOM regulatory archipelago (Lonfat et al., 2014; Montavon et al., 2011).
156 This element was strongly bound by both transcription factors and accessible as judged by
157 ATAC-Seq, suggesting that it is an active enhancer element in distal limb cells and that it
158 controls the expression of the most posterior *Hoxd* genes there (Figure 1A). Upon closer
159 inspection, we found that the peak from CUT&RUN experiments matched precisely with the
160 peak observed in a previously reported E11.5 whole limb bud ChIP-Seq experiment for
161 HOXA13 and HOXD13 (Figure 1B, dark grey lines)(Sheth et al., 2016). It also matched a
162 small DNA fragment that is highly conserved across tetrapods but thus far absent from all
163 evaluated fish genomes including that of coelacanth (Figure S1A; dashed box). This element
164 had all of the hallmarks of an active distal limb enhancer controlled by HOX13 transcription
165 factors and hence we called it enhancer element II1.

166 To test if this short element indeed carries the enhancer activity reported for the large
167 versions of Island II (Figures 1B and S1A; green rectangles below), we cloned the sequence
168 (532bp, mm10 chr2:74075311-74075843) and constructed a reporter transgene where the
169 enhancer is located 5' to the *HBB* promoter and the *LacZ* gene (Figures 1B and S1A; II1 TgN),
170 and injected it into mouse pronuclei. Five founder animals were identified and then crossed
171 with wild types. Embryos were collected at E12.5 and stained for *LacZ*. Four of the founders
172 transmitted the transgene and produced strong distal limb specific staining (Figure 1C, top).
173 One of the four transmitting founders also had low staining levels in nearly the entire embryo
174 and another founder produced staining in the central nervous system. The four animals
175 transmitting limb staining displayed a pattern closely matching the expression domains of both
176 *Hoxd13* and *Hoxa13* (Figure 1C). The variation observed in staining in other tissues likely
177 resulted from different integration sites.

178 **Deletion of the II1 enhancer sequence**

179 The C-DOM regulatory landscape contains multiple enhancer elements that are
180 necessary to produce a robust activation of *Hoxd* genes in the distal limbs and genitals

181 (Amândio et al., 2020; Montavon et al., 2011) and our ATAC-Seq and CUT&RUN
182 experiments showed that element II1 is among the most strongly bound and accessible elements
183 throughout the C-DOM (Figure 1A). Because of this, we anticipated that it may make a
184 measurable contribution to the transcription of *Hoxd* genes in the distal limb. To determine
185 what effect this element contributes to the global regulatory activity of C-DOM, we used
186 CRISPR/Cas9 to delete the II1 enhancer element (Figures 1B and S1A; green rectangle Del
187 II1). Several founders were obtained and we produced embryos homozygous for this
188 ($HoxD^{DelII1}$) deletion. We collected embryos at E12.5, measured their levels of distal limb *Hoxd*
189 mRNAs by RT-qPCR and looked at the transcript distribution by *in situ* hybridization. Using
190 both approaches, we did not observe any significant change, neither in the level of *Hoxd* gene
191 transcription in the distal limbs of embryos missing the II1 enhancer (Figure S1B, C), nor in
192 the spatial distribution. While this result was somewhat surprising given how strong the signal
193 for H3K27ac, ATAC and HOX13 binding are, a similar lack of effect was observed when
194 single genital enhancers were deleted within this same C-DOM TAD. However, removing
195 several such enhancers had a cumulative effect thus demonstrating the regulatory resilience of
196 this regulatory landscape (Amândio et al., 2020).

197 **Targeted insertion of a C-DOM enhancer into T-DOM**

198 The T-DOM TAD contains multiple enhancers (Figure 2A, orange rectangles), which
199 activate *Hoxd* genes in proximal limb cells, and the deletion of T-DOM abolishes all *Hoxd*
200 gene transcripts from the proximal pieces of the growing limb buds (Andrey et al., 2013; Bolt
201 et al., 2021). In distal limb cells these elements are no longer at work and the entire T-DOM is
202 switched off at the same time as enhancers within C-DOM are activated, including the II1
203 element. HOX13 proteins bind throughout the T-DOM and are associated with the
204 decommissioning of the proximal limb enhancers. When HOX13 factors are not present the
205 decommissioning does not occur and the 'proximal' enhancers continue operating in distal cells
206 (Beccari et al., 2016). In fact, the T-DOM proximal enhancers CS39 and CS65 also operate in
207 distal cells when introduced randomly into the genome as transgenes (Beccari et al., 2016),
208 suggesting that the T-DOM environment represses the function of such sequences in distal
209 cells, and that this mechanism may involve the binding of HOX13 factors.

210 To further challenge this repressive effect and see whether it would dominate over the
211 strong distal specificity of an enhancer that normally operates in distal cells, we set up to
212 relocate a single copy of the enhancer II1 into T-DOM. We used the same transgenic construct
213 (Figure 1, II1 TgN) used to test the enhancer element by random insertion transgenesis (Figure
214 1C), but we attached homology arms to the 5' and 3' ends to target insertion to a region of T-

215 DOM in between - yet at a distance from - two strong CS39 and CS65 proximal limb enhancers
216 (Andrey et al., 2013; Beccari et al., 2016). This region was also selected because it had low
217 levels of the Polycomb Group histone H3 modification H3K27me3, such that H3K27me3 short
218 distance spreading (see Cheutin and Cavalli, 2019) would not directly impact the inserted
219 element. Fertilized eggs were injected with both the targeting construct and various CRISPR
220 components (Figure 2A, Table S1).

221 We identified two founder lines that carried the construct and produced *LacZ* staining.
222 The $HoxD^{II1\text{-}T\text{-}DOM\text{-}542}$ founder line ('allele 542') showed a strong *LacZ* staining in the proximal
223 limb (Figure 2B, black arrows), with very weak staining appearing in the distal limb from E12.5
224 onwards (Figure 2B, white arrows), limited to a small region of the forming digits. Since the
225 II1 transgene was located within T-DOM, which hosts multiple proximal limb enhancers, the
226 strong proximal limb activity likely resulted from the transgene behaving as an enhancer sensor
227 in these cells. The second line identified ($HoxD^{II1\text{-}T\text{-}DOM\text{-}320}$) produced a strikingly different
228 staining pattern (Figure 2C). At E10.5 both the II1 T-DOM 542 and II1 T-DOM 320 lines had
229 strong staining in the proximal limb, but not in the distal limb (Figure 2B, C; compare white
230 and black arrows). Then, at E11.5, while both lines continued to produce strong proximal limb
231 staining, the II1 320 line also had strong distal limb staining (Figure 2C, white arrows). By
232 E12.5 the staining pattern between the lines had almost completely diverged, with II1 T-DOM
233 542 showing strong proximal and very weak distal staining. In contrast the II1 T-DOM 320
234 line produced very strong distal staining but it was nearly absent in the proximal limb (Figure
235 2B, C).

236 To try to understand this difference and to confirm that the II1 transgene was inserted
237 at the expected position, we performed long-read sequencing with the Oxford Nanopore
238 MinION in conjunction with the nCATS protocol (Gilpatrick et al., 2020) to enrich for
239 sequencing reads covering the region around the insertion site (Figure S2A). In the 542 sample,
240 we were able to detect a sequencing read that extended the length of the transgene, both
241 homology arms, and several kilobases of the region that flanks the insertion (Figure S2B), thus
242 confirming that the 542 allele was a single copy transgene inserted correctly at the right site.
243 Because this enhancer was inserted into T-DOM (hereafter II1 T-DOM) producing the correct
244 allele, we used it for all subsequent experiments. We were not able to completely sequence the
245 II1 T-DOM 320 allele because it contained multiple tandem copies of the insertion, yet we
246 could reconstruct a putative genome based on unique overlapping reads which indicated that a
247 minimum of four copies of the transgene were inserted as a tandem array with multiple
248 orientations (Figure S2C).

249 From the analysis of these two alleles, we conclude that when inserted as a single copy
250 at the selected position within T-DOM, the distal limb enhancer activity of II1 was almost
251 entirely repressed and thus this sequence behaved there like any other native proximal limb
252 enhancers located in T-DOM (Andrey et al., 2013; Beccari et al., 2016). Since the heterologous
253 promoter trapped the activity of the surrounding proximal limb enhancers, the final pattern
254 appeared exactly as the reverse pattern for this enhancer when operating from C-DOM. This
255 positive response in proximal limb cells also acted as an internal control; the transgene was
256 capable of transcriptional activity, with a capacity for tissue specific expression. The repression
257 from the chromatin environment in distal cells was completely alleviated when the transgene
258 was present in multiple tandem copies, suggesting a potential micro-structure capable of
259 escaping the negative effect of the local TAD environment. The decreased expression of this
260 allele in proximal cells suggested that the array of transgenes was somehow insulated from
261 receiving the influence of T-DOM proximal enhancers.

262 **HOX13 transcription factors bind to II1 T-DOM in distal limb cells**

263 Despite the observed repression, when the II1 transgene was inserted into T-DOM as a
264 single copy, we scored a very low level of *LacZ* staining in E12.5 distal limb cells, detected
265 only at a late stage and appearing as defined spots in the digit mesenchyme, i.e., in cells that
266 do not reflect well the wild type expression pattern which is more widespread at this stage
267 (Figure 1C and 2B). Because the native II1 enhancer activity seems to be dependent on HOX13
268 proteins (Desanlis et al., 2020), we looked for both the accessibility of- and HOX13 binding to
269 the II1 element within the context of the inactive T-DOM in distal limb cells, to see if HOX13
270 factors could still access and bind the relocated enhancer sequence.

271 We first performed an ATAC-seq on embryos homozygous for the II1 T-DOM single-
272 copy insertion along with wild type controls. We mapped the reads on the mutant genome and
273 used a high mapping quality score (MAPQ30) to ensure that all reads observed on the coverage
274 tracks mapped uniquely to the native II1 element in C-DOM (II1 C-DOM) and the II1 transgene
275 recombined into T-DOM (II1 T-DOM). In wild type control samples, as expected, no ATAC
276 signal was observed over the native II1 C-DOM, neither in proximal limb bud cells, nor in
277 forebrain cells (Figure 3A left). In contrast, a strong ATAC signal appeared over the II1
278 element in distal limb cells (Figures 1B, 3A; DFL). In the II1 T-DOM transgene, there was a
279 weak signal observed over the *HBB* promoter element in forebrain and proximal limb cells, the
280 latter signal in proximal forelimb cells (PFL) likely reflecting the *LacZ* staining in response to
281 nearby proximal limb enhancers. In the distal forelimb samples (DFL), there was a further
282 increase in signal over the *HBB* promoter, and the accessible region extended over the II1

283 enhancer portion of the transgene. This signal over the II1 enhancer was nevertheless weaker
284 than that observed on the II1 element located in its native context (Figure 3A, ATAC DFL,
285 compare left and right).

286 The increased accessibility over the enhancer portion of the relocated II1 transgene in
287 T-DOM suggested that HOX13 transcription factor may still be able to bind. We assessed this
288 by performing CUT&RUN experiments using HOXA13 and HOXD13 antibodies on distal
289 limb cells from embryos homozygous for the II1 transgene in T-DOM and wild type controls.
290 To discriminate between reads from II1 in C-DOM or in T-DOM, only those reads containing
291 sequences uniquely mapping to sequences outside the II1 element itself were considered (see
292 methods). As expected, HOX13 proteins bound strongly to the native II1 element in C-DOM,
293 with a peak over the 3' end of the enhancer in controls (Figure 3A, left). At the analogous
294 position of the II1 transgene in T-DOM, we observed a similar peak for both HOXA13 and
295 HOXD13, indicating that HOX13 transcription factors were able to bind to the II1 enhancer
296 element when inserted into the T-DOM (Figure 3A, right). However, the presence of these
297 factors was not able to drive robust transcription of the nearby *LacZ* transgene in distal limb
298 bud cells.

299 To confirm that the HOX13 peaks detected in the II1 enhancer element indeed
300 corresponded to the presence of the expected HOX13 binding motif(s), we performed a motif
301 search analysis (Heinz et al., 2010) for our CUT&RUN samples and for a previously reported
302 dataset using ChIP-Seq (Sheth et al., 2016). In both datasets, we found motifs for HOX13
303 factors within the II1 element at four different positions. Three of these sites are closely
304 clustered at the 3' end of the enhancer element (Figure 3B, pink bars) and match the peak
305 summit for HOXA13 and HOXD13 in both the native II1 enhancer environment in C-DOM
306 and in the transgene in T-DOM (Figure 3A, grey columns). An additional motif was found
307 within the II1 region, but locates outside the peak region (Figure 3B, asterisk). The three
308 clustered motifs match previously reported HOX13 motifs (Desanlis et al., 2020; Sheth et al.,
309 2016) and their position in relation to the position of CUT&RUN reads suggested that these
310 sites are, in large part, responsible for the distal limb enhancer activity of the native II1 element
311 in C-DOM.

312 **HOX13 binding sites are essential to II1 enhancer activity in distal limb buds**

313 The presence of HOX13 factors bound to the II1 enhancer sequence when integrated
314 into T-DOM suggested that they may be responsible for the weak remaining *LacZ* staining in
315 some specific distal limb cells, even though this staining was distinct from the strong and

316 general accumulation scored with the randomly integrated transgene (II1 TgN, compare Figure
317 1C and 2B). Instead, it could be caused by some other factor(s) taking advantage of the pioneer
318 activity of HOX13 factors (Amândio et al., 2020; Desanlis et al., 2020). We verified this by
319 testing the necessity of these three HOX13 binding sites in the II1 enhancer recombined into
320 T-DOM. We implemented a CRISPR approach *in vivo* to delete the HOX13 binding sites in
321 the II1 T-DOM transgene, by using guides to delete either two or three of the HOX13 binding
322 sites (Figure 4A). Embryos hemizygous for the II1 enhancer in T-DOM were electroporated
323 with CRISPR guides and Cas9 protein and the embryos were collected at E12.5 and stained for
324 *LacZ*. Subsequently, the induced mutations were confirmed by Sanger sequencing (Table S2).
325 While the *LacZ* staining in proximal limb cells was not affected by these mutations, in all cases
326 where the binding sites were deleted, the remaining *LacZ* staining found in distal limb cells of
327 II1 T-DOM transgenic embryos was completely ablated even after overstaining the samples
328 (Figures 4B, Del TFBS and S4-1A).

329 Of note, in several F0 embryos, we observed very strong proximal and distal limb *LacZ*
330 expression (Figures 4B, S4-1B, Del C-T). When we sequenced the genomic DNA of these
331 embryos, we found that they all contained a large deletion extending from the native II1
332 enhancer element within C-DOM up to the II1 transgene within the T-DOM region
333 (approximately 1.2 Mb in length), due to the use of guide target sequences present in both
334 copies of the II1 sequence (scheme in Figure S4-1C). In these embryos, the HOX13 binding
335 sites are deleted and the *LacZ* transgene, formerly located within the T-DOM, was fused with
336 a portion of the C-DOM (Figure S4-1C). In such a configuration, it is very likely that the
337 enhancer remaining 5' to the II1 site in C-DOM (island I) was then able to act on the transgene
338 driving expression in the distal limb (Figures 4B and S4-1C), a regulatory influence obviously
339 not permitted in the presence of an integral native T-DOM. Expression in proximal cells was
340 controlled by those proximal enhancers located telomeric to the deletion breakpoint (Figure
341 2A, CS65 and PLEs).

342 As a control for the requirement of HOX13 factors for the function of the II1 enhancer
343 sequence, we generated two variants of the II1 TgN transgene construct lacking either two (Del
344 2x13) or three (Del 3x13) HOX13 binding sites (Figure 4C, Table S1). We then injected these
345 constructs into embryos to produce random integration events, and then stained for *LacZ*. In
346 nearly all cases, neither the Del 2x13, nor the Del 3x13 variants were able to produce distal
347 limb staining, when compared with the embryos containing the complete II1 enhancer sequence
348 (Figure 4D). There were two exceptions to this: in one Del 2x13 embryo, we observed a clear
349 distal limb staining although the proximal boundaries were very different than that seen in the
350 normal II1 transgene (Figure S4-2, Del 2x13, yellow asterisk) and, in one Del 3x13 embryo,

351 there was strong proximal limb staining and a portion of this staining extended into the
352 posterior portion of the distal limb (approximately digit 5), yet most of the distal limb staining
353 was absent (Figure S4-2, Del 3x13, yellow asterisk). In the remaining embryos carrying the
354 transgene, twenty-seven embryos with the Del 2x13 transgene and seventeen embryos with the
355 Del 3x13 transgene did not produce any distal limb staining (Figures 4D and S4-2).

356 As a final control experiment that integrates the two experiments above, we targeted
357 the insertion of two additional variants of the II1 transgene into the same position of T-DOM
358 as the II1 T-DOM 542 allele. In the first variant, we used the Del 3x13 transgene construction
359 used to test the need for the three HOX13 binding sites when integrated at random locations
360 throughout the genome (Figure 4C-D). When we stained embryos carrying this transgene, we
361 again observed very strong proximal limb staining that matched the 542 allele, indicating that
362 the transgene behaved as a T-DOM enhancer sensor in proximal limb cells (Figure S4-3A), but
363 there was no staining in the distal limb. In the second variant we completely removed the II1
364 enhancer element from the transgene and inserted into the T-DOM. In this construction the
365 *LacZ* staining pattern also produced strong proximal limb staining and no staining in the distal
366 limb (Figure S4-3B).

367 Altogether, these results indicate that even in the presence of bound HOX13 proteins,
368 which are normally the essential factors for its activation, the II1 enhancer sequence
369 recombined within T-DOM was unable to express its full potential. Indeed, only a weak
370 remnant of a transcriptional activity was scored in distal cells, at a late stage and low level,
371 even though the reporter system could work at high efficiency in proximal cells. This suggested
372 that, in distal limb cells, the surrounding chromatin context of T-DOM exerted a dominant
373 effect to repress the activity that this sequence normally displays when positioned within C-
374 DOM, even if the binding of HOX13 factors was still observed.

375 **A TAD-driven repression of distal enhancers in proximal limb cells?**

376 In order to challenge this negative in-*cis* effect, we used CRISPR to delete portions of
377 the T-DOM adjacent to the introduced enhancer transgene and then evaluated the effect of
378 these deletions on transcription of the transgenic *LacZ* construct. Fertilized eggs hemizygous
379 for the II1 T-DOM recombined enhancer allele were electroporated with CRISPR guides and
380 Cas9 protein (Table S1). Since the DNA sequences targeted by guides are present on both the
381 wild type and the II1 T-DOM alleles (Figure 5A, Table S1), we used genotyping PCR to screen
382 for embryos carrying the expected deletion and used changes in *LacZ* staining to associate the

383 deletion with the II1 T-DOM chromosome. As a control, we used littermate embryos that
384 contained the same deletion but on the wild type (non-transgenic) chromosome.

385 The first deletion extended from the 3' end of the *Mtx2* gene, up to, but not including,
386 the II1 enhancer element in the T-DOM (Figure 5A; Del *Mtx2*-II1-T-DOM). This deleted
387 region of T-DOM contains the CS39 and CS93 proximal limb enhancers (Andrey et al., 2013;
388 Yakushiji-Kaminatsui et al., 2018). In this deletion, we only scored a slight reduction in the
389 extent of the *LacZ* staining in proximal limb cells (Figure 5B, black arrows), likely due to the
390 removal of some proximal limb enhancers. In distal limb cells, there was a clear increase in the
391 *LacZ* staining throughout the distal limb mesenchyme spanning almost the entire digital plate,
392 as compared to the same deletion on the wild type chromosome (Figure 5B, compare left and
393 right, Figure S5A).

394 The second deletion extended from the 3' end of the *LacZ* transgene to the telomeric
395 end of the T-DOM regulatory landscape (Figure 5A, Del II1-T-DOM-*Hnrnpa3*). This portion
396 of T-DOM contains the CS65 and PLEs proximal limb enhancer elements (Andrey et al., 2013;
397 Bolt et al., 2021). In this deletion we observed a severe loss of *LacZ* staining in proximal limb
398 cells (Figure 5C, Figure S5B). In distal limb cells, there was a slight difference in the
399 distribution of *LacZ* positive cells, yet no obvious increase in staining when compared to the
400 control chromosome, unlike in the former deletion (Figure 5C, Figure S5B). These results
401 showed that staining could be recovered when the centromeric flanking piece of T-DOM was
402 removed and hence that this chromatin segment somehow exerted a robust repressive effect on
403 the II1 transgene in distal limb bud cells.

404 **The II1 T-DOM transgene contacts the *HoxD* gene cluster**

405 In distal limb bud cells at E12.5, strong chromatin contacts are detected between the II1
406 enhancer sequence (or a larger sequence including it) and the 'posterior' part of the *HoxD*
407 cluster. Along with other C-DOM cis-regulatory regions, the II1-*Hox* cluster interactions
408 collectively sustain activation of *Hoxd13* to *Hoxd11* in the digital plate (Montavon et al., 2011).
409 Because of this, we wondered how this enhancer sequence would behave when relocated within
410 the 3D chromatin space of the neighboring TAD. In other words, would it maintain its contacts
411 with these 'distal' limb genes (*Hoxd13* to *Hoxd11*), not establish any contacts with the cluster
412 at all, or would it adopt the interaction tropism of its new T-DOM neighborhood for the more
413 'proximal' limb genes (*Hoxd10*, *Hoxd9, Hoxd8*)? We performed Capture Hi-C (CHi-C) on
414 E12.5 proximal and distal forelimb cells micro-dissected either from wild type embryos, or
415 from embryos homozygous for the recombined II1 T-DOM transgene. The captured reads were

416 mapped onto the mutant genome excluding all reads that would ambiguously map to both the
417 II1 enhancer in C-DOM and in T-DOM, i.e., sequence reads that would not extend outside of
418 the enhancer itself and hence could not be uniquely assigned to either one of the two sites.

419 In both the proximal and distal forelimb datasets, strong contacts were established
420 between the II1 enhancer element within T-DOM and the *HoxD* gene cluster, as revealed by
421 the subtraction of the mutant contact signal from the wild type (Figure 6A, black arrows; Figure
422 S6A, black arrows). This gain of contacts between the recombined II1 enhancer and the *HoxD*
423 cluster in both proximal and distal cells, were the only noticeable change induced by the
424 presence of the transgene on the general chromatin configuration of the locus (Figure 6A). In
425 order to evaluate these new contacts in greater detail, we plotted pairwise heatmaps between
426 the *HoxD* cluster and II1 T-DOM, as well as H3K27me3, H3K27ac, and CTCF ChIP-seq
427 datasets (Rodriguez-Carballo et al., 2017; Yakushiji-Kaminatsui et al., 2018).

428 These alignments revealed clear differences between contacts in proximal and distal
429 cells, in both their relative strengths and localization. In proximal cells, the II1 enhancer and
430 the *HBB* promoter formed contacts mostly concentrated over the *Hoxd8* to *Hoxd10* region
431 (Figure 6B, grey dashed box). However, even at this 5kb resolution, we were not able to resolve
432 if the contacts were being mediated by the II1 transcription factor binding sites or the *HBB*
433 promoter, because both portions of the transgene are within the same *DpnII* restriction fragment
434 (Figure 6B, C). In distal limb bud cells, the contact dynamic was quite different. While the
435 fragment containing the II1 enhancer and the *HBB* promoter continued to form the strongest
436 contacts with the cluster, the region of highest contact had shifted from the 3' end of *Hoxd8* to
437 the 5' end, and the robust contacts detected around *Hoxd10* in proximal cells had nearly
438 disappeared (Figure 6C). Overall, the contacts were more evenly distributed over the region
439 extending from *Hoxd1* to *Hoxd9*, but they were excluded from the *Hoxd12* to *Hoxd13* region
440 (Figure 6C, grey dashed line).

441 The contacts between the II1 T-DOM transgene and the *HoxD* cluster changed between
442 proximal and distal limb bud cells, exactly matching the distribution of active *versus* inactive
443 chromatin in these two developmental contexts, respectively. In proximal limbs, the region of
444 enriched contacts corresponded to a depletion in H3K27me3 marks and an enrichment in
445 H3K27ac, which exactly matches the region of *HoxD* that is actively transcribed in proximal
446 limb cells (Andrey et al., 2013; Tarchini and Duboule, 2006). In distal limb cells, the contacts
447 became more evenly distributed and correlated with the presence of H3K27me3, stopping
448 abruptly within the H3K27ac-positive portion of *HoxD* cluster, before reaching *Hoxd13* and
449 *Hoxd12* (Figure 6C). Therefore, the II1 enhancer sequence, when recombined into T-DOM

behaved spatially like a strong T-DOM proximal limb enhancers (Figure S6A), even though it had no intrinsic proximal limb specificity, as demonstrated by random transgenesis (Figure 1C, Figure 4D). Regardless of the context and the developmental time, when positioned into T-DOM, it never contacted the *Hoxd12* to *Hoxd13* region, which is the part of the cluster that this enhancer sequence normally contacts with the highest affinity in distal limb bud cells, nor did it influence in any way or in any cell type the chromatin structure that is normally found in T-DOM.

DISCUSSION

The importance and status of enhancer sequences has evolved considerably since their discovery (Banerji et al., 1981). Initially described as short non-coding sequences that can increase the transcription rate of a target gene at a distance and regardless of orientation, they are now known to modulate gene expression in many different ways (Schaffner, 2015). Enhancers have a particular importance for genes with specialized expression patterns, producing transcription in specific cell types and tissues, at precise times and quantities, either during development or subsequently (Long et al., 2016). These sequences are thought to have evolved along with the emergence of novel body structures, recruiting genes already functional elsewhere, to accompany or trigger the formation of these novelties (Carroll et al., 2008). In vertebrates, where multi-functionality is common for genes having important functions during development, enhancers have accumulated in the vicinity of the transcription units forming regulatory landscapes (Spitz et al., 2003) that sometimes extend over several megabases (see (Bolt and Duboule, 2020). Within these landscapes, gene activation can be distributed across multiple enhancers (Marinic et al., 2013; Montavon et al., 2011) often leading to functional redundancy between them or to more complex interactions (Amândio et al., 2020; Osterwalder et al., 2018). Alternatively, enhancer sequences can be grouped together in a more compact manner, either to ensure a coordinated function, as exemplified by the *Globin* genes (Grosveld et al., 2021; Liu et al., 2018; Oudelaar et al., 2021), or to maximize transcription in a given cellular context such as the compact regulatory structure referred to as super enhancers (Blobel et al., 2021; Hnisz et al., 2013).

Enhancers are often embedded into Topologically Associating Domains (TADs), which are regions where certain DNA-DNA interactions are favored while adjacent regions are excluded from the interaction space (Dixon et al., 2012; Nora et al., 2012; Sexton et al., 2012). As illustrated at the *HoxD* locus, the genomic dimensions of TADs sometimes correspond to the extents of regulatory landscapes (Andrey et al., 2013), with TADs providing boundaries to

484 the interaction space of enhancers within the three-dimensional organization of the genome.
485 Consequently, TADs have been considered to be permissive structures augmenting enhancer
486 and promoter interactions, while simultaneously providing borders that prevent enhancers from
487 interacting with elements outside the TAD, and hence to regulate genes in an inappropriate
488 manner (Lupianez et al., 2015, 2016). In this view, however, the enhancer sequence is
489 considered as a regulatory element that can explore and act within the nuclear space somewhat
490 freely unless TAD borders are present to frame its realm of action. Alternatively, there may be
491 some loci where the function of an enhancer sequence can be subordinated to the global
492 chromatin context of a given TAD thus introducing a level of regulatory control derived from
493 a large chromatin domain rather than by individual DNA sequences (Andrey et al., 2013;
494 Marinic et al., 2013; Rouco et al., 2021).

495 The transfer of enhancer II1 into T-DOM may illustrate such a case, where a potent and
496 highly penetrant enhancer sequence is inhibited precisely in those cells where it normally
497 functions, by placing it within a global environment that is not operational in these cells. In its
498 normal location among C-DOM enhancers, the II1 sequence is bound by HOX13 factors
499 (Beccari et al., 2016; Desanlis et al., 2020; Sheth et al., 2016), which are responsible for its
500 strong activation potential as demonstrated here by the deletion of these sites in the transgenic
501 context, which leads to the loss of *LacZ* staining in digit cells. When recombined into T-DOM,
502 the II1 enhancer is silenced yet it still recruits HOX13 factors and hence its silencing cannot
503 be attributed to the absence of the necessary activating factors. In fact, HOX13 factors are also
504 bound to T-DOM whenever this TAD becomes inactive in distal limb cells and covered by
505 H3K27me3 marks (Beccari et al., 2016), suggesting that the same factors may act in both a
506 positive and a negative manner in different chromatin contexts.

507 One potential explanation to this observation is that HOX13 factors may function with
508 more than one modality, depending on their context. On the one hand, these factors may bind
509 to and activate an enhancer-reporter transgene in a sequence-specific manner, as many
510 transcription factors do, leading to the pattern described herein and its absence in transgenes
511 lacking the binding sites. On the other hand, both HOXD13 and HOXA13 proteins contain a
512 large poly-alanine stretch (Bruneau et al., 2001), which may potentially drive the formation of
513 phase-separated globules by co-condensing with transcriptional co-activators/co-repressors, as
514 shown for these and other genes (Basu et al., 2020; Grosveld et al., 2021). It is thus possible
515 that, due to their high content in bound HOX13 proteins, which contain stretches of poly-
516 alanine (Bruneau et al., 2001), both C-DOM and T-DOM are used to form large transcription-
517 hub condensate (Grosveld et al., 2021)(Amândio et al., 2020), leading to a positive
518 transcriptional outcome for C-DOM in distal limb cells, and a negative outcome for T-DOM,

519 within the same cells, due to the inclusion of additional co-factors that are specific to each
520 domain (Karr et al., 2021). This latter explanatory framework would also account for why the
521 deletion of the enhancer II1 *in vivo* had essentially no detectable effect upon transcription of
522 target *Hoxd* genes in distal limbs, much like what was reported for C-DOM enhancers used for
523 external genitals (Amândio et al., 2020) as well as in other comparable instances (Osterwalder
524 et al., 2018). In both cases, removing a single component of the aggregate would not matter
525 too much, whereas removing several related enhancers would then have a measurable impact.

526 This situation contrasts those where a single enhancer is responsible for target gene
527 activation, the deletion of which usually seriously impairs the structure (see e.g. (Lettice et al.,
528 2003; Shapiro et al., 2004). However, the former mechanism would not preclude the capacity
529 for a single distal limb enhancer to trigger *LacZ* expression, as illustrated either by the weak
530 staining detected when II1 was relocated into T-DOM, or when a large deletion brought the
531 *LacZ* reporter close to Island 1, following the fusion between parts of C-DOM and T-DOM.

532 The II1 enhancer was selected because it is one of the strongest and most penetrant
533 distal limb cells enhancer reported thus far (Lonfat et al., 2014). Yet it was silenced when
534 introduced into T-DOM, along with several native T-DOM limb enhancers, which have a
535 strong proximal specificity when inside T-DOM, while they can work efficiently in distal cells
536 as well when randomly integrated as transgenes (Beccari et al., 2016). The repression of T-
537 DOM in distal limb cells is reflected by the presence of large arrays of H3K27me3 marks
538 (Andrey et al., 2013), and it is thus possible that the recombined II1 enhancer was included
539 into this negative chromatin domain. The analysis of the II1 T-DOM 320 line (Figure 2C)
540 suggests that this repression can be competed out by the presence of multiple copies of the
541 enhancer-reporter construct, perhaps due to the formation of a particular sub-structure escaping
542 the negative effect of T-DOM, or simply because of the accumulation of some transcription
543 factors. This line was nevertheless not analyzed further due to technical difficulties associated
544 with duplicated genomic sequences. We however conclude that special attention should be
545 given to copy number when interpreting results from transgenic experiments.

546 The negative effect of T-DOM over the II1 enhancer construct *in-cis* was further
547 suggested by our flanking deletions analyses. The (Del II1-T-DOM-*Hnrnpa3*) telomeric
548 deletion did not substantially change the weak distal staining of the transgene, yet it severely
549 reduced expression in the proximal domain, showing that the most potent proximal enhancers
550 were located in the deleted interval. In contrast, the (Del *Mtx2*-II1-T-DOM) centromeric
551 deletion consistently had the opposite effect, with only a slight reduction of the activity in the
552 proximal limb domain while distal expression was re-activated, thus mapping the main T-DOM

553 region carrying the repressive effect between the gene cluster and the integration site of the
554 enhancer-reporter construct. However, it was not assessed whether this reactivated distal *LacZ*
555 staining is due to the II1 enhancer itself or to the de-repression of other T-DOM enhancers,
556 which would then act on the *HBB* promoter due to the in-*cis* proximity.

557 Finally, the recombined II1 construct was able to specifically and rather strongly
558 contact the *HoxD* cluster, in both proximal (when T-DOM is active) and distal (when T-DOM
559 is inactive) limb bud cells. In these two instances, however, the contacts were distributed
560 differently. In proximal cells, contacts were established with the specific part of the cluster that
561 is heavily transcribed, following the behavior of other T-DOM located enhancers (Andrey et
562 al., 2013). In this case, these interactions were not directly driven by CTCF, for the integration
563 site was selected at a distance from such sites. The enhancer was likely included into a global
564 structure interacting with specific *Hoxd* promoters and possibly organized by CTCF sites
565 within both the gene cluster and T-DOM (Rodríguez-Carballo et al., 2020). In distal cells, the
566 interactions were less specific and likely reflected contacts between large H3K27me3
567 decorated chromatin segments (Noordermeer et al., 2011; Vieux-Rochas et al., 2015), as
568 suggested by the absence of contacts with those genes heavily transcribed in distal cells. There
569 again, the II1 enhancer behaved like its new neighboring proximal enhancers.

570 CONCLUSION

571 Altogether, we conclude that in distal limb bud cells, the necessary decommissioning
572 of all previously acting proximal enhancers is partly achieved -or secured- by a TAD-wide
573 silencing mechanism, as illustrated by the appearance of large domains of H3K27me3. This
574 mechanism is potent enough to prevent the expression of one of the strongest distal limb
575 enhancers, after its recombination within T-DOM. This silencing is partly alleviated when a
576 large piece of flanking chromatin is removed, suggesting the importance of neighboring
577 sequences within the TAD to achieve this effect. When this distal enhancer was introduced into
578 this 'proximal TAD', it behaved in all respects like its new neighbor proximal enhancers, thus
579 illustrating the potential of chromatin domains, in some cases, to impose another level of
580 coordinated regulation on top of enhancer sequence specificities.

581

582 **MATERIALS & METHODS**

583 **Animal work**

584 All experiments were approved and performed in compliance with the Swiss Law on Animal
585 Protection (LPA) under license numbers GE 81/14 and VD2306.2 (to D.D.). All animals were
586 kept in a continuous back cross with C57BL6 × CBA F1 hybrids. Sex of the embryos was not
587 considered in this study. Mice were housed at the University of Geneva Sciences III animal
588 colony with light cycle between 07:00 and 19:00 in the summer and 06:00 and 18:00 in winter,
589 with ambient temperatures maintained between 22 and 23 °C and 45 and 55% humidity, the air
590 was renewed 17 times per hour.

591 **Genotyping**

592 When samples were to be used directly for experiments, a rapid protocol was implemented:
593 Yolk sacs were collected and placed into 1.5 ml tubes containing Rapid Digestion Buffer
594 (10mM EDTA pH8.0 and 0.1mM NaOH) then placed in a thermomixer at 95º for 10 min with
595 shaking at 900 rpm. While the yolk sacs were incubating, the PCR master mix was prepared
596 with Z-Taq (Takara R006B)(see Table S1 for genotyping primers) and aliquoted into PCR
597 tubes. The tubes containing lysed yolk sacs were then placed on ice to cool briefly and quickly
598 centrifuged at high speed. The lysate (1ul) was placed into the reaction tubes and cycled 32X
599 (2s at 98º, 2s at 55º, 15s at 72º). 20ul of the PCR reaction was loaded onto a 1.5% agarose gel
600 and electrophoresis was run at 120V for 10 minutes. Alternatively, when samples could be kept
601 for some time, a more conventional genotyping protocol was applied; Tail Digestion Buffer
602 (10mM Tris pH8.0, 25mM EDTA pH8.0, 100mM NaCl, 0.5% SDS) was added to each yolk
603 sac or tail clipping at 250ul each along with 4ul Proteinase K at 20mg/ml (EuroBio
604 GEXPRK01-15) and incubated overnight at 55ºC. The samples were then incubated at 95º for
605 15 minutes to inactivate the Proteinase K and stored at -20ºC until ready for genotyping.
606 Genotyping primers (Table S1) were combined with Taq polymerase (Prospec ENZ-308) in
607 25ul reactions and cycled 2X with T_a = 64ºC and then cycled 32X with T_a = 62ºC.

608 ***LacZ* staining**

609 Embryos were collected in ice cold 1X PBS in a 12-well plate. They were then fixed for 5
610 minutes at room temperature in freshly prepared 4% PFA with gentle shaking on a rocker plate.
611 After fixing they were washed three times in washing solution (2mM MgCl2, 0.01% Sodium
612 Deoxycholate, 0.02% Nonidet P40, and 1X PBS) for 20 minutes each at room temperature on

613 a rocker plate. After approximately one hour of washing the wash solution was removed and
614 replaced with staining solution (5mM Potassium Ferricynide, 5mM Potassium Ferrocynide,
615 2mM MgCl2 hexahydrate, 0.01% Sodium Deoxycholate, 0.02% Nonidet P40, 1mg/ml X-Gal,
616 and 1X PBS). The plate was wrapped in aluminum foil and placed on a rocker plate over night
617 at room temperature. The following morning the staining solution was removed and the
618 embryos were washed three times in 1X PBS and then fixed in 4% PFA for long-term storage.
619 Images of embryos were collected with an Olympus DP74 camera mounted on an Olympus
620 MVX10 microscope using the Olympus cellSens Standard 2.1 software.

621 **Whole-mount *in situ* hybridization (WISH)**

622 Embryos were collected at E12.5 and processed following a previously reported WISH
623 procedure (Woltering et al., 2009). Briefly, embryos were fixed overnight in 4% PFA at 4ºC.
624 The following day they were washed and dehydrated through 3 washes in 100% methanol and
625 then stored at -20ºC until ready for processing. Each sample was prepared with Proteinase K
626 (EuroBio GEXPRK01-15) at 1:1000 for 10 min. Hybridizations were performed at 69ºC and
627 post-hybridization washes were performed at 65ºC. Staining was performed with BM-Purple
628 (Roche 11442074001). All WISH were performed at on at least three biological replicates.
629 Images of embryos were collected with an Olympus DP74 camera mounted on an Olympus
630 MVX10 microscope using the Olympus cellSens Standard 2.1 software.

631 **RT-qPCR**

632 Embryos were isolated from the uterus and placed into 1x DEPC-PBS on ice. The yolksacs
633 were collected for genotyping. The embryos were transferred into fresh 1x DEPC-PBS and the
634 distal limb portion was excised, placed into RNALater (ThermoFisher AM7020), and stored at
635 -80ºC until processing. Batches of samples were processed in parallel to collect RNA with
636 Qiagen RNEasy extraction kits (Qiagen 74034). After isolating total RNA, first strand cDNA
637 was produced with SuperScript III VILO (ThermoFischer 11754-050) using approximately
638 500ng of total RNA input. cDNA was amplified with Promega GoTaq 2x SYBR Mix and
639 quantified on a BioRad CFX96 Real Time System. Expression levels were determined by dCt
640 (GOI - Tbp) and normalized to one for each condition by dividing each dCT by the mean dCT
641 for each wild type set. Table S1 contains the primer sequences used for quantification. Box
642 plots for expression changes and two-tailed unequal variance t-tests were produced in
643 DataGraph 4.6.1.

644 **CUT&RUN**

645 Embryos were collected in ice-cold 1X PBS and yolk sacs were processed according to the
646 rapid genotyping protocol described above. Embryos with the correct genotype were
647 transferred to fresh PBS and dissected. The dissected tissue samples were transferred into 1X
648 PBS containing 10% FCS and then digested with collagenase (see ATAC-Seq protocol below).
649 For the HOXD13 and HOXA13 CUT&RUN, pools of cells from individual embryos were
650 processed. All samples were processed according to the CUT&RUN protocol (Skene et al.,
651 2018) using a final concentration of 0.02% digitonin (Apollo APOBID3301). Cells were
652 incubated with 0.5 ug/100ul of anti-HOXD13 antibody (Abcam ab19866), 0.5ug/100ul of anti-
653 HOXA13 (Abcam Ab106503) in Digitonin Wash Buffer at 4ºC. The pA-MNase was kindly
654 provided by the Henikoff lab (Batch #6) and added at 0.5ul/100ul in Digitonin Wash Buffer.
655 Cells were digested in Low Calcium Buffer and released for 30 minutes at 37ºC. Sequencing
656 libraries were prepared with KAPA HyperPrep reagents (07962347001) with 2.5ul of adaptors
657 at 0.3uM and ligated for 1 hour at 20ºC. The DNA was amplified for fourteen cycles. Post-
658 amplified DNA was cleaned and size selected using 1:1 ratio of DNA:Ampure SPRI beads
659 (A63881) followed by an additional 1:1 wash and size selection with HXB. HXB is equal parts
660 40% PEG8000 (Fisher FIBBP233) and 5M NaCl.

661 CUT and RUN libraries were sequenced paired-end on a HiSeq4000 sequencer and processed
662 as in (Bolt et al., 2021), mapped either on mm10 or on the II1TDOM-542 mutant genome. The
663 E11.5 whole-forelimb HOXA13 and HOXD13 ChIP-Seq datasets (SRR3498934 of
664 GSM2151013 and SRR3498935 of GSM2151014) as well as E12.5 distal and proximal
665 forelimb H3K27Ac and CTCF ChIP-Seq datasets (SRR5855214 of GSM2713703,
666 SRR5855215 of GSM2713704, SRR5855220 of GSM2713707 and SRR5855221 of
667 GSM2713708) were processed similarly to what has been previously published (Beccari et al.,
668 2021). Adapter sequences and bad quality bases were removed with Cutadapt [Martin et al.
669 https://doi.org/10.14806/ej.17.1.200.] version 1.16 with options -a
670 GATCGGAAGAGCACACGTCTGAACTCCAGTCAC -A
671 GATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATC
672 ATT -q 30 -m 15 (−A being used only in PE data sets). Reads were mapped with bowtie
673 (Langmead and Salzberg, 2012) 2.4.1 with default parameters on mm10. Alignments with a
674 mapping quality below 30, as well as discordant pairs for PE datasets, were discarded with
675 samtools view version 1.8 (Li, 2011; Li et al., 2009). Coverage and peak calling were computed
676 by macs2 (Zhang et al., 2008) version 2.1.1.20160309 with options --bdg --call-summits --
677 gsize '1870000000', and -f BAMPE for PE. The HOX13 motifs where identified with

678 findMotifsGenome.pl from the Homer tool suite (Heinz et al., 2010) using the narrowPeak of
679 HOXA13 and HOXD13 ChIP as well as the third replicate of the HOXA13 and HOXD13 CUT
680 and RUN with the option -size 50. The best motif of each of these four datasets was used to
681 scan the sequence of the II1 enhancer. Four motifs were identified, for the three displayed in
682 Figure 2B, the logo of the motif giving the best score is shown.

683 **ATAC-Seq**

684 Samples used for ATAC-Seq were processed following the original protocol. Cells were
685 collected in 1X PBS on ice and yolk sacs were collected for each sample. Embryos were rapidly
686 genotyped (see above) and those with the correct genotype were transferred to fresh 1X PBS
687 and dissected. Tissue samples were transferred into 300ul 1X PBS containing 10% FCS on ice
688 until ready for processing. To each sample, 8ul of collagenase (at 50mg/ml, Sigma C9697) was
689 added and tubes were placed in a Thermomixer at 37º with shaking at 900rpm for
690 approximately 5 minutes or until the samples were completely disaggregated. The samples
691 were then placed into a centrifuge at 4ºC and centrifuged at 500xg for 5 minutes. The
692 supernatant was removed and cells were gently resuspended in ice-cold 1X PBS. The cells in
693 each sample were counted with a Countess (ThermoFisher) using Trypan Blue and checked for
694 viability >90%. The volume needed to contain was 50,000 cells was determined and that
695 volume was transferred to a new tube and centrifuged at 4ºC at 500xg for 5 minutes. The
696 supernatant was removed and the cell pellet was gently resuspended in lysis buffer then
697 immediately centrifuged at 500g for 10 minutes at 4ºC. The lysis buffer was removed and the
698 cell pellet was gently resuspended in 50ul tagmentation mix (Nextera FC-121-1030) and then
699 incubated at 37º in a thermomixer at 300rpm for 30 minutes. The samples were then mixed
700 with Buffer QG from the Qiagen MinElute PCR Purification kit (28004) and processed
701 according to that protocol, and then eluted from the column in 11ul EB. Samples were stored
702 at -20ºC until ready for library preparation. For library preparations, the samples were
703 amplified with Nextera Index Primers (FC-121-1011) using NEBNext High-Fidelity 2x PCR
704 Master Mix (M0541) and cycled 11 times. After PCR the reactions were cleaned first with the
705 Qiagen MinElute PCR Purification Kit and then with AMPure XP beads (A63881) at a ratio of
706 1.8:1.0 followed by elution with 15ul EB.

707 Adapter sequences and bad quality bases were removed with Cutadapt (Martin, 2011) version
708 1.16 with options -a CTGTCTCTTATACACATCTCCGAGCCCACGAGAC -A
709 CTGTCTCTTATACACATCTGACGCTGCCGACGA -q 30 -m 15. Reads were mapped with
710 bowtie (Langmead et al., 2009) 2.4.1 with parameters -I 0 -X 1000 --fr --dovetail --very-
711 sensitive on mm10 or on the II1TDOM-542 mutant genome. Alignments with a mapping

712 quality below 30, discordant pairs, and reads mapping to the mitochondria, were discarded with
713 bamtools version 2.4.0 [https://github.com/pezmaster31/bamtools]. PCR duplicates were
714 removed with Picard[http://broadinstitute.github.io/picard/index.html] version 2.18.2 before
715 the BAM to BED conversion with bedtools (Quinlan, 2014) version 2.30.0. Coverage and peak
716 calling were computed by macs2 (Zhang et al., 2008) version 2.1.1.20160309 with options --
717 format BED --gsize 1870000000 --call-summits --keep-dup all --bdg --nomodel --extsize 200
718 --shift -100.

719 **Capture Hi-C**

720 Samples used in the Capture Hi-C were identified by PCR screening embryos at E12.5 as
721 described above. Collagenase treated samples were cross-linked with 1% formaldehyde
722 (ThermoFisher 28908) for 10 minutes at room temperature and stored at -80º until further
723 processing. The SureSelectXT RNA probe design used for capturing DNA was done using the
724 SureDesign online tool by Agilent. Probes cover the region mm9 chr2:72240000-76840000
725 producing 2x coverage, with moderately stringent masking and balanced boosting. Capture and
726 Hi-C were performed as previously reported. Sequenced DNA fragments were processed as
727 previously reported but the mapping was performed on a mutant genome reconstructed from
728 minion and Sanger sequencing (see below). A custom R (www.r-project.org) script based on
729 the SeqinR package (Charif and Lobry, 2007) was used to construct a FASTA file for the
730 mutant chromosome 2 from the wild-type sequence and the exact position and sequence of
731 breakpoints.

732 Subtraction of matrices was performed with HiCExplorer (Ramirez et al., 2016; Wolff et al.,
733 2020, 2011) version 3.6. Heatmaps were plotted using a custom version (Lopez-Delisle et al.,
734 2021) of pyGenomeTracks (Ramírez et al., 2018) based on 3.6.

735 **II1 TgN Cloning and transgenesis**

736 The II1 enhancer sequence (mm10 chr2:74075305-74075850) was amplified from the fosmid
737 clone WI1-109P4 using primers 001 and 002 (Table S1). The 001 primer contains a *XhoI* site
738 and LoxP sequence followed by sequence to the II1 enhancer. The 002 primer contains at its
739 5' end a *HindIII* site. This PCR product was gel purified with Qiagen Gel Extraction Kit
740 (28704). The PCR fragment and the pSKlacZ reporter construct were digested with *XhoI* and
741 *HindIII* and ligated together with Promega 2X Rapid Ligation kit (C6711) to produce pSK-II1-
742 LoxP-*LacZ* construct. This vector (15ug) was cut with *XbaI* and *XhoI* to release the enhancer-
743 reporter construct. The digest was separated on a 0.7% agarose gel for 90 minutes. The 4190bp

744 fragment was excised from the gel and purified with Qiagen Gel Extraction Kit (28704) and
745 eluted in 30ul EB followed by phenol-chloroform extraction and ethanol precipitation and then
746 the pellet was dissolved in 30ul TE (5mM Tris pH7.5, 0.5mM EDTA pH8.0). DNA was
747 injected at 3ng/ul into pronuclei. Five founder animals were identified carrying the transgene
748 by PCR. Four male founders with the transgene were put into cross with wild type females and
749 embryos were collected at E12.5 to test for *LacZ* staining. All four male founder lines (*LacZ*/40,
750 41, 44, and 46) produced distal limb staining but *LacZ*/40 was chosen for amplification of the
751 breeding line due to high transmission of the transgene.

752 **II1 T-DOM targeted insertion**

753 The II1 TgN transgenic construct, outlined above, was used for the targeted insertion but
754 homology arms (HA) were attached (Left HA: mm10 chr2:75268556-75269591, Right HA:
755 mm10 chr2:75269617-75270665). The cloning vector was linearized with *KpnI*, separated on
756 a 0.7% agarose gel for 90 minutes at 90 volts, then the 9.0kb band was extracted and purified
757 two times with Qiagen Gel Extraction kit (28704). The DNA was quantified by Qubit dsDNA
758 and diluted to 5ng/ul with IDTE (11-05-01-05). Wildtype fertilized eggs were injected with the
759 construct and the supercoiled pX330 (Addgene #42230) expression vector containing the
760 sgRNA sequence (r4g9: mm10 chr2:75269597-75269616). Animals were genotyped (Table
761 S1) to identify founders, and then sequenced with minION (below).

762 **minION Sequencing**

763 Long-read sequencing was performed on the II1 T-DOM alleles (II1 T-DOM 320 and II1 T-
764 DOM 542) following the nCATS protocol with minor changes (Gilpatrick et al., 2020). Yolk
765 sacs were isolated from embryos containing the II1 T-DOM transgene and digested with Tail
766 Digestion Buffer (see above) and Proteinase K overnight at 55ºC with no shaking. The
767 following day the samples were incubated at 95ºC for 10 minutes to inactivate the Proteinase
768 K followed by ethanol precipitation and eluted in 200ul of 10mM Tris pH7.5. CRISPR guides
769 (Table S1) were designed in CHOPCHOP v3.0 (Labun et al., 2019) and synthesized as Alt-R
770 RNAs by IDT. CRISPR crRNAs were duplexed with tracrRNAs according to the IDT protocol
771 (Alt-R CRISPR-Cas9 System: In vitro cleavage of target DNA with ribonucleoprotein
772 complex, version 2.2). Two master mixes of guide RNAs and Cas9 protein (1081059) were
773 prepared (see Supplementary Figure 3 for sequence and cutting locations with the locus map),
774 containing either SCS-12 and SCS-13 or SCS-14. The gDNA (9ug) from the yolk sac was
775 dephosphorylated with NEB Quick CIP (M0510) for 10 minutes at 37ºC followed by 2 minutes

776 at 80ºC to inactivate the CIP. The gDNA was split into two equal pools and each pool was then
777 combined with the guideRNP master mixes to cut the gDNA for 30 minutes at 37ºC followed
778 by 5 minutes at 72ºC. The samples were then A-Tailed and AMX adaptors were ligated (Oxford
779 Nanopore SQK-LSK109). The reactions were size selected with 0.3X Ampure SPRI beads
780 (A63881) followed by two washes with Long Fragment Buffer and then eluted for 30 minutes
781 in 15ul EB. The DNA libraries were then prepared according to the Oxford Nanopore protocol
782 for sequencing on a minION (ENR_9084_v109_revP_04Dec2018). The sequencing ran for
783 approximately 24 hours and was stopped for processing after all nanopores were depleted.

784 Bases were called from the fast5 files using Guppy base-caller (Oxford Nanopore
785 Technologies) for CPU version 5.0.16+b9fcd7b. Reads were mapped on mm10 with minimap2
786 (Li, 2018) version 2.15 with parameter -ax map-ont. Only primary alignments were kept with
787 samtools view version 1.10 (Li, 2011; Li et al., 2009) and reads mapping to II1
788 (mm10:chr2:74073413-74076528) or to the insertion region (mm10:chr2:75262998-
789 75286118) were further analyzed. Read sequences were compared to the wild-type genome,
790 the expected mutant genome as well as the sequence of the cloning vector using a Perl script
791 as in (Schmidl et al., 2015) with the following modification: 20 bp of the MinION reads were
792 tested against the reference for 5 bp-sliding windows and only 20-mers completely identical to
793 unique 20-mers in the reference were kept. The output was then processed in R (www.r-
794 project.org) to display dot plots. The in-depth analysis of reads for allele 320 allowed to
795 propose a configuration that would match all reads containing in total four times the II1-*LacZ*
796 construct.

797 **Del II1 TFBS TgN cloning and transgenesis**

798 The enhancer constructs with mutations (Figure 4C) in the transcription factor binding sites
799 were constructed *in silico* and synthesized by TWIST Bioscience (San Francisco, CA). The
800 enhancer sequences are available in Table S1. The enhancer sequences were synthesized with
801 *BglII* and *AgeI* restriction sites at the 5' and 3' ends respectively. The mutant enhancer
802 sequences (Del 2x13 and Del 3x13) were restriction digested along with pSKlacZ and ligated
803 together with Promega 2X Rapid Ligation kit to produce pSK-II1Del2X13lacZ and pSK-
804 II1Del3X13lacZ. The enhancer-reporter fragments were released from the vector with *BglII*
805 and *XhoI* and purified as above. Pro-nuclear injections were performed by the transgenic
806 platform of the University of Geneva, medical school (CMU). Embryos were collected at
807 approximately E12.5 and stained for *LacZ*.

808 **Deletions of transcription factor binding sites within T-DOM *in vivo***

809 Guide sequences were selected from the UCSC mm10 genome browser track CRISPR/Cas9 -
810 NGG Targets. The crRNA Alt-R guides were synthesized by IDT. Males homozygous for the
811 II1 T-DOM 542 allele were crossed with super-ovulated wild type females (BL6XCBA-F1)
812 and fertilized eggs were collected. The embryos were electroporated with CRISPR guides
813 (12ug of each guide) and TrueCut Cas9 v2 protein (Thermo Fisher A36497) with a NEPA21
814 (NEPA GENE Co. Ltd, Chiba, Japan) and then reimplanted into surrogate females. Embryos
815 were collected at E12.5. Yolk sacs were digested and the II1:*HBB:LacZ* transgene was PCR
816 amplified and Sanger sequenced to identify transgenic embryos containing the mutagenized
817 enhancer element. Embryos that were mosaic for the mutation were not included in this
818 analysis. The embryos were *LacZ* stained (see above) at 37º for 16 hours, washed in PBS and
819 post-fixed, then stored in 70% ethanol for photographing.

820 **Generation of Del *Mtx2*-II1-T-DOM and Del II1-T-DOM-*Hnrnpa3* alleles *in vivo***

821 The same methodology was used for this experiment as in the Del II1 T-DOM TFBS
822 experiment above. The sequences for CRISPR guides used in this experiment are listed in Table
823 S1. The embryos were genotyped for the presence of the deletion using primers in Table S1.
824 Embryos with ambiguous genotyping results were not used in these results.

825 **Targeted insertion of Del 3x13 and *HBB*:*LacZ* transgene in T-DOM of ESCs.**
826 The two targeted insertions of control transgenes into T-DOM, genetic editing, and cellular
827 culture were performed as previously reported (Figure S4-3) (Andrey and Spielmann, 2017;
828 Kraft et al., 2015). The guide sequence (r4g9, Table S1) was cloned into pX459 vector from
829 Addgene (#62988) and 8ug of the vector was used for mESC transfection. The pX459 vector
830 was co-transfected with 4ug of the vector containing one of the two transgene cassettes. The
831 two transgene cassettes (Del 3x13 and HBB:lacZ) contain the same vector backbone with the
832 HBB promoter, lacZ gene, and SV40 polyA signal and homology arms used to target the r4
833 region of T-DOM (mm10 chr2:75269597-75269616). This is the same insertion site as the II1
834 T-DOM 542 and 320 alleles. The Del 3x13 contains the II1 enhancer element but with the three
835 HOX13 binding sites removed (see Figure 4C, D). The HBB:lacZ transgene is the same but
836 does not contain any portion of the II1 enhancer element so that it is strictly an enhancer sensor
837 in the T-DOM. These constructs were co-transected into G4 mESCs obtained from the Nagy
838 laboratory (George et al., 2007). After genotyping to confirm the insertion, the desired mESCs
839 were thawed, seeded on male and female CD1 feeders and grown for 2 days before the

840 aggregation procedure. ESCs were then aggregated with tetraploid (c57bl6J x B6D2F1)
841 morula-stage embryos and let developed until blastula prior to transfer into CD1 foster females
842 by the transgenic mouse platform at the University of Geneva Medical School (Artus and
843 Hadjantonakis, 2011).
844

845 **ACKNOWLEDGEMENTS**

846 We thank Jozsef Zakany for his help in an initial phase of this work and other colleagues from
847 the Duboule laboratories for discussions This work was supported in part using the resources
848 and services of the Gene Expression Research Core Facility (GECF) at the School of Life
849 Sciences of EPFL and the transgenic platform at the medical school, University of Geneva.

850 **ETHICS APPROVAL**

851 All experiments involving animals were performed in agreement with the Swiss Law on
852 Animal Protection (LPA), under license No. GE 81/14 (to DD).

853 **DATA AVAILABILITY**

854 All raw and processed datasets are available in the Gene Expression Omnibus (GEO) repository
855 under accession number XXXXXX. All scripts necessary to reproduce figures from raw data
856 are available at https://github.com/lldelisle/scripts\.

857 **COMPETING INTERESTS**

858 The authors declare that they have no competing interests.

859 **FUNDING**

860 C.C.B was supported by the Eunice Kennedy Shriver National Institute of Child Health &
861 Human Development of the National Institutes of Health, under Award Number
862 F32HD093555. This work was supported by funds from the Ecole Polytechnique Fédérale
863 (EPFL, Lausanne), the University of Geneva, the Swiss National Research Fund (No.
864 310030B_138662 and 310030_196868 to D.D. and No. PP00P3_176802 to G.A.) and the
865 European Research Council grant Regul*Hox* (No 588029) (to D.D.). Funding bodies had no
866 role in the design of the study and collection, analysis and interpretation of data and in writing
867 the manuscript.

AUTHOR CONTRIBUTIONS

C.C.B.: Designed and conducted experiments, analyzed datasets, formalized results and wrote the paper.

L.L-D. : Analyzed and evaluated the statistical significance of datasets. Wrote the paper.

B.M.: Produced and maintained all of the mouse lines.

A.H.: Performed the Del 2x13 and 3x13 experiment

A.R.: Performed the Del 3x13 T-DOM experiment

G.A.: Designed and performed the *HBB*:*LacZ* T-DOM experiment

D.D.: Designed experiments, transported mice, dissected some limb buds and wrote the paper.

879 **LEGENDS TO FIGURES**

880 **Figure 1**. **Identification of a distal limb bud specific enhancer**. **A**. ATAC-seq profile (top)
881 and binding profiles of both the HOXA13 (middle) and HOXD13 (bottom) transcription factors
882 by CUT&RUN, using E12.5 wildtype distal forelimb cells and covering the entire *HoxD* locus,
883 including the two flanking TADs C-DOM and T-DOM (mm10 chr2:73950000-75655000).
884 Green rectangles below are distal limb enhancers in C-DOM and orange rectangles are
885 proximal limb enhancers in T-DOM. The *Hoxd* gene cluster is indicated with a box at the center
886 and the *Lnpk* and *Mtx2* genes are indicated as rectangles with black borders. The *Hoxd13* gene
887 is on the centromeric end of the cluster and indicated by a purple box with a circle on the top,
888 whereas *Hoxd1* is telomeric and indicated by a square on top. The C-DOM Island II enhancer
889 (Montavon et al., 2011) is a green rectangle with a black border and its corresponding signals
890 are indicated by a dashed vertical rectangle. **B**. Magnification of the same tracks as above
891 centered around the Island II enhancer with, in addition, the H3K27ac ChIP-seq signal (top) in
892 E12.5 distal forelimbs (Rodriguez-Carballo et al., 2017). The profiles indicated by the dark
893 grey lines are from ChIP-seq using E11.5 whole limb buds (Sheth et al., 2016) for comparison
894 and are shown for comparison. Below the CUT&RUN profiles are the MACS2 peak summits
895 for the corresponding CUT&RUN samples. The green rectangles below indicate the regions
896 described as Island II in (Montavon et al., 2011) or in (Lonfat et al., 2014). The Del II1 shows
897 the region deleted in this work (Figure S1) and the II1 TgN is the transgene used panel in C.
898 **C.** *LacZ* staining pattern produced by the II1:*HBB:LacZ* (II1 TgN) enhancer reporter transgene
899 at E12.5, showing high specificity for distal limb cells (top). Below are whole-mount *in situ*
900 hybridizations for *Hoxa13* and *Hoxd13* in wild type E12. 5 forelimbs for comparison.
901

902 **Figure 2. Targeted recombination of the II1 reporter construct into T-DOM**. **A**. General
903 scheme of the *HoxD* locus with the two triangles on top showing the extents of both the C-
904 DOM (green) and T-DOM (orange) TADs. Various enhancers are shown either in green (distal
905 limbs) or yellow (proximal limbs) rectangles and the *HoxD* cluster is boxed. The large arrow
906 on top indicates the origin and new location of the island II enhancer transgene into T-DOM.
907 Below is a map of the II1:*HHB:LacZ* construct containing both left (L-HA) and right (R-HA)
908 homology arms, the II1 enhancer element and the *HBB* promoter with a *LacZ* reporter gene. **B**.
909 *B-galactosidase* staining time course and *LacZ* mRNAs (right panel) of the single-copy II1 T-
910 DOM 542 founder line. At E10.5 and 11.5, staining is very strong in the proximal limb (white
911 arrow) while absent in the distal portion (black arrow). By E12.5 weak staining appears in the
912 digit mesenchyme of the distal limb. The WISH for *LacZ* mRNA confirms that the distal limb
913 staining comes from transcription in distal limb cells rather than from stable *B-galactosidase*
914 activity. **C**. *B-galactosidase* staining of the multi-copy II1 T-DOM 320 founder line. Staining

915 is much stronger in this allele and similar to line 542 in E10.5 limb buds. However, by E11.5,
916 strong staining is gained in distal limb cells whereas it disappears from the proximal domain.
917
918 **Figure 3. HOX13 proteins bind to the II1 enhancer in T-DOM**. **A**. ATAC-Seq and
919 CUT&RUN reads mapped to the II1 enhancer sequence, either in its native environment within
920 C-DOM (left column; mm10 chr2:74074674-74076672), or after its targeted recombination
921 within T-DOM (line 542, right column; mm10 chr2:75268925-75270923). The green
922 rectangles below indicate the extent of the region used for the II1 enhancer element with, in
923 pink, the position of the three HOX13 binding sites. The II1 C-DOM element is not accessible
924 by ATAC-Seq in E12.5 forebrain (FB), nor in proximal forelimb cells (PFL) samples. At E12.5
925 it becomes highly accessible in distal forelimb cells (DFL) and is strongly bound by HOX13
926 proteins. The II1 element in T-DOM has low accessibility in the FB and PFL samples, even
927 though there is high transcription of the transgene in PFL. Similar to the II1 element in C-
928 DOM, the II1 enhancer in T-DOM is occupied by HOX13 proteins in distal limb cells. It also
929 shows an additional peak over the *HBB* promoter (orange arrow). This peak is likely a non-
930 specific signal resulting from promiscuous MNase activity used in the CUT&RUN technique.
931 In all samples but HOXD13 in the II1 T-DOM allele, experiments were performed in duplicate.
932 One replicate is plotted as a solid color and the other is shown as a superimposed black line. **B**.
933 On top is a schematic of the II1 enhancer element with the four HOX13 motif indicated as pink
934 bars. The pink bar with an asterisk indicates the motif position that is not in the HOX13 and
935 ATAC peak. At the bottom are the three HOX13 motifs identified by HOMER motif discovery
936 in the CUT&RUN experiments here and the E11.5 whole forelimb ChIP-seq (Sheth et al.,
937 2016).
938
939 **Figure 4. Deletions of HOX13 binding sites. A**. On top are shown the HOX13 motifs in the
940 II1 element, as extracted from the transcription factor binding (Figure 3B), with their positions
941 indicated below (light blue boxes with orientations). The dark blue boxes below the DNA
942 sequence indicate the positions and orientations of the CRISPR guides. Combinations of guides
943 were used to generate small deletions (g2 and g3 or g2 and g6). The yellow crosses indicate
944 the approximate cutting position of the Cas9. **B**. E12.5 F0 embryos stained for *LacZ*. The left
945 panel is a II1 T-DOM limb showing the staining pattern with the recombined II1 enhancer (no
946 CRISPR cutting), with weak staining in the distal limb, particularly in the digit mesenchyme.
947 The two central panels are representative embryos (n indicated in upper left corners) with the
948 indicated deletions (see Figure S4-1 for images of all embryos including these two #5083 and
949 #5146). Embryos carrying either the g2+g3 or g2+g6 deletions lost all staining in distal limb
950 cells. Several embryos showed strong distal limb staining (as shown in the right panel, see also

951 Figure S4-1B). They contained a large deletion that extends from the native II1 element in C-
952 DOM to the II1 transgenic element in T-DOM, due to the guide RNA sequences present at both
953 sites (scheme in Figure S4-1C). Deletions in all embryos were sequenced (Table S2). Embryos
954 with ambiguous sequencing results or mosaicism were not used. **C**. Sequence map of part of
955 the II1 enhancer elements used to generate randomly integrated transgenic embryos. The top
956 track is the wild type sequence around the three HOX13 binding sites for II1 used in control
957 embryos (II1 TgN ctrl). The Del 2x13 sequence lacks the two centromeric HOX13 binding
958 sites while the Del 3x13 sequence lacks all three HOX13 binding. **D**. On top are *LacZ* stained
959 E12.5 transgenic embryos and the table on the bottom reports the number of embryos that were
960 positive for the transgene by PCR (PCR+) followed but the number of embryos that stained in
961 the distal forelimb (DFL) as well as the embryos with no staining in the DFL (No DFL Stain).
962 The p-values are determined by Fisher's exact test. This indicates that the 2x13 and 3x13
963 deletions are likely to be responsible for the loss of distal limb staining in these embryos
964 compared with the II1 control embryos. The embryo on the left is a II1 control embryo
965 containing the wild type II1 sequence. The embryos in the center and right are representatives
966 of the staining (if any) obtained with the 2x13 and 3x13 constructs, with an absence of distal
967 limb staining. Pictures of all embryos generated in this experiment containing some *LacZ*
968 staining are in Figure S4-2, including the three shown here.
969
970 **Figure 5. Inhibition of distal enhancer activity by the T-DOM chromatin environment**.
971 **A.** Schematic of T-DOM with the *Hoxd* gene cluster on the left (purple boxes), *Hoxd1* with a
972 square and *Hoxd13* with a circle on top. The *Mtx2* gene is next to *Hoxd1* and *Hnrnpa3* gene is
973 the small white box with black border on the right of T-DOM. The position of the II1 transgene
974 insertion into T-DOM is indicated by a green rectangle with black border (mm10
975 chr2:75269597-75269616). Orange rectangles are known proximal limb enhancers. The two
976 regions deleted by CRISPR are indicated above the genomic map (Del *Mtx2*-II1-T-DOM and
977 Del II1-T-DOM-*Hnrnpa3*). **B**. Effect of deleting the centromeric portion of T-DOM (Del *Mtx2*-
978 II1-T-DOM) on *LacZ* staining, with a light loss of staining in the proximal domain (black
979 arrows) and a gain in the distal domain (see also Figure S5A). **C**. Effect of deleting the
980 telomeric portion of T-DOM (Del II1-T-DOM-*Hnrnpa3*) on *LacZ* staining, with an almost
981 complete loss of staining in the proximal domain (arrows) and no substantial impact on the
982 distal domain (see also Figure S5B). The embryos shown here are also displayed in Figure S5
983 to show the complete series of stained embryos.
984
985 **Figure 6. The II1 enhancer in T-DOM contacts the *Hoxd* gene cluster. A.** Capture Hi-C
986 maps at 5kb bin resolution over the entire *HoxD* locus (mm10 chr2:73950000-75655000),

987 displayed as the subtraction of wild type signal (blue) from the II1 T-DOM 542 allele signal
988 (red). The II1 enhancer T-DOM insertion site (arrow on the green rectangle at the bottom)
989 produces increased contacts with the *Hoxd* gene cluster (black arrow pointing to the red bins).
990 The contacts are established in both proximal forelimb (PFL) and distal forelimb (DFL) cells.
991 **B.** Contacts between the *Hoxd* gene cluster (x-axis, genes are indicated below) and the region
992 covering the II1 T-DOM reporter transgene (y-axis). The II1 T-DOM construct is schematized
993 on the y-axis for clarity, with a solid black line indicating the boundary between bins. The II1
994 enhancer is shown in green, the *HBB* promoter is black, and the *LacZ* gene is in grey. The
995 panels below are the H3K27ac, H3K27me3 (from (Yakushiji-Kaminatsui et al., 2018) and
996 bound CTCF (from (Rodriguez-Carballo et al., 2017) from wild type PFL cells, aligned with
997 the interaction matrix above. The strongest contacts are between the II1 T-DOM reporter
998 transgene and the region around *Hoxd8* and *Hoxd10*, matching genes transcribed in proximal
999 limb cells, indicated by a grey dashed box. **C.** Same as in B but using distal forelimb cells
1000 (DFL). The II1 T-DOM insert establishes more diffuse contacts, extending from *Hoxd1* to
1001 *Hoxd11* and stopping abruptly before those genes highly expressed in distal cells (*Hoxd12*,
1002 *Hoxd13*).
1003
1004

1005 **LEGENDS TO SUPPLEMENTARY FIGURES**

1006 **Supplementary Figure S1. Deletion of the II1 enhancer does not alter *Hoxd* gene**
1007 **expression in the distal limb bud cells**. **A**. Magnified representation of the Island II1 enhancer
1008 element (mm10 chr2:74072912-74077028, (Montavon et al., 2011). HOXA13 (top) and
1009 HOXD13 (bottom) E12.5 CUT&RUN profiles are indicated by light or dark blue, whereas
1010 HOXA13 and HOXD13 E11.5 forelimb ChIP-seq profiles (Sheth et al., 2016) are shown as
1011 dark grey lines super-imposed for comparison. The DNA sequence conservation track MultiZ
1012 from UCSC (mm10) is shown below, with the most conserved region indicated with a bracket.
1013 **B**. RTqPCR for multiple *Hoxd* genes using distal forelimb bud cells from E12.5 wild type and
1014 homozygous Del II1 mutant specimen. P-value is indicated above, no significant difference in
1015 expression was detected for any *Hoxd* genes in the absence of the II1 enhancer (two-tailed
1016 unequal variance t-test). **C**. Whole-mount *in situ* hybridizations for *Hoxd13* and *Hoxd11* in
1017 embryos homozygous for the II1 enhancer deletion also show no change in their expression
1018 domains when this enhancer is deleted from C-DOM.

1019 **Supplementary Figure S2**. **Long-read sequencing of the genomic structure after targeted**
1020 **insertion. A**. Map of the T-DOM after homologous recombination of the II1:*HBB:LacZ*
1021 enhancer-reporter transgene, indicating the location of CRISPR guides used in the nCATS

1022 protocol for enrichment of sequencing reads (Gilpatrick et al., 2020). The red crosses indicate
1023 the location of the CRISPR cutting guide (see Supplementary Table S1). Two guides were used
1024 outside the II1 transgene (SCS12 and 13) and one guide within the II1 transgene (SCS14).
1025 Below the map of the CRISPR cutting position is a map of the transgene construct. The colors
1026 indicated for different portions of the construct match to the sequencing alignments below. **B**.
1027 Dotplot maps of sequencing reads recovered from the II1 T-DOM 542 allele and showing a
1028 clear one-copy recombination at the expected site. The x-axis is the position along the mutant
1029 II1 T-DOM 542 chromosome and the values on the y-axis represent the base pair position of
1030 the transgene. Each circle represents a 20bp alignment (see methods) so multiple adjacent 20bp
1031 matching reads appear as a line. The best matching read is drawn in black, while shorter reads
1032 are drawn in grey. **C**. Dotplot maps of five sequencing reads recovered from the II1 T-DOM
1033 320 allele, showing insertion of multiple copies.

1034 **Supplementary Figure S4-1**. **Photos of all embryos containing the deletions of the HOX13**
1035 **binding sites present in the II1 enhancer-reporter construct targeted into T-DOM**.
1036 Individual embryos are numbered. Embryos #5083 and #5146 are already shown in Figure 4B,
1037 but are reproduced here for easier comparison. The g2 + g3 and g2 + g6 deletions removed
1038 only binding sites within the II1 enhancer. **B**. Staining of embryos containing the Del C-T
1039 deletion. These embryos are positive in both proximal and distal limb bud cells. Embryo #5150
1040 is already shown in Figure 4B but is added here for comparison. **C**. Schematic of the Del C-T
1041 deletion. The Del C-T created a large deletion fusing parts of T-DOM and C-DOM due to the
1042 presence of the sequence targeted by the guides RNAs on both native (C-DOM) and transgenic
1043 (T-DOM). After deletion, the II1 *LacZ* reporter transgene is flanked by a centromeric distal
1044 limb enhancer (green) and several telomeric proximal enhancers (CS65, PLEs, orange), thus
1045 accounting for its expression in both limb domains.

1046 **Supplementary Figure S4-2**. **Photos of randomly integrated transgenic embryos stained**
1047 **for *LacZ* expression when HOX13 binding sites were deleted from the II1 enhancer**. The
1048 left two columns (II1 TgN ctrl) are control embryos containing the entire II1 enhancer
1049 sequence. The two columns in the center (Del 2x13) show stained embryos containing the II1
1050 enhancer element lacking the two centromeric HOX13 binding sites (see Figure 4A, C). The
1051 two columns in the right (Del 3x13) show stained embryos lacking all three HOX13 binding
1052 sites. The orange arrows indicated the location of *LacZ* staining when it is not detected in the
1053 distal limb buds. The unique embryo ID is shown at the bottom of each picture. The three
1054 embryos # cb09 04, cb17 05 and cb15 01 are those also displayed in Figure 4D. There are
1055 reproduced here for easier comparison. The two outlier embryos showing staining in distal limb
1056 cells are shown with a yellow asterisk.

1057 **Supplementary Figure S4-3. *LacZ* staining pattern in control transgenes inserted into T-**
1058 **DOM. A**. The *LacZ* staining in a control embryo containing the Del 3x13 variant of the II1
1059 enhancer. This variant contains the same targeting construct as the II1 T-DOM 542 allele, but
1060 the II1 enhancer has been replaced with the same variant used to produce the randomly
1061 integrated Del 3x13 (Figure 4D). The staining is very strong in the proximal limb and
1062 completely absent in the distal limb. This corroborates the observation from Figure 4, that the
1063 three HOX13 binding sites in II1 are necessary for the distal limb staining either as a randomly
1064 integrated transgene or a targeted insertion transgene in the T-DOM. **B**. The *LacZ* staining in
1065 a control embryo that does not contain the II1 enhancer element. This transgene contains the
1066 same targeting construct as the II1 T-DOM 542, but without the enhancer. The staining is very
1067 strong in the proximal limb and completely absent in the distal limb.

1068 **Supplementary Figure S5**. **Photos of all embryos carrying large deletions flanking the**
1069 **insertion of the of II1 enhancer-reporter construct within T-DOM** (see Figure 5B, C). All
1070 embryos were genotyped for the expected deletion. All embryos with the expected deletions
1071 are represented here unless they produce ambiguous PCR results or were mosaic for the
1072 deletion. The embryos #4470, 5023, 4492, 4491 from Figure 5 are reproduced here for easier
1073 comparison.

1074 **Supplementary Figure S6**. **Capture Hi-C maps**. Capture Hi-C maps covering the entire
1075 *HoxD* locus for wild type (wt) and the II1 enhancer recombined within T-DOM (II1 T-DOM),
1076 using both proximal (PFL) and distal (DFL) forelimb cells samples (mm10 chr2:73950000-
1077 75655000). The new contacts formed between the II1 enhancer-promoter sequence and the
1078 *Hoxd* gene cluster are indicated by the black arrow. The contacts were scored in both PFL and
1079 DFL cells, yet with a difference in resolution, being more diffuse in the DFL cells (see Figure
1080 6). The black arrows below points to the location of the II1 enhancer within C-DOM and the
1081 transgene integration site within T-DOM. The *HoxD* cluster position is indicated at the center.
1082 The *Hoxd13* gene is indicated by a purple pin with a circle and the *Hoxd1* is indicated with a
1083 square.

1084

1085 **REFERENCES**

1086 Acemel, R.D., Tena, J.J., Irastorza-Azcarate, I., Marletaz, F., Gomez-Marin, C., de la Calle-
1087 Mustienes, E., Bertrand, S., Diaz, S.G., Aldea, D., Aury, J.M., et al. (2016). A single three-
1088 dimensional chromatin compartment in amphioxus indicates a stepwise evolution of
1089 vertebrate Hox bimodal regulation. Nat Genet *48*, 336–341.

1090 Amândio, A.R., Lopez-Delisle, L., Bolt, C.C., Mascrez, B., and Duboule, D. (2020). A
1091 complex regulatory landscape involved in the development of mammalian external genitals.
1092 ELife *9*, e52962.
1093 Andrey, G., and Spielmann, M. (2017). CRISPR/Cas9 Genome Editing in Embryonic Stem
1094 Cells. Methods Mol. Biol. Clifton NJ *1468*, 221–234.
1095 Andrey, G., Montavon, T., Mascrez, B., Gonzalez, F., Noordermeer, D., Leleu, M., Trono,
1096 D., Spitz, F., and Duboule, D. (2013). A switch between topological domains underlies HoxD
1097 genes collinearity in mouse limbs. Science *340*, 1234167.
1098 Artus, J., and Hadjantonakis, A.-K. (2011). Generation of Chimeras by Aggregation of
1099 Embryonic Stem Cells with Diploid or Tetraploid Mouse Embryos. In Transgenic Mouse
1100 Methods and Protocols, M.H. Hofker, and J. van Deursen, eds. (Totowa, NJ: Humana Press),
1101 pp. 37–56.
1102 Banerji, J., Rusconi, S., and Schaffner, W. (1981). Expression of a beta-globin gene is
1103 enhanced by remote SV40 DNA sequences. Cell *27*, 299–308.
1104 Basu, S., Mackowiak, S.D., Niskanen, H., Knezevic, D., Asimi, V., Grosswendt, S.,
1105 Geertsema, H., Ali, S., Jerković, I., Ewers, H., et al. (2020). Unblending of Transcriptional
1106 Condensates in Human Repeat Expansion Disease. Cell *181*, 1062-1079.e30.
1107 Beccari, L., Yakushiji-Kaminatsui, N., Woltering, J.M., Necsulea, A., Lonfat, N., Rodriguez-
1108 Carballo, E., Mascrez, B., Yamamoto, S., Kuroiwa, A., and Duboule, D. (2016). A role for
1109 HOX13 proteins in the regulatory switch between TADs at the HoxD locus. Genes Dev *30*,
1110 1172–1186.
1111 Beccari, L., Jaquier, G., Lopez-Delisle, L., Rodriguez-Carballo, E., Mascrez, B., Gitto, S.,
1112 Woltering, J., and Duboule, D. (2021). Dbx2 regulation in limbs suggests interTAD sharing
1113 of enhancers. Dev. Dyn. Off. Publ. Am. Assoc. Anat. *250*, 1280–1299.
1114 Berlivet, S., Paquette, D., Dumouchel, A., Langlais, D., Dostie, J., and Kmita, M. (2013).
1115 Clustering of tissue-specific sub-TADs accompanies the regulation of HoxA genes in
1116 developing limbs. PLoS Genet *9*, e1004018.
1117 Bernstein, B.E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., Bailey, D.K., Huebert, D.J.,
1118 McMahon, S., Karlsson, E.K., Kulbokas, E.J., Gingeras, T.R., et al. (2005). Genomic maps
1119 and comparative analysis of histone modifications in human and mouse. Cell *120*, 169–181.
1120 Bolt, C.C., and Duboule, D. (2020). The regulatory landscapes of developmental genes.
1121 Development *147*, dev171736.
1122 Bolt, C.C., Lopez-Delisle, L., Mascrez, B., and Duboule, D. (2021). Mesomelic dysplasias
1123 associated with the HOXD locus are caused by regulatory reallocations. Nat. Commun. *12*,
1124 5013.
1125 Bruneau, S., Johnson, K.R., Yamamoto, M., Kuroiwa, A., and Duboule, D. (2001). The

1126 mouse Hoxd13(spdh) mutation, a polyalanine expansion similar to human type II
1127 synpolydactyly (SPD), disrupts the function but not the expression of other Hoxd genes. Dev.
1128 Biol. *237*, 345–353.
1129 Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013).
1130 Transposition of native chromatin for fast and sensitive epigenomic profiling of open
1131 chromatin, DNA-binding proteins and nucleosome position. Nat. Methods *10*, 1213–1218.
1132 Carroll, S.B., Prud'homme, B., and Gompel, N. (2008). Regulating evolution. Sci Am *298*,
1133 60–67.
1134 Charif, D., and Lobry, J.R. (2007). SeqinR 1.0-2: A Contributed Package to the R Project for
1135 Statistical Computing Devoted to Biological Sequences Retrieval and Analysis. In Structural
1136 Approaches to Sequence Evolution: Molecules, Networks, Populations., U. Bastolla, M.
1137 Porto, H.E. Roman, and M. Vendruscolo, eds. (New York, USA: Springer Verlag), pp. 207–
1138 232.
1139 Cheutin, T., and Cavalli, G. (2019). The multiscale effects of polycomb mechanisms on 3D
1140 chromatin folding. Crit. Rev. Biochem. Mol. Biol. *54*, 399–417.
1141 Davis, A.P., Witte, D.P., Hsieh-Li, H.M., Potter, S.S., and Capecchi, M.R. (1995). Absence
1142 of radius and ulna in mice lacking hoxa-11 and hoxd-11. Nature *375*, 791–795.
1143 Desanlis, I., Kherdjemil, Y., Mayran, A., Bouklouch, Y., Gentile, C., Sheth, R., Zeller, R.,
1144 Drouin, J., and Kmita, M. (2020). HOX13-dependent chromatin accessibility underlies the
1145 transition towards the digit development program. Nat. Commun. *11*, 2491.
1146 Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B.
1147 (2012). Topological domains in mammalian genomes identified by analysis of chromatin
1148 interactions. Nature *485*, 376–380.
1149 Dolle, P., Dierich, A., LeMeur, M., Schimmang, T., Schuhbaur, B., Chambon, P., and
1150 Duboule, D. (1993). Disruption of the Hoxd-13 gene induces localized heterochrony leading
1151 to mice with neotenic limbs. Cell *75*, 431–441.
1152 Fromental-Ramain, C., Warot, X., Messadecq, N., LeMeur, M., Dolle, P., and Chambon, P.
1153 (1996). Hoxa-13 and Hoxd-13 play a crucial role in the patterning of the limb autopod.
1154 Development *122*, 2997–3011.
1155 Gentile, C., Berlivet, S., Mayran, A., Paquette, D., Guerard-Millet, F., Bajon, E., Dostie, J.,
1156 and Kmita, M. (2019). PRC2-Associated Chromatin Contacts in the Developing Limb Reveal
1157 a Possible Mechanism for the Atypical Role of PRC2 in HoxA Gene Expression. Dev. Cell
1158 *50*, 184-196.e4.
1159 George, S.H.L., Gertsenstein, M., Vintersten, K., Korets-Smith, E., Murphy, J., Stevens,
1160 M.E., Haigh, J.J., and Nagy, A. (2007). Developmental and adult phenotyping directly from
1161 mutant embryonic stem cells. Proc. Natl. Acad. Sci. U. S. A. *104*, 4455–4460.

1162 Gilpatrick, T., Lee, I., Graham, J.E., Raimondeau, E., Bowen, R., Heron, A., Downs, B.,
1163 Sukumar, S., Sedlazeck, F.J., and Timp, W. (2020). Targeted nanopore sequencing with
1164 Cas9-guided adapter ligation. Nat. Biotechnol. *38*, 433–438.
1165 Grosveld, F., van Staalduinen, J., and Stadhouders, R. (2021). Transcriptional Regulation by
1166 (Super)Enhancers: From Discovery to Mechanisms. Annu. Rev. Genomics Hum. Genet. *22*,
1167 127–146.
1168 Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C.,
1169 Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription
1170 factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell
1171 *38*, 576–589.
1172 Karr, J.P., Ferrie, J.J., Tjian, R., and Darzacq, X. (2021). The transcription factor activity
1173 gradient (TAG) model: contemplating a contact-independent mechanism for enhancer-
1174 promoter communication. Genes Dev.
1175 Kraft, K., Geuer, S., Will, A.J., Chan, W.L., Paliou, C., Borschiwer, M., Harabula, I., Wittler,
1176 L., Franke, M., Ibrahim, D.M., et al. (2015). Deletions, Inversions, Duplications: Engineering
1177 of Structural Variants using CRISPR/Cas in Mice. Cell Rep. *10*, 833–839.
1178 Labun, K., Montague, T.G., Krause, M., Torres Cleuren, Y.N., Tjeldnes, H., and Valen, E.
1179 (2019). CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing.
1180 Nucleic Acids Res. *47*, W171–W174.
1181 Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat
1182 Methods *9*, 357–359.
1183 Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-
1184 efficient alignment of short DNA sequences to the human genome. Genome Biol *10*, R25.
1185 Lettice, L.A., Heaney, S.J., Purdie, L.A., Li, L., de Beer, P., Oostra, B.A., Goode, D., Elgar,
1186 G., Hill, R.E., and de Graaff, E. (2003). A long-range Shh enhancer regulates expression in
1187 the developing limb and fin and is associated with preaxial polydactyly. Hum Mol Genet *12*,
1188 1725–1735.
1189 Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association
1190 mapping and population genetical parameter estimation from sequencing data. Bioinformatics
1191 *27*, 2987–2993.
1192 Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics *34*,
1193 3094–3100.
1194 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis,
1195 G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence
1196 Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.
1197 Liu, N., Hargreaves, V.V., Zhu, Q., Kurland, J.V., Hong, J., Kim, W., Sher, F., Macias-

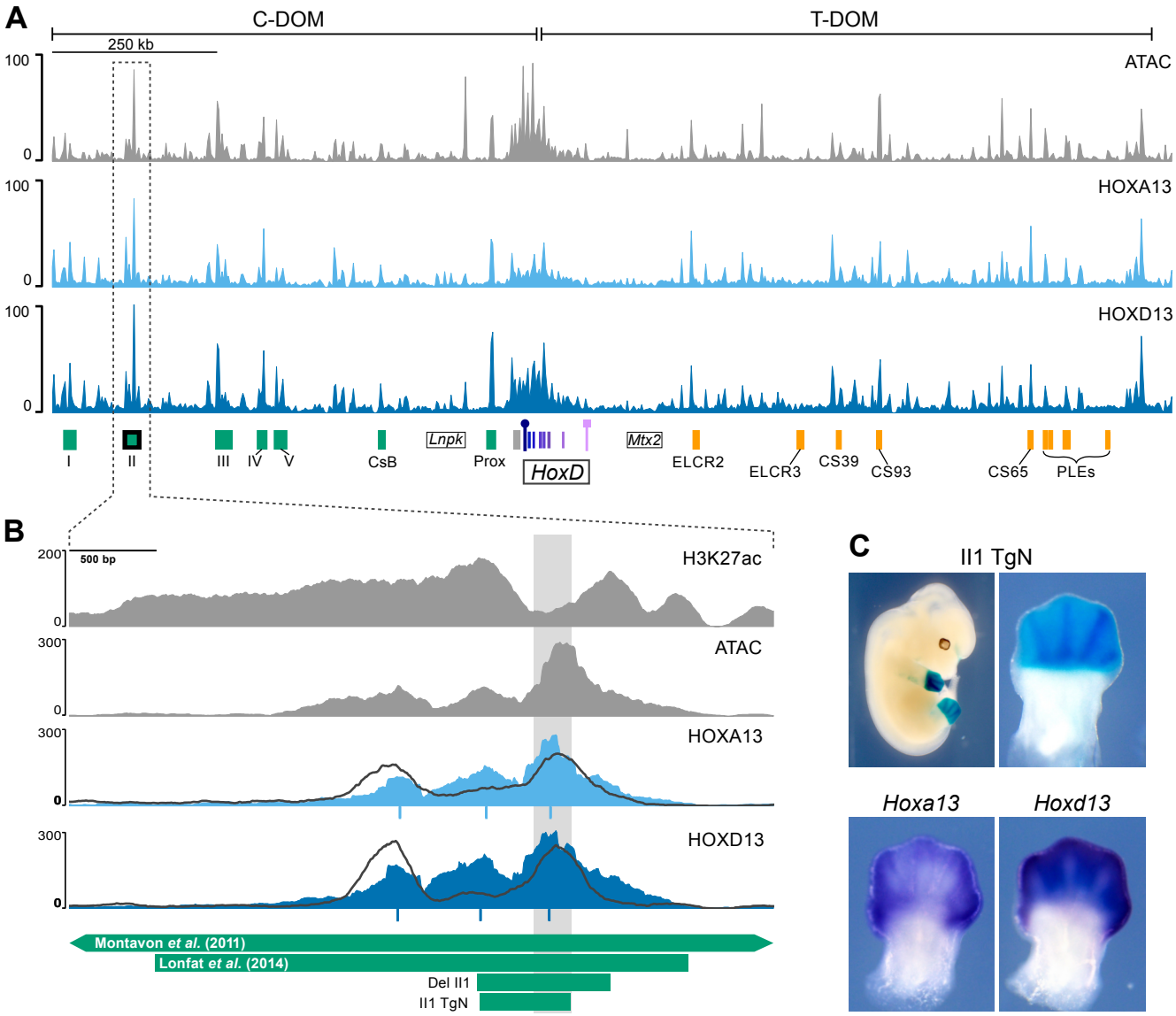
1198 Trevino, C., Rogers, J.M., Kurita, R., et al. (2018). Direct Promoter Repression by BCL11A
1199 Controls the Fetal to Adult Hemoglobin Switch. Cell *173*, 430-442.e17.
1200 Lonfat, N., Montavon, T., Darbellay, F., Gitto, S., and Duboule, D. (2014). Convergent
1201 evolution of complex regulatory landscapes and pleiotropy at Hox loci. Science *346*, 1004–
1202 1006.
1203 Long, H.K., Prescott, S.L., and Wysocka, J. (2016). Ever-Changing Landscapes:
1204 Transcriptional Enhancers in Development and Evolution. Cell *167*, 1170–1187.
1205 Lopez-Delisle, L., Rabbani, L., Wolff, J., Bhardwaj, V., Backofen, R., Grüning, B., Ramírez,
1206 F., and Manke, T. (2021). pyGenomeTracks: reproducible plots for multivariate genomic
1207 datasets. Bioinformatics *37*, 422–423.
1208 Marinic, M., Aktas, T., Ruf, S., and Spitz, F. (2013). An integrated holo-enhancer unit
1209 defines tissue and gene specificity of the Fgf8 regulatory landscape. Dev. Cell *24*, 530–542.
1210 Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing
1211 reads. EMBnet.Journal *17*.
1212 Montavon, T., Soshnikova, N., Mascrez, B., Joye, E., Thevenet, L., Splinter, E., de Laat, W.,
1213 Spitz, F., and Duboule, D. (2011). A regulatory archipelago controls Hox genes transcription
1214 in digits. Cell *147*, 1132–1145.
1215 Noordermeer, D., Leleu, M., Splinter, E., Rougemont, J., De Laat, W., and Duboule, D.
1216 (2011). The dynamic architecture of Hox gene clusters. Science *334*, 222–225.
1217 Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van
1218 Berkum, N.L., Meisig, J., Sedat, J., et al. (2012). Spatial partitioning of the regulatory
1219 landscape of the X-inactivation centre. Nature *485*, 381–385.
1220 Osterwalder, M., Barozzi, I., Tissières, V., Fukuda-Yuzawa, Y., Mannion, B.J., Afzal, S.Y.,
1221 Lee, E.A., Zhu, Y., Plajzer-Frick, I., Pickle, C.S., et al. (2018). Enhancer redundancy
1222 provides phenotypic robustness in mammalian development. Nature *554*, 239–243.
1223 Oudelaar, A.M., Beagrie, R.A., Kassouf, M.T., and Higgs, D.R. (2021). The mouse alpha-
1224 globin cluster: a paradigm for studying genome regulation and organization. Curr. Opin.
1225 Genet. Dev. *67*, 18–24.
1226 Ramirez, F., Ryan, D.P., Gruning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S.,
1227 Dundar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-
1228 sequencing data analysis. Nucleic Acids Res *44*, W160-5.
1229 Ramírez, F., Bhardwaj, V., Arrigoni, L., Lam, K.C., Grüning, B.A., Villaveces, J.,
1230 Habermann, B., Akhtar, A., and Manke, T. (2018). High-resolution TADs reveal DNA
1231 sequences underlying genome organization in flies. Nat. Commun. *9*, 189.
1232 Rodriguez-Carballo, E., Lopez-Delisle, L., Zhan, Y., Fabre, P.J., Beccari, L., El-Idrissi, I.,
1233 Huynh, T.H.N., Ozadam, H., Dekker, J., and Duboule, D. (2017). The HoxD cluster is a

1234 dynamic and resilient TAD boundary controlling the segregation of antagonistic regulatory
1235 landscapes. Genes Dev. *31*, 2264–2281.
1236 Rodríguez-Carballo, E., Lopez-Delisle, L., Willemin, A., Beccari, L., Gitto, S., Mascrez, B.,
1237 and Duboule, D. (2020). Chromatin topology and the timing of enhancer function at the
1238 *HoxD* locus. Proc. Natl. Acad. Sci. *117*, 31231–31241.
1239 Rouco, R., Bompadre, O., Rauseo, A., Fazio, O., Peraldi, R., Thorel, F., and Andrey, G.
1240 (2021). Cell-specific alterations in Pitx1 regulatory landscape activation caused by the loss of
1241 a single enhancer. Nat. Commun. *12*, 7235.
1242 Schaffner, W. (2015). Enhancers, enhancers – from their discovery to today's universe of
1243 transcription enhancers. Biol. Chem. *396*, 311–327.
1244 Schmidl, C., Rendeiro, A.F., Sheffield, N.C., and Bock, C. (2015). ChIPmentation: fast,
1245 robust, low-input ChIP-seq for histones and transcription factors. Nat Methods *12*, 963–965.
1246 Schuettengruber, B., Bourbon, H.-M., Di Croce, L., and Cavalli, G. (2017). Genome
1247 Regulation by Polycomb and Trithorax: 70 Years and Counting. Cell *171*, 34–57.
1248 Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello,
1249 H., Tanay, A., and Cavalli, G. (2012). Three-dimensional folding and functional organization
1250 principles of the Drosophila genome. Cell *148*, 458–472.
1251 Shapiro, M.D., Marks, M.E., Peichel, C.L., Blackman, B.K., Nereng, K.S., Jonsson, B.,
1252 Schluter, D., and Kingsley, D.M. (2004). Genetic and developmental basis of evolutionary
1253 pelvic reduction in threespine sticklebacks. Nature *428*, 717–723.
1254 Sheth, R., Barozzi, I., Langlais, D., Osterwalder, M., Nemec, S., Carlson, H.L., Stadler, H.S.,
1255 Visel, A., Drouin, J., and Kmita, M. (2016). Distal Limb Patterning Requires Modulation of
1256 cis-Regulatory Activities by HOX13. Cell Rep *17*, 2913–2926.
1257 Skene, P.J., Henikoff, J.G., and Henikoff, S. (2018). Targeted in situ genome-wide profiling
1258 with high efficiency for low cell numbers. Nat. Protoc. *13*, 1006–1019.
1259 Spitz, F., Gonzalez, F., and Duboule, D. (2003). A global control region defines a
1260 chromosomal regulatory landscape containing the HoxD cluster. Cell *113*, 405–417.
1261 Tarchini, B., and Duboule, D. (2006). Control of Hoxd genes' collinearity during early limb
1262 development. Dev Cell *10*, 93–103.
1263 Vieux-Rochas, M., Fabre, P.J., Leleu, M., Duboule, D., and Noordermeer, D. (2015).
1264 Clustering of mammalian Hox genes with other H3K27me3 targets within an active nuclear
1265 domain. Proc Natl Acad Sci U A *112*, 4672–4677.
1266 Wolff, J., Rabbani, L., Gilsbach, R., Richard, G., Manke, T., Backofen, R., and Grüning,
1267 B.A. (2020). Galaxy HiCExplorer 3: a web server for reproducible Hi-C, capture Hi-C and
1268 single-cell Hi-C data analysis, quality control and visualization. Nucleic Acids Res. *48*,
1269 W177–W184.

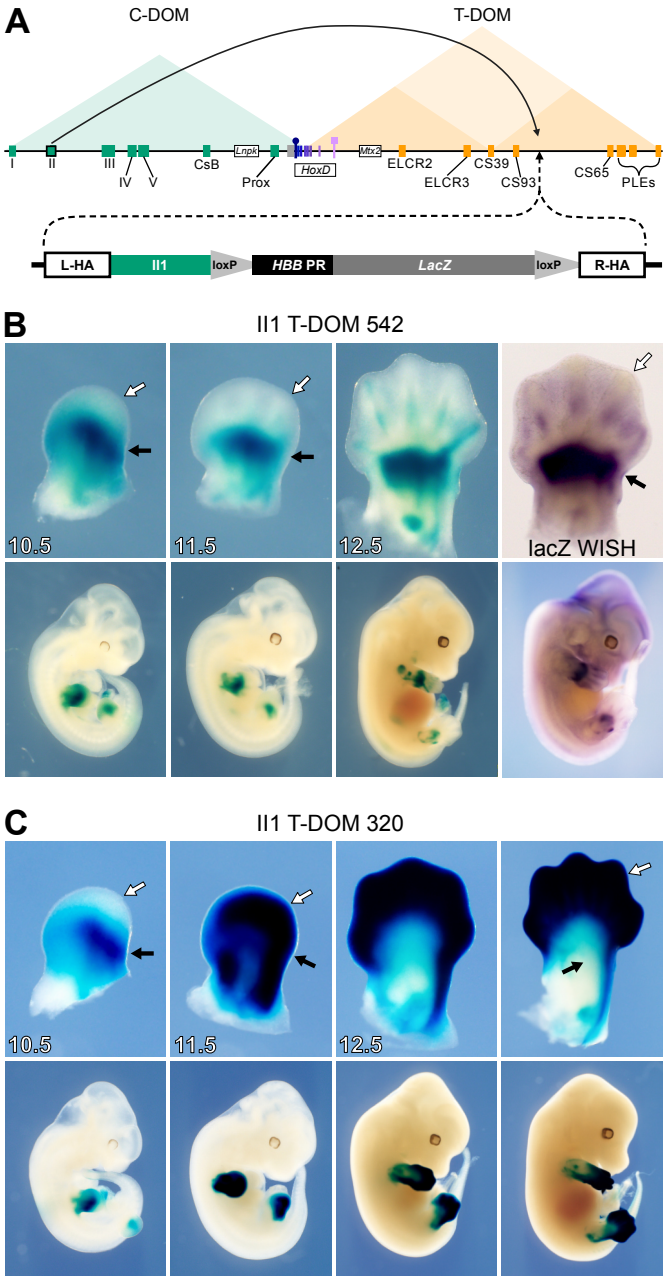
1270 Wolff, P., Weinhofer, I., Seguin, J., Roszak, P., Beisel, C., Donoghue, M.T., Spillane, C.,
1271 Nordborg, M., Rehmsmeier, M., and Kohler, C. (2011). High-resolution analysis of parent-
1272 of-origin allelic expression in the Arabidopsis Endosperm. PLoS Genet *7*, e1002126.
1273 Woltering, J.M., and Duboule, D. (2010). The origin of digits: expression patterns versus
1274 regulatory mechanisms. Dev Cell *18*, 526–532.
1275 Woltering, J.M., Vonk, F.J., Muller, H., Bardine, N., Tuduce, I.L., de Bakker, M.A.,
1276 Knochel, W., Sirbu, I.O., Durston, A.J., and Richardson, M.K. (2009). Axial patterning in
1277 snakes and caecilians: evidence for an alternative interpretation of the Hox code. Dev Biol
1278 *332*, 82–89.
1279 Woltering, J.M., Noordermeer, D., Leleu, M., and Duboule, D. (2014). Conservation and
1280 divergence of regulatory strategies at Hox Loci and the origin of tetrapod digits. PLoS Biol
1281 *12*, e1001773.
1282 Yakushiji-Kaminatsui, N., Lopez-Delisle, L., Bolt, C.C., Andrey, G., Beccari, L., and
1283 Duboule, D. (2018). Similarities and differences in the regulation of HoxD genes during
1284 chick and mouse limb development. PLoS Biol. *16*, e3000004.
1285 Zakany, J., and Duboule, D. (2007). The role of Hox genes during vertebrate limb
1286 development. Curr. Opin. Genet. Dev. *17*, 359–366.
1287 Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C.,
1288 Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS).
1289 Genome Biol *9*, R137.

1290

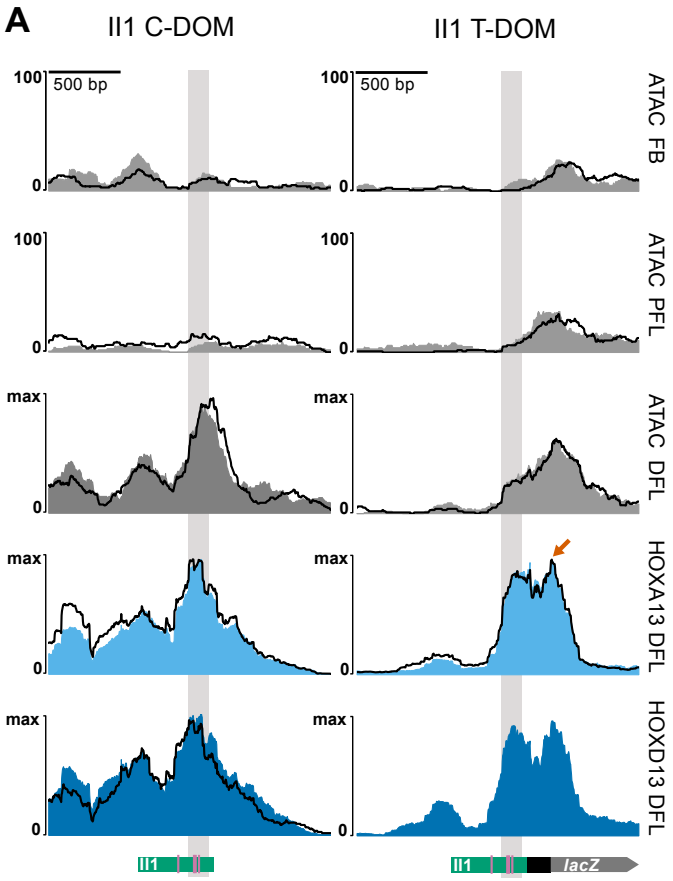
1291

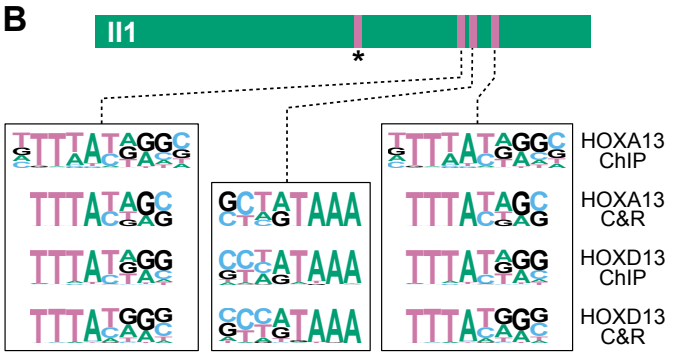


Bolt 2021, II1TDOM Figure 1

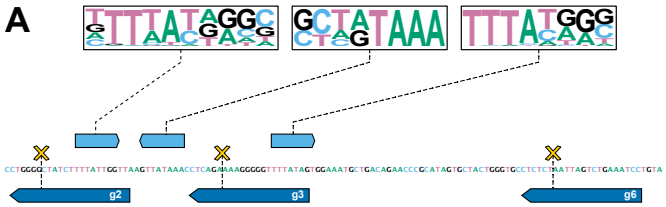


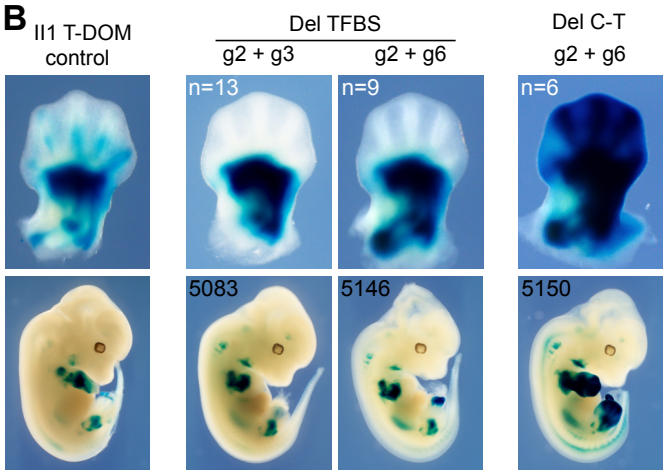
Bolt 2021, II1TDOM Figure 2

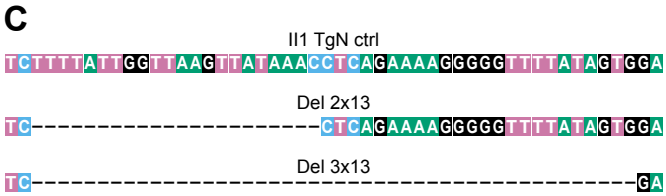


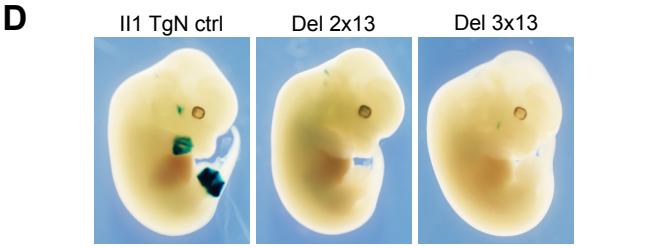


Bolt 2021, II1TDOM Figure 3



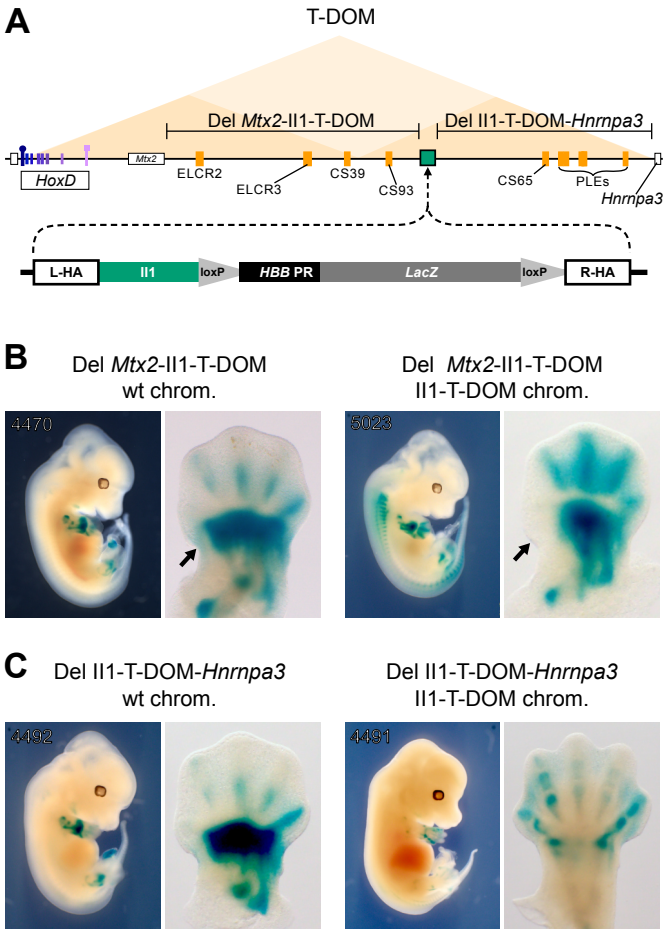




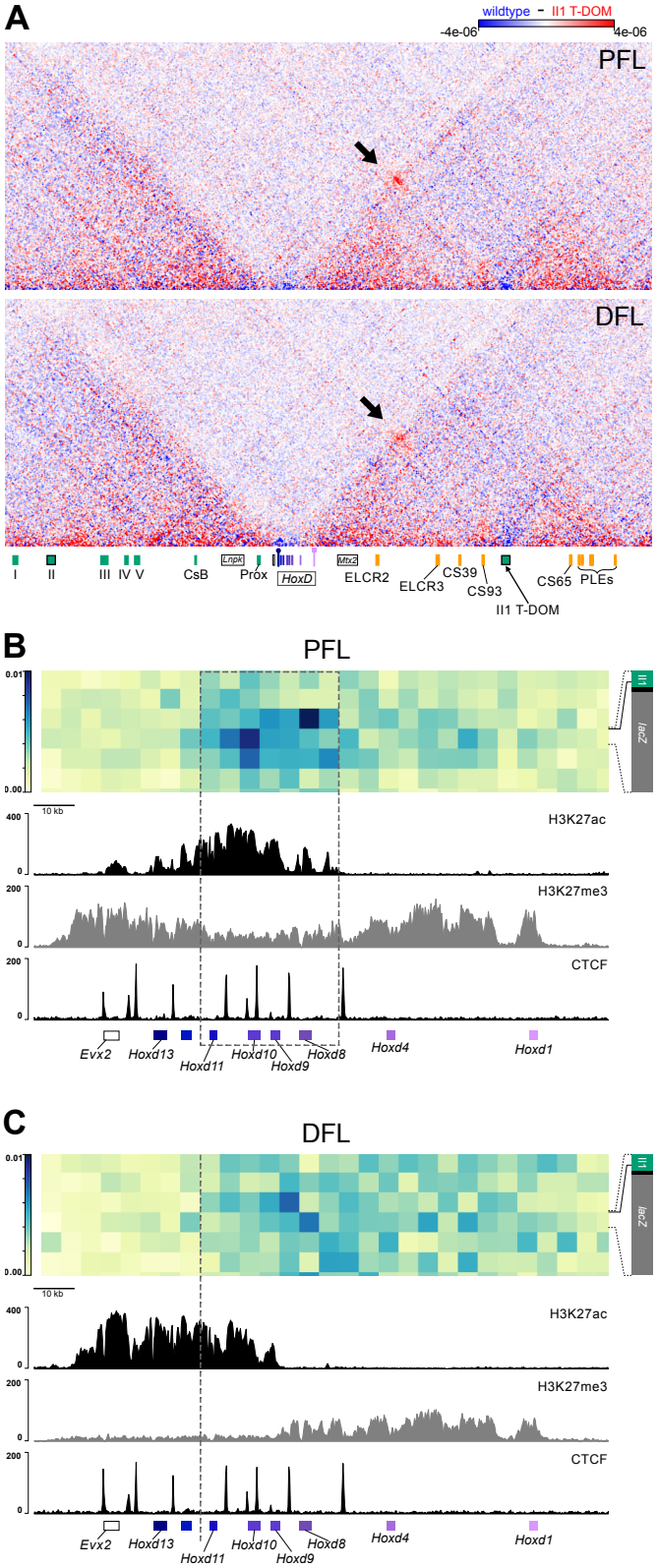


Construct	PCR +	DFL (Stained)	No DFL Stain	p-value
Il1 TgN ctrl	12	7	5	0.00029
Del 2x13	28	1	27	0.00252
Del 3x13	18	1	17	

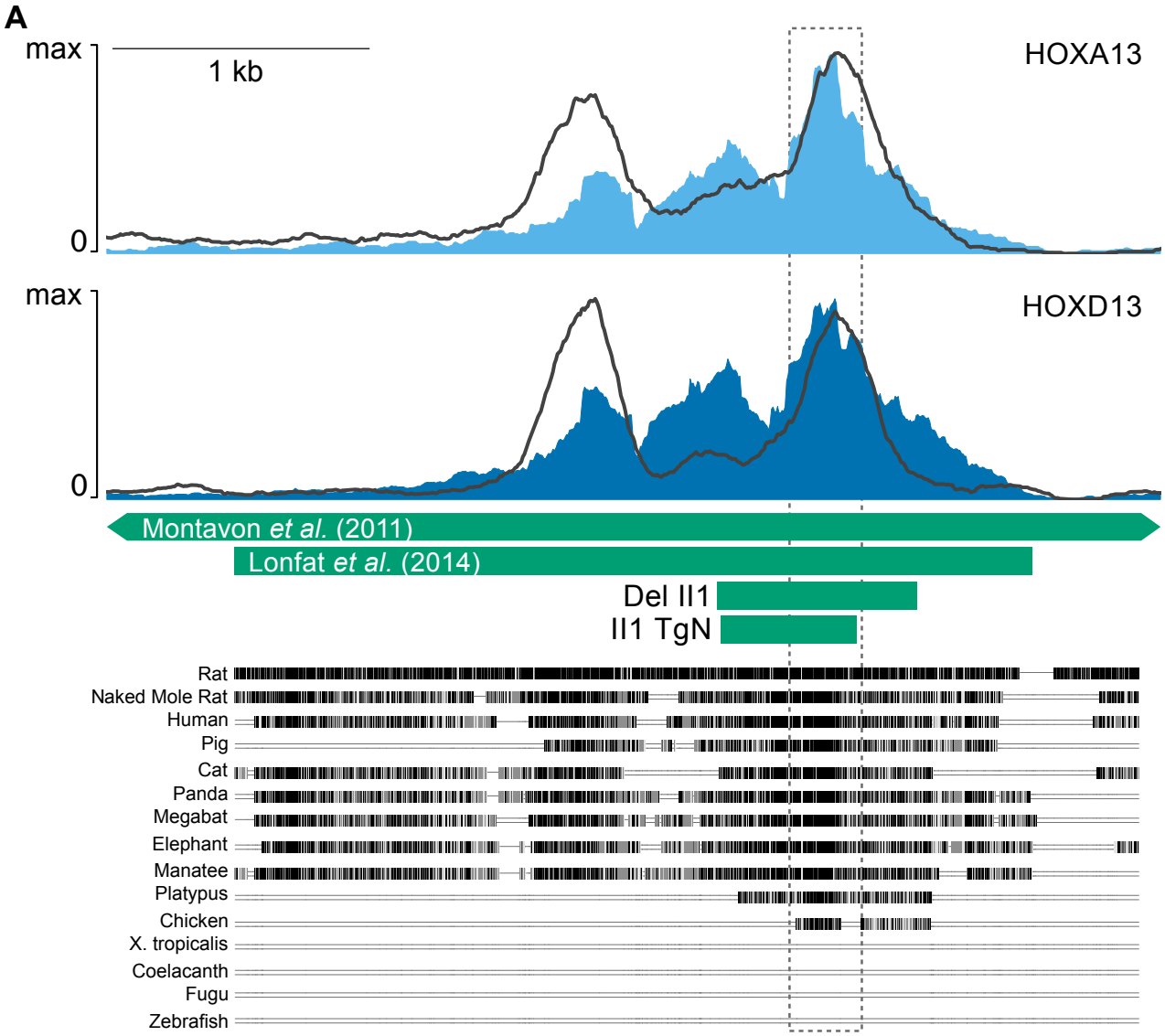
Bolt 2021, Il1TDOM Figure 4

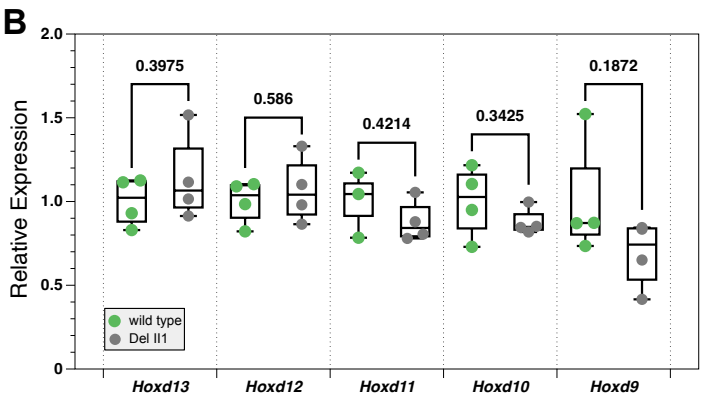


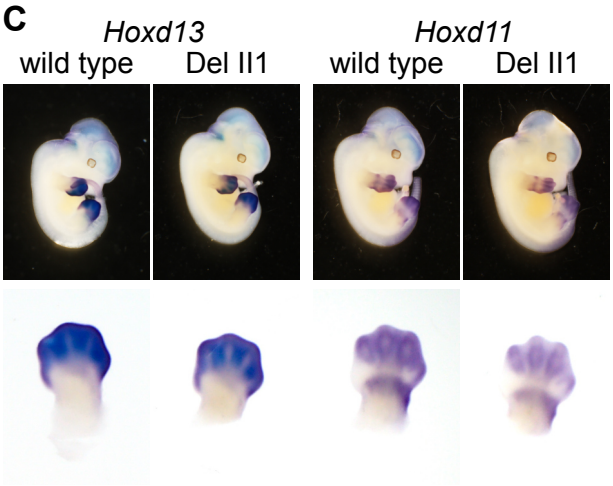
Bolt 2021, II1TDOM Figure 5



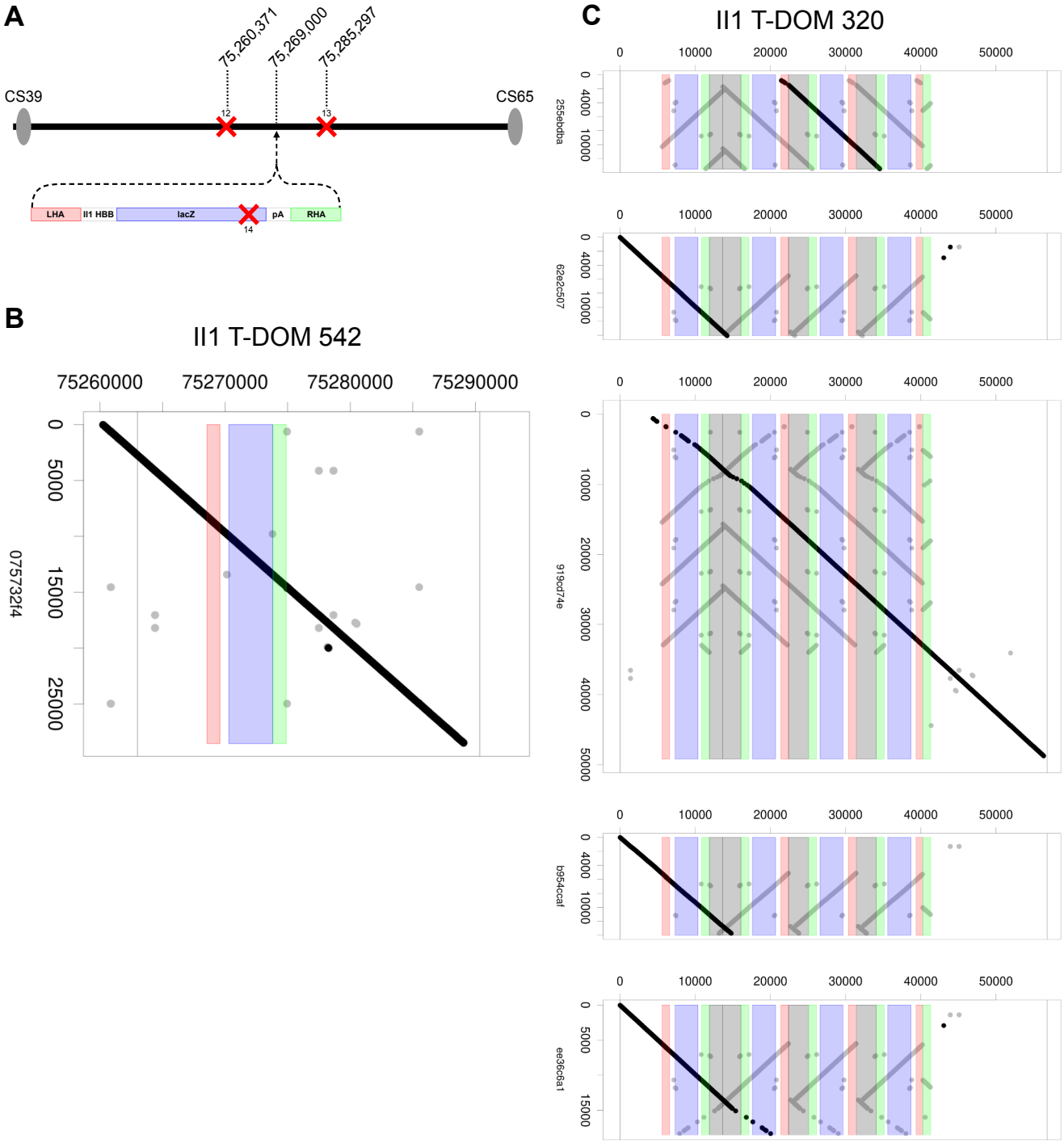
Bolt 2021, II1TDOM Figure 6



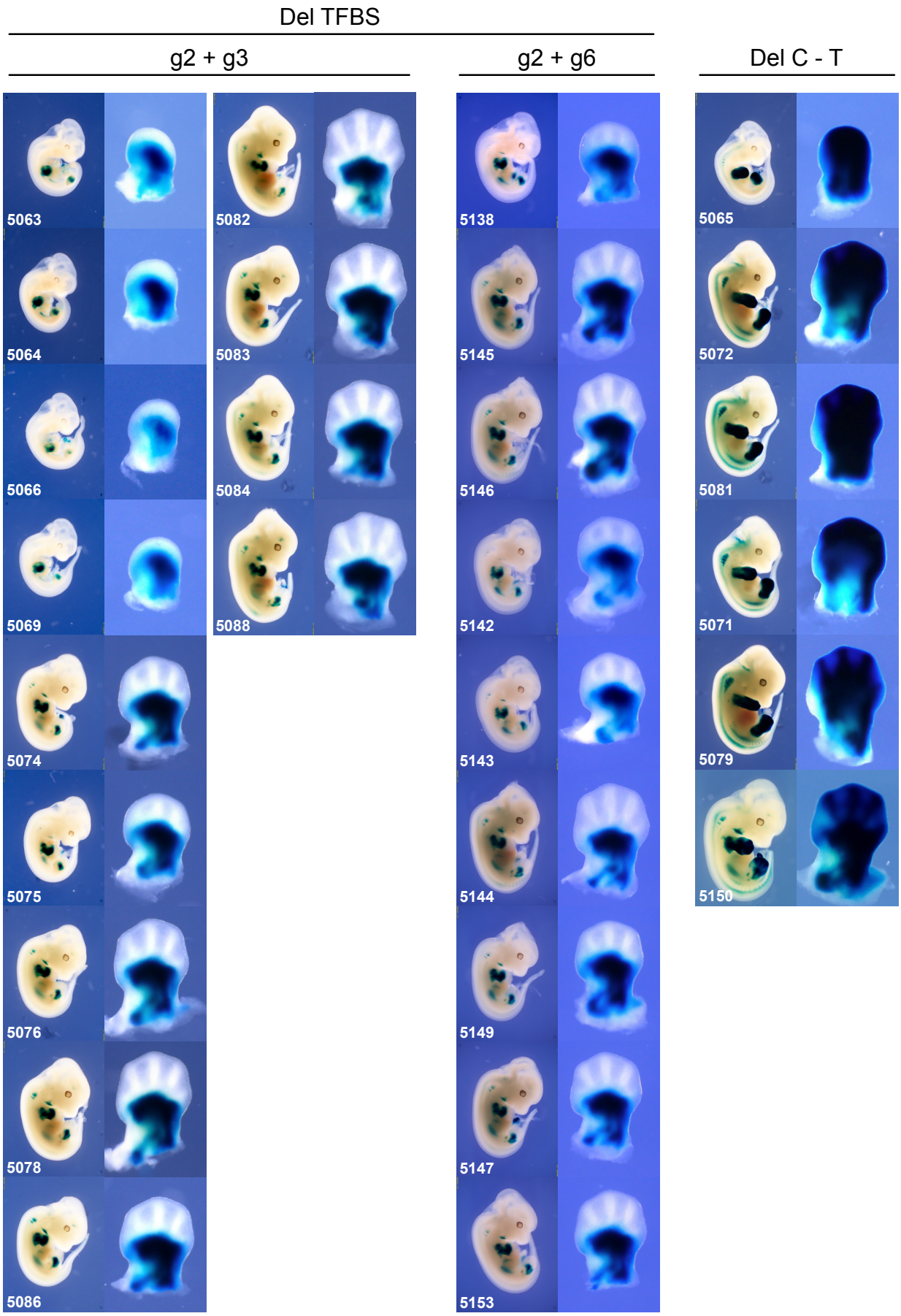




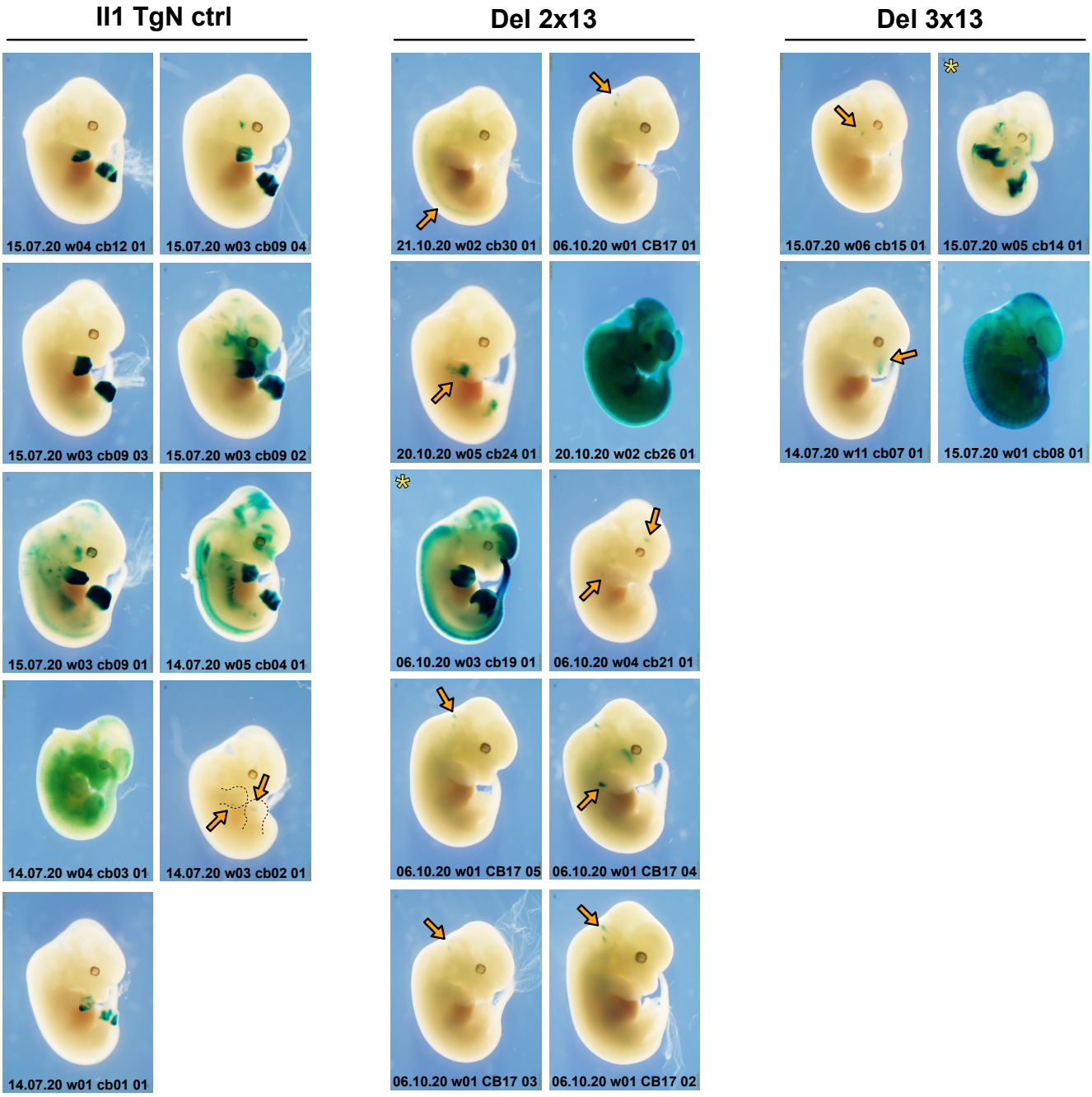
Bolt 2021, II1TDOM Figure S1



Bolt 2021, II1TDOM Figure S2

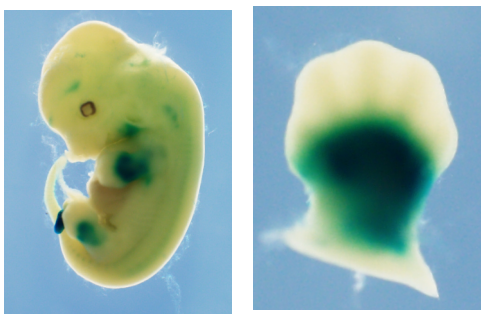


Bolt 2021, Il1TDOM Figure S4-1

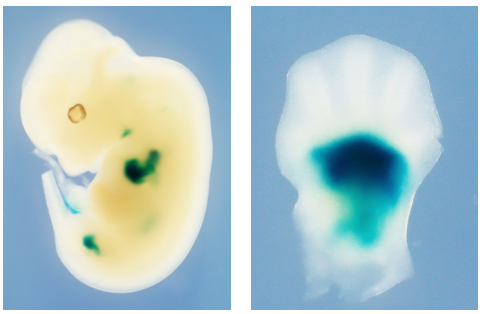


Bolt 2021, Il1TDOM Figure S4-2

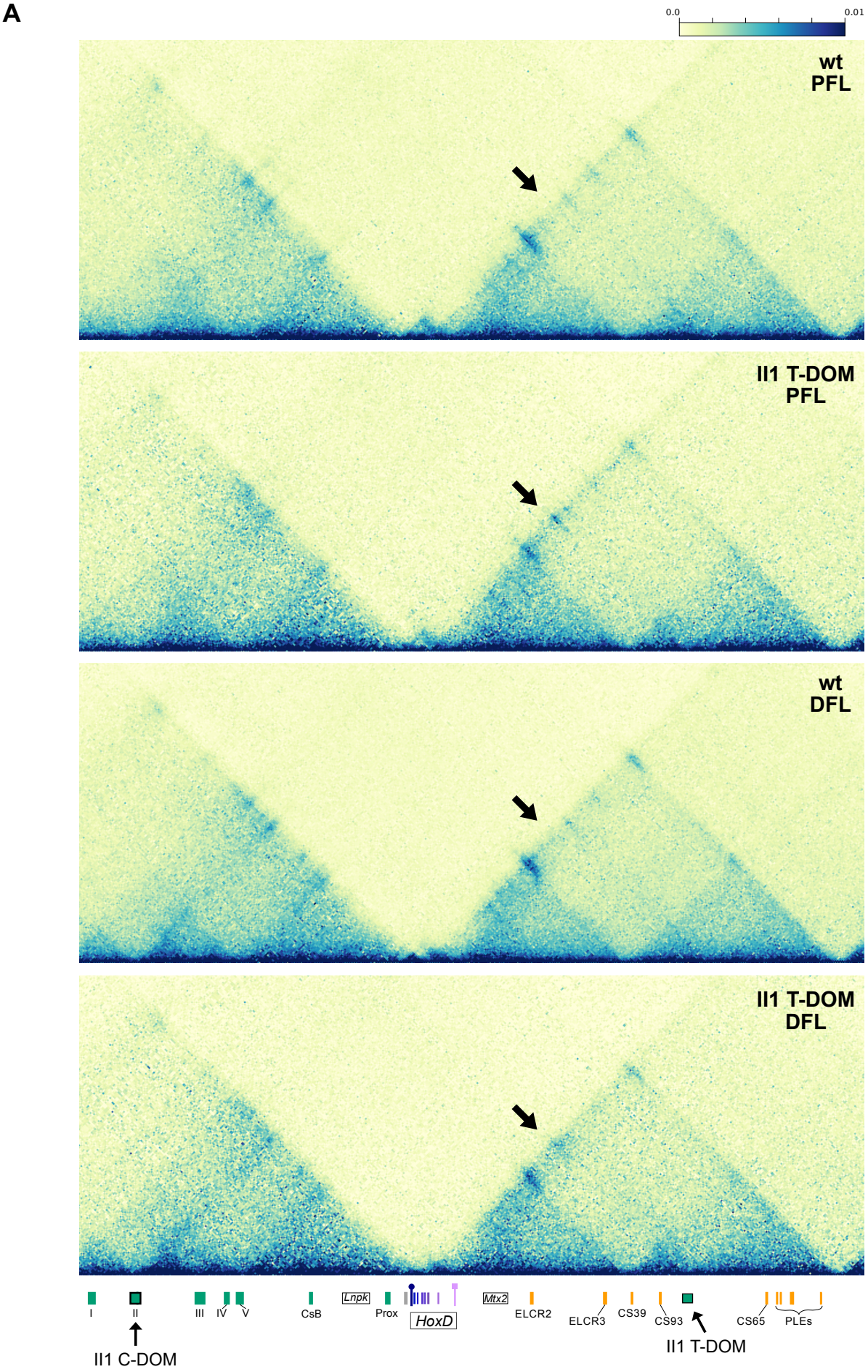
A Del 3x13 T-DOM



B *HBB:lacZ* T-DOM



Bolt 2021, Il1TDOM Figure S4-3



Bolt 2021, II1TDOM Figure S6