1 CONSERVATION AND VARIATIONS OF BIMODAL *HoxD* GENE REGULATION

2 DURING TETRAPOD LIMB DEVELOPMENT

3

4 Nayuta Yakushiji-Kaminatsui[1], Lucille Lopez-Delisle[1,ᵌ], Christopher Chase Bolt[1,ᵌ],

5 Guillaume Andrey[1,∞], Leonardo Beccari[2] and Denis Duboule[1,2,*]

6

7 [1]School of Life Sciences, Federal Institute of Technology, Lausanne, 1015 Lausanne,

8 Switzerland, [2]Department of Genetics and Evolution, University of Geneva, 1211 Geneva

9 4, Switzerland.

10 ∞Present address: Max Planck Institute for Molecular Genetics, RG Development &

11 Disease, 14195 Berlin, Germany,

12

13 ᵌThese authors contributed equally.

14

15 *Corresponding author: Denis.Duboule@epfl.ch

16

17 Keywords: Forelimb, hindlimb, genome topology, TADs, Polycomb, tetrapod evolution,

18 enhancers

19

20 Short title: Comparative regulation of *Hoxd* genes during tetrapod limb development

21

22   **ABSTRACT**

23   In all tetrapods examined thus far, the development and patterning of limbs

24   require the activation of gene members of the *HoxD* cluster. In mammals, they are

25   controlled by a complex bimodal regulation, which controls first the proximal patterning,

26   then the distal structure, allowing at the same time the formation of the wrist and ankle

27   articulations. We analyzed the implementation of this regulatory mechanism in chicken,

28   i.e. in an animal where large morphological differences exist between fore- and hindlimbs.

29   We report that while this bimodal regulation is globally conserved between mammals and

30   avian, some important modifications evolved at least between these two model systems,

31   in particular regarding the activity of specific enhancers, the width of the TAD boundary

32   separating the two regulations and the comparison between the forelimb *versus* hindlimb

33   regulatory controls. Some aspects of these regulations seem to be more conserved

34   between chick and bats than with the mouse situation, which may relate to the extent to

35   which forelimbs and hindlimbs of these various animals differ in their functions.

36

37   **AUTHOR SUMMARY**

38   The morphologies of limbs largely vary either amongst tetrapod species, or even

39   between the fore- and hindlimbs of the same animal species. In order to try and evaluate

40   whether variations in the complex regulation of *Hoxd* genes during limb development

41   may contribute to these differences, we compared their transcriptional controls during

42   both fore- and hindlimb buds development in either the mouse, or the chicken embryos.

43   We combined transcriptome analyses with 3D genome conformation, histone

2

44   modification profiles and mouse genetics and found that the regulatory mechanism

45   underlying *Hoxd* gene expression was highly conserved in all contexts, though with some

46   clear differences. For instance, we observed a variation in the TAD boundary interval

47   between the mouse and the chick, as well as differences in the activity of a conserved

48   enhancer element (*CS93*) situated within the T-DOM regulatory landscape. In contrast to

49   the mouse, the chicken enhancer indeed displayed a stronger activity in fore- than in

50   hindlimb buds, coinciding with the observed striking differences in the mRNA levels.

51   Altogether, differences in both the timing and duration of TAD activities and in the width

52   of their boundary may parallel the important decrease in *Hoxd* gene transcription in chick

53   hindlimb *versus* forelimb buds. These differences may also account for the slightly

54   distinct regulatory strategies implemented by mammals and birds at this locus, potentially

55   leading to substantial morphological variations.

56

57   **INTRODUCTION**

58   Tetrapod limbs are organized into three parts bearing skeletal elements; the

59   stylopodium (humerus/femur), the zeugopodium (radius/fibula, ulna/tibia) and the

60   autopodium, the latter including the acropod (phalanges, metacarpals/metatarsals) and the

61   mesopodium (carpals and tarsals)[1]. Limbs can display large variations in their

62   morphologies, either between tetrapod species or within the same species, between fore-

63   and hindlimbs, as a result of their adaptation to different functions and ecological niches.

64   For example, frogs display particular shapes of carpal and tarsal elements, with an

65   elongated proximal tarsal whenever detectable[2], whereas geckos forelimb skeletal

66   elements somehow resemble those of their hindlimbs[3]. Another example of this

67   morphological flexibility are the forelimbs of bats, which do have digits early on similar

68   to those of other mammals but that subsequently elongate to make flight possible[4].

69   In this context, birds are a fascinating animal group as they evolved forelimbs

70   (wings) and hindlimbs (legs) specialized for flying or for terrestrial locomotion,

71   respectively[5]. Recent studies using comparative genomics approaches either amongst

72   birds or between bats and mice, have revealed that some DNA sequences potentially

73   involved in limb development and which are highly conserved can display differential

74   enhancer activities as compared to their mouse orthologous sequences[6,7]. Furthermore,

75   the analysis of several domestic pigeons displaying variations in foot feathering within

76   the same species, suggested that changes in cis-regulatory elements in the genes encoding

77   forelimb- or hindlimb-specific transcription factors may contribute to a partial

78   transformation from hindlimb to forelimb identity[8]. Taken together, these observations

79   suggest that both the gain of species-specific enhancers, the different activities of the

80   same regulatory sequences and alterations in DNA sequences amongst various species

81   and/or within the same species, contributed to generate these large morphological

82   differences.

83   In addition to their essential role during axial patterning and organogenesis in

84   vertebrates[9,10], *Hox* genes are required for proper growth and skeletal patterning of

85   tetrapod limbs. In particular, genes belonging to the *HoxA* and *HoxD* clusters for both

86   fore- and hindlimbs, as well as some genes of the *HoxC* cluster during hindlimb

87   development[11,12]. Regarding both *Hoxd* and *Hoxa* genes, recent development in

4

88   chromatin conformation techniques have made it possible to associate previously defined

89   limb regulatory landscapes to large chromatin interaction domains referred to as TADs

90   (topologically associating domains)[13–15]. Therefore, multiple limb-specific DNA

91   regulatory sequences were identified on either side of the *HoxA* and *HoxD* clusters,

92   grouped either into TADs and smaller sub-megabase large domains (sub-TADs)[16–19].

93   At the murine *HoxD* locus, two overlapping subsets of genes are controlled by the

94   two adjacent regulatory TADs located on the telomeric side (T-DOM) and the centromeric

95   side (C-DOM) of the cluster[17]. The region of the cluster extending from *Hoxd1* to

96   *Hoxd8* generates constitutive interactions with T-DOM, whereas the 5' region of the

97   cluster, which includes *Hoxd13* to *Hoxd12,* predominantly contacts C-DOM. Interestingly,

98   the genes from *Hoxd9* to *Hoxd11,* located within the boundary between these two TADs,

99   interact with both T-DOM and C-DOM, yielding their transcripts in both the future

100   zeugopod and acropod domain. In this situation, the expression of *Hoxd1* to *Hoxd11* is

101   driven by contacts with enhancer elements situated in T-DOM in the prospective

102   zeugopod. Then, as the future acropod domain is generated, *Hoxd* genes start to weaken

103   their interactions with T-DOM and *Hoxd9* to *Hoxd11* switch to establish interactions

104   within C-DOM. This switch is partly controlled by HOX13 proteins, which inhibit T-

105   DOM activity while re-enforcing C-DOM-located enhancers function[20]. This bimodal

106   regulatory mechanism allows the production of a domain of low *Hoxd*-expression in

107   which both T-DOM and C-DOM regulations are functionally inert, giving rise to the

108   future mesopodium. While this complex system seems to be globally conserved

109   throughout evolution[21,22], some slight modifications could readily lead to important

110   changes in expression specificities.

111   In different tetrapod species, the morphological diversifications seen in particular in

112   both the mesopod and the zeugopod between fore- and hindlimbs were suggested to result

113   from heterochronic variations in *Hox* gene expression[2][23]. Both gain of function and

114   loss of function experiments have revealed that either the ectopic expression or the mis-

115   regulation of *Hoxa13* and *Hoxd13* can induce a substantial reduction and malformation

116   of the zeugopod, similar to mesomelic dysplasia conditions in human families (e.g.[24]).

117   This is due to the potential of these particular HOX13 proteins to antagonize the function

118   of other HOX proteins to control and stimulate the ossification of limb skeletal

119   elements[25]. In this view, HOX protein production controlled by the T-DOM stimulate

120   bone growth, whereas C-DOM enhancers upregulate *Hoxd13* to antagonize this property,

121   leading to both smaller bones (phalanges) and the termination of the structure, in a dose-

122   dependent manner[26–30].

123   In this context, it was recently reported that a bat regulatory sequence located within

124   T-DOM and controlling *Hoxd* genes displays differential enhancer activity in the limbs,

125   when compared to its mouse orthologous sequence[6]. These findings suggest that some

126   variations in limb morphology may be associated with the mechanism of bimodal gene

127   regulation described at the *HoxD* locus. However, this mechanism was reported to be

128   implemented during the development of forelimb buds only and hence it remained unclear

129   as to how much the important differences between fore- and hindlimbs either amongst

130   various species or within the same species, may be related to variations of this mechanism.

131   To tackle this issue, we used a comparative regulatory approach involving chick and

6

132　　mouse embryonic fore- and hindlimbs, mostly for two reasons. First chicken embryos,

133　　unlike mice, display striking differences between the morphologies of their adult wings

134　　(forelimbs) and legs (hindlimbs) (Fig. 1A, B, left). Secondly, it was reported that *Hoxd*

135　　gene expression domains during chick wing- and leg buds development showed important

136　　deviations when compared to their mouse counterparts[23,31]. These features suggested

137　　that the bimodal regulatory system at work at the mouse *HoxD* locus may be operating

138　　slightly differently during the development of the avian appendicular skeletons.

139　　　　Here we combine analyses of transcriptome, 3D genome conformation, histone

140　　modification and mouse genetics to show that this bimodal regulatory mechanism is

141　　highly conserved in birds. However, in chicken leg buds, the duration of the T-DOM

142　　regulation is much reduced, as suggested by dynamic changes in histone modifications,

143　　along with a slight difference in chromatin interaction profiles at specific enhancer

144　　regions, which accounts for the concurrent reduction in *Hoxd* genes expression in the

145　　zeugopod. Finally, by using mutant mouse embryos lacking a large part of T-DOM, we

146　　uncovered differences in the effect of T-DOM deletion for *Hoxd* genes between fore- and

147　　hindlimbs. Therefore, while these regulatory mechanisms are virtually the same amongst

148　　tetrapod species and even within the same species between the fore- and hindlimbs, slight

149　　differences are scored, which may partly contribute to the observed morphological

150　　differences.

151

152　　**RESULTS**

153　　**Transcription of *Hoxd* genes in mouse and chick limb buds**

7

154   We first compared the expression patterns of *Hoxd* genes in mouse fore- and

155   hindlimbs at E12.5 (Fig. 1A) with those observed in chick at either HH28 (equivalent to

156   E12.25 to E12.5) (Fig. 1B) or HH30 (equivalent to E13 to E13.5) (Fig. 1D). In mouse

157   fore- and hindlimbs, expression of *Hoxd13* and *Hoxd12* was strong in the prospective

158   acropod region (hereafter termed "distal"), whereas the domains of *Hoxd11* and *Hoxd10*

159   transcripts were split into the distal and zeugopod (hereafter termed "proximal") regions,

160   except the future mesopod, which was labelled by a *Col2a1*-expressing domain (Fig. 1A,

161   arrowheads). These expression patterns were similar in both fore- and hindlimbs, except

162   for a clearly weaker expression level in the hindlimb proximal domain.

163   When compared with *Hoxd* gene expression in the mouse counterparts, at least two

164   dramatic differences were confirmed in chick limbs. First, unlike the *Hoxd12* expression

165   pattern observed in murine limbs, the chick *Hoxd12* gene was strongly expressed in chick

166   proximal forelimb (wing) (Fig. 1B). Secondly, the expression of all *Hoxd* genes was

167   significantly reduced in the chick proximal hindlimb (leg) by stage HH28, when

168   compared to chick proximal wing and mouse proximal limbs[23,31]. As a result, the

169   expression domains of the chick *Hoxd12* in wing buds was much like that of *Hoxd11* or

170   *Hoxd10*, in contrast with the mouse where *Hoxd12* was never strongly expressed in the

171   proximal domain. However, the transition between the two *Hoxd*-expressing domains

172   also labelled the future wing mesopod (Fig.1B, arrowheads). Of note, expression of all

173   *Hoxd* genes was weak in proximal leg buds, again in contrast to what was observed during

174   mouse limb bud development (Fig. 1).

175   To characterize these differences in more detail, we performed RNA-seq analyses

8

176   by using HH30 limb buds in order to more easily micro-dissect the various domains and

177   thus exclude a potential contamination of the future mesopod from the distal domain.

178   RNA profiles confirmed the differences detected by WISH. First, *Hoxd11* to *Hoxd8* were

179   expressed at lower levels in the mouse proximal hindlimb when compared to forelimb, a

180   situation re-enforced in chick proximal leg where only a weak *Hoxd12* signal was

181   detected (Fig. 1C, D, upper tracks). In contrast, more reads were scored for *Hoxa10* to

182   *Hoxa11* in both mouse and chick proximal hindlimb when compared to forelimb (S1 Fig).

183   In the distal domains, transcription patterns and profiles from mouse and chick were

184   similar between fore- and hindlimbs, at both *HoxA* and *HoxD* clusters (Fig. 1C, D, lower

185   tracks, S1 Fig). However, the chick profile revealed a consistently higher transcription of

186   *Hoxd12*, which was as strong as *Hoxd13*, whereas the amount of transcripts in the mouse

187   counterpart was much lower than that for *Hoxd13* (Fig. 1C, D, lower tracks). Taken

188   together, these initial results indicated that both the expression status and the transcript

189   domains of *Hoxd* genes displayed significant differences, either amongst species or within

190   the developing fore- and hindlimb buds, in particular in chicken.

191

192   **Bimodal regulation in both fore- and hindlimb buds**

193   To determine to what extent these differences could result from variations in the

194   implementation of the bimodal regulatory mechanism, we performed comparative 4C-

195   seq analyses. We used a variety of viewpoints located at comparable positions to reveal

196   potential interactions in both mouse and chicken limb buds. To do this, we cross-

197   annotated *Hoxd* genes regulatory sequences identified in the mouse genome onto the

9

198   chick genome by using the LiftOver tool in UCSC. These annotations were then used for

199   all following experiments. In both fore- and hindlimbs, interactions were scored between

200   *Hoxd* genes and the regulatory sequences *island III* and *Prox*, as hallmarks of C-DOM

201   transcriptional activity, or with the *CS39* sequence as a proxy for T-DOM activity in either

202   the distal or the proximal region, respectively[17,18]. As seen in mouse forelimbs,

203   *Hoxd11* mainly contacted T-DOM sequences, especially *CS39*, in mouse proximal

204   hindlimb cells, i.e., in cells where T-DOM was fully active and where C-DOM was silent

205   (Fig. 2A, top). In contrast, in mouse distal hindlimb cells, it preferentially interacted with

206   C-DOM sequences such as *island III* and *Prox* (Fig. 2A, bottom). Quantification of

207   contacts indicated 75% of telomeric contacts in proximal forelimb cells and 50% in distal

208   forelimb cells, showing that *Hoxd11* had reallocated 25% of its global interactions toward

209   the C-DOM TAD in distal cells. Likewise, mouse hindlimb cells showed the same

210   interaction profiles, with 70% of telomeric contacts in proximal hindlimb cells and 45%

211   in distal hindlimb cells (Fig. 2A). This comparison indicated that the bimodal regulation

212   is virtually identical between fore- and hindlimbs in mouse.

213   We then examined the related interaction patterns in chick wing and leg cells by

214   using a region between *Hoxd11* and *Hoxd10* as a viewpoint (termed '*Hoxd10-11*'), i.e. a

215   sequence located as close as possible to the bait used in the mouse experiments. In chick

216   proximal wing cells, *Hoxd10-11* mostly interacted with the *CS39* and *CS93* regions

217   located in T-DOM, as well as with a region near the *Hnrnpa3* gene, where the distal TAD

218   border is observed in the murine locus (Fig. 2B, black arrowhead). These predominant

219   contacts with T-DOM were reduced in chick distal wing cells, as in the mouse, and 25%

10

220   of contacts were indeed reallocated to C-DOM sequences such as the chicken *island III*

221   and *Prox* counterparts. When compared to chick proximal wing cells, the global

222   interaction with the T-DOM was decreased from 81 to 70% in proximal leg cells. In

223   particular, the interactions between the *Hoxd10-11* bait and the *CS93* sequence in T-DOM

224   were lost in proximal leg cells, which may account for the significant reduction of *Hoxd*

225   expression in chick proximal leg buds (Fig. 2B, red arrows). In contrast, the interaction

226   established by the *Hoxd10-11* bait in chick wing and leg distal cells were comparable, as

227   expected from transcripts analyses, and interactions were observed up to around the

228   *Atp5g3* gene where the border of C-DOM TAD had been mapped in mouse (Fig. 2B,

229   white arrowheads).

230

231   **Different enhancer activity of mouse and chick CS93 in fore- *versus* hindlimbs**

232   The mouse *CS93* sequence contains the former *CS9* sequence[17], which was unable

233   to elicit any reporter gene expression in a mouse transgenic context (Fig. 3A). Likewise,

234   a larger murine sequence encompassing *CS9* and referred to as mouse BAR116 (Bat

235   Accelerated Region 116) did not show any enhancer activity in the limbs[6] (Fig. 3A). In

236   contrast, the bat BAR116 sequence was able to drive strong expression in transgenic

237   mouse forelimb buds whereas a weak activity was detected in hindlimb buds, correlating

238   with the different expression levels of *Hoxd10* and *Hoxd11* observed between bat fore-

239   and hindlimb buds[6] (Fig. 3A). This sequence was thus proposed as having evolved a

240   bat-specific function.

241   Therefore, low expression of *Hoxd* genes in proximal hindlimbs seems to be a

11

242   common feature of bats and chicken, suggesting that the chick *CS93* sequence may have

243   a limb enhancer activity similar to that reported for the bat BAR116. We examined the

244   enhancer activity of chick *CS93* using a transgenic mouse *lacZ* reporter system and

245   compared it to the activity of full-length mouse *CS93* sequence by using lentivector-

246   mediated transgenesis[32,33]. We initially cloned a 2kb large sequence containing chick

247   *CS93* and those surrounding sequences (Fig. 3B), which showed higher interactions with

248   *Hoxd10* to *Hoxd11* in the 4C profiles obtained from proximal wing cells (Fig. 2B, track

249   1). We noted that the surrounding sequences are not particularly conserved among species,

250   whereas the *CS93* region of the chick genome is more conserved with the bat than with

251   the mouse counterpart (Fig. 3B). By using the BLAT search tool in UCSC, we also found

252   that most of conserved regions from the bat BAR116 and the chick *CS93* sequences can

253   be aligned onto the mouse *CS9* region, respectively (Fig. 3A).

254   We assessed their enhancer activities and, unlike for the mouse BAR116, the full-

255   length mouse *CS93* triggered *lacZ* transcription in E10.5 limb buds with an expression

256   localized to the prospective stylopod and zeugopod at E12.5 (Fig. 3C). The staining was

257   weaker in hindlimb than in forelimb buds, possibly due to the delay in limb

258   development[34,35]. Accordingly, the 380bp localized in 5' of the mouse *CS93* seem to

259   be essential for expression. On the other hand, we found that the 2kb large sequence

260   containing the chick *CS93* displayed limb enhancer activity in mouse limb buds at E12.5

261   (Fig. 3D). The transgene expression driven by chick *CS93* came under two different

262   patterns. The first one displayed *lacZ* staining throughout the forelimbs (n=2/5), which

263   was similar to the staining observed when the bat BAR116 sequence was assessed in

264   mouse forelimb bud (Figs 3 and 4 in[6]). The second pattern concerned mostly the

265   proximal forelimb buds (n=3/5), as seen when the mouse *CS93* was used (Fig. 3C). In

266   both cases, a weaker expression was observed in hindlimb bud, as in the case of bat

267   BAR116. These results suggest that the downregulation of *Hoxd* genes in chick leg bud

268   involves a weaker activity of- and less interactions with the *CS93* sequence. Taken

269   together with previous research [6], these results indicate that, unlike for the bat BAR116

270   and the chick *CS93*, the mouse *CS9* is devoid of limb enhancer activity. Accordingly, the

271   380bp localized in 5' of the mouse *CS93* seem to be essential for expression.

272

273   **Implementation of the regulatory switch between TADs in mouse and chicken**

274   We observed that the differences in *Hoxd12* expression were more substantial

275   between chick fore- and hindlimb buds than in the murine counterparts (Fig. 1C, D).

276   Therefore, we wondered whether the regulatory switch from the T-DOM to the C-DOM

277   regulations would occur differently in the two species. We produced and examined 4C

278   interactions profiles by using *Hoxd12* itself as a bait. Similar to the profiles obtained with

279   the *Hoxd10-11* bait, we observed weaker interactions between *Hoxd12* and both the *CS39*

280   and *CS93* regions in T-DOM in chick proximal leg cells than in proximal wing cells (S2A

281   Fig, top). The profiles with the *Hoxd10-11* bait showed strong and stable interactions with

282   T-DOM, when compared with C-DOM, in both proximal and distal limbs (Fig. 2B). We

283   also found that *Hoxd12* mainly contacted T-DOM in both chick proximal wing and leg

284   cells, whereas it established more interactions with C-DOM in both chick distal cells (S2A

285   Fig, top and bottom).

13

286   The murine *Hoxd9* to *Hoxd11* genes, but not *Hoxd12*, are located within the TAD

287   boundary and interact both with T-DOM and with C-DOM. In contrast, in chicken limb

288   buds, *Hoxd12* was able to switch contacts from T-DOM to C-DOM, suggesting that the

289   TAD boundary in chick could be located between *Hoxd12* and *Hoxd13* (see also S2B and

290   S2C Fig) thus suggesting that the exact genomic interval where the switch in contacts

291   occurs between the genes and the TADs is displaced towards a more 5' position within

292   the chicken gene cluster, when compared to murine limb buds where this switch region

293   was localized around the *Hoxd11* locus[17,19]. This same switch was observed in both

294   chick wing- and leg bud cells, regardless of the various expression levels of *Hoxd* genes

295   in the proximal region, revealing that the switch between TADs is independent of *Hoxd*

296   expression itself in proximal cells.

297   These results showed that the bimodal regulatory mechanism and the sequential

298   transition from the proximal to the distal global controls are implemented during chick

299   limb development similarly to what was described in mice. Therefore, the differences in

300   gene expression observed both between mice and chicken and between chick wing- and

301   leg buds cannot be solely explained by visible variations in the respective interaction

302   profiles. Instead, they ought to involve the distinct use of enhancers (or group thereof-)

303   within an otherwise globally conserved chromatin architecture.

304

305   **Premature termination of the proximal TAD activity in chick leg buds**

306   Since the chromatin architecture at the *HoxD* locus is seemingly comparable

307   between mouse and chicken in both fore- and hindlimb buds, we looked for what may

14

308   cause the drastic reduction of *Hoxd* expression observed in chick proximal leg. Within

309   the T-DOM TAD structure, the interaction profiles obtained from chick leg proximal cells

310   showed reduced contacts between *Hoxd* promoters and enhancers in T-DOM (Fig. 2 and

311   S2 Fig). We complemented these observations by assessing the functional state of T-DOM

312   sequences by comparing particular histone modifications profiles between chick wing-

313   and leg buds, at several developmental stages (Fig. 4). We looked at the acetylation of

314   H3K27 (H3K27ac), a modification associated with transcriptional and enhancer activity

315   and at the trimethylation of the same residue (H3K27me3), a mark associated with

316   Polycomb-dependent silencing[36]. In both wing- and leg buds at stage HH19, a stage

317   which corresponds to ca. E9.5 in mouse, enrichments of H3K27ac were detected over

318   both T-DOM and the *HoxD* cluster itself, showing that the activation of *Hoxd* genes by

319   the T-DOM enhancers had been properly initiated in leg buds (Fig. 4A, tracks 1 and 2).

320   Of note, higher levels of this mark were scored over the *Hoxd11* to *Hoxd13* region in leg-

321   than in wing buds (S3A Fig).

322   At stage HH20 (approximatively E10 in mouse), the H3K27ac enrichment in T-

323   DOM was still substantial in wing buds. In marked contrast, however, this level appeared

324   dramatically reduced in leg buds (Fig. 4A, tracks 3 and 4, region *a*), thus coinciding with

325   low gene expression (S3A Fig). The accumulation of H3K27ac observed near the TAD

326   border was specific for the early wing-bud (Fig. 4A, tracks 1 and 3, region *b*).

327   Furthermore, H3K27ac signals over C-DOM were not yet observed at these stages

328   (except around the island I region), in agreement with the fact that the regulatory switch

329   had not yet occurred. At a later stage (HH28, i.e. approximately E12.5 in mouse),

15

330 enrichment of H3K27ac within the *HoxD* cluster was significantly lost in proximal leg

331 bud cells where *Hoxd* expression was weak (Fig. 4A, track 6, S3A Fig). In contrast,

332 H3K27ac accumulation over the T-DOM in proximal wing bud cells remained, yet it

333 started to slowly decrease as observed in mouse proximal forelimb at E12.5 (Fig. 4A,

334 track 5). At the same time, H3K27ac was finally enriched over both C-DOM and the

335 *HoxD* cluster, in both wing- and leg distal cells, as scored in mouse distal forelimb buds

336 (Fig. 4A, tracks 7 and 8, S3A Fig)[17,20]. These various profiles showed that in chick leg

337 bud cells, the functional switch between T-DOM and C-DOM had occurred normally,

338 except that after its initial onset, T-DOM activity was terminated much more rapidly than

339 in the wing bud, followed by a decrease in accumulation of H3K27ac at the target *HoxD*

340 cluster itself.

341   We complemented these observations by analyzing H3K27me3 marks, which

342 antagonize the acetylation of the same residue[36]. At stage HH20, no H3K27me3 signal

343 was detected over T-DOM in both wing- and leg buds, except around *CS39* in leg buds

344 (Fig. 4B, tracks 1 and 2, region '*a*'), in agreement with the H3K27ac profiles (compared

345 with Fig. 4A, tracks 3 and 4). In contrast, strong levels of H3K27me3 enrichment were

346 observed over the C-DOM regions, where H3K27ac peaks were not detected (Fig. 4B,

347 tracks 1 and 2), suggesting that the activation of *Hoxd* genes by C-DOM regulation had

348 not yet occurred at this stage.

349   At the *HoxD* cluster itself, stronger levels of H3K27me3 enrichment were clearly

350 detected in leg buds, as compared with wing buds, from the pseudo-*Hoxd1* gene to *Hoxd8*,

351 a DNA interval controlled by T-DOM regulation (S3B Fig). At later stages, H3K27me3

16

352   marks were observed over C-DOM in proximal wing bud cells, where C-DOM is inactive,

353   whereas the levels of H3K27me3 marks over T-DOM, in both proximal and distal wing

354   bud cells, were comparable to those seen in the H3K27ac profiles (Fig. 4B, tracks 3, 5).

355   We had previously suggested that the binding of HOX13 proteins to sequences located in

356   T-DOM could cooperate with Polycomb complexes to terminate T-DOM regulatory

357   activity[20].

358   Altogether, the distribution of both H3K27ac and H3K27me3 marks in chicken limb

359   buds were in agreement with the observed expression status of *Hoxd* genes profiles. A

360   major difference was scored, however, when compared to their mouse counterparts. In

361   proximal leg bud cells at HH28 indeed, where *Hoxd* gene expression is quite weak (Fig.

362   4B, track 4; regions '*a*' and '*b*'), T-DOM and the *HoxD* cluster were heavily decorated

363   with H3K27me3 marks, in addition to the C-DOM TAD. The latter profile resembled that

364   obtained from distal leg bud cells at the same stage, i.e. cells where T-DOM is inactive

365   and completely shut down (Fig. 4B, track 6; regions '*a*' and '*b*'), suggesting that T-DOM

366   was not operational in proximal leg cells, unlike in the mouse forelimb proximal

367   situation[17,20].

368   We also examined the distribution of both H3K27ac and H3K27me3 marks over

369   the *HoxA* cluster and its limb regulatory landscape that maps within a sub-TAD adjacent

370   to *Hoxa13* [16](S4 Fig). H3K27ac distribution suggested an earlier activation of *Hoxa13*

371   in leg- than in wing buds at HH19 and HH20 (S4 Fig), which corroborated the

372   transcriptomes established at HH20 (S4C Fig). At HH28 however, the profiles of

373   H3K27ac and H3K27me3 marks over the *HoxA* cluster and within the flanking regions

17

374   were comparable between wing- and leg bud cells (S4 Fig) and did not show any of the

375   differences reported for the *HoxD* cluster and their regulatory landscapes. H3K27ac

376   profiles were similar over the *HoxA* limb enhancers in wing- and leg tissues at all stages,

377   except perhaps for a slight decrease in H3K27ac in leg bud at HH20 (S4A Fig).

378

379   **Chromatin conformations of the chicken *HoxD* cluster in wing- and leg buds**

380   Gene expression usually occurs concomitantly with enhancer-promoter contacts[37–

381   39]. Because of the dramatic difference in T-DOM activity observed in chick leg- *versus*

382   wing buds at stage HH20 (Fig. 4A, tracks 3 and 4, region '*a*'), we looked for potential

383   parallel differences in chromatin contacts by performing high-resolution capture Hi-C

384   analyses (CHi-C)[40,41] using wing- and leg buds at HH20. Since such a global

385   chromatin assessment had not been done during chick development, it also allowed us to

386   compare it with mouse counterpart cells and see to what extent these complex regulatory

387   systems were conserved in distinct groups of tetrapods.

388   The CHi-C profiles of chicken cells confirmed that the *HoxD* cluster is positioned at

389   the boundary between two TADs, similar to what was proposed in mouse limb bud

390   tissues[17,19] and the two sub-TADs seen in the murine T-DOM were detected as well

391   (S5A and S5B Fig). To position the boundary between the two TADs, we applied the

392   TopDom algorithm[42], which determined the border around the *Hoxd13* locus in both

393   wing- and leg bud cells at HH20 (Fig. 5A). This extended the conclusion reached after

394   the 4C analyses, i.e. that the TAD boundary region in chick was displaced towards the 5'

395   part of the gene cluster when compared to mouse limb bud cells[19].

396   In mouse limb cells, this TAD boundary falls within a region where multiple CTCF

397   sites are occupied[43–45]. CTCF is an architectural protein that both helps defining

398   constitutive domains of interaction and triggers enhancer-promoter contacts[46]. We thus

399   examined the presence of bound CTCF at the chick *HoxD* locus and surrounding TADs

400   and found that the profiles were comparable between wing- and leg buds at HH20 (S5

401   Fig). As for the mouse *HoxD* cluster[19], the orientations of the CTCF motives located

402   on either sides of the TAD boundary were facing those sites found in their flanking TADs,

403   suggesting the possibility for long-range loops to be established (S5C Fig, top)(e.g.[47]).

404   Noteworthy, the orientation of the CTCF motives were conserved between mouse and

405   chick. However, we found less bound CTCF in the chicken *HoxD* cluster than in the

406   mouse counterpart, which could affect the strength and/or stability of the TAD boundary

407   in chick (S5C Fig, bottom).

408   Using CHi-C at 5kb resolution, the distribution of contacts was relatively similar

409   between wing- and leg bud cell populations, despite the reduced level of H3K27ac in the

410   T-DOM and near the TAD border in leg bud cells described above. The contacts between

411   the *HoxD* cluster (galGal5, chr7:16,315,000-16,395,000) and the T-DOM region '*a*'

412   (galGal5, chr7: 16,045,000-16,255,000) were indeed comparable between wing- and leg

413   bud cells (Wilcoxon sum test, W=222,690, p-value=0.663) (Fig. 5B). However, when the

414   comparison was further limited to a 30kb large region including *CS93* (galGal5,

415   chr7:16,090,000-16,120,000) rather than to the entire region '*a*', the contact intensity in

416   leg bud cells was clearly below that scored in wing bud cells (W=5,495, p-value=0.0213).

417   These reduced contacts between *Hoxd* genes and the surroundings of region *CS93*

19

418   confirmed the observation made by looking at the 4C profiles obtained using the later

419   stage (HH30) (Fig. 2).

420   The fact that bound CTCF was not detected around the *CS93* region suggests that

421   CTCF-independent variations in enhancer-promoter interactions may participate to the

422   important decrease in *Hoxd* gene expression levels in leg bud cells. The decreased

423   contacts seen in leg bud cells were also scored between the *HoxD* cluster and around the

424   T-DOM border region *'b'* (galGal5, chr7: 15,910,000-15,985,000; W=32,746, p-

425   value=0.009415). The decrease in contacts between the *HoxD* cluster and both the *CS93*

426   and TAD border regions in leg bud cells was somehow compensated for by an increase

427   of interactions between the *HoxD* cluster and the 5' region of *CS39* where several bound

428   CTCF sites were observed. The 4C profiles obtained when using the *Hoxd10-11* bait

429   showed interactions with the 5' region of *CS39* in both chick proximal and distal cells,

430   suggesting that these stable contacts are associated with CTCF, as described in mouse

431   developing limb buds[48].

432

433   **Regulation of T-DOM by HOXA13**

434   We asked what may cause both the robust reduction in H3K27ac marks in chicken

435   T-DOM and the decrease in contacts between *Hoxd* genes and the *CS93* region in leg bud

436   cells at HH20. As it was reported that HOX13 proteins bind T-DOM-located sequences

437   concomitantly to its inactivation and that the absence of these proteins prevented T-DOM

438   to terminate operating[20,49], we re-assessed the expression dynamics of *Hoxa13*. We

439   found it expressed earlier in chick leg bud than in wing buds[31](S4C Fig), suggesting

20

440   that the exact timing of *Hoxa13* transcriptional activation may fix the duration of T-DOM

441   activity during limb development.

442   We examined this possibility by performing time-course *in situ* hybridization and

443   RT-qPCR experiments, using chick and mouse entire fore- and hindlimb buds from HH20

444   to HH22 and E10.5 to E10.75, respectively (Fig. 6). While these developmental stages

445   are not strictly equivalent between chick and mouse[50], they were selected because the

446   size difference between chick wing- and leg buds is not yet too large between HH20 and

447   HH22[51]. Also, *Hoxa13* starts to be expressed in mouse forelimb buds at around

448   E10.5[52]. While the onset of *Hoxa13* expression was detected in chick wing bud at

449   HH22, *Hoxa13* transcripts were already well present in chick leg bud at HH20-21 (Fig.

450   6A, left). Also, the expression level of this gene in leg buds was markedly stronger than

451   in wing buds (Fig. 6A, right). *Hoxa11* expression was also higher in chick leg buds than

452   in forelimb buds, as was also observed in the RNA-seq dataset (S4C and S6A Fig),

453   suggesting that the entire chicken *HoxA* cluster was activated in leg buds well before it

454   was switched on in wing buds. This was nevertheless not a general phenomenon for *Hox*

455   genes, for the expression of *Hoxd13* was comparable between wing- and leg buds (Fig.

456   6A, bottom and S6B, S6C Fig, left).

457   In the mouse, the development of the forelimb bud precedes that of hindlimb buds

458   by about half a day. In contrast, the initiation of both wing- and leg bud in chicken is

459   almost concomitant and the growth of the leg bud exceeds that of the wing bud[34,51].

460   However, even when considering these developmental differences, the dramatic

461   variations we scored between both the timing of *Hoxa13* activation and its transcript

21

462   levels between the chick wing- and leg buds were different from the situation observed

463   in murine fore- and hindlimb bud (Fig. 6B, top). Also, expression of other *Hoxd* genes

464   increased in both mouse fore- and hindlimb buds, as development proceeded (Fig. 6B

465   bottom, S6B and S6C Fig. right), unlike in chicken leg buds. Therefore, an inverse

466   correlation exists between the activation of *Hoxa13* on the one hand, and the

467   downregulation of *Hoxd* genes such as *Hoxd11* in chick leg bud on the other hand. This

468   correlation was observed neither in chick wing bud, nor in mouse limb buds, supporting

469   the idea that an early activation of *Hoxa13* induces a premature termination of T-DOM

470   activity in chick leg bud.

471   We looked whether the profiles of H3K27ac ChIP-seq and Hi-C data obtained from

472   chick limb tissues and covering the *HoxA* cluster would reveal traces of this early and

473   strong activation of *Hoxa13* seen in chick leg buds at HH20 (S4A and S6D Fig). While

474   this activation was consistent with enriched H3K27ac marks over the *HoxA* cluster itself,

475   it was not fully consistent with the distribution of chromatin marks over those enhancers

476   previously described to regulate *Hoxa13* in developing mouse limbs[16](S4A Fig). Hi-C

477   heat maps, however, showed increased contacts between the 5' region of the *HoxA* cluster

478   and its limb enhancer regions in leg bud at HH20, when compared to wing buds,

479   suggesting this may contribute to an early activation of *Hoxa* genes in leg buds (S6D Fig).

480

481   **Different impacts of T-DOM upon mouse fore- and hindlimb buds developments**

482   The importance of the T-DOM TAD for mouse proximal limb development was

483   initially assessed in forelimbs exclusively. The fact that birds displayed this striking

22

484 difference in T-DOM-dependent regulations in wing and leg buds suggested that tetrapod

485 fore- and hindlimbs may use the T-DOM enhancers in a slightly different manner. We

486 investigated this possibility by looking at the effect of a full deletion of T-DOM upon

487 *Hoxd* gene regulations in both murine fore- and hindlimb buds, in order to identify a

488 potential difference in the genetic requirements for T-DOM regulations between the two

489 contexts. We analyzed mouse limb buds carrying the $HoxD^{Del(attp-SB3)}$ allele, where an

490 approximatively 1Mb large region including the entire T-DOM as well as its distal border

491 was deleted. *Hoxd* transcripts produced in E12.5 proximal limbs by the $HoxD^{Del(attp-SB3)}$

492 allele (Fig. 7A left, *Del(attp-SB3)/Δ*) were scored by using both *in situ* hybridization and

493 RT-qPCR (Fig. 7A right, S7A Fig. left). In such mutant proximal forelimb buds, *Hoxd11*

494 to *Hoxd8* transcripts were depleted for more than 90% when compared to control

495 proximal forelimbs (Fig. 7A right, S7A Fig. left). Surprisingly however, *Hoxd11*

496 transcripts were not as dramatically affected in proximal mutant hindlimbs, and

497 transcripts produced by the *Hoxd10* to *Hoxd8* genes were decreased in amount by 50 to

498 60% when compared to control animals (Fig. 7A right, S7A Fig. left). The reduced level

499 of *Hoxd* gene expression elicited by the mouse T-DOM deletion in the forelimb bud thus

500 mimics the situation observed in chick proximal leg bud (S7A Fig. right) and these

501 significant differences also exist in the way T-DOM operates in murine forelimb *versus*

502 hindlimb buds.

503 Interestingly, the sustained expression of *Hoxd* genes in T-DOM deletion mutant

504 proximal hindlimb buds completely disappeared when a larger deletion was engineered

505 between the *Mtx* and *Titin (Ttn)* genes (Fig. 7A), indicating that the genomic regions

23

506   between SB3 and *Ttn*, i.e. telomeric to the T-DOM TAD, contributes to the difference in

507   gene expression observed between the mouse fore- and hindlimb buds when T-DOM is

508   removed.

509   To identify potential forelimb to hindlimb differences in chromatin re-

510   organization after the deletion of T-DOM, we generated 4C profiles from the mutant allele

511   using the *Hoxd11* promoter as a viewpoint (Fig. 7B). In control proximal fore- and

512   hindlimb cells, *Hoxd11* mostly contacted the intact T-DOM, with a particularly strong

513   interaction with and around the CS39 region (Fig. 7B, tracks 1 and 3). In proximal cells

514   deleted for T-DOM, interactions within the *HoxD* cluster were increased and ectopic

515   contacts were established (or strongly re-enforced) with the newly fused neighboring

516   telomeric TAD (Fig. 7B, tracks 2 and 4). We used this 4C-seq dataset to determine three

517   candidate regions of potential enhancer activity, referred to as hindlimb enhancer (HE) 1

518   to 3 (Fig. 7B, C, red arrows). We cross-checked this selection either with DNaseI

519   hypersensitive sites (HS) data from E11.5 hindlimb buds (GSM1014179)[53], with

520   potential enhancer regions as defined by the Limb-Enhancer Genie tool[54] and with

521   H3K4me1 ChIP-seq datasets obtained from control and mutant hindlimb proximal

522   domains (Fig. 7C and S7B Fig). According to these criteria, HE1 turned out to be the

523   most promising region and we thus assessed its enhancer potential in developing

524   transgenic limb buds.

525   By using a transgenic enhancer reporter system, the HE1 region reproducibly drove

526   *lacZ* expression in proximal fore- and hindlimb buds, lateral plate and somatic mesoderm

527   at E12.5 (Fig. 7D). This result indicated that the HE1 enhancer activity is not specific for

24

528   the proximal hindlimb, even though it was potentially active in a hindlimb-specific

529   manner after deletion of T-DOM. To confirm this observation, we performed a 4C-seq

530   analysis by using the HE1 sequence as a bait (S7C and S7D Fig), using both control and

531   T-DOM deleted mutant proximal hindlimb cells. The 4C-profiles revealed that the HE1

532   sequence established slightly more contacts with the *Hoxd9* to *Hoxd12* DNA interval in

533   mutant than in control hindlimb cells.

534   Finally, we looked at potential genetic interactions between the limb-specific

535   differences in *Hoxd* gene expression and the *HoxC* gene cluster. Both *Hoxc10* and *Hoxc11*

536   are indeed strongly transcribed in proximal cells of hindlimb buds (S7E Fig), whereas

537   these transcripts were not detected in the equivalent forelimbs territories[55].

538   Furthermore, in proximal hindlimbs where *Hoxc* genes are expressed, the amount of *Hoxd*

539   transcripts is decreased when compared to forelimb buds (Fig. 1C). Also, the deletion of

540   *Hoxc11* on top of *Hoxa11/Hoxd11* double knockout mice exacerbated the observed

541   hindlimb malformations[11][56], suggesting that HOXC proteins in hindlimb buds may

542   help maintaining *Hoxd* transcripts, a function that would become visible in the absence

543   of T-DOM, when *Hoxd* genes are no longer transcribed in forelimb proximal cells. We

544   thus performed *in situ* hybridization for *Hoxd11* by deleting the entire *HoxC* cluster[57]

545   on the top of the deletion of T-DOM. In these combined mutant limb buds, expression of

546   *Hoxd11* was still detected in hindlimb proximal cells, indicating that the persistence of

547   *Hoxd11* expression in hindlimb buds in the absence of T-DOM did not depend upon the

548   presence of hindlimb proximal cells-specific *Hoxc* transcripts.

549

550 **DISCUSSION**

551 **Conservation of the bimodal regulation in birds**

552   While the expression of *Hox* genes belonging to the *HoxA*, *HoxC* and *HoxD* clusters

553 during limb development are globally comparable between mammals and birds, some

554 clear differences are nevertheless apparent, such as the quasi absence of *Hoxd* genes

555 transcription in the proximal part of the developing leg buds in birds, while their

556 expression there is required for proper mouse hindlimb development[56,58]. Also, while

557 *Hoxd12* is expressed in the mouse limb buds like *Hoxd13*, i.e. mostly under the control

558 of C-DOM only, its expression in the avian wing buds resembles that of *Hoxd11*, i.e.

559 involving also the control of T-DOM, in future zeugopod cells. Because these expression

560 specificities depend on the implementation of global regulations located within the two

561 TADs flanking the *HoxD* cluster, we wondered whether the structures of these TADs were

562 somehow modified in birds, or at least whether they would show some variations either

563 between the two species, or between the bird wing and leg buds.

564   A global analysis of 4C and capture Hi-C datasets did not reveal any salient

565 differences between mammals and birds regarding the way they implement this complex

566 bimodal limb regulation at the *HoxD* locus. The TADs appeared well conserved between

567 the two species, as well as the presence in chick of most -if not all- regulatory elements

568 that had been described in the mouse counterparts, on both sides on the gene

569 cluster[17,18], even though the chick TADs were reduced in size. We thus conclude that

570 the bimodal regulatory strategy described in mammals (see[59]) is implemented in a

571 similar manner during bird development. This re-enforces the idea that the function of

26

572   *Hox* genes at these early steps of limb development is mostly to set up and organize the

573   basic plan of the future appendages, rather than elaborating or fine-tuning on a pre-

574   patterned structure.

575

576   **Interspecies comparison of the TAD boundary at *HoxD***

577   The bimodal regulatory mechanism and the transition between the two global

578   controls at *HoxD* are thus well conserved amongst tetrapods. However, the distinct

579   expression of *Hoxd12* in proximal limbs between mouse and chick suggests that the width

580   of the TAD boundary at the *HoxD* locus may slightly vary between the two species. By

581   using transcriptome, 4C and Hi-C datasets, we previously observed slightly different

582   positions of this boundary in mouse distal *versus* proximal limb cells, due to the fact that

583   *Hoxd10* and *Hoxd11* respond to both T-DOM and C-DOM regulations. We thus proposed

584   that the boundary was located between *Hoxd11* and *Hoxd12* in proximal cells, while

585   between *Hoxd9* and *Hoxd10* in distal cells[17,19] (Fig. 8A).

586   In contrast, the chick *Hoxd12* is strongly expressed in proximal wing buds,

587   suggesting that the TAD boundary, in this context, expands towards the 5' part of the gene

588   cluster, close to the *Hoxd13* locus (Fig. 8B). Our CHi-C analysis re-enforced this view

589   and positioned this boundary around the *Hoxd13* gene in chick limb buds at early stages

590   (HH20), i.e. when T-DOM is active and controls the first phase of *Hoxd* genes

591   transcription (S5 Fig). Subsequently (HH30), the boundary region was slightly reduced

592   from *Hoxd13* to an area between *Hoxd13* and *Hoxd12* in chick limbs. Of note, *Hoxd12* is

593   expressed in proximal limbs in geckos as in chicken[23], suggesting that the TAD

27

594   boundary at the *HoxD* locus in proximal buds may have been slightly shifted during

595   tetrapod evolution between birds and squamates on the one hand, and mammals on the

596   other hand.

597   Different tetrapod species may thus display various positions of their TAD

598   boundaries at *Hox* loci, acting as morphological cursors that would redistribute the

599   various sub-sets of *Hox* genes responding to either proximal or distal enhancers. The

600   distinct topologies of these boundaries may rely upon differences in the distribution

601   and/or usage of CTCF binding sites. In the mouse, subsets of genes responding to either

602   proximal or distal limb enhancers are delimited by different sets of bound CTCF sites

603   [19] (Fig. 8A). Here, we show that chicken wing bud cells have less bound CTCF sites

604   and of reduced intensity in the *HoxD* cluster than their murine counterparts, which could

605   modulate the positioning of the boundary. In this case, the decrease in the overall strength

606   of the boundary effect as a result of having less sites occupied by CTCF, may explain the

607   extension of interactions up to *Hoxd12-Hoxd13* established by proximal enhancers (Fig.

608   8B). This hypothesis could nevertheless not be verified on chicken leg proximal cells as

609   these cells do not strongly express the genes controlled by T-DOM.

610

611   **Distinct T-DOM regulations in mouse, chick and bat fore- and hindlimb buds**

612   During bat limb development, *Hoxd10* and *Hoxd11* transcripts are progressively lost

613   throughout the hindlimbs only, in part due to the distinct enhancer activity of the BAR116

614   sequence located within T-DOM[6]. When the mouse BAR116 cognate sequence was

615   used in a transgenic context, no activity was detected in any limb cells. Likewise, when

616   we used mouse *CS9*, i.e. a shorter fragment of the *CS93* sequence, staining was not

617   observed. However, when the full length *CS93* sequence was injected, a robust enhancer

618   activity was scored in a proximal limb region (Fig. 3). This discrepancy between two

619   experiments involving almost the same sequences may be caused by the positions of the

620   regions used for the mouse transgenic enhancer assays, the mouse sequence extending a

621   bit more towards one of its extremities. Either the enhancer activity was provided by this

622   sub-fragment, or this fragment may be required for the expression of a more widespread

623   activity of the full DNA sequence. It remains that the BAR116 sequence may not be a bat

624   specific enhancer.

625   However, while the bat BAR116 showed different enhancer activities between the

626   mouse transgenic fore- and hindlimb buds, i.e. strong in forelimb and weak in hindlimb,

627   the mouse equivalent appeared to have similar enhancer activities between fore- and

628   hindlimbs, in agreement with the continuous expressions of *Hoxd10* and *Hoxd11* in both

629   types of limbs. To further validate this correspondence, we looked at the behavior of the

630   chick *CS93* sequence in the same mouse transgenic context. While two sets of patterns

631   were obtained with various distal to proximal distributions of the *lacZ* staining, a clear

632   imbalance was scored between forelimb and hindlimb cells, with a much stronger

633   expression in the former than in the latter. Therefore, the chick enhancer sequence

634   behaved more like the bat sequence than like their murine counterparts, an effect that was

635   re-enforced by the sequence alignments, which revealed more similarities between chick

636   and bats than between the two mammalian species. This similarity correlates with *Hoxd*

637   gene expression and may relate to the salient distinctions in morphologies between fore-

29

638   and hindlimbs.

639

640   **Premature termination of T-DOM regulation in chick hindlimb buds**

641   The termination of the T-DOM regulatory activity in proximal limb cells coincides

642   with the binding of the HOXA13 protein at various sites within the TAD. Also, the

643   removal of both *Hoxa13* and *Hoxd13* functions lead to the continuation of T-DOM

644   regulation and to the failure of C-DOM activation, suggesting that HOX13 proteins are

645   necessary to terminate T-DOM function and to implement the bimodal switch[20,49]. The

646   chick *Hoxd13* gene starts to be expressed at around stage HH18-19[31] when H3K27ac

647   enrichment is not yet detected over C-DOM (except for *island I*)(Fig. 4). Instead,

648   H3K27me3 marks are still present over C-DOM at this early stage, unlike in the early

649   mouse limb buds (Fig. 4 and S3 Fig)[17], suggesting that *Hoxd13* early activation in chick

650   may be driven by the T-DOM regulation until the C-DOM regulation is implemented and

651   takes it over. This idea is well supported by our CHi-C analysis showing that the TAD

652   boundary is expanded towards the *Hoxd13* locus in the early chick limb buds.

653   In addition, a major difference was observed in the activation of *Hoxa13* between

654   chick and mouse hindlimb buds, with an earlier and stronger activation of *Hoxa13* in

655   chick leg buds at HH20, when compared to both mouse hindlimb buds and chick wing

656   buds. This indicates that the T-DOM activity may be readily terminated by the early

657   presence of HOXA13 protein and that the C-DOM regulation is ready to be active earlier

658   in chick leg buds than in wing buds. The potential causes for both this early activation of

659   *Hoxa13* in chick leg buds and the strong level of H3K27me3 observed over C-DOM in

30

660   chick wing and leg buds remain to be determined.

661

662   **Enhancer reallocation and anterior-posterior (AP) position of the limb buds**

663   In the mouse, we show that the deletion of the telomeric TAD has different effects

664   upon forelimb and hindlimb proximal cells, as fair levels of *Hoxd* gene expression remain

665   only in the proximal hindlimb domain. Since a larger deletion including more telomeric

666   sequences totally abrogated *Hoxd* expression in proximal hindlimb, we concluded that

667   additional hindlimb-specific enhancers may be located telomeric of the T-DOM TAD.

668   Interaction profiles established after the deletion of T-DOM revealed novel hindlimb-

669   specific contacts between target *Hoxd* genes and the newly identified HE1 sequence,

670   which is normally located outside T-DOM but brought to the vicinity of the cluster after

671   the deletion. The HE1 region is located near the *Agps* and *Pde11a* genes and our 4C

672   profile using HE1 as a bait showed strong interactions with both genes under normal

673   condition. Noteworthy, *Agps* is involved in the rhizomelic chondrodysplasia punctate 3

674   (RCDP3) condition, with a shortening of proximal limbs[60,61], suggesting that HE1

675   may be involved in the regulation of the *Agps* gene. The deletion of T-DOM may thus

676   reallocate part of the HE1 proximal limb enhancer activity towards *Hoxd* promoters.

677   Our genetic approach however makes it difficult to assess whether this sequence is

678   used for *Hoxd* regulation under normal circumstances, or whether the interactions

679   observed are mostly due to its new proximity to the target genes induced by the deletion

680   of T-DOM. In the former case, this may indicate that as in chick and bats, the global C-

681   DOM regulation may be more active in forelimb than in hindlimb buds and hence the

31

682   HE1 enhancer may not be necessary. In the case of the mouse, this deficit of regulation

683   during proximal hindlimb development could have been compensated for by evolving

684   additional enhancers outside the TAD (Fig. 8C). The HE1 sequence is bound by several

685   factors, such as the YY1, proposed to mediate enhancer-promoter contacts at distance in

686   ESCs[62], or PITX1, a hindlimb-specific factor[63], suggesting that multiple

687   transcriptional factors may contribute to this interaction.

688   Finally, the sustained expression of *Hoxd* genes observed in T-DOM deleted mutant

689   proximal hindlimb cells may merely reflect the history of early limb bud cells. In the wild

690   type condition indeed, the anterior bud emerges from a field of lateral plate mesoderm

691   (LPM) devoid of transcripts for *Hoxd9*, *Hoxd10* or *Hoxd11*. In contrast, posterior limb

692   buds derive from LPM cells already expressing these genes, due to their more posterior

693   AP position along the trunk mesoderm. In the absence of T-DOM, expression of these

694   genes would not occur in the anterior buds due to their repressed state and the lack of

695   appropriate enhancers, whereas expression could be inherited and maintained in the

696   posterior buds through a mechanism independent of T-DOM, perhaps involving the HE1

697   sequence.

698

699   **Materials and Methods**

700   **Animal experimentation**

701   All experiments involving animals were performed in agreement with the Swiss law on

702   animal protection (LPA), under license No GE 81/14 (to DD). Chick embryos from a

703   white Leghorn strain were incubated at 37.5°C and staged according to ref[51].

32

704

705   ***In situ* hybridization and *lacZ* staining**

706   Whole mount *in situ* hybridization was performed as described previously[64]. For *lacZ*

707   staining, embryos were fixed in 1x PBS pH7.39-7.41, 2mM $MgCl_2$, 4% PFA/PBS, 0.2%

708   glutaraldehyde and 5mM EDTA for 20 min at room temperature and washed three times

709   for 20 min in 1x PBS, 2mM $MgCl_2$, 0.2% NP40 and 0.01% sodium deoxycholate.

710   Samples were stained in 5mM potassium ferrocyanide, 5mM potassium ferricyanide and

711   0.5 mg/ml X-gal at room temperature overnight, followed by wash with the wash solution

712   once and refixation with 4%PFA/PBS at 4°C overnight.

713

714   **Micro-dissection chick limbs at HH30**

715   Limb tissues at HH30 were micro-dissected into acropod (distal), mesopod and zeugopod

716   (proximal) regions, respectively. Only distal and proximal regions were used for RNA-

717   seq and 4C-seq.

718

719   **RNA-seq and Data analysis**

720   Total RNA was extracted from mouse and chick limb bud tissues using the RNeasy Micro

721   Kit (QIAGEN) following the manufacturer instructions. Libraries were prepared with at

722   least 200 ng of total RNA following Illumina TruSeq stranded mRNA sample preparation

723   guide. Sequencings were performed with 100-bp or 75-bp single-end. The data was

724   mapped onto either GRCm38 (mm10) or the International Chicken Genome Reference

725   Consortium Gallus_gallus-5.0 (galGal5) using the Tophat2 (Version 2.0.9)[65] and

33

726   unique mapped reads were extracted. The number of uniquely mapped reads was

727   calculated using FLAGSTAT (SAMtools, Version 0.1.18)[66] and this value was used for

728   the subsequent normalization of each coverage data to be the million reads number. All

729   analyses were processed by our Galaxy server (the Bioteam Appliance Galaxy Edition,

730   https://bioteam.net, https://bioteam.net/products/galaxy-appliance)[67].

731

732   **RNA extraction and RT-qPCR**

733   Total RNA was extracted using the RNeasy Micro Kit (QIAGEN), following the

734   manufacturer instructions. 1 μg total RNA was used for cDNA synthesis with SuperScript

735   VILO (Invitrogen). RT-qPCR was performed on a CFX96 real-time system (BIORAD)

736   using the GoTaq qPCR Master Mix (Promega). Each RT-qPCR was carried out with at

737   least two biological replicates and each experimental information is described in S1 Table.

738   Primer sequences for qPCR are listed in S2 Table.

739

740   **4C-seq and Data analysis**

741   Each mouse and chick limb tissue were fixed separately with 2% formaldehyde, lysed

742   and stored at -80°C. Samples were digested with *NlaIII* and *DpnII* as primary and

743   secondary restriction enzymes, respectively, and ligation steps were performed using high

744   concentrated T4 DNA ligase (Promega)[68]. Inverse PCRs for amplification were carried

745   out using primers for each viewpoint[69](S3 Table). PCR products were multiplexed and

746   sequenced with 100bp single-end, followed by post-processing (de-multiplexing,

747   mapping, and 4C analysis) using the HTS station (http://htsstation.epfl.ch)[70]. Fragment

34

748   scores were normalized to the mean score of fragments falling into a region defined as

749   the bait coordinated ± 1 Mb, and the data were smoothed by using a running mean with

750   a window size of 11 fragments. The information regarding fragments excluded during the

751   procedure is provided in S3 Table. A relative quantification of the signal spanning both

752   the *HoxD* telomeric or centromeric domains was performed. Genomic coordinates for

753   calculations are:

754   mm10, chr2: 73,921,943-74,648,943 / chr2: 74,765,943-75,601,943;

755   galGal5, chr7: 15,920,642-16,318,067 / chr7: 16,414,183-16,699,172.

756   This quantification is not absolute and only reflects the balance of contacts between the

757   two domains, for each sample.

758

759   **ChIP-seq and Data analysis**

760   Chromatin   immunoprecipitation   experiments   were   performed   as   previously

761   described[20]. Micro-dissected limb tissues from mouse and chick embryos were cross-

762   linked with 1% formaldehyde/PBS for 15 min at room temperature. Chromatin was

763   sheared and used for each immunoprecipitation with anti-H3K27ac (ab4729, Abcam),

764   anti-H3K27me3 (07-449, Merck Millipore), anti-H3K4me1 (ab8895, Abcam) and anti-

765   CTCF (61311, Active Motif). Libraries were prepared with at least 5 ng of purified DNA

766   following Illumina TruSeq ChIP library preparation guide. Sequencing was performed

767   with 100bp single-end. Demultiplexed ChIP-seq reads were mapped onto the galGal5 or

768   mm10 using Bowtie (Version 0.12.7)[71], with parameters "-m1 –strata –best" according

769   to conditions described previously[72], and PCR duplicates were removed from mapped

35

770   reads using SAMtools (Version 0.1.18)[66]. By using bamCompare (Version 2.5.0), the

771   ChIP data from H3K27ac, H3K27me3 and H3K4me1 and the input data were normalized

772   and compared to compute the log2 of the number of reads ratio. The CTCF ChIP data was

773   normalized to obtain 1x depth of coverage by using bamCoverage (Version 2.5.0)[73,74].

774   The CTCF motif orientation analysis was performed as previously described[19]. All

775   analysis was done with our Galaxy server (the Bioteam Appliance Galaxy Edition,

776   https://bioteam.net, https://bioteam.net/products/galaxy-appliance)[67].

777

778   **SureSelect probe design and Capture Hi-C (CHi-C)**

779   The library of SureSelect enrichment probes were designed over the genomic interval

780   (galGal3: chr7:15,990,001-19,170,000) using the SureDesign online tool of Agilent.

781   Probes cover the entire genomic region (galGal5, chr.7: 14,946,000-17,870,000) and were

782   not designed specifically in proximity of *DpnII* sites. Dissected tissues were dissociated

783   in 10% FCS/PBS with collagenase (C7657, Sigma) to a final concentration of about 1.3

784   µg/µL, and samples were incubated in Thermomixer at 37°C at 800 rpm for 20 min. After

785   discarding the supernatant, cells were crosslinked with 1% formaldehyde/PBS at RT for

786   10 min, quenched with glycine and centrifuged to discard the supernatant. Cells were

787   resuspended with PBS containing proteinase inhibitor and then centrifuged again. After

788   removing supernatant, cells were kept at -80°C before use. Hi-C library preparation was

789   performed as described in[75], with the following changes: (i) Resuspended crosslinked

790   cells in ice-cold Lysis buffer were placed on a rotation wheel at 4°C at 30 rpm for 30 min

791   for cell lysis. (ii) For chromatin digestion, 400U of *DpnII* (R0543M, New England

36

792 Biolabs) were added to the samples and incubated at 37°C at 700 rpm for 4 hrs. Another

793 400U of *DpnII* were added and samples were incubated overnight. (iii) Blunt-end ligation

794 of biotin filled-in DNA was performed at RT at 30 rpm on rotating wheel for 4 hrs. (iv)

795 No removal of biotin from un-ligated ends was performed. (v) DNA was sheared to a size

796 of 200 to 800bp by using COVARIS E220, with the following conditions; 175W, 10%

797 Duty factor, 200 Cycles per Burst, 60 sec. (vi) DNA pulldown was performed by using

798 Dynabeads MyOne Streptavidin T1 (65601, Thermo Fisher). (vii) DNA was measured by

799 Qubit and 200ng were used for further treatment followed by manufacturer's protocol

800 (SureSelect^{XT} Target Enrichment System for Illumina Paired-End Multiplexed

801 Sequencing Library).

802

803 **CHi-C data analysis**

804 Paired-end sequencing data were processed as follows. First, adapters were removed

805 using   cutadapt   version   1.6[76]   with   the   following   parameters: -a

806 AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC   for   R1   and   -a

807 AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCG

808 TATCATT for R2. They were then processed by using hicup version 6.1.0 with the

809 bowtie2 version 2.2.6[77] and SAMtools version 1.2[66], with galGal5 as reference

810 genome and GATC as restriction enzyme recognition sequence. The pairs were next

811 converted from bam to tabulated files, with the position of the middle of the fragment to

812 which hicup assigned the read, by using an ad-hoc python script (available upon request).

813 Only valid pairs with both MAPQ above 30 were kept. Then, pairs with both mates in the

37

814   capture region (galGal5, chr7:14,946,000-17,870,000) were extracted and processed with

815   cooler to obtain a balance matrix of the capture region with 5kb bins. The Fig. 5 and S5

816   Fig were obtained with personal R scripts (available upon request) (http://www.R-

817   project.org). S5A and S5B Fig are the balanced matrices with linear scale. Fig 5 was

818   obtained by subtracting the two balanced matrices. To assess the significance of increased

819   contact between two regions, a Wilcoxon rank sum test was performed using R with the

820   values of the bins in the region of the two balanced matrices. Because 75% of valid pairs

821   MAPQ30 do not involve the capture region, all valid pairs were also processed with

822   cooler to obtain a balance matrix of the whole chromosome 2 at 40kb. These matrices

823   were used in S6 Fig. To define TAD borders (Fig 5), the TopDom algorithm[42] was run

824   with a window size of 28 from the 10kb binned balanced matrices, as gaps were too

825   numerous at a 5kb resolution.

826

827   **Mutant stocks**

828   The $HoxD^{Del(8-13)}$ and $HoxD^{Del(attp-SB3)}$ alleles have been previously described[17,78]. The

829   $HoxD^{Del(Mtx-Ttn)}$ allele was produced by TAMERE using the TiE2 (*Titin* exon 2) allele[79]

830   (kindly provided by Dr. Michael Gotthardt) and a *Mtx2* gene trap allele

831   (https://igtc.org/cgi-bin/annotation.py?cellline=CSI574). The sequences of genotyping

832   primers are indicated in S2 Table. All embryos analyzed in Fig 7 and S7 Fig were

833   heterozygotes, balanced by the $HoxD^{Del(8-13)}$ allele.

834

835   **Analysis of sequence alignment and limb enhancer prediction**

38

836   mVista, tools for comparative genomics was used for comparison between sequences of

837   the mouse *CS93* (mm10, chr2: 75,208,103-75,210,328), the bat BAR116 (Myoluc2,

838   GL429772: 6,606,808-6,608,652) and the 2kb region containing the chick *CS93* (galGal5,

839   chr7: 16,104,863-16,106,863), with LAGAN alignment program with default parameter

840   (http://genome.lbl.gov/vista/index.shtml). Potential limb enhancer regions were

841   identified by using Limb-Enhancer Genie tool, with the following condition: (i) analysis

842   type: Scan for top, (ii) method: Combined Model (https://leg.lbl.gov/)[54].

843

844   **Enhancer transgenic assays**

845   For the enhancer assays, embryos carrying the mouse *CS93/lacZ* and *HE1/lacZ* were

846   generated by lentivirus-mediated transgenesis and pronuclear injection, respectively. The

847   mouse *CS93* (mm10, chr2: 75,208,104-75,210,328) was amplified from C57BL/6

848   genomic DNA and cloned into the *pRRL-lacZ* vector, as described previously[17].

849   Lentiviruses were produced and injected into the peri-vitelline space of mouse

850   zygotes[32]. The mouse *HE1* (mm10, chr2: 75,959,179-75,960,378) and the region

851   containing the chick *CS93* sequence (galGal5, chr7: 16,104,863-16,106,863) were

852   obtained from B6CBAF1/J and White Leghorn genomic DNA, respectively, and cloned

853   into a *βglobin-lacZ* vector. The construct was injected into mouse oocytes. All transgenic

854   embryos were harvested between E10.5 and E12.5 and used for *lacZ* staining.

855

856   **Accession Numbers**

857   RNA-seq, 4C-seq, ChIP-seq and CHi-C datasets are available from the NCBI Gene

39

858   Expression Omnibus repository under accession number GSE115563. The public datasets

859   used in this research are listed in S4 Table.

860

861   **ACKNOWLEDGEMENTS**

871

872   **Author Contributions**

873   Conceptualization: N.Y.-K., G.A. and D.D.

874   Methodology: N.Y.-K., G.A., C.C.B. and D.D.

875   Investigation: N.Y.-K., C.C.B., G.A. and L.B.

876   Formal Analysis: N.Y.-K. and L.L.-D.

877   Writing-Original Draft: N.Y.-K., L.L.-D. and D.D.

878   Writing-Review and Editing: L.L.-D., G.A., L.B. and C.C.B.

879   Supervision and funding: D.D.

40

880   **LEGENDS TO FIGURES**

881   **Fig. 1. *Hoxd* genes expression in mouse and chick limb buds.** (A, B) *In situ*

882   hybridization analysis of E12.5 mouse and HH28 (equivalent to E12.25 to E12.5) chick

883   fore- and hindlimb buds showing the expression of *Hoxd* genes and *Col2a1* or *Aggrecan*,

884   which are markers for chondrocyte differentiation. (A, left) Schemes showing

885   morphologies of forelimb (red) and hindlimb (yellow) in mouse. (A, right) Expression

886   patterns of *Hoxd* genes in forelimb buds are comparable to those in hindlimb buds. The

887   expression domain of *Col2a1* (opened arrowheads) corresponds to the low *Hoxd*

888   expression region leading to the future mesopodium. (B, left) Schemes representing

889   morphologies of forelimb (wing, red) and hindlimb (leg, yellow) buds in chick. (B, right)

890   Expression of *Hoxd* genes in proximal leg is significantly reduced and restricted to the

891   presumptive ulna. (C, D) Transcription profiles of *Hoxd* genes in micro-dissected

892   proximal and distal domains, either from E12.5 mouse (C) or HH30 (equivalent to E13

893   to E13.5) chick (D) fore- and hindlimb buds. Each right limb in (A, B) is oriented with

894   proximal to the bottom and distal to the top. FL, forelimb; HL, hindlimb. The *Y* axis

895   represents the strand-specific RNA-seq read counts, normalized by the total number of

896   million mapped reads.

897

898   **Fig. 2. Conserved bimodal TAD regulatory structure at *HoxD* locus in tetrapod.**

899   (A, B) 4C-seq tracks showing contacts established by mouse *Hoxd11* and chick *Hoxd10-*

900   *11* viewpoint in mouse and chick proximal and distal cells from fore- and hindlimb at

901   E12.5 and HH30, respectively. (A) The interactions between *Hoxd11* and around *CS39*

41

902   were mainly observed in proximal cells, whereas those between *Hoxd11* and *island III* or

903   *Prox*, which are hallmarks of the C-DOM activity, were increased in distal region. (B)

904   Each contact extended to the predicted TAD border located on either side at *HoxD* cluster

905   (C-DOM side, opened arrowheads; T-DOM side, closed arrowheads). In addition to the

906   interactions between *Hoxd10-11* and *CS39*, contacts were also observed with *CS93* in

907   proximal wing bud cells, whereas it decreased in proximal leg bud cells where *Hoxd*

908   expression is strongly reduced (red arrows). FL, forelimb; HL, hindlimb.

909

910   **Fig. 3. Differential enhancer activities of mouse and chick *CS93* in fore- and**

911   **hindlimb buds.** (A) Genomic coordinates and sequence alignment using either the bat or

912   the chick sequence onto the mouse genome, and schematics of enhancer activities for

913   each of the identified sequences[6,17] and this work. Both mouse BAR116 and mouse

914   *CS9* showed no enhancer activity in limbs[6,17], whereas the bat BAR116 displayed

915   different patterns between mouse fore- and hindlimb[6]. The sequences of both the bat

916   BAR116 (Myoluc2, GL429772: 6,606,808-6,608,652) and the chick *CS93* (galGal5,

917   chr7:16,104,952-16,105,803) were aligned onto the mouse genome. (B) Coordinates of

918   either the chick *CS93* (green rectangle), or the 2kb large region used for enhancer assay

919   (blue domain). The 2kb sequence contains the chick *CS93* region and the region of high

920   interactions with the *Hoxd10* to *Hoxd11* region in proximal wing bud cells at HH30.

921   Conservation plot of mouse *CS93* and bat BAR116 using the 2kb region of chick *CS93*

922   as a reference. The peaks represent a conservation higher than 50%. Pink regions are

923   conserved non-coding sequences. The phylogenetic tree shows the highest conservation

42

924   of the chick CS93 with the bat BAR116 sequences. (C) Enhancer activities of mouse

925   *CS93* and chick *CS93* in mouse fore- and hindlimb buds from E10.5 to E12.5. The *lacZ*

926   expression pattern (left) showed that mouse *CS93* has an enhancer activity in the proximal

927   region of developing limb buds at E10.5 to E12.5. In contrast to the mouse, the chick

928   *CS93* (right) showed differential enhancer activity between fore- and hindlimb buds at

929   E12.5, as was also reported for the bat BAR116 sequence. The numbers of *lacZ* positive

930   embryos over total transgene integrated are indicated.

931

932   **Fig. 4. Premature termination of T-DOM activity in chick leg buds.** (A, B)

933   Comparison of H3K27ac and H3K27me3 ChIP-seq profiles in either whole, proximal or

934   distal wing- and leg buds at HH19 (equivalent to mouse E9.5), HH20 (equivalent to

935   mouse E10) and HH28. (A) In chick leg bud, enrichment of H3K27ac at region '*a*' in T-

936   DOM was initially detected at HH19, whereas it was significantly decreased at HH20.

937   Few H3K27ac marks were scored in region '*b*' in hindlimb bud at both HH19 and HH20,

938   as compared with those in wing buds. At HH28, the accumulation of H3K27ac marks was

939   quite low in both the *HoxD* cluster and the T-DOM region in proximal leg, when

940   compared to distal wing cells, whereas the profiles of H3K27ac in the distal region where

941   *Hoxd* genes are strongly expressed were similar between wing and leg buds at HH28. (B)

942   In whole wing- and leg buds at HH20, several C-DOM regions were covered by

943   H3K27me3, whereas T-DOM was either not yet labelled, or started to be enriched around

944   *CS39* in forelimb bud and hindlimb buds, respectively. In proximal leg buds where *Hoxd*

945   expression was reduced, H3K27me3 enrichment was observed at the *HoxD* cluster and

43

946   over T-DOM, when compared to proximal forelimb buds. Both regions '*a*' and '*b*' in T-

947   DOM seemed to be specific targets of H3K27me3 enrichment. Enrichment (*Y-axis*) of

948   ChIP is shown as the $\log_2$ ratio of the normalized number of reads between ChIP and input

949   samples. FL, forelimb; HL, hindlimb.

950

951   **Fig. 5. Subtraction of CHi-C matrices between wing- and leg buds with plotted TADs.**

952   (A) The black lines demarcate the TADs in wing- and leg buds as identified using the

953   TopDom algorithm. Subtraction are those of CHi-C matrices shown in S5A and S5B Fig,

954   with wing bud cells in red and leg bud cells in blue. (B) Subtraction of the CHi-C matrices

955   shown in S5A and S5B Fig at regions between the *HoxD* cluster and the area ranging

956   from region '*a*' to region '*b*' within T-DOM, with a 5kb resolution. A significant decrease

957   in contacts was detected between the *HoxD* cluster and the *CS93* region, which

958   corresponded to the reduction in H3K27ac levels (black rectangle). The increased

959   contacts involving the *HoxD* cluster shown in red and blue represent wing- and leg buds

960   cells, respectively. ChIP-seq profiles of CTCF and H3K27ac from wing- and leg buds at

961   HH20 are shown in red and blue, respectively. Opened and closed arrowheads indicate

962   the orientation of the CTCF motives. Enrichments (*Y-axis*) of CTCF and H3K27ac ChIP

963   are shown at the normalized 1x sequencing depth or the $\log_2$ ratio of the normalized

964   number of reads between ChIP and input samples, respectively. FL, forelimb; HL,

965   hindlimb.

966

967   **Fig. 6. Leg bud-specific early activation of the chick *Hoxa13*.** (A) Expression patterns

44

968   and levels of *Hoxa13* and *Hoxd11* in chick wing- and leg buds from HH20 to HH22.

969   Stronger expression of *Hoxa13* was observed in chick leg bud as compared to that in wing

970   bud (top). Expression level of *Hoxd11* increased in wing bud as development proceeded,

971   yet it decreased in leg bud (bottom). Expression levels were normalized to *Gapdh* and

972   depicted as fold change relative to wing bud at HH20-21. Error bars indicate standard

973   deviation of three biological replicates. **P<0.01, Student's *t*-test. (B) Expression of

974   *Hoxa13* and *Hoxd11* in mouse fore- and hindlimb buds, from E10.5 to E10.75. Both genes

975   were increasingly transcribed in fore- and hindlimb buds as development proceeded.

976   Expression levels were normalized to *Gapdh* and depicted as fold change relative to

977   forelimb buds at E10.5. Error bars indicate standard deviation of two or four biological

978   replicates. **P<0.01, NS, P>0.05, Student's *t*-test.

979

980   **Fig. 7. Different effects of a T-DOM deletion on fore- and hindlimb buds.** (A) The

981   *HoxD$^{Del(attp-SB3)}$* and *HoxD$^{Del(Mtx-Ttn)}$* alleles were constructed by deleting T-DOM through

982   a ca. 1Mb or 2.1Mb large deletion, respectively (left, dashed line). *Hoxd11* expression in

983   E12.5 fore- and hindlimbs from either control (*HoxD$^{Del(8-13)/+}$*) animals (indicated as

984   '+/Δ') or mutants (*HoxD$^{Del(attp-SB3)/Del(8-13)}$, HoxD$^{Del(Mtx-Ttn)/Del(8-13)}$, HoxD$^{Del(8-13)/+}$;HoxC$^{-/-}$,*

985   *HoxD$^{Del(attp-SB3)/Del(8-13)}$;HoxC$^{-/-}$*) littermates (indicated as '*Del(attp-SB3)/Δ*', '*Del(Mtx-*

986   *Ttn)/Δ*', '*+/Δ/HoxC$^{-/-}$*', '*Del(attp-SB3)/Δ/HoxC$^{-/-}$*'). In *Del(attp-SB3)/Δ* mutant, *Hoxd11*

987   expression was dramatically reduced in proximal forelimbs (arrowhead). Unexpectedly

988   however, expression remained robust in proximal hindlimbs (arrowhead). In *HoxD$^{Del(Mtx-}$*

989    $^{TiE2)}$ mutant, *Hoxd11* expression disappeared in both proximal fore- and hindlimb buds

45

990   (arrowhead). The absence of both T-DOM and the *HoxC* cluster did not affect *Hoxd11*

991   expression. (B) HiC data adapted from ref[19] showing the two TADs on either side of

992   the *HoxD* cluster and the TAD next to T-DOM. 4C profiles represent contacts established

993   by *Hoxd11* in proximal fore- and hindlimb buds, from control or *Del(attp-SB3)/Δ* mutant

994   animals. In mutant cells deleted for T-DOM, (tracks 2 and 4), ectopic contacts between

995   *Hoxd11* and regions in the neighboring TAD appeared due to their new proximity to *HoxD*.

996   The shaded region with red arrows shows the domains where more contacts were detected

997   in mutant proximal hindlimb versus proximal forelimb buds. (C) Enlargement of 4C

998   profiles dispatched in (B), DNaseI, hyper-sensitive profiles using E11.5 embryos and

999   potential limb enhancer regions (pink rectangles), as identified by using Limb-Enhancer

1000   Genie tool. Potential hindlimb-specific enhancers (HE) are shown by red arrows. (D)

1001   Mouse HE1 is active in the proximal fore- and hindlimb buds and in the trunk at E12.5.

1002   The number indicate stained embryos over total number of integrations. PFL, proximal

1003   forelimb; PHL, proximal hindlimb; HL, hindlimb.

1004

1005   **Fig. 8. TAD boundaries at the mouse and chicken *HoxD* cluster and effect of T-DOM**

1006   **deletion.** (A, B) TADs boundaries at the *HoxD* locus in mouse (A) and chick (B) limb

1007   buds. (A) In the mouse, the boundary moves along a few genes, in a window likely

1008   determined by a series of CTCF sites. In chick, the boundary appears less robust and

1009   slightly expanded towards the *Hoxd13* locus. This latter situation may enable T-DOM

1010   enhancers to interact with *Hoxd13* more efficiently in chick than in murine cells (B).

1011   Black and white arrowheads indicate the orientation of CTCF motives. (C) The different

1012   effects of T-DOM deletion in mouse proximal fore- and hindlimb cells. The

1013   transcriptional status of *Hoxd* genes is different between forelimb and hindlimb buds due

1014   to their prior expression or non-expression in lateral plate mesoderm, respectively. In the

1015   absence of T-DOM, H3K27me3 marks may accumulate over the silenced *Hoxd* genes in

1016   mutant forelimb buds as described in [17], making them poorly responsive to the HE1

1017   enhancer. In contrast, mutant hindlimb bud proximal cells have their *Hoxd* genes poised

1018   for activation and *de novo* interactions with the HE1 enhancer may occur. H3K27me3

1019   marks are in blue and presumptive factors, such as a potential mediator of DNA looping

1020   and a hindlimb-specific transcription factor, are shown by a green oval and a yellow star.

1021

1022   **SUPPORTING INFORMATION**

1023   **S1 Fig (related to Fig 1).** *Hoxa* genes expression in mouse and chick limb buds.

1024   **S2 Fig (related to Fig 2).** The regulatory switch between TADs in mouse and chick limb

1025   buds.

1026   **S3 Fig (related to Fig 4).** H3K27ac, H3K27me3 and RNA-seq at *HoxD* in chick limbs.

1027   **S4 Fig (related to Fig 4).** H3K27ac and H3K27me3 profiles of ChIP-seq and RNA-seq

1028   at the *HoxA* locus in chick limbs.

1029   **S5 Fig (related to Fig 5).** Chromatin conformation at the chick *HoxD* locus in wing- and

1030   leg buds and conservation of CTCF sites orientations.

1031   **S6 Fig (related to Fig 6).** Expression of *Hoxa* and *Hoxd* genes in chick and mouse limb

1032   buds.

1033   **S7 Fig (related to Fig 7).** A T-DOM deletion induces interactions between HE1 and *Hoxd*

1034   genes.

1035   **S1 Table**

1036   Information about samples.

1037   **S2 Table**

1038   DNA sequences of primers used for both RT-qPCR analyses and genotyping.

1039   **S3 Table**

1040   DNA sequences of primers used for 4C analyses. Custom barcodes (4bp shown by

1041   NNNN) were introduced in between the Illumina adapter sequence and the specific

1042   viewpoint sequences to do multiplexing using different samples with the same viewpoint.

1043   **S4 Table**

1044   Public datasets used in this research.

1045

1046   **References**

1047   1.   Zeller R, Lopez-Rios J, Zuniga A. Vertebrate limb bud development: moving
1048   towards integrative analysis of organogenesis. Nature reviews Genetics. 2009;10: 845–
1049   58. doi:10.1038/nrg2681

1050   2.   Blanco MJ, Misof BY, Wagner GP. Heterochronic differences of Hoxa-11
1051   expression in Xenopus fore- and hind limb development: evidence for lower limb identity
1052   of the anuran ankle bones. Dev Genes Evol. 1998;208: 175–87.

1053   3.   Noro M, Uejima A, Abe G, Manabe M, Tamura K. Normal developmental stages
1054   of the Madagascar ground gecko Paroedura pictus with special reference to limb
1055   morphogenesis. Dev Dyn. 2009;238: 100–109. doi:10.1002/dvdy.21828

1056   4.   Sears KE, Behringer RR, Rasweiler JJ, Niswander LA. Development of bat
1057   flight: morphologic and molecular evolution of bat wing digits. Proc Natl Acad Sci USA.
1058   2006;103: 6581–6586. doi:10.1073/pnas.0509716103

1059   5.   Seki R, Kamiyama N, Tadokoro A, Nomura N, Tsuihiji T, Manabe M, et al.

1060 Evolutionary and developmental aspects of avian-specific traits in limb skeletal pattern.

1061 Zoolog Sci. 2012;29: 631–44. doi:10.2108/zsj.29.631

1062 6.   Booker BM, Friedrich T, Mason MK, VanderMeer JE, Zhao J, Eckalbar WL, et

1063 al. Bat Accelerated Regions Identify a Bat Forelimb Specific Enhancer in the HoxD

1064 Locus. PLoS Genet. 2016;12: e1005738. doi:10.1371/journal.pgen.1005738

1065 7.   Seki R, Li C, Fang Q, Hayashi S, Egawa S, Hu J, et al. Functional roles of Aves

1066 class-specific cis-regulatory elements on macroevolution of bird-specific features. Nat

1067 Commun. 2017;8: 14229. doi:10.1038/ncomms14229

1068 8.   Domyan ET, Kronenberg Z, Infante CR, Vickrey AI, Stringham SA, Bruders R,

1069 et al. Molecular shifts in limb identity underlie development of feathered feet in two

1070 domestic avian species. Elife. 2016;5: e12115. doi:10.7554/eLife.12115

1071 9.   Mallo M. Reassessing the Role of Hox Genes during Vertebrate Development

1072 and Evolution. Trends Genet. 2018;34: 209–217. doi:10.1016/j.tig.2017.11.007

1073 10.   Wellik DM. Hox patterning of the vertebrate axial skeleton. Dev Dyn. 2007;236:

1074 2454–2463. doi:10.1002/(ISSN)1097-0177

1075 11.   Wellik DM, Capecchi MR. Hox10 and Hox11 genes are required to globally

1076 pattern   the   mammalian   skeleton.   Science.   2003;301:   363–7.

1077 doi:10.1126/science.1085672

1078 12.   Zakany J, Duboule D. The role of Hox genes during vertebrate limb development.

1079 Current   Opinion   in   Genetics   &   Development.   2007;17:   359–66.

1080 doi:10.1016/j.gde.2007.05.011

1081 13.   Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, et al. Spatial

1082 partitioning of the regulatory landscape of the X-inactivation centre. Nature. 2012;485:

1083 381–5. doi:10.1038/nature11049

1084 14.   Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, et al. Topological domains in

1085 mammalian genomes identified by analysis of chromatin interactions. Nature. 2012;485:

1086 376–80. doi:10.1038/nature11082

1087 15.   Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M, et al.

1088 Three-dimensional folding and functional organization principles of the Drosophila

1089 genome. Cell. 2012;148: 458–72. doi:10.1016/j.cell.2012.01.010

1090 16.   Berlivet S, Paquette D, Dumouchel A, Langlais D, Dostie J, Kmita M. Clustering

1091 of tissue-specific sub-TADs accompanies the regulation of HoxA genes in developing

1092 limbs. PLoS genetics. 2013;9: e1004018. doi:10.1371/journal.pgen.1004018

1093 17.   Andrey G, Montavon T, Mascrez B, Gonzalez F, Noordermeer D, Leleu M, et al.
1094 A switch between topological domains underlies HoxD genes collinearity in mouse limbs.
1095 Science. 2013;340: 1234167. doi:10.1126/science.1234167

1096 18.   Montavon T, Soshnikova N, Mascrez B, Joye E, Thevenet L, Splinter E, et al. A
1097 regulatory archipelago controls Hox genes transcription in digits. Cell. 2011;147: 1132–
1098 45. doi:10.1016/j.cell.2011.10.023

1099 19.   Rodriguez-Carballo E, Lopez-Delisle L, Zhan Y, Fabre PJ, Beccari L, El-Idrissi
1100 I, et al. The HoxD cluster is a dynamic and resilient TAD boundary controlling the
1101 segregation of antagonistic regulatory landscapes. Genes Dev. 2017;31: 2264–2281.
1102 doi:10.1101/gad.307769.117

1103 20.   Beccari L, Yakushiji-Kaminatsui N, Woltering JM, Necsulea A, Lonfat N,
1104 Rodriguez-Carballo E, et al. A role for HOX13 proteins in the regulatory switch between
1105 TADs at the HoxD locus. Genes Dev. 2016;30: 1172–86. doi:10.1101/gad.281055.116

1106 21.   Woltering JM, Noordermeer D, Leleu M, Duboule D. Conservation and
1107 divergence of regulatory strategies at Hox Loci and the origin of tetrapod digits. PLoS
1108 Biol. 2014;12: e1001773. doi:10.1371/journal.pbio.1001773

1109 22.   Guerreiro I, Gitto S, Novoa A, Codourey J, Nguyen Huynh TH, Gonzalez F, et
1110 al. Reorganisation of Hoxd regulatory landscapes during the evolution of a snake-like
1111 body plan. eLife. 2016;5. doi:10.7554/eLife.16087

1112 23.   Kamiyama N, Seki R, Yokoyama H, Tamura K. Heterochronically early decline
1113 of Hox expression prior to cartilage formation in the avian hindlimb zeugopod. Dev
1114 Growth Differ. 2012;54: 619–32. doi:10.1111/j.1440-169X.2012.01359.x

1115 24.   Flöttmann R, Kragesteen BK, Geuer S, Socha M, Allou L, Sowińska-Seidler A,
1116 et al. Noncoding copy-number variations are associated with congenital limb
1117 malformation. Genet Med. 2017; doi:10.1038/gim.2017.154

1118 25.   Villavicencio-Lorini P, Kuss P, Friedrich J, Haupt J, Farooq M, Turkmen S, et al.
1119 Homeobox genes d11-d13 and a13 control mouse autopod cortical bone and joint
1120 formation. The Journal of clinical investigation. 2010;120: 1994–2004.
1121 doi:10.1172/JCI41554

1122 26.   Yokouchi Y, Nakazato S, Yamamoto M, Goto Y, Kameda T, Iba H, et al.
1123 Misexpression of Hoxa-13 induces cartilage homeotic transformation and changes cell
1124 adhesiveness in chick limb buds. Genes & development. 1995;9: 2509–22.

1125 27.   Tschopp P, Duboule D. A regulatory "landscape effect" over the HoxD cluster.

1126   Developmental biology. 2011;351: 288–96. doi:10.1016/j.ydbio.2010.12.034

1127   28.   Spitz F, Montavon T, Monso-Hinard C, Morris M, Ventruto M-L, Antonarakis S,

1128   et al. A t(2;8) balanced translocation with breakpoints near the human HOXD complex

1129   causes mesomelic dysplasia and vertebral defects. Genomics. 2002;79: 493–8.

1130   doi:10.1006/geno.2002.6735

1131   29.   Gonzalez-Martin MC, Mallo M, Ros MA. Long bone development requires a

1132   threshold of Hox function. Developmental biology. 2014;392: 454–65.

1133   doi:10.1016/j.ydbio.2014.06.004

1134   30.   Herault Y, Fraudeau N, Zakany J, Duboule D. Ulnaless (Ul), a regulatory

1135   mutation inducing both loss-of-function and gain-of-function of posterior Hoxd genes.

1136   Development. 1997;124: 3493–500.

1137   31.   Nelson CE, Morgan BA, Burke AC, Laufer E, DiMambro E, Murtaugh LC, et

1138   al. Analysis of Hox gene expression in the chick limb bud. Development. 1996;122:

1139   1449–66.

1140   32.   Friedli M, Barde I, Arcangeli M, Verp S, Quazzola A, Zakany J, et al. A

1141   systematic enhancer screen using lentivector transgenesis identifies conserved and non-

1142   conserved functional elements at the Olig1 and Olig2 locus. PLoS ONE. 2010;5: e15741.

1143   doi:10.1371/journal.pone.0015741

1144   33.   Lois C, Hong EJ, Pease S, Brown EJ, Baltimore D. Germline transmission and

1145   tissue-specific expression of transgenes delivered by lentiviral vectors. Science.

1146   2002;295: 868–72. doi:10.1126/science.1067081

1147   34.   Martin P. Tissue patterning in the developing mouse limb. Int J Dev Biol.

1148   1990;34: 323–336.

1149   35.   Fernández-Terán MA, Hinchliffe JR, Ros MA. Birth and death of cells in limb

1150   development: a mapping study. Dev Dyn. 2006;235: 2521–2537.

1151   doi:10.1002/dvdy.20916

1152   36.   Tie F, Banerjee R, Stratton CA, Prasad-Sinha J, Stepanik V, Zlobin A, et al. CBP-

1153   mediated acetylation of histone H3 lysine 27 antagonizes Drosophila Polycomb silencing.

1154   Development. 2009;136: 3131–41. doi:10.1242/dev.037127

1155   37.   Bonev B, Mendelson Cohen N, Szabo Q, Fritsch L, Papadopoulos GL, Lubling

1156   Y, et al. Multiscale 3D Genome Rewiring during Mouse Neural Development. Cell.

1157   2017;171: 557-572.e24. doi:10.1016/j.cell.2017.09.043

1158   38.   Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT,

51

1159   et al. A 3D map of the human genome at kilobase resolution reveals principles of

1160   chromatin looping. Cell. 2014;159: 1665–80. doi:10.1016/j.cell.2014.11.021

1161   39.   Deng W, Rupon JW, Krivega I, Breda L, Motta I, Jahn KS, et al. Reactivation of

1162   developmentally silenced globin genes by forced chromatin looping. Cell. 2014;158:

1163   849–860. doi:10.1016/j.cell.2014.05.050

1164   40.   Dryden NH, Broome LR, Dudbridge F, Johnson N, Orr N, Schoenfelder S, et al.

1165   Unbiased analysis of potential targets of breast cancer susceptibility loci by Capture Hi-

1166   C. Genome Res. 2014;24: 1854–1868. doi:10.1101/gr.175034.114

1167   41.   Hughes JR, Roberts N, McGowan S, Hay D, Giannoulatou E, Lynch M, et al.

1168   Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-

1169   throughput experiment. Nat Genet. 2014;46: 205–212. doi:10.1038/ng.2871

1170   42.   Shin H, Shi Y, Dai C, Tjong H, Gong K, Alber F, et al. TopDom: an efficient and

1171   deterministic method for identifying topological domains in genomes. Nucleic acids

1172   research. 2016;44: e70. doi:10.1093/nar/gkv1505

1173   43.   Soshnikova N, Montavon T, Leleu M, Galjart N, Duboule D. Functional analysis

1174   of CTCF during mammalian limb development. Dev Cell. 2010;19: 819–30.

1175   doi:10.1016/j.devcel.2010.11.009

1176   44.   Narendra V, Bulajić M, Dekker J, Mazzoni EO, Reinberg D. CTCF-mediated

1177   topological boundaries during development foster appropriate gene regulation. Genes

1178   Dev. 2016;30: 2657–2662. doi:10.1101/gad.288324.116

1179   45.   Darbellay F, Duboule D. Topological Domains, Metagenes, and the Emergence

1180   of Pleiotropic Regulations at Hox Loci. Current topics in developmental biology.

1181   2016;116: 299–314. doi:10.1016/bs.ctdb.2015.11.022

1182   46.   Merkenschlager M, Nora EP. CTCF and Cohesin in Genome Folding and

1183   Transcriptional Gene Regulation. Annual review of genomics and human genetics.

1184   2016;17: 17–43. doi:10.1146/annurev-genom-083115-022339

1185   47.   Sanborn AL, Rao SSP, Huang S-C, Durand NC, Huntley MH, Jewett AI, et al.

1186   Chromatin extrusion explains key features of loop and domain formation in wild-type and

1187   engineered   genomes.   Proc   Natl   Acad   Sci   USA.   2015;112:   E6456-6465.

1188   doi:10.1073/pnas.1518552112

1189   48.   Andrey G, Schöpflin R, Jerković I, Heinrich V, Ibrahim DM, Paliou C, et al.

1190   Characterization of hundreds of regulatory landscapes in developing limbs reveals two

1191   regimes   of   chromatin   folding.   Genome   Res.   2017;27:   223–233.

1192   doi:10.1101/gr.213066.116

1193   49.   Sheth R, Barozzi I, Langlais D, Osterwalder M, Nemec S, Carlson HL, et al.
1194   Distal Limb Patterning Requires Modulation of cis-Regulatory Activities by HOX13.
1195   Cell Rep. 2016;17: 2913–2926. doi:10.1016/j.celrep.2016.11.039

1196   50.   Towers M, Tickle C. Growing models of vertebrate limb development.
1197   Development. 2009;136: 179–90. doi:10.1242/dev.024158

1198   51.   Hamburger V, Hamilton HL. A series of normal stages in the development of the
1199   chick embryo. 1951. Developmental dynamics : an official publication of the American
1200   Association of Anatomists. 1992;195: 231–72. doi:10.1002/aja.1001950404

1201   52.   Mercader N, Selleri L, Criado LM, Pallares P, Parras C, Cleary ML, et al. Ectopic
1202   Meis1 expression in the mouse limb bud alters P-D patterning in a Pbx1-independent
1203   manner. Int J Dev Biol. 2009;53: 1483–94. doi:10.1387/ijdb.072430nm

1204   53.   ENCODE Project Consortium. An integrated encyclopedia of DNA elements in
1205   the human genome. Nature. 2012;489: 57–74. doi:10.1038/nature11247

1206   54.   Monti R, Barozzi I, Osterwalder M, Lee E, Kato M, Garvin TH, et al. Limb-
1207   Enhancer Genie: An accessible resource of accurate enhancer predictions in the
1208   developing   limb.   PLoS   Comput   Biol.   2017;13:   e1005720.
1209   doi:10.1371/journal.pcbi.1005720

1210   55.   Peterson RL, Papenbrock T, Davda MM, Awgulewitsch A. The murine Hoxc
1211   cluster contains five neighboring AbdB-related Hox genes that show unique spatially
1212   coordinated expression in posterior embryonic subregions. Mech Dev. 1994;47: 253–60.

1213   56.   Davis AP, Witte DP, Hsieh-Li HM, Potter SS, Capecchi MR. Absence of radius
1214   and ulna in mice lacking hoxa-11 and hoxd-11. Nature. 1995;375: 791–5.
1215   doi:10.1038/375791a0

1216   57.   Suemori H, Noguchi S. Hox C cluster genes are dispensable for overall body
1217   plan of mouse embryonic development. Dev Biol. 2000;220: 333–42.
1218   doi:10.1006/dbio.2000.9651

1219   58.   Carpenter EM, Goddard JM, Davis AP, Nguyen TP, Capecchi MR. Targeted
1220   disruption of Hoxd-10 affects mouse hindlimb development. Development. 1997;124:
1221   4505–14.

1222   59.   Rodrigues AR, Tabin CJ. Developmental biology. Deserts and waves in gene
1223   expression. Science. 2013;340: 1181–2. doi:10.1126/science.1239867

1224   60.   Itzkovitz B, Jiralerspong S, Nimmo G, Loscalzo M, Horovitz DDG, Snowden A,

53

1225 et al. Functional characterization of novel mutations in GNPAT and AGPS, causing

1226 rhizomelic chondrodysplasia punctata (RCDP) types 2 and 3. Hum Mutat. 2012;33: 189–

1227 197. doi:10.1002/humu.21623

1228 61. Liegel RP, Ronchetti A, Sidjanin DJ. Alkylglycerone phosphate synthase

1229 (AGPS) deficient mice: models for rhizomelic chondrodysplasia punctate type 3

1230 (RCDP3) malformation syndrome. Mol Genet Metab Rep. 2014;1: 299–311.

1231 doi:10.1016/j.ymgmr.2014.06.003

1232 62. Weintraub AS, Li CH, Zamudio AV, Sigova AA, Hannett NM, Day DS, et al.

1233 YY1 Is a Structural Regulator of Enhancer-Promoter Loops. Cell. 2017;171: 1573-

1234 1588.e28. doi:10.1016/j.cell.2017.11.008

1235 63. Infante CR, Park S, Mihala AG, Kingsley DM, Menke DB. Pitx1 broadly

1236 associates with limb enhancers and is enriched on hindlimb cis-regulatory elements. Dev

1237 Biol. 2013;374: 234–244. doi:10.1016/j.ydbio.2012.11.017

1238 64. Woltering JM, Vonk FJ, Muller H, Bardine N, Tuduce IL, de Bakker MA, et al.

1239 Axial patterning in snakes and caecilians: evidence for an alternative interpretation of the

1240 Hox code. Developmental biology. 2009;332: 82–9. doi:10.1016/j.ydbio.2009.04.031

1241 65. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2:

1242 accurate alignment of transcriptomes in the presence of insertions, deletions and gene

1243 fusions. Genome biology. 2013;14: R36. doi:10.1186/gb-2013-14-4-r36

1244 66. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence

1245 Alignment/Map format and SAMtools. Bioinformatics. 2009;25: 2078–9.

1246 doi:10.1093/bioinformatics/btp352

1247 67. Afgan E, Baker D, van den Beek M, Blankenberg D, Bouvier D, Čech M, et al.

1248 The Galaxy platform for accessible, reproducible and collaborative biomedical analyses:

1249 2016 update. Nucleic Acids Res. 2016;44: W3–W10. doi:10.1093/nar/gkw343

1250 68. Noordermeer D, Leleu M, Schorderet P, Joye E, Chabaud F, Duboule D.

1251 Temporal dynamics and developmental memory of 3D chromatin architecture at Hox

1252 gene loci. eLife. 2014;3: e02557. doi:10.7554/eLife.02557

1253 69. Noordermeer D, Leleu M, Splinter E, Rougemont J, De Laat W, Duboule D. The

1254 dynamic architecture of Hox gene clusters. Science. 2011;334: 222–5.

1255 doi:10.1126/science.1207194

1256 70. David FP, Delafontaine J, Carat S, Ross FJ, Lefebvre G, Jarosz Y, et al.

1257 HTSstation: a web application and open-access libraries for high-throughput sequencing

54

1258   data analysis. PloS one. 2014;9: e85879. doi:10.1371/journal.pone.0085879

1259   71.   Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient

1260   alignment of short DNA sequences to the human genome. Genome biology. 2009;10: R25.

1261   doi:10.1186/gb-2009-10-3-r25

1262   72.   Riising EM, Comet I, Leblanc B, Wu X, Johansen JV, Helin K. Gene silencing

1263   triggers polycomb repressive complex 2 recruitment to CpG islands genome wide.

1264   Molecular cell. 2014;55: 347–60. doi:10.1016/j.molcel.2014.06.005

1265   73.   Ramirez F, Dundar F, Diehl S, Gruning BA, Manke T. deepTools: a flexible

1266   platform for exploring deep-sequencing data. Nucleic acids research. 2014;42: W187-91.

1267   doi:10.1093/nar/gku365

1268   74.   Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, et al.

1269   deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic

1270   acids research. 2016;44: W160-5. doi:10.1093/nar/gkw257

1271   75.   Belaghzal H, Dekker J, Gibcus JH. Hi-C 2.0: An optimized Hi-C procedure for

1272   high-resolution genome-wide mapping of chromosome conformation. Methods.

1273   2017;123: 56–65. doi:10.1016/j.ymeth.2017.04.004

1274   76.   Martin M. Cutadapt removes adapter sequences from high-throughput

1275   sequencing reads. 2011. 2011;17. doi:10.14806/ej.17.1.200

1276   77.   Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat

1277   Methods. 2012;9: 357–9. doi:10.1038/nmeth.1923

1278   78.   Tarchini B, Huynh TH, Cox GA, Duboule D. HoxD cluster scanning deletions

1279   identify multiple defects leading to paralysis in the mouse mutant Ironside. Genes &

1280   development. 2005;19: 2862–76. doi:10.1101/gad.351105

1281   79.   Radke MH, Peng J, Wu Y, McNabb M, Nelson OL, Granzier H, et al. Targeted

1282   deletion of titin N2B region leads to diastolic dysfunction and cardiac atrophy. Proc Natl

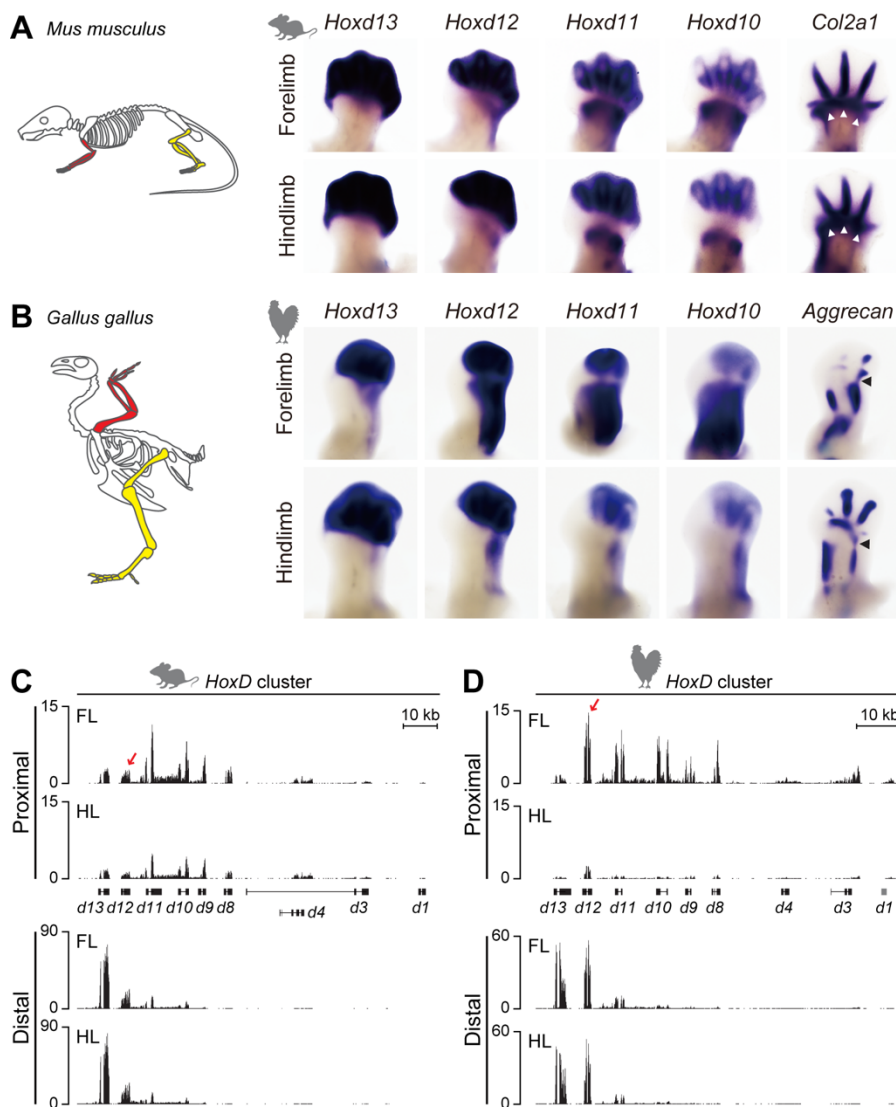1283   Acad Sci USA. 2007;104: 3444–3449. doi:10.1073/pnas.0608543104
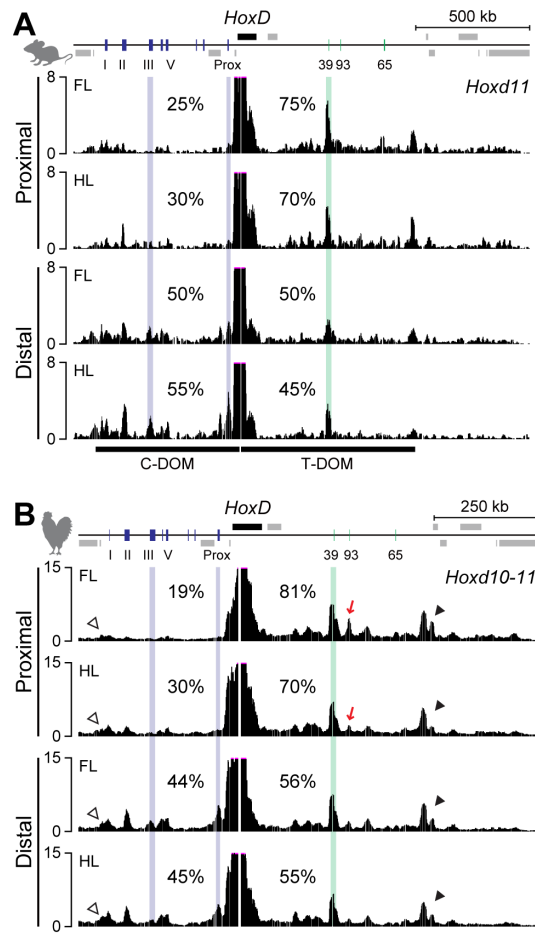
1284

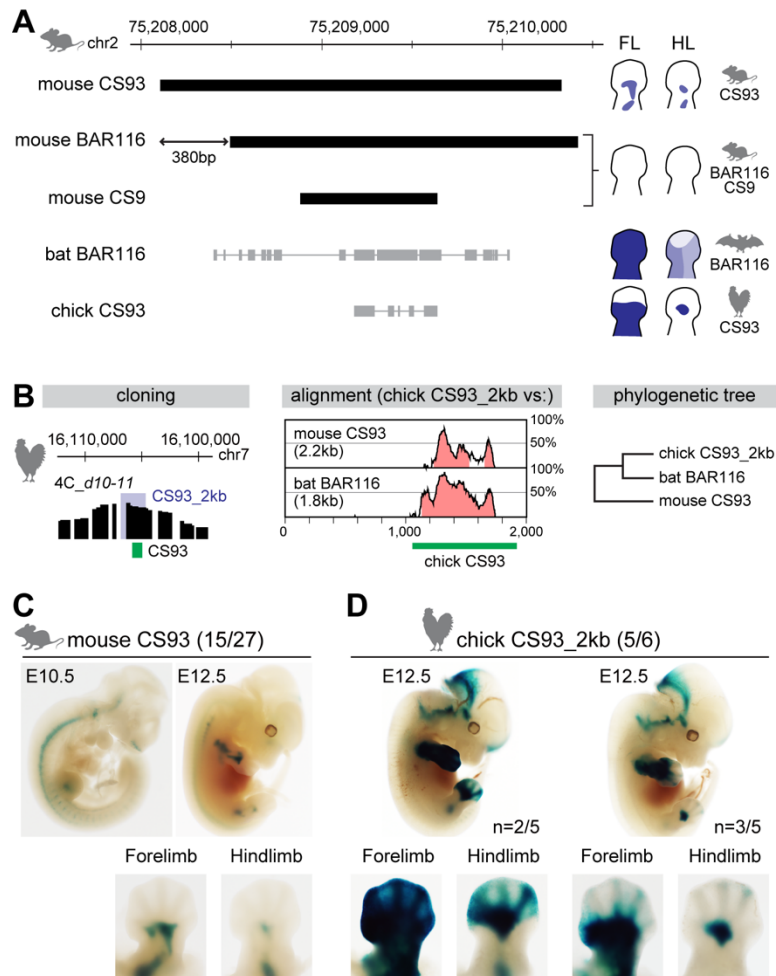Fig 1. Yakushiji-Kaminatsui et al.

Fig 2. Yakushiji-Kaminatsui et al.

A

chr2   75,208,000   75,209,000   75,210,000   FL   HL

mouse CS93

mouse BAR116   380bp

mouse CS9

bat BAR116   BAR116

chick CS93   CS93

B

| cloning | alignment (chick CS93_2kb vs:) | phylogenetic tree |

16,110,000   16,100,000   chr7

4C_d10-11   CS93_2kb

CS93

mouse CS93 (2.2kb)   100% 50%

bat BAR116 (1.8kb)   100% 50%

0   1,000   2,000

chick CS93

chick CS93_2kb

bat BAR116

mouse CS93

C   mouse CS93 (15/27)

E10.5   E12.5

Forelimb   Hindlimb

D   chick CS93_2kb (5/6)

E12.5   E12.5

n=2/5   n=3/5

Forelimb   Hindlimb   Forelimb   Hindlimb

Fig 3. Yakushiji-Kaminatsui et al.

A

HoxD

250 kb

I II III V   Prox   39 93   65

H3K27ac

Whole limb
- HH19FL
- HH19HL
- HH20FL
- HH20HL

Proximal
- HH28FL
- HH28HL

Distal
- HH28FL
- HH28HL

a   *b*

B

HoxD

250 kb

I II III V   Prox   39 93   65

H3K27me3

Whole limb
- HH20FL
- HH20HL

Proximal
- HH28FL
- HH28HL

Distal
- HH28FL
- HH28HL

a   *b*

Fig 4. Yakushiji-Kaminatsui et al.

A

FL 0.003
0
-0.003
HL

250 kb

I II III V Prox *HoxD* 39 93 65

B

16.25M   16.15M   16.05M   15.95M

16.4M

d13
10 8
4
d1

16.3M

CTCF

39 93 65
a *b*

FL 0.003
0
-0.003
HL

CTCF

K27ac

Fig 5. Yakushiji-Kaminatsui et al.

Fig 6. Yakushiji-Kaminatsui et al.

Fig 7. Yakushiji-Kaminatsui et al.
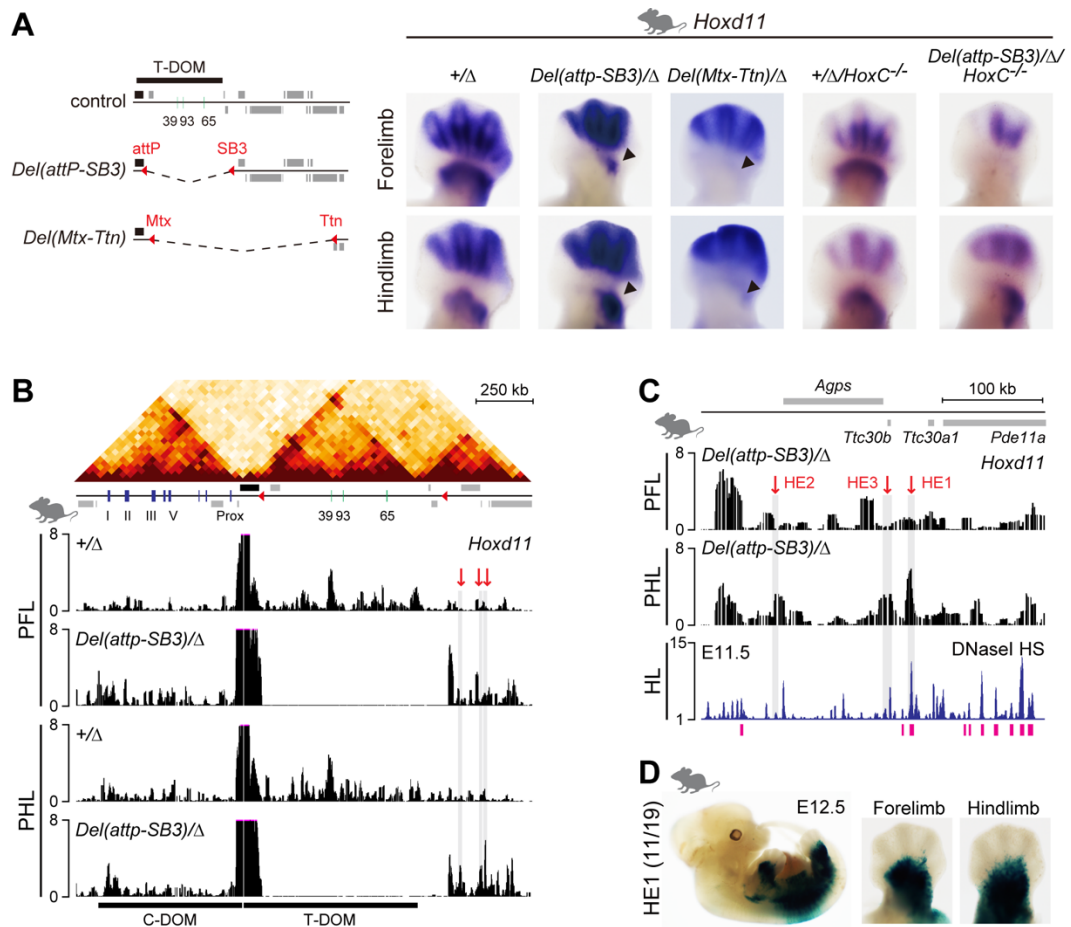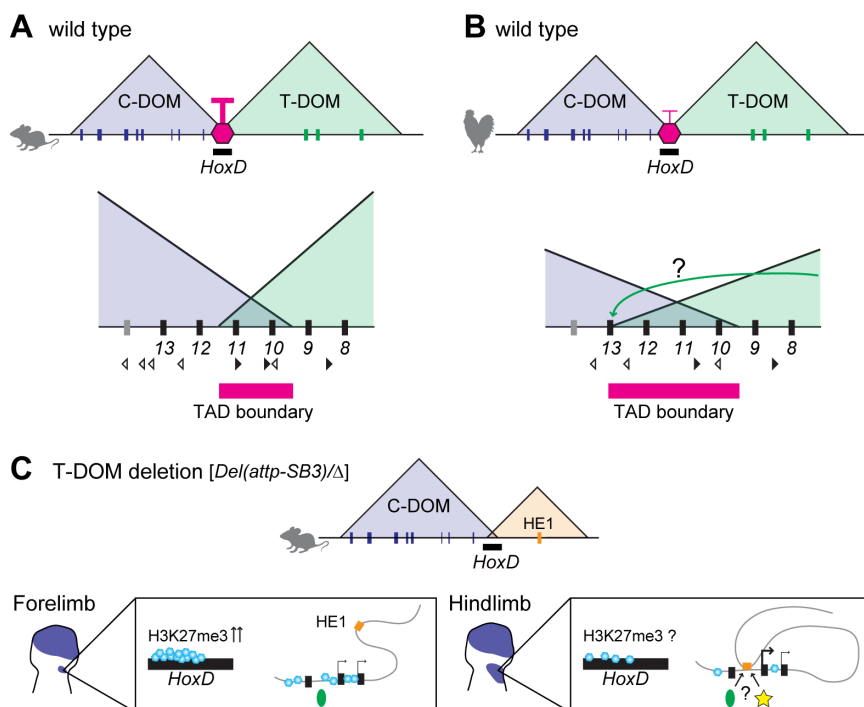
A wild type

C-DOM
T-DOM
HoxD

13 *12* *11* *10* *9* *8*
TAD boundary

B wild type

C-DOM
T-DOM
HoxD

?

13 *12* *11* *10* *9* *8*
TAD boundary

C T-DOM deletion [*Del(attp-SB3)/Δ*]

C-DOM
HE1
HoxD

Forelimb
H3K27me3 ↑↑
HE1
HoxD

Hindlimb
H3K27me3 ?
HoxD
?

Fig 8. Yakushiji-Kaminatsui et al.

1   CONSERVATION AND VARIATIONS OF BIMODAL *HoxD* GENE REGULATION

2   DURING TETRAPOD LIMB DEVELOPMENT

3

4   Nayuta Yakushiji-Kaminatsui[1], Lucille Lopez-Delisle[1,3], Christopher Chase Bolt[1,3],

5   Guillaume Andrey[1,∞], Leonardo Beccari[2] and Denis Duboule[1,2,*]

6

7   [1]School of Life Sciences, Federal Institute of Technology, Lausanne, 1015 Lausanne,

8   Switzerland, [2]Department of Genetics and Evolution, University of Geneva, 1211 Geneva

9   4, Switzerland.

10   [∞]Present address: Max Planck Institute for Molecular Genetics, RG Development &

11   Disease, 14195 Berlin, Germany,

12

13   [3]These authors contributed equally.

14

15   *Corresponding author: Denis.Duboule@epfl.ch

16

SUPPORTING INFORMATION

17

Supplemental Figure S1 to S7

18

Supplemental Table S1 to S4

19

20

1

21   **LEGENDS TO SUPPLEMENTAL FIGURES**

22   **S1 Fig (related to Fig 1).** *Hoxa* **genes expression in mouse and chick limb buds.**

23   (A, B) *In situ* hybridization analysis of E12.5 mouse and HH28 chick fore- and hindlimb

24   buds showing the expression of *Hoxa* genes. (A) Expression patterns of *Hoxa11* and

25   *Hoxa13* in mouse forelimb are similar to those in hindlimb at E12.5. (B) Stronger

26   expression of *Hoxa11* is observed in the chick leg bud than in the wing bud. (C, D)

27   Transcription profiles of *Hoxa* genes in micro-dissected proximal and distal domains from

28   either E12.5 mouse (C) or HH30 chick (D) fore- and hindlimb buds. Right limbs in (A,

29   B) are oriented with proximal to the bottom and distal to the top. FL, forelimb; HL,

30   hindlimb. The *Y* axis represents the strand-specific RNA-seq read counts, normalized by

31   the total number of million mapped reads.

32

33   **S2 Fig (related to Fig 2). The regulatory switch between TADs in mouse and chick**

34   **limb buds.** (A-C) The 4C interaction profiles with chick *Hoxd12* (A), mouse *Hoxd13* (B)

35   and chick *Hoxd13* (C), in mouse (E12.5) and chick (HH30) fore- and hindlimbs. (A) In

36   addition to the *CS93* region, contacts between *Hoxd12* and the *CS39* region were also

37   reduced in chick proximal leg bud cells. In the distal wing- and leg bud cells, *Hoxd12*

38   mainly contacted C-DOM, in contrast to the profile observed with the *Hoxd10-11* bait.

39   (B, C) Both mouse *Hoxd13* and chick *Hoxd13* promoters constitutively interacted with

40   C-DOM. The interaction between *Hoxd13* and either island III or *Prox* specifically

41   increased in both mouse and chick distal limbs. FL, forelimb; HL, hindlimb.

42

2

43   **S3 Fig (related to Fig 4). H3K27ac, H3K27me3 and RNA-seq at *HoxD* in chick limbs.**

44   (A) H3K27ac marks (tracks 1-2 and 5-10) and transcription profiles (tracks 3 and 4) at

45   the *HoxD* locus either in whole, proximal or distal wing- and leg buds. 5' *Hoxd* genes are

46   covered by H3K27ac in the leg bud at HH19 and HH20. However, the level of *Hoxd*

47   transcripts was reduced at HH20 (see also S3B Fig, track 4). In proximal leg buds at

48   HH28, a significant decrease in H3K27ac enrichment was detected, which corresponded

49   to the reduction in *Hoxd* expression (track 8). (B) H3K27me3 distribution in either whole,

50   proximal or distal wing- and leg buds at HH20 and HH28. Stronger enrichments were

51   observed in both whole leg buds at HH20 and proximal leg buds at HH28, when compared

52   to the corresponding samples from wing buds. The *Y* axis represents the strand-specific

53   RNA-seq read counts, normalized by the total number of million mapped reads.

54   Enrichment (*Y-axis*) of ChIP is shown as the $\log_2$ ratio of the normalized number of reads

55   between ChIP and input samples. FL, forelimb; HL, hindlimb.

56

57   **S4 Fig (related to Fig 4). H3K27ac and H3K27me3 profiles of ChIP-seq and RNA-**

58   **seq at the *HoxA* locus in chick limbs.** (A, B) Distributions of H3K27ac and H3K27me3

59   marks over the *HoxA* cluster and its regulatory elements in either whole, proximal or

60   distal wing- and leg buds at HH19, HH20 and HH28. (A) Stronger enrichment of

61   H3K27ac around the 5' *Hoxa* genes were observed in leg buds at both HH19 and HH20,

62   whereas less marks were scored at HH20, in the region covering the *e10* to *e16* enhancers,

63   when compared to wing- and leg buds at HH19. At HH28, profiles established from

64   proximal or distal region were comparable between wing- and leg buds. (B) 3' *Hoxa*

3

65   promoters were not labelled by H3K27me3 marks in wing buds at HH20 (track 1). Strong

66   enrichments of H3K27me3 over the *HoxA* regulatory elements were not scored, unlike in

67   both C-DOM and T-DOM at the *HoxD* locus (see also Fig. 4B). (C) H3K27ac marks

68   (tracks 1-2 and 5-10) and transcription profiles (tracks 3 and 4) at the *HoxA* locus in either

69   whole, proximal or distal wing- and leg buds. More H3K27ac marks were detected at 5'

70   *Hoxa* genes in whole leg buds at both HH19 and HH20, corresponding to higher levels

71   of *Hoxa* gene transcripts in leg buds than in wing buds (red arrows in tracks 3 and 4). (D)

72   H3K27me3 profiles in either whole, proximal or distal wing- and leg buds at HH20 and

73   HH28. The *HoxA* regulatory elements at the chick locus were identified by using mouse

74   coordinates and the LiftOver function of the UCSC genome browser. The *Y* axis

75   represents the strand-specific RNA-seq read counts, normalized by the total number of

76   million mapped reads. Enrichment (*Y-axis*) of ChIP is shown as the $\log_2$ ratio of the

77   normalized number of reads between ChIP and input samples. FL, forelimb; HL, hindlimb.

78

79   **S5 Fig (related to Fig 5). Chromatin conformation at the chick *HoxD* locus in wing-**

80   **and leg buds and conservation of CTCF sites orientations.** (A, B) Capture Hi-C (CHi-

81   C) heat maps at 5kb resolution, transcription profiles and CTCF ChIP-seq by using either

82   whole wing- or leg buds at HH20. Both CHi-C heat maps and CTCF distributions were

83   relatively similar between wing- and leg buds, although some slight differences were

84   detected. A noticeable downregulation of *Hoxd* gene expression was observed in leg buds

85   when compared to wing buds. (C) Comparison of CTCF sites orientations at the *HoxD*

86   cluster between mouse distal forelimb at E12.5 (top) and chick wing bud at HH20

4

87   (bottom). Opened and closed arrowheads indicate the orientation of the CTCF motives.

88   The *Y* axis represents the strand-specific RNA-seq read counts, normalized by the total

89   number of million mapped reads. Enrichment (*Y*-axis) is shown at the normalized 1x

90   sequencing depth of CTCF ChIP. FL, forelimb; HL, hindlimb.

91

92   **S6 Fig (related to Fig 6). Expression of *Hoxa* and *Hoxd* genes in chick and mouse**

93   **limb buds.** (A) *Hoxa11* expression was stronger in chick leg buds than in wing buds (left),

94   whereas it seems increased in both mouse fore- and hindlimb buds, as development

95   proceeded (right). (B) Expression of *Hoxd12* in both chick wing buds and mouse limb

96   buds displayed a similar trend, except for the chick leg buds. (C) Expression of *Hoxd13*

97   in both chick limb buds and mouse forelimb buds were similar, except in mouse hindlimb

98   buds. (D) Hi-C data at *HoxA* with 40kb resolution using wing- and leg buds at HH20.

99   More contacts were scored between the *HoxA* cluster and its regulatory regions in leg

100   buds than that in wing buds (black rectangle). Expression levels were normalized to

101   *Gapdh* and are shown as fold change relative to forelimb buds at either E10.5, or HH20-

102   21. Error bars indicate standard deviation of either three (chick), two (E10.5) or four

103   (E10.75) biological replicates. **P<0.01, *P<0.05, NS, P>0.05, Student's *t*-test.

104

105   **S7 Fig (related to Fig 7). A T-DOM deletion induces interactions between HE1 and**

106   ***Hoxd* genes.** (A) Relative expression levels for each *Hoxd* gene in mouse and chick

107   proximal fore- and hindlimbs. Expression levels in mouse and chick proximal fore- or

108   hindlimb buds were normalized to *mGapdh* and *chGapdh*, respectively and are shown as

5

109   fold change relative to mouse control or chick proximal forelimbs at E12.5 or HH28.

110   Error bars indicate standard deviation of three (control), two (mutant) or three (chick)

111   biological replicates. **P<0.01, NS, P>0.05, Student's *t*-test. (B) H3K4me1 profiles

112   obtained from proximal fore- and hindlimb buds of either control or *Del(attp-SB3)/Δ*

113   mutant embryos at E12.5. The putative HE1 enhancer was covered by H3K4me1 marks

114   and merged with a predicted enhancer region. (C, D) 4C-seq profiles with HE1 as a

115   viewpoint, by using either control, or *Del(attp-SB3)/Δ* mutant proximal hindlimb buds at

116   E12.5. The region from *Hoxd8* to *Hoxd13* is shaded (C). (D) Enlargement of (C) focusing

117   on the *HoxD* cluster. Within the *Hoxd8* to *Hoxd13* interval, more contacts between HE1

118   and *Hoxd* promoters were detected in the mutant material where *Hoxd* expression

119   remained as compared to the control situation where *Hoxd* expression was severely down-

120   regulated. (E) *Hoxc11* expression from control and *Del(attp-SB3)/Δ* mutant at E12.5.

121   (left) Expression of *Hoxc11* in proximal hindlimb buds partly overlapped with that of

122   *Hoxd11*. The deletion of T-DOM did not affect *Hoxc11* expression. Enrichment (*Y*-axis)

123   of ChIP is shown at the $\log_2$ ratio of the normalized number of reads between ChIP and

124   input samples. FL, forelimb; HL, hindlimb.

125

126   **S1 Table**

127   Information about samples.

128

129   **S2 Table**

130   DNA sequences of primers used for both RT-qPCR analyses and genotyping.
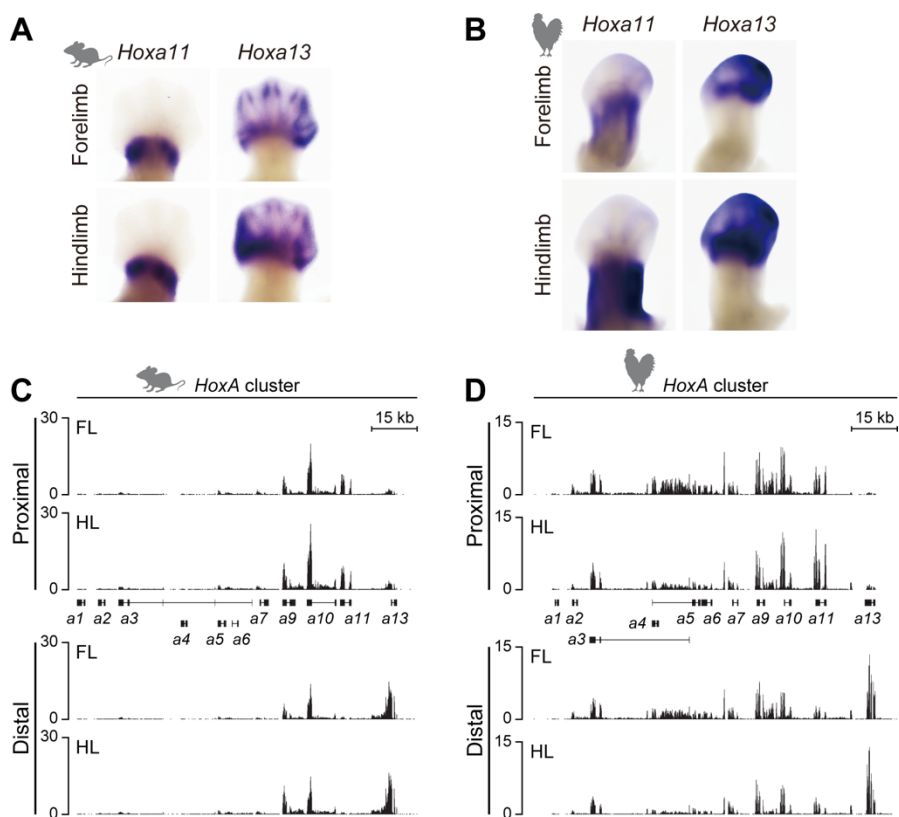
6

131   **S3 Table**

132   DNA sequences of primers used for 4C analyses. Custom barcodes (4bp shown by

133   NNNN) were introduced in between the Illumina adapter sequence and the specific

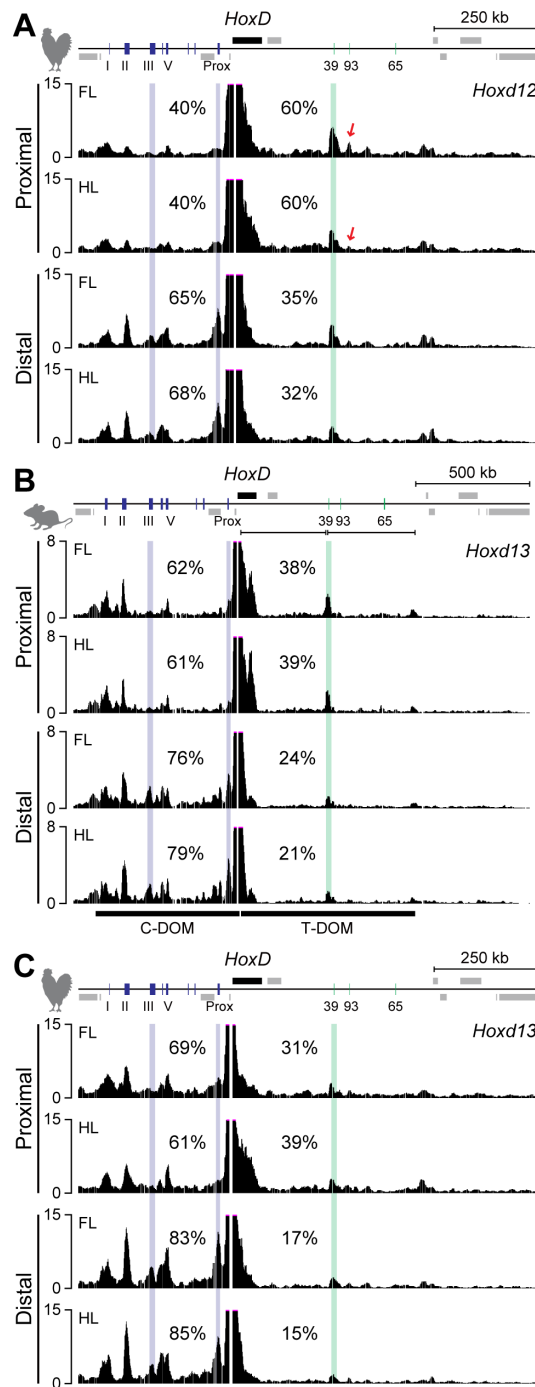134   viewpoint sequences to do multiplexing using different samples with the same viewpoint.

135

136   **S4 Table**

137   Public datasets used in this research.

A

Hoxa11 *Hoxa13*

Forelimb

Hindlimb

B

Hoxa11 *Hoxa13*

Forelimb

Hindlimb

C *HoxA* cluster

15 kb

Proximal — FL, HL (30)

a1 a2 a3   a4 a5 a6   a7 a9 a10 a11 a13

Distal — FL, HL (30)

D *HoxA* cluster

15 kb

Proximal — FL, HL (15)

a1 a2 a3   a4 a5 a6 a7 a9 a10 a11 a13

Distal — FL, HL (15)

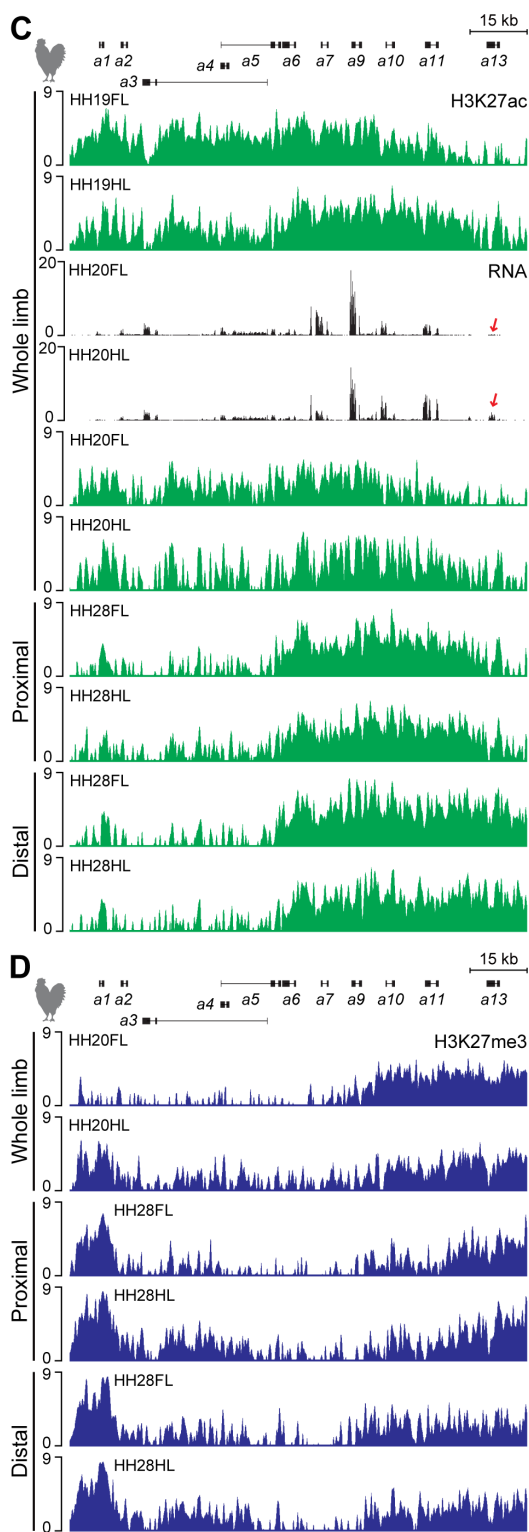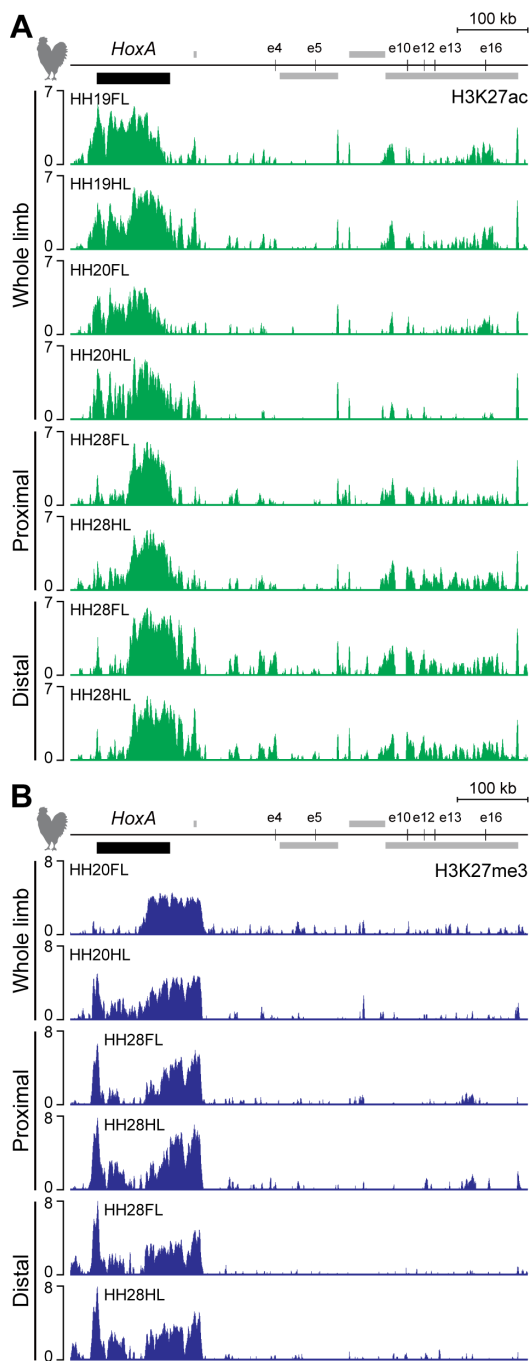S1 Fig (related to Fig 1). Yakushiji-Kaminatsui et al.

S2 Fig (related to Fig 2). Yakushiji-Kaminatsui et al.
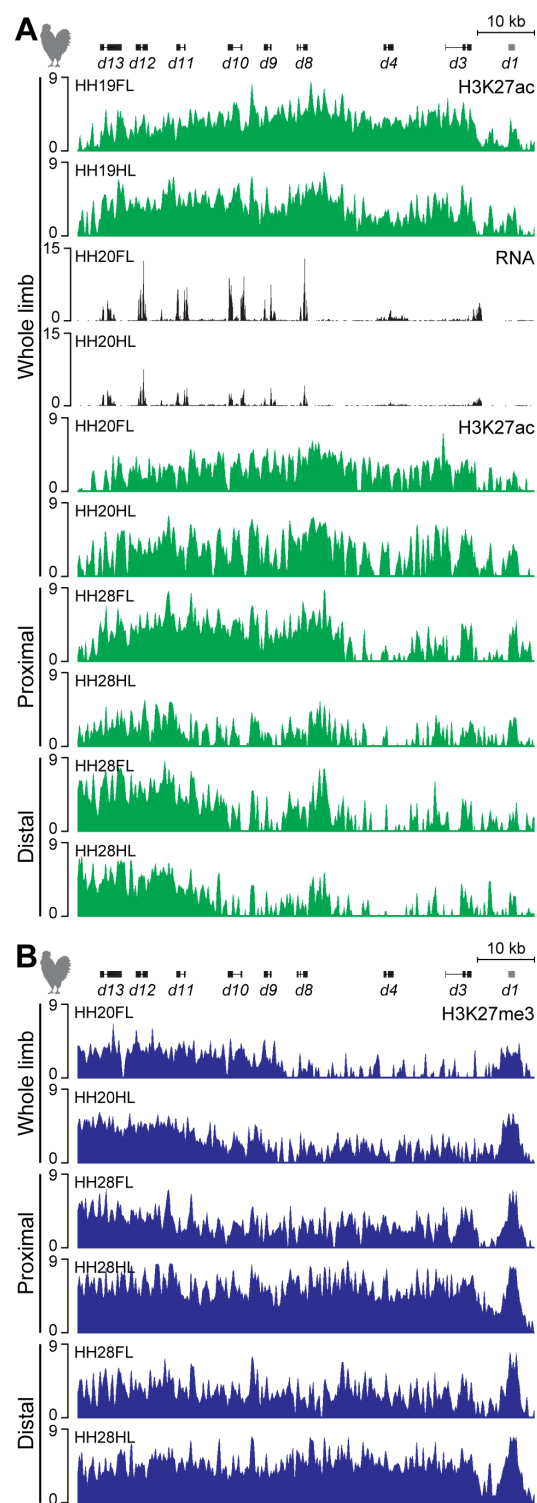
S3 Fig (related to Fig 4). Yakushiji-Kaminatsui et al.

S4 Fig (related to Fig 4). Yakushiji-Kaminatsui et al.
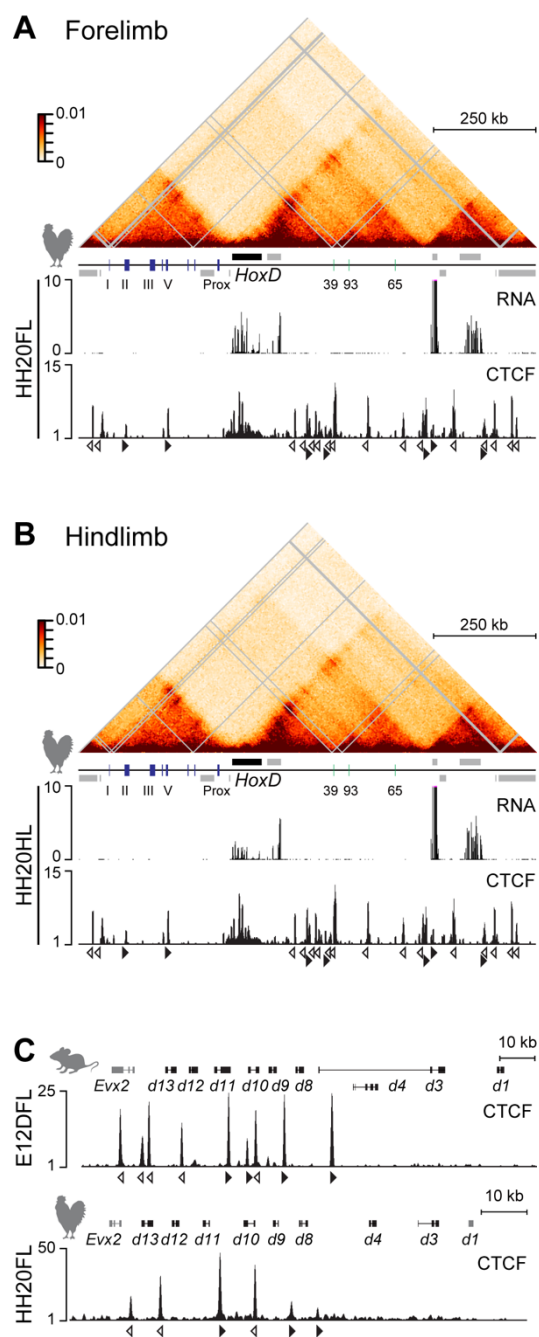
A Forelimb

B Hindlimb

C

S5 Fig (related to Fig 5). Yakushiji-Kaminatsui et al.

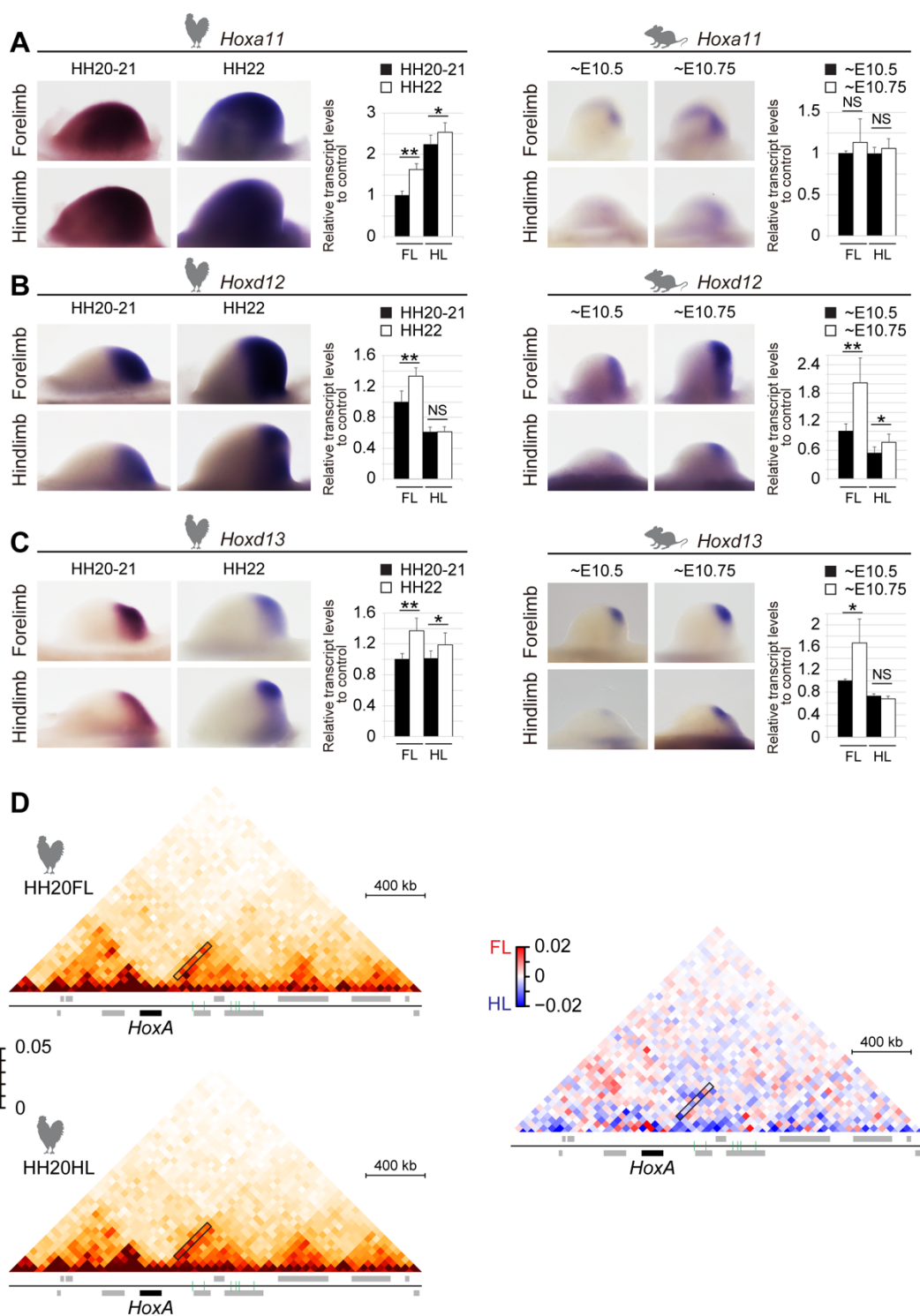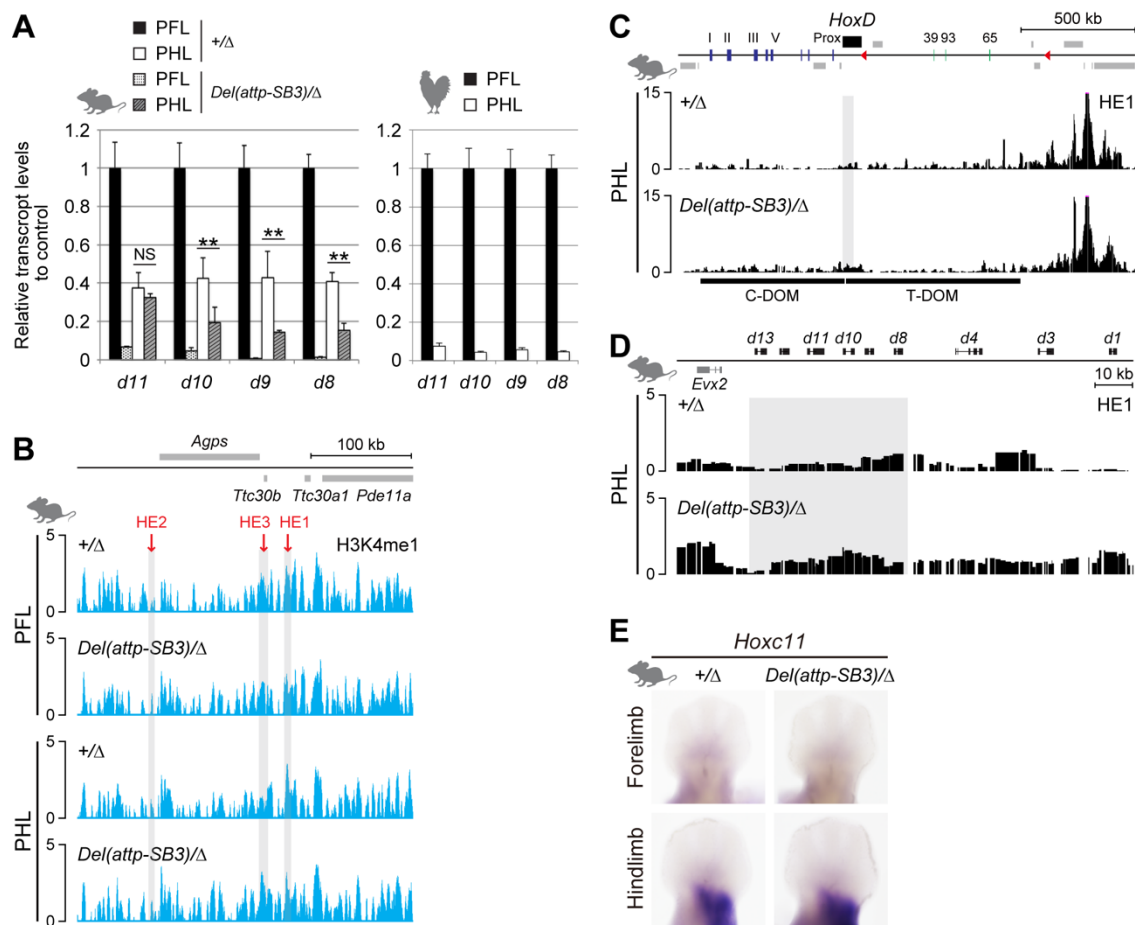S6 Fig (related to Fig 6). Yakushiji-Kaminatsui et al.

S7 Fig (related to Fig 7). Yakushiji-Kaminatsui et al.

S1 Table. Sample information

Experiment	Animal	Stage	genotype	Tissue	Nb of pair of limb (Cell number)	Biological replicate	antibody
RNA-seq	mouse	E12.5	wild type	proximal hindlimb	1 pair	1	
				distal hindlimb	1 pair	1	
	chick	HH20	wild type	whole forelimb bud	4 pairs	1	
				whole hindlimb bud	4 pairs	1	
		HH30		proximal forelimb	2 pairs	1	
				distal forelimb	2 pairs	1	
				proximal hindlimb	2 pairs	1	
				distal hindlimb	2 pairs	1	
RT-qPCR	mouse	~E10.5	wild type	whole forelimb bud	4 pairs	2	
				whole hindlimb bud	4 pairs	2	
		~E10.75		whole forelimb bud	3 pairs	4	
				whole hindlimb bud	3 pairs	4	
		E12.5	+/Del(8-13)	proximal forelimb	2 pairs	3	
				proximal hindlimb	2 pairs	3	
			Del(attp-SB3)/Del(8-13)	proximal forelimb	2 pairs	2	
				proximal hindlimb	2 pairs	2	
	chick	HH20-21	wild type	whole forelimb bud	3 pairs	3	
				whole hindlimb bud	3 pairs	3	
		HH22		whole forelimb bud	2 pairs	3	
				whole hindlimb bud	2 pairs	3	
		HH28		proximal forelimb	1 pair	3	
				proximal hindlimb	1 pair	3	
4C-seq	mouse	E12.5	wild type	proximal hindlimb	9 pairs	1	
				distal hindlimb	9 pairs	1	
			+/Del(8-13)	proximal forelimb	9 pairs	1	
				proximal hindlimb	9 pairs	1	
			Del(attp-SB3)/Del(8-13)	proximal forelimb	9 pairs	1	
				proximal hindlimb	9 pairs	1	
	chick	HH30	wild type	proximal forelimb	14 pairs	1	
				distal forelimb	24 pairs	1	
				proximal hindlimb	6 pairs	1	
				distal hindlimb	13 pairs	1	
CHi-C	chick	HH20	wild type	whole forelimb bud	51 pairs (c.a. 8.8x10^6)	1	
				whole hindlimb bud	51 pairs (c.a. 1.24x10^7)	1	
ChIP-seq	mouse	E12.5	+/Del(8-13)	proximal forelimb	14 pairs	1	H3K4me1 (Abcam, ab8895, Lot: GR312093-1, 3 ug each)
				proximal hindlimb	14 pairs	1	
			Del(attp-SB3)/Del(8-13)	proximal forelimb	15 pairs	1	
				proximal hindlimb	15 pairs	1	
	chick	HH19	wild type	whole forelimb bud	32 pairs	1	H3K27ac (Abcam, ab4729, Lot: GR184558-1, 3 ug each)
				whole hindlimb bud	32 pairs	1	
		HH20		whole forelimb bud	58 pairs	1	H3K27ac (Abcam, ab4729, Lot: GR288020-1, 3 ug each) and H3K27me3 (Merk Millipore, 07-449, Lot: 2919706, 25 uL each)
				whole hindlimb bud	58 pairs	1	
				whole forelimb bud	42 pairs	1	CTCF (Activ motif, #61311, Lot: 34614003, 5 uL each)
				whole hindlimb bud	42 pairs	1	
		HH28		proximal forelimb	27 pairs	1	H3K27ac (Abcam, ab4729, Lot: GR184558-1, 3ug each) and H3K27me3 (Merk Millipore, 07-449, Lot: 2382150, 3uL each)
				distal forelimb	27 pairs	1	
				proximal hindlimb	27 pairs	1	
				distal hindlimb	27 pairs	1	

S2 Table. Primer information

RT-qPCR

Name of gene		Sequence (5' -> 3')
mGapdh	forward	TGCGACTTCAACAGCAACTC
	reverse	ATGTAGGCCATGAGGTCCAC
mHoxd8	forward	GAGCTCGTCTCCTTCTCAAATGTT
	reverse	AGGTTTGTCTTCCTCTCCGTCTAC
mHoxd9	forward	CTGAAGGAGGAGGAGAAGCA
	reverse	GTGTAGGGACAGCGCTTTTT
mHoxd10	forward	GCCAGGAGCCCACTAAAGTC
	reverse	TGCTTTCCTTCTCCTGCACT
mHoxd11	forward	CCGCAGCCTCTAACTTCTACA
	reverse	GCCTCGTAGAACTGATCAAAGC
mHoxd12	forward	CTCTTGCCTGCGATCTTCACT
	reverse	GAATTCATTGACCAGGAATTCGTT
mHoxd13	forward	GGTGTACTGTGCCAAGGATCAG
	reverse	CACATGTCCGGCTGGTTT
mHoxa11	forward	GCCGTCTCGTCCAATTTCTA
	reverse	GTAGTCGGAGGAAGCGAGGT
mHoxa13	forward	CTGGAACGGCCAAATGTACT
	reverse	CCTATAGGAGCTGGCGTCTG
chGapdh	forward	GGTGGTGCTAAGCGTGTTA
	reverse	CCCTCCACAATGCCAA
chHoxd8	forward	GTCTCCCGCTCAAATGTTTC
	reverse	CCTCGATTCGTCTCTTCCTG
chHoxd9	forward	GCCATGAAAGCGAAGAAGTT
	reverse	GACCAGGAGGTGGAGAACAC
chHoxd10	forward	TGAACTCGTTTCAGGGCTATT
	reverse	GACGTCCCTTCCTACCAGAG
chHoxd11	forward	GGTGAGGCTGTTGCAGAGAA
	reverse	TCCTCGATCGCTGAGGAACT
chHoxd12	forward	GCAGTCCAGTCATGCCTCAG
	reverse	ATTCGGCGATTTGCTGTTTT
chHoxd13	forward	CACAGAGCTCGCACTTCTGG
	reverse	ATCTCGGGCTGGTTTAGTGC
chHoxa11	forward	GCAGAGGAGATCATGCACAG
	reverse	GCCCACCGTGCTATAGAAAT
chHoxa13	forward	AAGTGTACTGCCCCAAGGAG
	reverse	ACCTTTGTGTAGGGCACTCG

Genotyping primer

Name of strain		Sequence (5' -> 3')
Del(Mtx-Ttn): wt	271	CCAGTGTTCTGACATTAAAGGGCAG
	272	CACTGGCTTACAGACAGGAAAAATTAGG
Del(Mtx-Ttn): mutant	290	CCTGGGGTACCCTATTGGAGTC
	272	CACTGGCTTACAGACAGGAAAAATTAGG

S3 Table. 4C primer information

Name of viewpoint		Sequence (5' -> 3')	Genomic coordinates for excluded fragments during analysis	Barcode sequence (NNNN)
mHoxd11	inverse forward	AATGATACGGCGACCACCGAACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNAAGCATACTTCCTCAGAAGAGGCA	chr2: 74,683,503-74,688,501 (mm10)	Bc14: AATT, Bc34: AGCC, Bc54: TTAA, Bc84: CGAT
	inverse reverse	CAAGCAGAAGACGGCATACGACTAGGAAAATTCCTAATTTCAGG		
mHoxd13	inverse forward	AATGATACGGCGACCACCGAACACTCTTTCCCTACACGACGCTCTTCCGATCTAAAATCCTAGACCTGGTCATG	chr2: 74,663,180-74,670,260 (mm10)	
	inverse reverse	CAAGCAGAAGACGGCATACGAGGCCGATGGTGCTGTATAGG		
mHE1	inverse forward	AATGATACGGCGACCACCGAACACTCTTTCCCTACACGACGCTCTTCCGATCTTGTCTTCCCTCCGGTATG	chr2: 75,953,598-75,958,250 (mm10)	
	inverse reverse	CAAGCAGAAGACGGCATACGACTCCGTCAGGCAACTGAT		
chHoxd10-11	inverse forward	AATGATACGGCGACCACCGAACACTCTTTCCCTACACGACGCTCTTCCGATCTGAGTCGCCGGCTCTTTCTATG	chr7: 16,370,905-16,378,558 (galGal5)	
	inverse reverse	CAAGCAGAAGACGGCATACGAAGGACTTTGGCGTGCAGAC		
chHoxd12	inverse forward	AATGATACGGCGACCACCGAACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNAATACCGGACTCGCAAGG	chr7: 16,382,849-16,387,653 (galGal5)	A14: ATCG, C44: CTGT
	inverse reverse	CAAGCAGAAGACGGCATACGAGCTGAAGTCCTTTGTTGGA		
chHoxd13	inverse forward	AATGATACGGCGACCACCGAACACTCTTTCCCTACACGACGCTCTTCCGATCTAGAAATTCACATTTCAAAGGGTTGAC	chr7: 16,391,744-16,396,736 (galGal5)	
	inverse reverse	CAAGCAGAAGACGGCATACGAGACGAAGTTTACTCAGTTGCTCCT		

S4 Table. Public datasets used in this research

Dataset	Accession number	Reference
RNA-seq, proximal forelimb, E12.5	GSM2090082	Beccari L et al., A role for HOX13 proteins in the regulatory switch between TADs at the HoxD locus. **2016** Genes Dev 30, 1172-86
RNA-seq, distal forelimb, E12.5	GSM2090083	
4C-seq, proximal forelimb, E12.5, Hoxd11 as bait	GSM2092702	
4C-seq, distal forelimb, E12.5, Hoxd11 as bait	GSM2092703	
4C-seq, proximal forelimb, E12.5, Hoxd13 as bait	GSM2092704	
4C-seq, distal forelimb, E12.5, Hoxd13 as bait	GSM2092705	
DNaseI HS seq, hindlimb bud, E11.5	GSM1014179	ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. **2012** Nature 489, 57-74