LEGENDS TO SUPPLEMENTARY FIGURES

Figure S1. Comparison of *HoxD* regulatory landscapes in mammals and fishes. a. *Hoxd* gene expression and regulation in mouse limb buds at E12.5. The *HoxD* cluster is flanked by two gene deserts, named according to their relative position (3' or 5') with respect to *Hoxd* gene orientation. The 3DOM regulatory landscape activates *Hoxd4* to *Hoxd11* in the proximal limb territory (green). The 5DOM activates *Hoxd10* to *Hoxd13* in the distal limb territory (blue). Schemes are based on ref.[15]. **b.** Gene expression in fin buds at 40-60 hpf in the cognate zebrafish *hoxda* cluster. The fish cluster is also flanked by two gene deserts but their regulatory potentials are unknown (question marks). Fish *hoxd9a* to *hoxd11a* are expressed in the preaxial fin territory (purple) whereas *hoxd11a* to *hoxd13a* are expressed in a postaxial domain (orange). Schemes and WISH are inspired from[14,53,54].

Figure S2. 3D chromatin conformation at the mouse and fish *HoxD* loci. Contact frequency heatmaps at the mouse *HoxD* (E18.5 male UGS, one representative replicate out of two) and fish *hoxda* (24 hpf and 48 hpf total embryos[41,55]) loci (top and bottom, respectively). The similarities in the constitutive structural organization of the mouse and the fish loci are underlined either by the position and relative extents of TADs (thick black lines), the presence of a sub-TAD boundary within 3DOM (asterisk), as well as by the positions and orientation of CTCF binding sites (red and blue arrowheads). *Hox* genes are in purple-scale rectangles and other genes are grey rectangles. Bin size is 10 kb. The scales on the *x* axes were adjusted to comparable sizes for ease of comparison, yet the fish locus is more compact. Scale bars in both cases; 100 kb.

Figure S3. The *HoxD* locus is part of a large syntenic interval. a. The mouse *HoxD* locus (mm39) is on top and the zebrafish *hoxda* locus (danRer11) is shown below. *Hox* genes are in purple-scale rectangles and annotated mouse enhancers are shown as either blue (5DOM) or green (3DOM) rectangles. Conserved sequences between the two gene deserts are shown as vertical black bars. Those conserved sequences overlapping with known murine enhancers were used to annotate the corresponding elements in zebrafish (blue rectangles). **b.** Synteny plot representing sequences conserved between the mouse and the zebrafish *HoxD* loci. On the *x* axis is the mouse locus (mm10, chr2: 73605690-75662521) and on the *y* axis is the zebrafish locus (danRer11, chr9: 1639965-2393397, inverted *y* axis). Despite a mouse locus that is in average 2.6 times larger than its zebrafish counterpart, the order of most conserved sequences is maintained, showing the absence of substantial genomic rearrangement at these gene deserts. **c.** Size comparisons between different regions of the zebrafish *hoxda* and the mouse *HoxD* loci. The left panel shows that the *Hox* clusters have maintained a similar size over time, while gene deserts have expanded in mouse and/or contracted in zebrafish. The right panel shows that the ratio between the sizes of 5DOM *versus* 3DOM is inverted in the two species.

Figure S4. Chromatin profiles in zebrafish embryos. a. Expression of *hoxd13a*, *hoxd10a* and *hoxd4a* in control zebrafish embryos by WISH. Stages are indicated on top of the panels. Scale bars; 200 µm. **b.** Dissection plan used for panel (**c**). PT, posterior trunk. **c.** ATAC-seq profile and both H3K27ac and H3K27me3 CUT&RUN profiles over the zebrafish *hoxda* locus in 16 hpf dissected heads (grey, one representative condition out of three) and 16 hpf posterior trunk cells (PT, blue, one representative replicate out of three). Both the *hoxda* cluster and 3DOM show specific open sites in posterior trunk, where *hoxda* genes are expressed, when compared to forebrain. The CUT&RUN profiles in posterior trunk cells show enrichment for H3K27ac (green coverage) on the central and anterior parts of the *hoxda* cluster, while H3K27me3 (red coverage) is enriched on the posterior part and on *evx2*. Scale bar; 100 kb.

Figure S5. *Hoxd13a* expression in control, heterozygous and homozygous mutant fin buds at 60 and 72 hpf. a. Schematic of the deletion and spatial orientation of the fin buds**. b.** Various samples are shown to illustrate the variability observed. While a clear tendency is observed in the loss of the distal most expression in homozygous mutants, expression is still observed in some samples as well as in post-axial cells, unlike the situation in developing limb buds where expression is entirely absent in the comparable deletion. Scale bar: 50 µm.

Figure S6. WISH of *hoxd13a*, *hoxd10a* or *hoxd4a* in zebrafish embryos lacking either 3DOM (a), or 5DOM (b). The genotypes (in red, top) and genes analyzed (left) are shown as well as the stages (up left). **a.** Deletion of 3DOM. Black arrowheads (empty for no expression and full for expression) indicate differential gene expression in the cloacal region, whereas red arrowheads (empty for no expression and full for expression) point to the pectoral fin buds. Control and homozygote mutant embryos are shown side by side for each condition, except for *hoxd4a* where a heterozygous (Het) mutant is shown. Wild-type and homozygous embryos originate from the same clutch of eggs and were processed together. In Del(3DOM) mutant embryos (**a**), *Hoxd10a* and *hoxd4a* transcription is lost in fin buds, whereas *hoxd13a* transcripts in the cloaca are not affected. B, branchial arches; R, rhombomeres; Scale bars; 200 µm. **b**. In contrast, *hoxd13a* mRNAs are lost from the cloacal region in Del(5DOM) mutant animals at 36 hpf (red arrowheads), while still clearly detected in the fin buds, indicating that the 5DOM is necessary for *hoxd13a* transcription in the pseudo-cloacal region.

Figure S7. *Hoxd* gene expression in the mouse urogenital system. a. Schematic representations of male and female urogenital systems. K: Kidney, B: Bladder, O: Ovary, T: Testis. The urogenital sinus (UGS) is indicated with a red circle. **b.** WISH of *Hoxd* genes in representative female and male urogenital systems. All *Hoxd* genes are expressed in anterior portions of the UGS including kidneys, the uterus, deferens ducts and the bladder, except *Hoxd13* transcripts, which are restricted to the UGS

(see Fig. 3b). **c-d.** Schematic representation of two *HoxD* genomic configurations, The first one is a deletion of the entire *HoxD* cluster (**c**), whereas the second one is a random integration of a transgene carrying the same *HoxD* transgene plus some flanking sequences in 5' (**d**, thick red bar). *Hox* genes as in shades of purple and the deletion breakpoints are shown as vertical dashed red lines. Scale bar; 100 kb. **c.** WISH of *Hoxd13* in UGS from a transgenic *HoxD* cluster (TgBAC), while lacking both endogenous copies of the *HoxD* cluster. Expression is not detected from the transgenic cluster. **d.** Likewise, ß-gal staining of UGS transgenic for the *HoxD* cluster containing a *LacZ* reporter displays no activity in the UGS. By contrast, *LacZ* staining of mutant *Inv(nsi-Itga6)d11lac* embryos, which also includes a *lacZ* reporter confirms that 5DOM is necessary and sufficient to drive expression in the UGS. Scale bar: 1 mm.

Figure S8. Regulatory potential of sub-regions within 5DOM. a. RNA-seq FPKM values for various mouse *Hoxd* genes in E18.5 UGS obtained from either wild-type or *Inv(nsi-Itga6)d11lac* mutant embryos (see schematic in Fig. 3c). Data are shown separately for females (n=3, dots) and males (n=3, triangles). Drastic decreases are observed for *Hoxd10, Hoxd12* and *Hoxd13* when 5DOM is disconnected from the *HoxD* cluster. *Hoxd11* could not be assessed due to the presence of a transgenic copy of this gene in the *LacZ* reporter cassette. **b.** On top is a scheme of the 5DOM regulatory landscape on mm39 with *Hox* genes in purple. Blue rectangles indicate previously described 5DOM enhancers. The red arrowheads delimit the serial deletion breakpoints. The three consecutive deletions are depicted by red dashed lines. Below are RT-qPCR quantifications of expression levels relative to wild-type (n=4) in three mutant lines carrying serial deletions of 5DOM (n=3). The horizontal red line represents the value of 1 for reference. Severe reductions are observed for both the *Del(Rel5-SB1)* and *Del(Rel1-Rel5)* conditions, unlike in the *Del(SB1-Atf2)* deletion. Scale bar; 100 kb.

Figure S9. Sequence conservation in vertebrates of the GT2, islE and CsB UGS enhancers. All three sequences are comprised in the box highlighted in Fig. 4a. The ATAC-seq and H3K27ac ChIP-seq profiles are shown with, below, their sequence conservation from fishes and mammals. The thick blue lines below the H3k27ac profiles indicate the extent of the transgenes assayed in Fig. 4. Scale bars; 1 kb.

Figure S10. *hox13* gene expression in the Daniocell atlas. a. UMAP of endoderm cells using matrices extracted from ref.[56], colored by tissue. The black rectangle indicates the UMAP region which contains cellular clusters from the cloacal region. All other panels in the figure correspond to this rectangle. **b.** UMAP of endodermal cells and identities of their clusters. The colors indicate the identities of cells from both the cloacal region (red arrow) and the posterior intestine (dark green, arrow). **c.** UMAP of selected endoderm, clustered by developmental stages (color code below). **d.** UMAP as in panel b, with the expression in red of the various *hox13* paralogous genes. All cells with a normalized expression

level above 2 are displayed in red. In panels **c** and **d**, arrowheads indicate *hox13* expressing cells in the cloacal region either at early (red) or late (black) timepoints. The black arrows point to *hox13* expression in intestinal cells.

Figure S11. Cloacal region phenotypes in *hox13* mutant zebrafish. a-f. Confocal micrographs of mutant cloacal regions at 6 dpf shown in single channel (**a-c**) and pseudo color (**d-f**). **d.** Triple *hoxa13a*;*hoxa13b*;*hoxd13a* heterozygotes (n=6) exhibit wild-type patterning with separate openings for the hindgut (blue) and pronephric duct (yellow). **e.** Homozygous *hoxa13a* single mutants show wild-type patterning (n=4). **f.** Homozygous *hoxa13b* single mutants have wild-type patterning (n=4). **g-h.** Length and width of the hindgut and pronephric duct in wild-type and *hoxa13* mutant zebrafish embryos at 3 dpf. **g.** The length (red dotted lines) and width (white dotted lines) of the hindgut and pronephric complex at the median fin fold level were quantified in wild-type (n=4, left) and *hoxa13a*;*hoxa13b* double mutant embryos (n=5, right). **h.** The length difference of the hindgut and pronephric complex between wild-type and *hoxa13* double mutant embryos is statistically significant (*p = 0.0101, two-sided Welch's t-test). The error bars indicate the standard error of the mean. Scale bar length is 30 μm in **a-f** and 100 μm in **g**.

Table S1. Extent of the mouse and zebrafish domains. Sizes are indicated in base pairs and were determined based on the transcription start sites of genes.

Table S2. Accession numbers for re-analyzed data. SRA accession numbers and reference of publications for the re-analyzed data when previously published data was used.

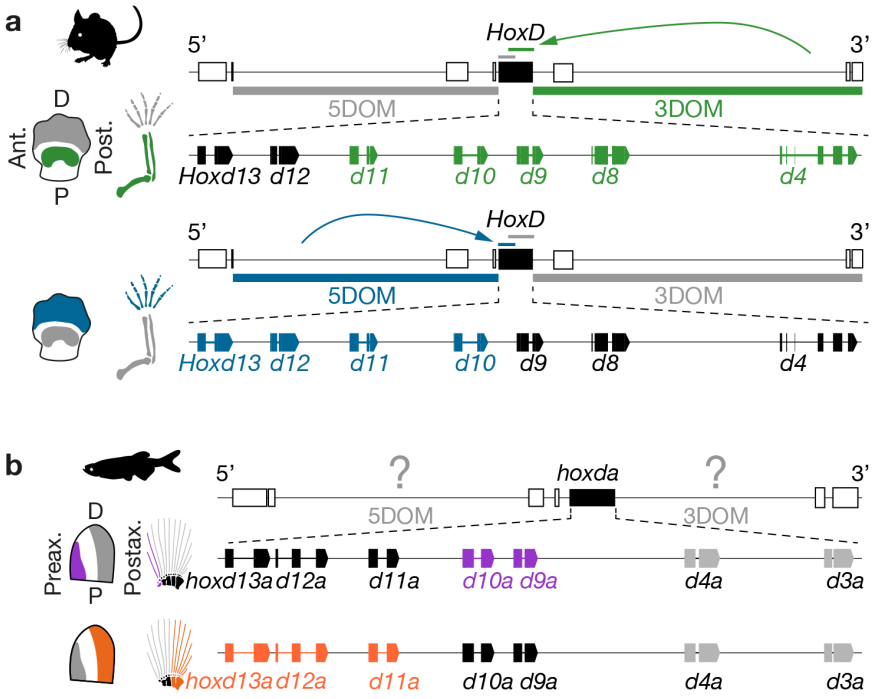
Table S3. RT-qPCR primers. Lists of primers used in RT-qPCR experiments.

File S1. Sequences of the zebrafish probes used for WISH

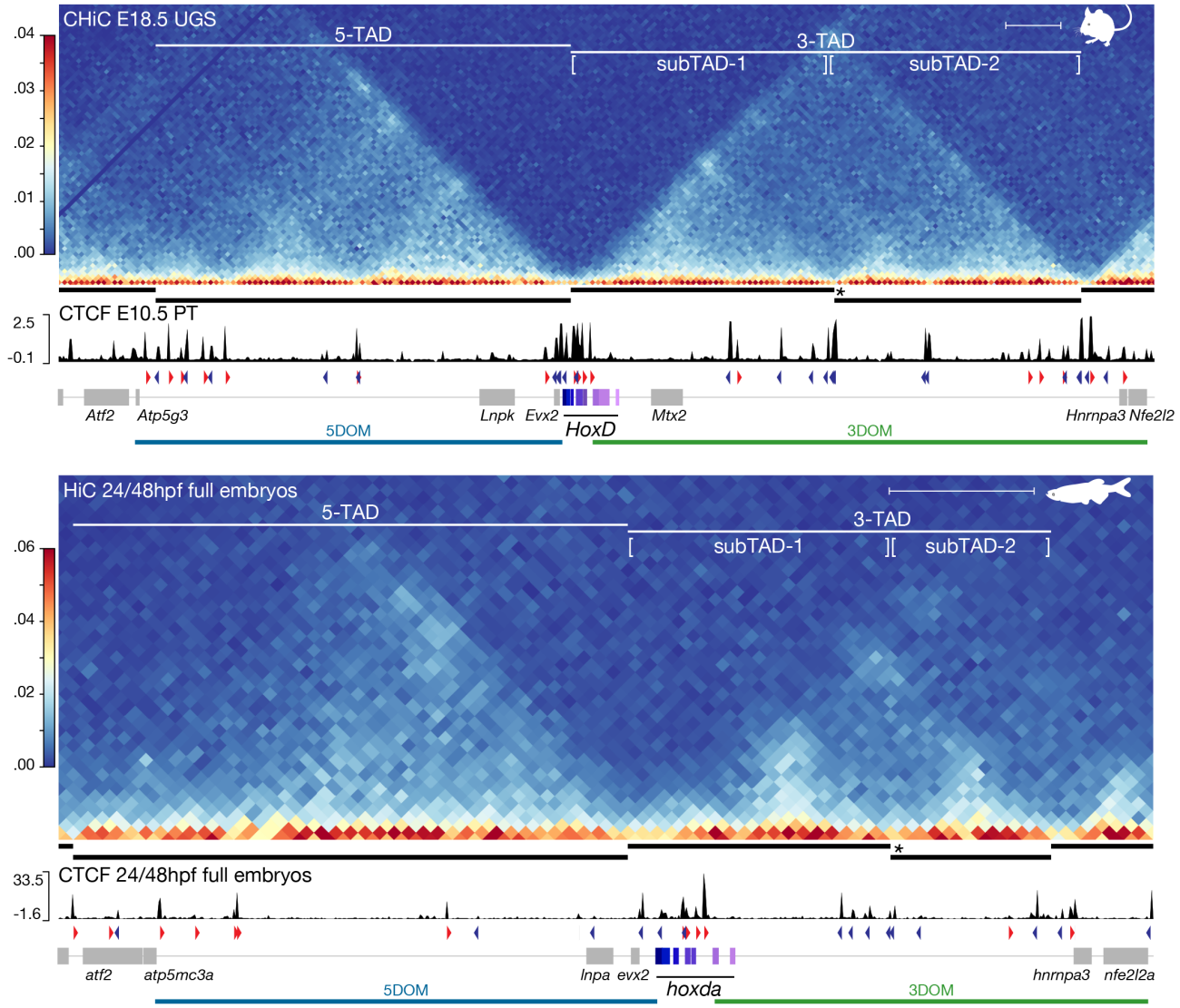
File S2. Sequences of the mouse probes used for WISH

File S3. Sequences of the zebrafish *hoxda*$^{Del(3DOM)}$ and *hoxda*$^{Del(5DOM)}$ founder alleles

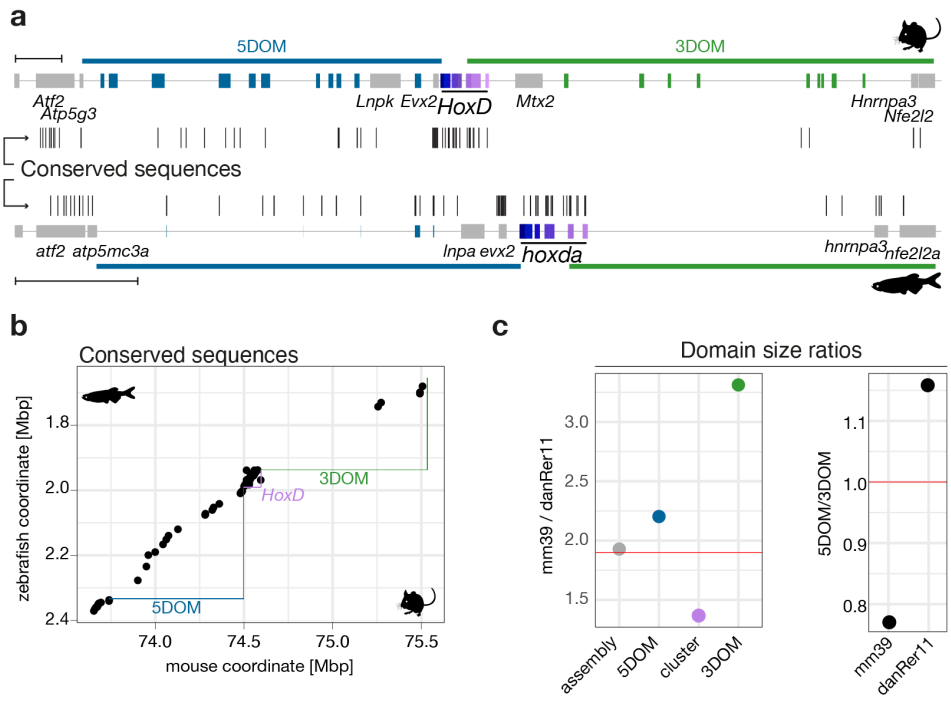
Hintermann*, Bolt* et al.
Figure S1



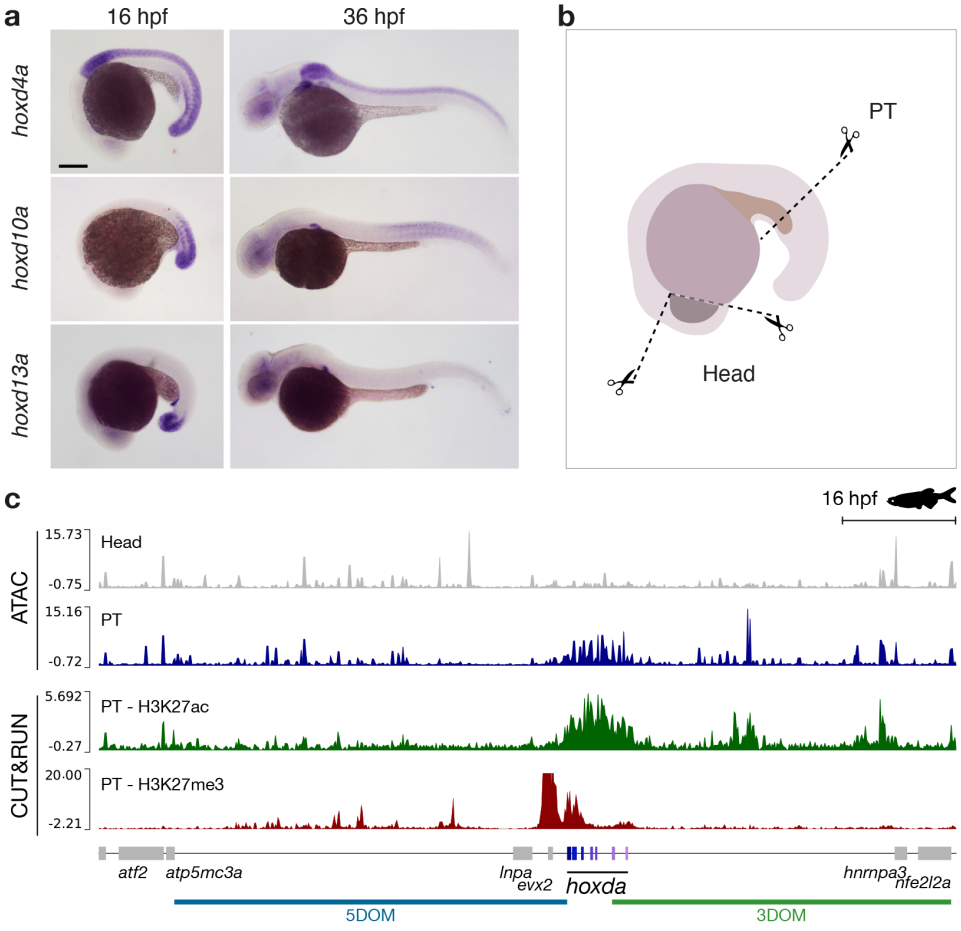
Hintermann*, Bolt* et al.
Figure S2



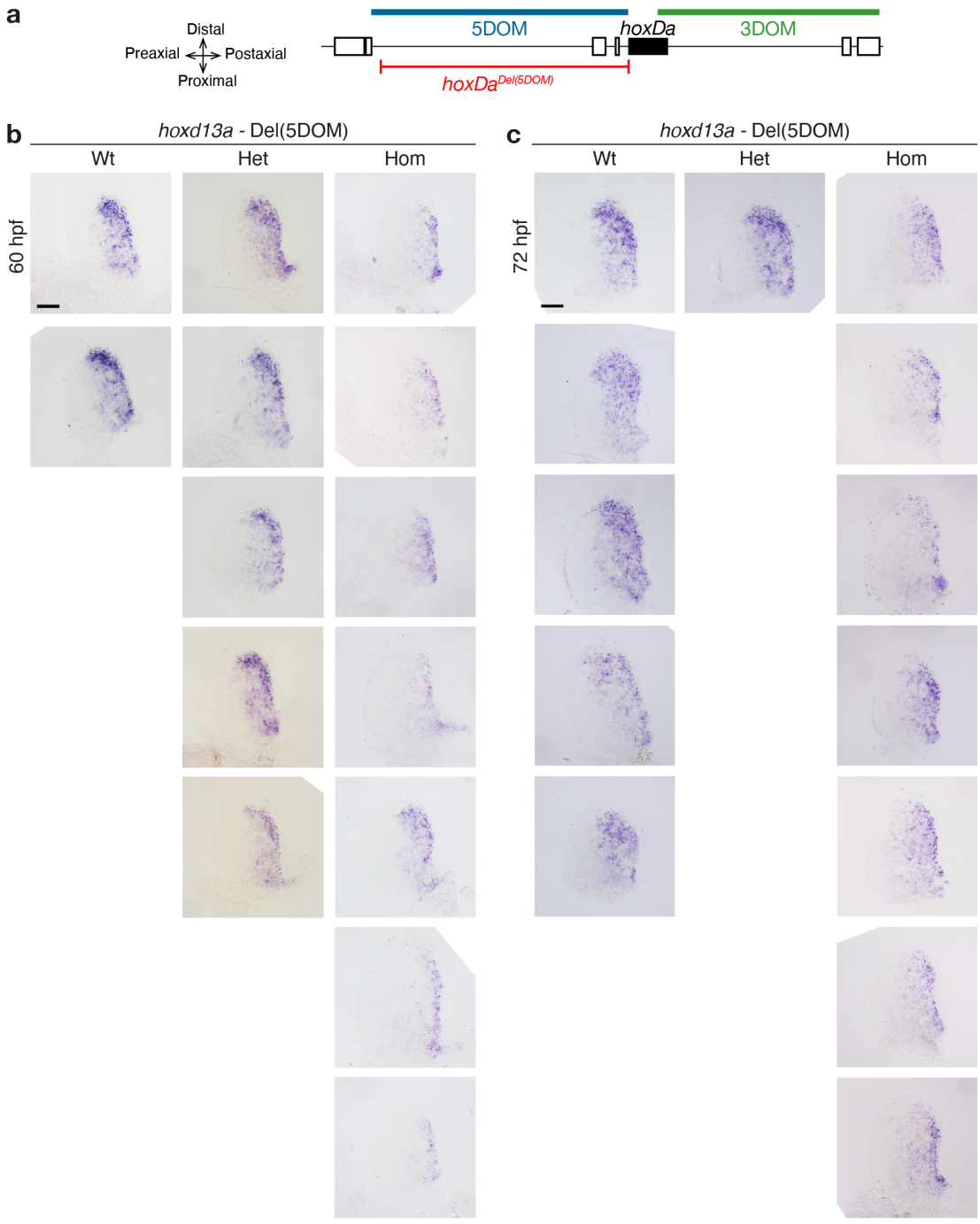
Hintermann*, Bolt* et al.
Figure S3



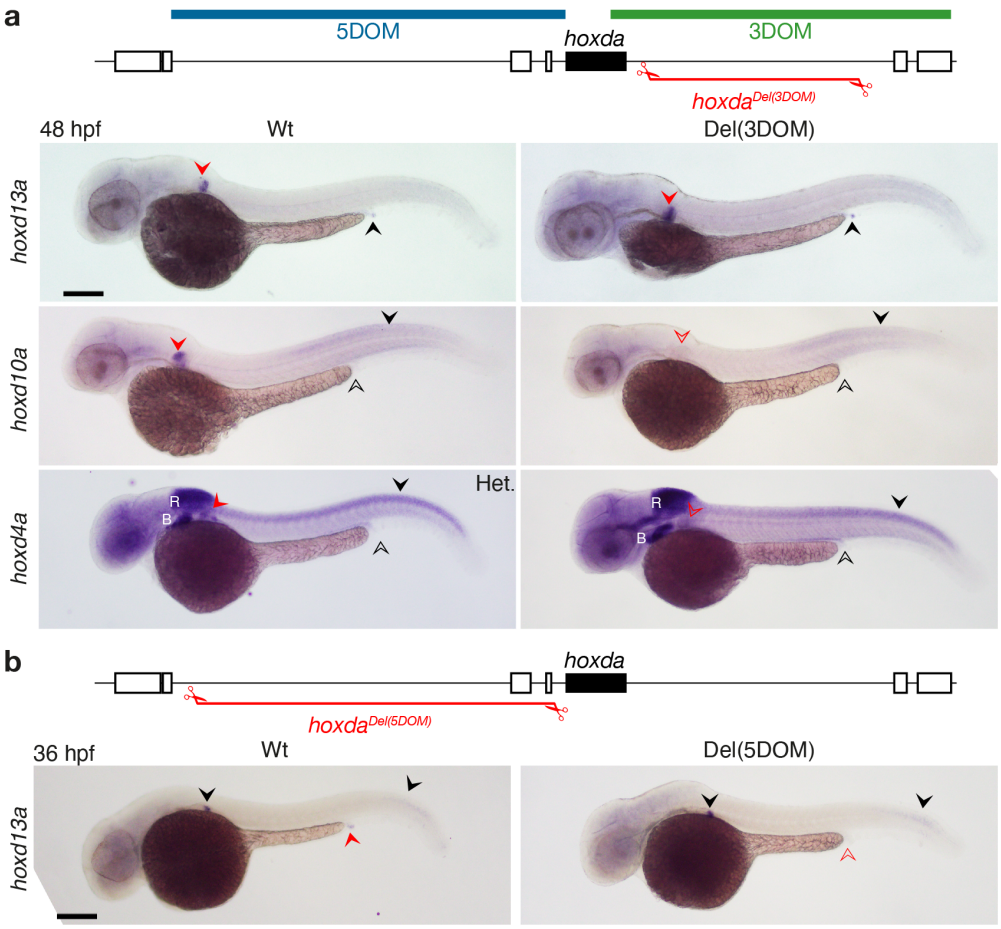
Hintermann*, Bolt* et al.
Figure S4



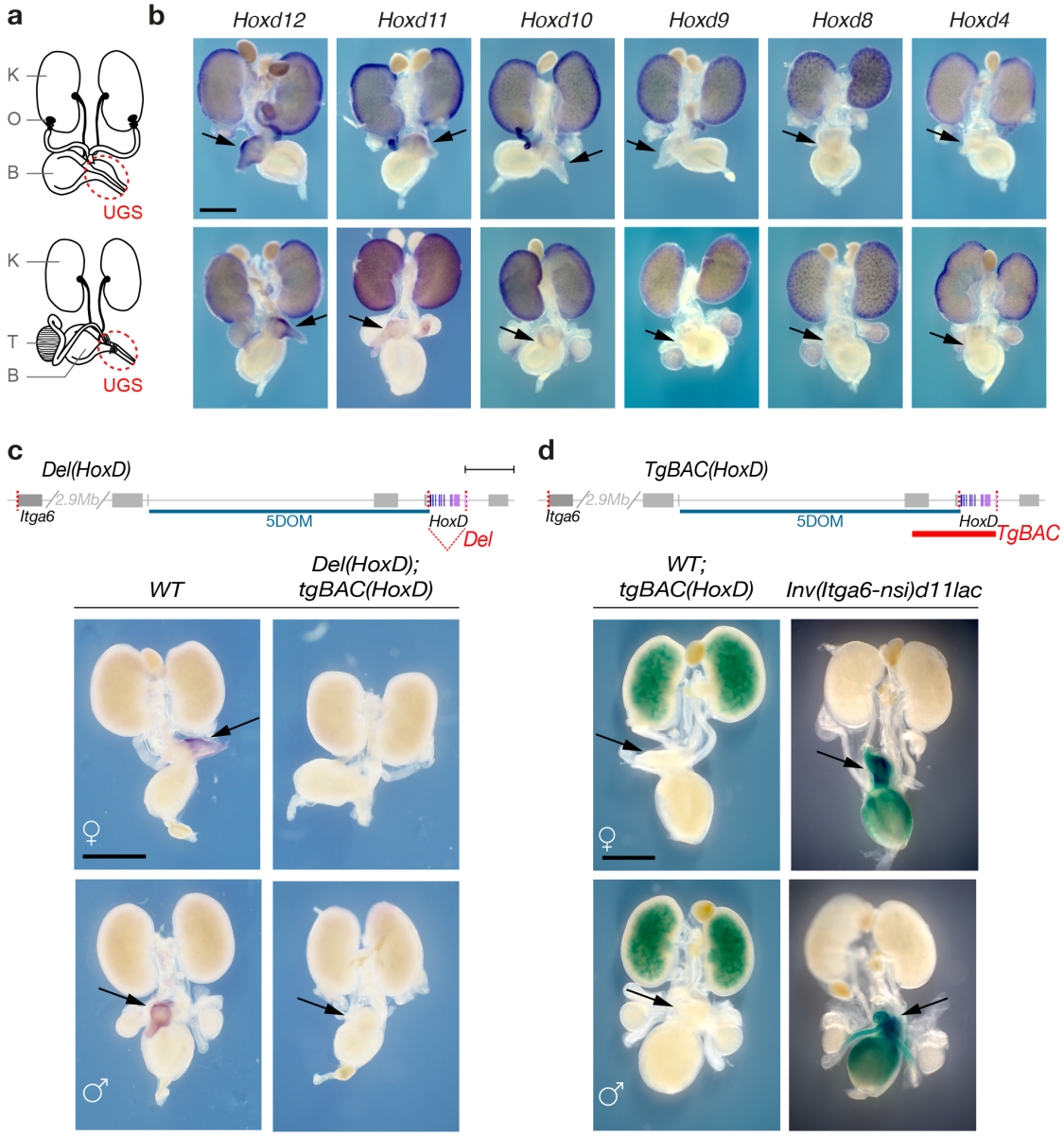
Hintermann*, Bolt* et al.
Figure S5



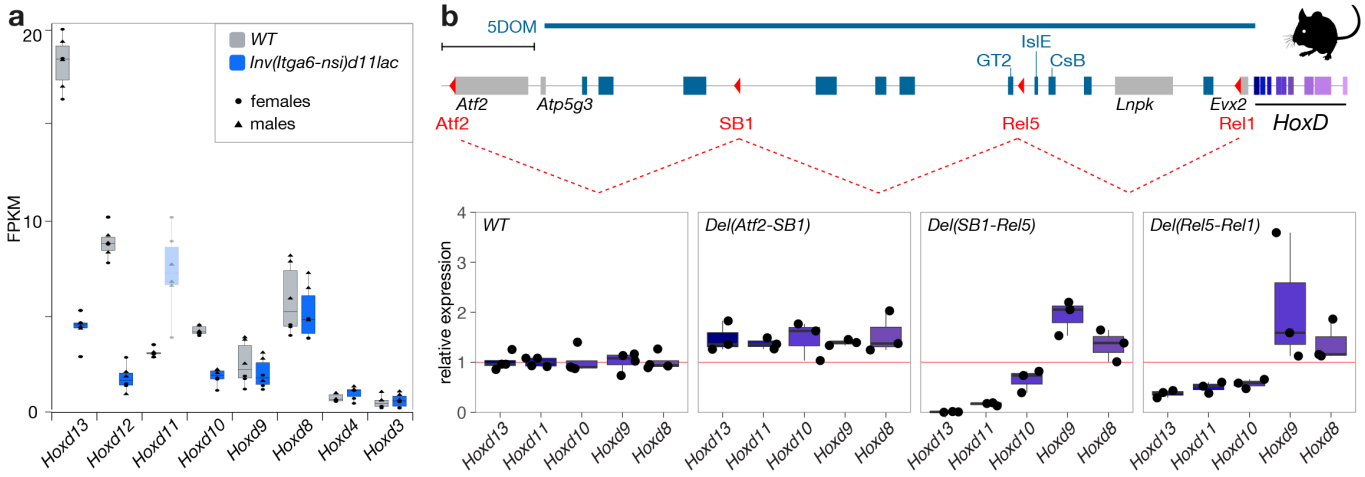
Hintermann*, Bolt* et al.
Figure S6



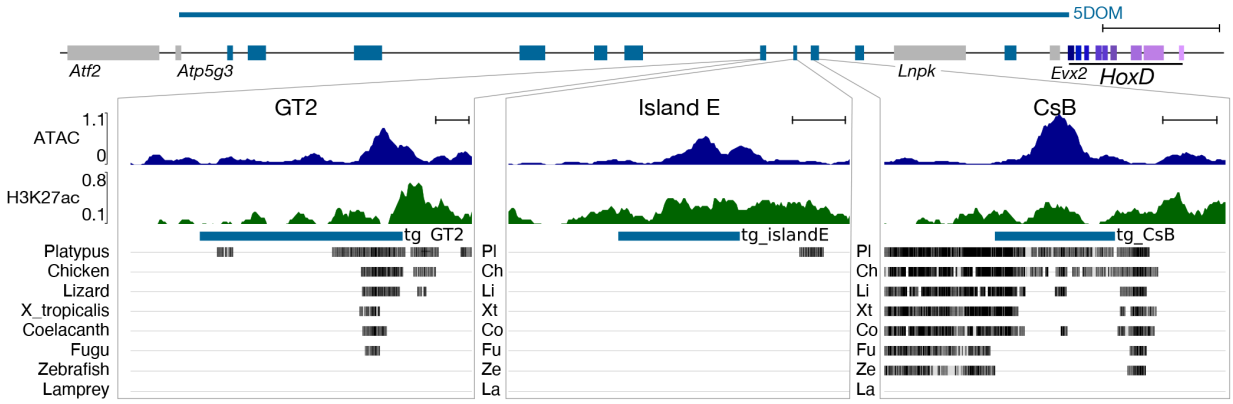
Hintermann*, Bolt* et al.
Figure S7



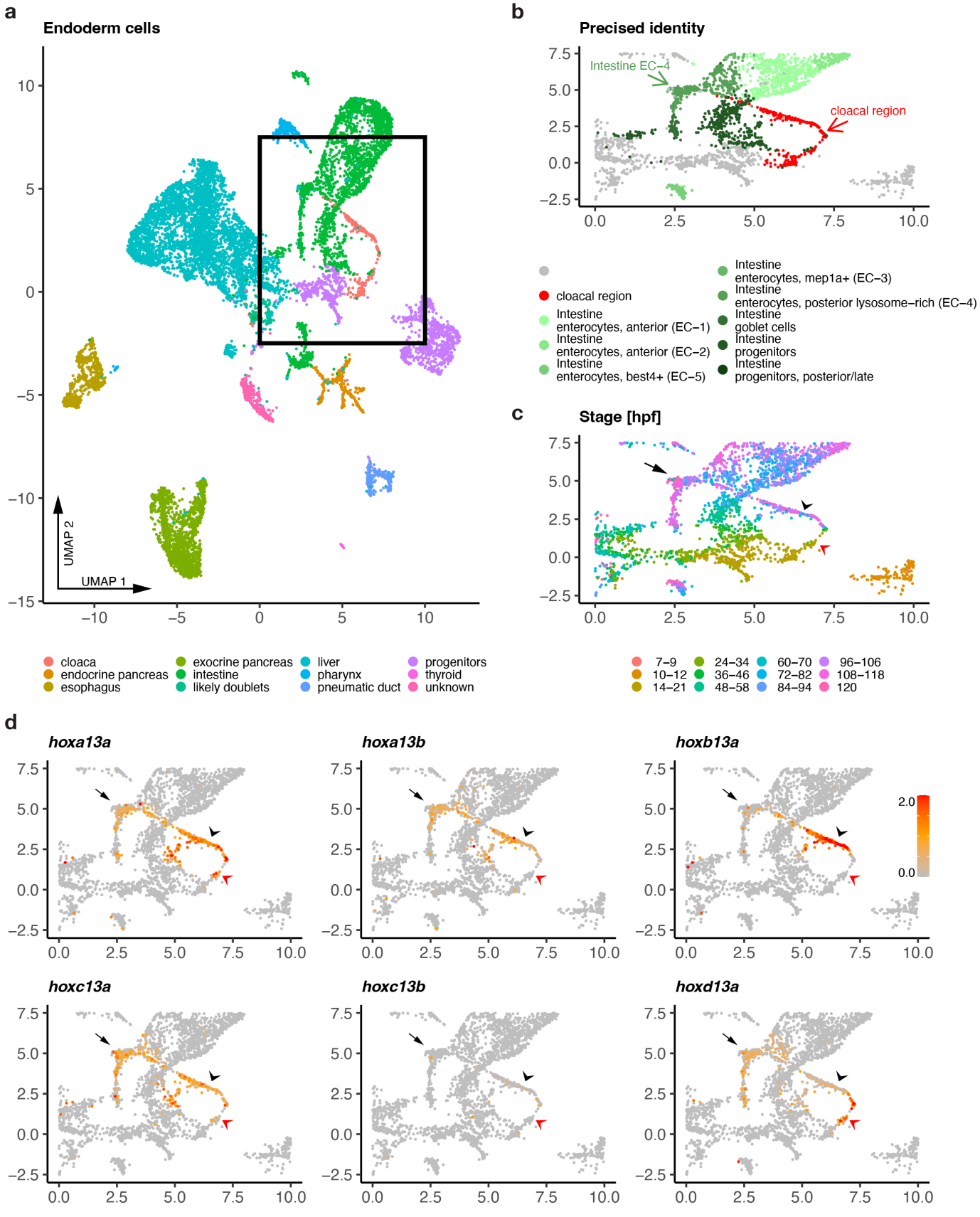
Hintermann*, Bolt* et al.
Figure S8



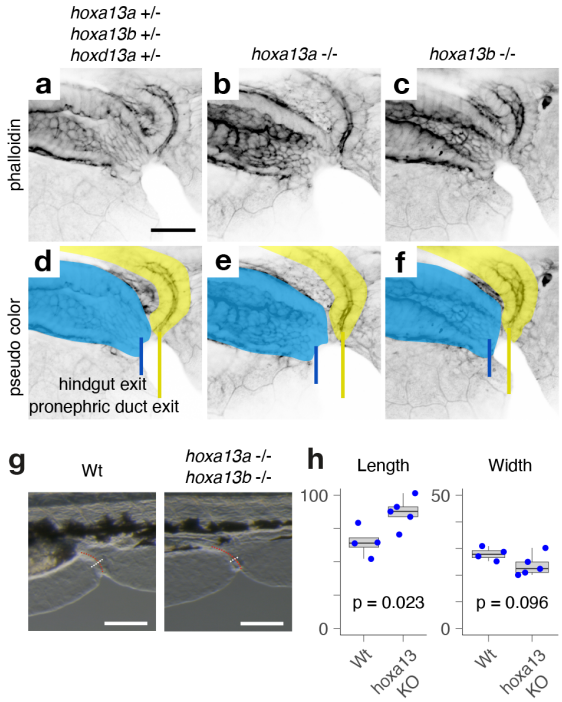
Hintermann*, Bolt* et al.
Figure S9



Hintermann*, Bolt* et al.
Figure S10



Hintermann*, Bolt* et al.
Figure S11



Hintermann*, Bolt* et al.

Sequences for zebrafish ISH probes

>hoxD4a_probe_danRer
GAGTGTGTGCGCGATCTCGATGCGACGCCGTCTTGTTAGATACCTGTTAAAATGAAACTCCTTCTCTAGTTCGAGAACCTGCTGTC
TTGTGTAGGCTGTCCGAGAACGTTTGGGTTCAGGTCCTGTGTAATCCGGGTTCACCGTAGTAACGTGCACTTTCTTCATCCACGG
GTAAACTACAGCAGGCTGCTTCGTTGGTATCCCGTTCTGGGTCTTTGTGTTTTGTTGTTGTCCACCAGTCCTCGATCCGGAAATCT
GGACCGCGGGACACTGCTCTGTCTGTGCAGGGAAAGGGCTAGGGGTGCTGGCTTGATCCTGCACATGACCCCGCGGCTGCACC
GACGAGCCCTGGACAGTGCTACAACTGTAAGGCTGTTCAGAGTAGTTTGACCGTGAATAGATTCCGGGATGCTGGAAATCAGTG
TCCTGCGACGGACTGTAGTAGCCTGGGCTCTGTTCAGGTATATAGCTGTTCTGAGAATATTCCTCGCAAGGAGGAAATTTGGGATC
CACATACTTGGAGTTCACCATGTACGAACTCATGGCCATTAATTTCTGAAGGTAGGAAATACTAATTTTTCTCGAGTTGTCTTTTTTT
CCTCCCTCCATAAAGCCCTCC

>hoxD10a_probe_danRer
CAGAACATGATTCGGGAATCAGGCTGGAGTACGCGGGTATCTCAGACGAACTGACCTTTGCTCTCTTGTCGGAGTACATACAACA
GTTAGTCTCTTCTTTTATGCTTTGTGAAAATGTGCAGGTGGACATCTGGTTGAGCGGTTGCTCTAATCGGCAGGACCTGCTCGGAT
CTGCCCAGGTATCAGTTTGAGGTATGTACGAGTGGACGGTCATGTCCATGTTTTGGTGGTTTACTTCCCCCCGTTTGCCAAGAGCC
GGCAGGAGTCCACAGGTTTGCATTCCATAGGTTCCCATTTCTGCAGTGGCTGCTGGCATGTACATGTTACTGCTGGAATAAAAATC
TGTTCTGCACGCCCCAATCAAGGAATCTACCAAAAAAGTGTTTGCTGCAGGAGAGCTGTTGGGAAAGGACATTTTGGATTTTTTT
TTGTTTTCTTTTTTTGTTTTAAGAAGCAGAGAAGTGTCCTCTGTAAGCTGACATATCTAGCAAAACAATCCCTGTATAACCCAGCGA
GTCTCTCTGCAACCACATGACACTCTCACCAATGAAATTTGAAAATGGCCTTGAGACGCAGG

>hoxD13a_probe_danRer
TGCAGACGTCAGGTCTCTTCTTATCCTTGACCCGTCGGTTTTGAAACCATATAGTTACTTGTCTCTCAGACAGGTTGGTAGAAGAA
GCGATCCGCCGTCTGTTCTCCTTTGTAATGAACTTAGTGGTGTTGTATTCACGCTCGAGCTCTTTCAGCTGAAATTTTGTGTAAGGA
ACCCGCTTCTTTCTCCCGCGCTGACAGAATGAAGCCGCTGCTGCTTCCTCTGTTAATGATGGTTTCCAGATATGCGGGCTCCGCGT
CTGGTCTTGAAAGCAATAAAGTTGACTGCTGCAATTGTTTGACCAGTCCCAATGTTGGTGGCCTTCCATTGTCAAACAAGTCTCAT
GTCTAAGATCTCCCATCATTGCTCTCTGAACCACAGGTAGATCAATGAAGGCAGGGATTCTCGGATAGGAGCGCGCGCAGCCTTG
ATAAAAGGCGAACTCTTTAAGCCAGCTGGAGGCTTCACCGTGGTCCACGGGTTTATTGGCGAACTGTCCGTTTTGCTTCGCAACA
CCAGGCGGGAACACGGCGTTCTGTGGAACTTTGTAACTGTTTTCCAAATGGCATCCAAACGTGAAGGAGCCAGAGCCGATATTG
GATGGAAAAGAAAAGTAAGGACTGATAGACTCATTGCATACAGAAGTCGGTCGATCAACAGCGGAGAGCACCGGCGCTCCTGAT
GTACACCGGCTTGAGTGAGTCC

Hintermann*, Bolt* et al.

Sequences for mouse ISH probes

>Hoxd4_probe_mouse
TAAGACTATTCACCATCTTAAGGTTTTTGGCACGCAGGTGTCGGTTATTTTCTGCAGTCCATTCCCCGA
ACAAATAAATATTTAAGCAAATATAAACTAAACATAAAATAACCCGTTTCCTTCTTGGATTTATTTCTT
TGGGCATGGGCCACTGGAAGCAAACAATTTCTTTCTTTCTTTTTTTTTTTTTTTTAAATCACAGAAGAAA
TTCTATAGATAATCCAGAAAGTTCTGGGATTCTGGCTCAGCCTCTCCTGCTTGCACTGAAAGCACCTAG
CACTGGCTGGCCTTCTGCCTCAACTCCCCCACAGTTACTAAGCCACACAGCTCCCTGAAGGGGGTGGT
GTGGAGGAGACCATCCCGGTTCCCACTGGGGTAAAGCCTCCAACCTTCATTTGCAAATTCAATCCACC
AGCAGAGATTTTCAGGAGTTTAATGACTCGCCAAGGGTTTATGTTCTCAGTAGCTCCTGCCATATGATA
ATCTACTTAAAGAAGCCCGCCCACCCTCCCACCCTTCCCCACTTTAGGGAGGGGGCTGCCCTCGGGCA
GGAAGGTAACCTAGTCCGAGGGCGGCCTGCAGGTCCCAGGTCCACCCCCACAGTACCTTAGAGCCCAG
TGAGGTGACAGCAGGCCCCGCCTGCACCCGCAGCTTCGCCTCAACCTGGAGTGCAAGAAGGGATAGG
CCCAGGGTCCCCACTTCTATAAGGTCGTCAGGTCCGTATGGTGGTCCTTGGCCATTGGCTGCAAATGCT
GGCCCGGGGCAGCTGAAGAGGAGCAGGAAGATGAGGAAGAAGACCTGCCCTTGGTGTTGGGCAGTTT
GTGGTCTTTTTTCCACTTCATCCTCCGGTTCTGGAACCAGATCCCCGGGCGAGCTCGAATTC

>Hoxd8_probe_mouse
TGCAGGCATAATTTTTAAAAACATTATTAATAAACTTCTTTTAATTTACTCACCATAAAGATTTAAAAT
GTTCAAGTGCACTTGTTTGATTAATAAAGCATTTTATATTCAATCATCTAGCGCTGATCATTAGGAAGT
TTATGTTGTTGAGGCAAACCACCAGCTCAGACGTTAAAAATTACCATTATAAAAATATAGCCACAAGG
TCCGTATCAAATTACCCTTCTATAGGACAGCTCAAGGACTGCTTCTATGCGATACTACAAACTATTCTC
TTTAAAAACACAGAAATCCCTTTCAAAGCGCAGAAACCTCAAGAGCAGTGACATTAAGGTAGCTTGAA
ACACAGGTAAAGTCCGTTTCCAAAACACAAGAGGGAGGACCGGTGGTGTCCACCAGTATATACTTATA
CTGATCATCAGTCTTAATAGTCTGTGGTAGAAGTGTAAAGGTAGAGTTTAATTTGTGGGGCAGCCTTCA
GCTCCGTCTTCTTCCAGCCCCGAGACTTCCTTTTTGGGGTCTCCATCCTTTGCCTCCGGCCGGGACGCA
GGAAACTTGTCTTTGTTGTTCTCCTTTTTCCATTTCATTCTCCTGTTCTGGAACCAGATTTTTACCTGTCT
CTCCGTGAGGGCCAGAGTATGGGAGACCTCGATTCTCCTCTTCCTGGTCAGATAAGGGTTAAAAAGGA
ATTC

>Hoxd9_probe_mouse
CTGAGTTGGATTTTAGTTGGGAGAGCAGGGTGGGGGTGTCCACACTCTGTCTTCCTTCAGATTCTCGGA
ATTAGATCGTTGGCCCTGTTTGGGAGGCTGGATAGGGTGAGACTAAACTACAGCACTTTTCAGAAACA
TGAGGGACATTTACATAGAATAGGGCTCCATCAAGCCTGCGTGCACTCACAAAATATAGCATTTCCTA
AGGAATCAGGGGAGCAGGGAAGGGAGGGAGTGGGACAGGACAAGACTCCAAAGCCACTTGGTTTCTC
CAATACATATGCGAGCAAATACAGTGTCCCCCAGATAACTCACAGAGACCAACCACACAGTCACTTCT
ACCTAAAAAAAGAAGAAAGTCCGAGTCGCTGCAGAGTTTCTGAATCAAGCACCCACAAAGAAAACAA
ACAACAACAAAAAACAAAAAACAAAACAAACAAACAAACAAACAAAAAAACCCGCTGCCTTTTGAG
TGCTGCAGGCCAGACCCGCACTGGGTCAGTCTCCTTTAGGGCACTTCTCCTTGCTCATCTTTTTCATTTT
CATCCTACGGTTCTGGAACCAGATTTTGACTTGTCTCTCTGTAAGGTTCAGAATCCTGGCCACCTCGTA
GCGCCGGTCCCGGGTGAGGTACATGTTGAAGAGGAATTCCTTCTCCAGCTCTAGCGTCTGGTATTTGGT
GTAGGGACAGCGCTTTTTCCGGGTGGAGCGAGCGTGGATC

>Hoxd10_probe_mouse
GGTTTTCTTAATCTTTTGAAATATATATAGATATAGGTTGTTTGTTTTTTTAAATACTGGATGTTTTCAA
AATGCCATGAGTTCATGGGCTAACCACACACAGAGAACTTTTATGTTAGGAAAAACATGGAAAAGTCA
CATTGCTGGCGGGATAGATGGGTGGGATGCGATGGATACGTTTAGATAGATGTTCTACAGTTCCAATA
AGTTAAGACCCAAATGAAAATGCGCAGTAATAGTTACCCTGCATTACGATGAAAAAAAATAAATTAC
ACGTGCTAAGTTTTTTTAATATATATATATATGTATACTCATAGGGATTACAGATCACTTGCAGCACGA
ACAGAATGACCCCAAAGTCACATTCTAGCAGGAAAACAGAACAGAACAGAACAAAACAAAACAAAA
AAACCAAAGAGCAATGAAGGGTTTGTTTCCGCCTTGCCAGTCCTCCAGCACTCAGGCCTGGAACTCGG
TGCCCCCTCTCGGATCCTGGCCTCACATGGGCCTCAGACCTAAGAAAAGGTGAGGTTGGCGGTCAGTT
CTCGGATTCGATTCTCTCGGCTCATCTTCTTGAGTTTCATTCGGCGGTTTTGAAACCAAATCTTGACCTG

CCTGTCGGTGAGGTTAACGCTCTTACTGATCTCTAGGCGGCGCTCGCGGGTGAGGTACATATTGAACA
AGAACTCTTTTTCCAATTCCAGCGTTTGGTGCTTGGTGTAAGGGCACCTCTTCTTTCTGCCACTCTTTGC
AGTGAGCCAATTGCTGGTTGGAGTATCAGACTTGATTTCCTCTTTGCTTTCCTTCTCCTGCACTTCGGGA
CTGGACACGGAAACCTCAGCCAGGCAGCTCCTGTCTTCTGGAAGGCCGCCTTTGGCCTCGGGGCTCTC
CACTTGGGAGACTTTAGTGGGCTCCTGGCCGCTCGCGCTTTCGTTCATCTTCTTTTCCATCTGCAGCTGG
GCAGCCGAGAGCTGCGGCTTGGCCGCGCCACGAGGGTTGAGCTGGAGCATGACAGTGGAGCTGCCTT
CGGGGCTGTTATTGCACTCTTGGGTTTCCCGGTGGCGTAGGTCTGACTCAGTCTAAA

>Hoxd11_probe_mouse
TTTTGCTAGAAAACATTTATTCACACACAAAAAGACAGTGACTCATGCCCAAAGGTACATTTCCCAGA
GTTTAAATAATCTGTGTTCACTGAAGGGAAAGAAAGGGATTTCTTAAACTCAAAACAAAACACTGGCT
CCTTCTAGCCCCATGGCCTAACTTCCTACAGATCTCGCCGCGTTGAGGACTCAGAGAGGGAGGAGAGG
AACAACAGCGCATTGGCGGTAAAATTGGAAAGGGATGTTTACATTTACCATGTTCCCCTTATTAGGAT
CCGCACAGCCAAGTTCCGGGCCTCTTCCTCAGGCTGTAGTGGTCGCTGGGAGGTCCCGAGTTGAAGGC
GAGAAGAAATCGTGAAGTTCCACAGAGAAGAGGCGTCATTGAACCCAAGGACATTGGGCCTGCTGAG
AGCAGGCTGGTGGGGAGGAGGGTGGGTGAGTGGGGCTCTGGGGAGGGGGTACATCCTGGAGTTCTCA
AAATAAGGGGTTTCCAGTGAAATATTGCAGACGGTCCCTGTTCAGTTTCTTTTCTTTCATCCTGCGATT
CTGGAACCAGATTTTGACTTGCCGGTCAGTGAGGTTGAGCATCCGAGAGAGTTGGAGTCTTTTCTCTTT
GTTTATGTATACATTAAAGAAAAACTCGCGTTCCAGTTCGCGGATCTGGTACTTGGTGTAGGGACAGC
GCTTTTTCCGGGACCTCTGGGGGGCCACTGTGCCGCCGCTCTTTTCGGCCCCCGCCTCCCCCGGGGGGC
CCTCGCCTTCACCGCCACCGCCTTCC

>Hoxd12_probe_mouse
TGGACAGATTCTATAACAGGCGTAGAGCAACTTCATTCACAATTCAGAAAACCTAGGTTTCTCAAGGT
AAACCTGGCAGGGCTTGGGGCCTTAAAAGCTTTACAAGAGCTGACCACCAGTCTCTGAGAAGAGACA
CCCTAGCTAAGTTCTGATGGAGTTGGTGGGTGGTGGAGAGGATGGGGAAGAAGTGCTGGTTTTCTGTC
CTTCAAGAGCAGCAACTACAGTCTGAGAATCTGAGAGGTTTCTTCTCCCACCCTTCAGACTCCTCAAGG
AGGCATTTTTTTTGGGGGGGGGGTCTCTTCCTGTGGTGATACATTAAGGAAATTCAAGAAAGACAGAC
ATATGGATAACTTCATTCATTCAAGAGGAGAGTCGTGGTTACAGAGTGCAGATCCTCCTCCGCCTCGCT
TACTCTTCTTTCTTCCCTGTGCAGCCAAGCCTCAAAACAGGCCAGCAGCTCCCCTCACCCATAGCTCAG
TTCTGGGGATCTAGTTACAGAGTGCTTGGCCTTGGCTCAGGGATAGGTGAGGCTGGAGCAGGGGAATT
ATGACAGTCTAGAGACCATACTTTGTTTTGGAGTCCCAGTTTCTCTCCCTGTCTCTGAGAAGGAGATCT
TCCAGCCTGTTTTACTCAGCAGGAGAACCTTTGAAGCGCTGAGGTTCCAGGGTATCCTTTCCTCCCCCC
CCCCCCACAAGTCTGAAGGTCTAGGGTAGCTTTTGTCCTCCCAAATCATTCAGGGCAGATGGGGGTTG
TGGAATCAGGCCCTTTCCTTCCTGCAGAGTGGAGAGCCCAGCTTTGTCACAAAGGGCTTCTGCTGCGA
AGGGGCTAGAGTTATCCCCAGCCCCTTGCACTGCAGCGGGGTCTCAGGGCCTTTCTACCTGCACTTATA
CCCGGAGCTCTAGCTAGGCTCCTGTTTCATGCAGAAAGAGCTGGATAGAGAAAGAAAAAGAAAAAGA
ACACGTTGAAAACCGGAAAAAACAAACCCTCATACAGTGTTTCAATAGTGAGCCCCGGATGTAAACAT
CATAGACAAGGACAGGTGACCCCCAGACACCATGCTGAATGTTTAAAGCCAGTGTTAGATTGCAATTC
CCAAACACCTCTCAGGAGGGTCCCAAAGAGAGCTGAAAGCGAAGGGGGCTCCACTGGCCTCATC

>Hoxd13_probe_mouse
TCTAGAGGCAGCACAAGCTTGAAAAACACAAATGGATGAAGACTCAGTGGAGACAAGAGAGTTCTTT
TTATTTATATCTCCCTAGGAGTAAAATGTGATTTTTAAATGCTCGAAGCACTGCTCACACACACTGTAT
CTACAAAAATAATCTTTCAGTCCCATAAATACGCAACATTTTTGAAAGTAGTTCATCACTTTCAGAAGT
TTCTTTTTCCCTCTTGGTCCTTAGAGCTCCATATTTAAAAAATGTGTTAGATATTTTAAAGCTCTGAGAG
GCCCATGCATTCAGACAATTAACTTCTTCACTGAGCTCTGGACAGAAGTTCCCACCCATGGCCACATTG
GCTGAGTCTCTTTCCAAGGATTGAGGAGATGGCCTTAGCTCCATTCCCACAGACCCAAGGAACGGGAT
GAATTTATACTATCTCCAGTCACTCTGCCAAGTGTAAGTATTTTAATAGCTGCATGTAGAGTTGGAGAA
ATAAAGGCATCCACTAGGAATGCTGTTTTCAGGTTCAAAAATCACAGCTACAAGACATGTTAAAACCA
CACAAAACAGCAAAGGAGACTGCCAAAGATAAGGGTGATTGTCTTAATTTCAATATTAATAAGGTAGC
TGCTTTTAAAGTCATCACGTGCCTTCAACCTCCAAGTTTCAGACTCTCCCAGCGGTCCTAAATCACCTA
ATGCCTAAGAGCTGTTTCCACTTGCCTTGGGGAAACTTCTGAGAGCTAATGATTCAATTACACAATCAC
AGGGTATGAAAGGCTAAAAAAAACGATTCTGGCCTAATGGATTATAAGCATAATCAGTGCTGGAAAC
ACATATAAAATCCTGTTAAAAAAATAAACCCCTAAATACAAGGGCACTTCTGAAATTCATTTGGGGGG
ATTATTGAAAAAAATACTTCAAGTTTTTTTTCCTCATACTGGCCTGGAGG

Hintermann*, Bolt* et al. 2024

Sanger sequences of the zebrafish founders for *hoxda*$^{Del(3DOM)}$ and *hoxda*$^{Del(5DOM)}$

>del3DOM_founder1
NNNNNNNNNNNNGNNNANTNNNCTCNTGACGACTACAAAGTTATTTAATCTACAAACCAAGTTGATTTTATATACTCCATTG
GGGTTTTGTAGCGTATTATTGACGTCTTGAATCCTTTTAGAGCGTGTGTGGTAATGTGTAGCCTCTTTACAGACTAATAAATGCT
ACAGGCCTCTCACATGATATGATTAATAACTGTAAAAGAAAACGTTGCGCCACCAGAAGTGTCTAAAAATGATAATCTAGTCTC
CCTATTGGAGGCCATAATGTAATGTGATGATGTATGGAGACCTGCTAACATTAACACAAGCCACACTACTCCTTTCACGTGCTA
CACTAAAAAGAGAGAAATTGGTGGAAGGGCATTGCTCTTCAACTAAACATCACAAATGTGCCGTTTCAGCAGTGTTACAGCAA
NTNC

>del3DOM_founder2
NNNNNNNNNNNNNNNNNTNNNCTCATGACGACTACNAAGTTATTTAATCTACAAACCAAGTTGATTTTATATACTCCATTGG
GGTTTTGTAGCGTATTATTGACGTCTTGAATCCTTTTAGAGCGAGTGTGGTAATGTGTAGCCTCTTTACAGACTAATAAATGCTA
CAGGCCTCTCACATGATATGATTAATAACTGTAAAAGAAAACGTTGCGCCACCAGAAGTGTCTAAAAATGATAATCTAGTCATT
ATGGCCATGGAGGCCATAATGTAATGTGATGATGTATGGAGACCTGCTAACATTAACACAAGCCACACTACTCCTTTCACGTGC
TACACTAAAAAGAGAGAAATTGGTGGAAGGGCATTGCTCTTCAACTAAACATCACAAATGTGCCGTTTCAGCAGTGTTACAGC
AAANTNCCGGGAGA

>del5DOM_founder1
GGNNNNNNNNNNNNNNNNNGGANTTCCNTGCATGCCNGAACATTTATGCTTCCTTTTCCCAAATGCCTGTAAAGCACAGA
CATTTTTCCTCCTTAATTACTTCCCGAATTTACCCTACTTCTGTCCCAATTTTTTTAAACTTACTTCAAATATTTGAAGTGTGATTTT
TGTGGCTGTCCTGAAGCGTTCAATCCATCTATAGCACAATCGCGTGTGCGTTTTATTTGAGTTTTGGGGATTTTCTGAGAGGAT
CCCATTGCGAGGTCATGTGTCTTCCATTATCTGGGTGTTTTCTAATCAGCCTGCAGTGGCTGANATTAAACAGGGAAGAATTCT
CCCATCAAACCGGACTGTCGTAAATCCAGCTCACCTTGGAGAGCAGAAGGGAAGTGCAGATCTGTAAACCCATCATCGCTCCG
CTCCCCCACACACACTTATGCTCCACAANAGGACNNNNN

>del5DOM_founder2
NNNNNNNNNNNNNNNNNNNANNNNNNGCATGCCAGAACATTTATGCTTCCTTTTCCCAAATGCCTGTAAAGCACAGACAT
TTTTCCTCCTTAATTACTTCCCGAATTTACCCTACTTCTGTCCCAATTTTTTAAAACTTACTTCAAATATTTGAAGTGTGATTTTTGT
GGCTGTCCTGAAGCGTTCAGTCCATCTATAGCACAATCGCGTGTGCGTTTTATTTGAGTTTTGGGGATTTTCTGAGAGGATCCC
ATTGCGAAGGTCATGTGTCTTCCATTATCTGGGTGTTTTCTAATCAGCCTGCAGTGGCTGANATTAAACAGGGAAGAATTCTCC
CATCAAACCGGGCTGTCGTAAATCCAGCTCACCTTGGAGAGCAGAAGGGAAGTGCAGATCTGTAAACCCATCATCGCTCCAAC
TCCCCCACACACACTTATGCTCCACAANAGGACCNNNN

Hintermann*, Bolt* et al.

Table S1: Sizes of mouse and zebrafish domains

mm39_short_name	mm39_RefSeq_ID	mm39_transcript_start_chr2	danRer11_short_name	danRer11_RefSeq_ID	danRer11_transcript_start_chr9
Hoxd4	NM_010469.2	74552322	*hoxd4a*	NM_001126445.2	1951004
Hoxd13	NM_008275.4	74498569	*hoxd13a*	NM_131169.3	1990311
Nfe2l2	NM_010902.5	75534860	*nfe2l2a*	NM_182889.1	1654399
Atp5g3	NM_001301721.1	73741670	*atp5mc3a*	NM_201176.1	2333895

Domain	Domain_name	mm39	danRer11	mm39/danRer11
Atp5g3-Hoxd13	5DOM	756899	343584	2.2
Hoxd4-Nfe2l2	3DOM	982538	296605	3.3
Hoxd13-Hoxd4	cluster	53753	39307	1.4
whole genome	assembly_Gb	2.7	1.4	1.9

Ratios	mm39	danRer11
5DOM/cluster	14.1	8.7
3DOM/cluster	18.3	7.5
5DOM/3DOM	0.8	1.2

Name	Assembly	Ratio
5DOM/cluster	mm39	14.1
3DOM/cluster	mm39	18.3
5DOM/3DOM	mm39	0.8
5DOM/cluster	danRer11	8.7
3DOM/cluster	danRer11	7.5
5DOM/3DOM	danRer11	1.2
5DOM	mm39/danRer11	2.2
3DOM	mm39/danRer11	3.3
cluster	mm39/danRer11	1.4
assembly	mm39/danRer11	1.9

Assembly	Annotation source	Release
danRer11	NCBI RefSeq genes, curated subset	Annotation Release NCBI Danio rerio Annotation Release 106 (2019-10-28)
mm39	NCBI RefSeq genes, curated subset (NM_*, NR_*, NP_* or YP_*)	Annotation Release NCBI RefSeq GCF_000001635.27-RS_2023_04 (2023-04-11)

Hintermann*, Bolt* et al.

Table S2: Accession numbers for re-analysed data

Name	SRA number	Publication	DOI
CTCF_ChIP_E105_PT_rep1	SRR17750150	(Hintermann et al. 2022)	https://doi.org/10.1242/dev.200594
Franke_CTCF_ChIP_24hpf_rep1	SRR12435909	(Franke et al. 2021)	10.1038/s41467-021-25604-5
Franke_CTCF_ChIP_24hpf_rep2	SRR14670351	(Franke et al. 2021)	10.1038/s41467-021-25604-5
Franke_CTCF_ChIP_48hpf_rep1	SRR14670354	(Franke et al. 2021)	10.1038/s41467-021-25604-5
Franke_CTCF_ChIP_48hpf_rep2	SRR14670355	(Franke et al. 2021)	10.1038/s41467-021-25604-5
Franke_48hpf_wt_rep1	SRR12435867	(Franke et al. 2021)	10.1038/s41467-021-25604-5
Franke_48hpf_wt_rep2	SRR12435868	(Franke et al. 2021)	10.1038/s41467-021-25604-5
Franke_24hpf_wt_rep1	SRR14670388	(Franke et al. 2021)	10.1038/s41467-021-25604-5
Franke_24hpf_wt_rep2	SRR14670389	(Franke et al. 2021)	10.1038/s41467-021-25604-5
Wike_24hpf_seq1	SRR12044304	(Wike et al. 2021)	10.1101/gr.269860.120
Wike_24hpf_seq2	SRR12044305	(Wike et al. 2021)	10.1101/gr.269860.120
Wike_24hpf_seq3	SRR12044306	(Wike et al. 2021)	10.1101/gr.269860.120
Wike_24hpf_seq4	SRR12044307	(Wike et al. 2021)	10.1101/gr.269860.120
Wike_24hpf_seq5	SRR12044308	(Wike et al. 2021)	10.1101/gr.269860.120
Wike_24hpf_seq6	SRR12044309	(Wike et al. 2021)	10.1101/gr.269860.120

Hintermann*, Bolt* et al.

Table S3: Sequences for primers and guides

ZEBRAFISH

Genotyping primers

Name	Sequence (5' to 3')
5DOM_WT_f	GAAAATGGCTGGGCAGGACA
5DOM_WT_r	GACGGTGTGTTCAATCGGGT
5DOM_Del_f	AATGGCTGGGCAGGACATAC
5DOM_Del_r	GTGGTCCTGTTGTGGAGCAT
3DOM_WT_f	GACACACAATGACCCACAATTC
3DOM_WT_r	ACGGCACATTTGTGATGTTTAG
3DOM_Del_f	CCTTCAAAACTCAAGGCCCATC
3DOM_Del_r	CTCCCGGATTTGCTGTAACAC
evx2_f	CGCACTGGCATTCCTCTGTTTT
evx2_r	GGAAGTGTTGTCGTTGTGGTGG
zf_CsA_f	CAGCCCGCAAAGCCTCATTTTA
zf_CsA_r	GTGTCAACGAGAGGAGAAGGCT
zf_CsB_f	ACCAGGAGAAACACCACACACA
zf_CsB_r	TGACCAACTGATAACCCCACCC
zf_islandV_f	CTCATTTGCGCCGCTGTCTTTA
zf_islandV_r	GGTTAGATGTGGGGTTTGGGGA
zf_islandII_f	AGCAAAGCCCGGCTAATAGACA
zf_islandII_r	TGACGCGTGGGCTTAAAATCAC
hoxa13a_8_del_f	GCCAAGGAGTTTGCCTTGTA
hoxa13a_8_del_r	TGACGACTTCCACACGTTTC
hoxa13b_14_ins_f	GATTGACCCGGTGATGTTTC
hoxa13b_14_ins_r	TACACTGGTTCGCAGCAAAA
hoxd13a_f	AAGCCGGTGTACATCAGGAG
hoxd13a_r	GTGGCCTTCCATTGTCAAAC

crRNA

Name	Sequence
hoxd13a_crRNA	CTGAGAGGATCCCATTGCGAAACACCTGGG
atp5g3a_crRNA	AACCATATCCACTCTTCAGGAGGTCATGTG
hoxd3a_crRNA	TGATGCTGCACCCTAAATGG
hnrnpa3_crRNA	ATAATCTAGTCATAGCTGGA

MOUSE

Genotyping primers

primer name	Sequence	Reference
Inv(Itga6-nsi)d11lac		
Inv(Itga6-nsi)d11lac_WT_f	GCAAGCCACTTGGAAACAACTGTTAATGG	(Tschopp and Duboule 2011)
Inv(Itga6-nsi)d11lac_WT_r	CCGTCCAATGTGCGTGTTTTCC	
Inv(Itga6-nsi)d11lac_Inv_f	GAGTTTCTCTTTGCTGTAATGAAGAGCTG	
Inv(Itga6-nsi)d11lac_Inv_r	CCGTCCAATGTGCGTGTTTTCC	
Inv(Itga6-attP)		
Inv(nsi-itga6)d11lac_WT_f	GCAAGCCACTTGGAAACAACTGTTAATGG	(Schep et al. 2016)
Inv(nsi-itga6)d11lac_WT_r	CCGTCCAATGTGCGTGTTTTCC	
Inv(nsi-itga6)d11lac_Inv_f	GAGTTTCTCTTTGCTGTAATGAAGAGCTG	
Inv(nsi-itga6)d11lac_Inv_r	CCGTCCAATGTGCGTGTTTTCC	
Del(HoxD)		
Del(HoxD)_WT_f	GAGCCCGACGCATCGAGATAGC	(Spitz et al. 2001)
Del(HoxD)_WT_r	CAAGGTCCTCAGCCTTAAGAGTGG	
Del(HoxD)_Del_f	AGGGATCCGGAGCATACCACTG	
Del(HoxD)_Del_r	CTCTCTCTACGAGGGAATGTGGAG	
tgBAC(HoxD), tg(GT2), tg(islandE), tg(CsB)		
tgLacZ_PCRb_f	CCTGCTGATGAAGCAGAACA	tgBAC(HoxD) in (Schep et al. 2016)
tgLacZ_PCRb_r	CAGCGACCAGATGATCACAC	
Del(Atf2-SB1)		
Del(Atf2-SB1)_WT_f	GACAATCGTATGCATGGCATACTCGG	(Montavon et al. 2011)
Del(Atf2-SB1)_WT_r	GATAGGAGTGACATTCAGACACGGC	
Del(Atf2-SB1)_Del_f	GTTTTCCCAGTCACGACGTTG	
Del(Atf2-SB1)_Del_r	GCCACTGGCCGAATATTACCTATTTTGTG	
Del(SB1-Rel5)		
Del(SB1-Rel5)_WT_f	GACAATCGTATGCATGGCATACTCGG	(Montavon et al. 2011)
Del(SB1-Rel5)_WT_r	GATAGGAGTGACATTCAGACACGGC	
Del(SB1-Rel5)_Del_f	CAGACTAGGCTTGCCTTACGG	
Del(SB1-Rel5)_Del_r	CCTGCTGCAGGGGTTGGAG	
Del(Rel5-Rel1)		
Del(Rel5-Rel1)_WT_f	CTAGAGAGTACAGCAATGACTTTTGGGC	(Montavon et al. 2011)
Del(Rel5-Rel1)_WT_r	CAGACTAGGCTTGCCTTACGG	
Del(Rel5-Rel1)_Del_f	ACGTGGAGTGGAGTGATGGTTG	
Del(Rel5-Rel1)_Del_r	GGCTGCTTTGGACAATGCTGG	

RT-qPCR primers

Name	Sequence
Hoxd13_F	AAGGATCAGCCACAGGGGTCCC
Hoxd13_R	GTAGACGCACATGTCCGGCTGG
Hoxd12_F	CTATGTGGGCTCGCTTCTGAA
Hoxd12_R	GGCTCTCAGGTTGGAAAAGTAG
Hoxd11_F	AAAAGACTCCAACTCTCTCGGA
Hoxd11_R	AGACGGTCCCTGTTCAGTTTC
Hoxd10_F	GCTGGTCCCCGAGTCTTGTCCT
Hoxd10_R	CCGGTGGCGTAGGTCTGACTCA
Hoxd9_F	CTCCACCCGGAAAAAGCGCTGT
Hoxd9_R	CGGTCCCGGGTGAGGTACATGT
Hoxd8_F_	TTCCCTGGATGAGACCACAAG
Hoxd8_R_	CTAGGGTTTGGAAGCGACTGT
Tbp_F	CCTTGTACCCTTCACCAATGAC
Tbp_R	ACAGCCAAGATTCACGGTAGA

Primers to clone transgenes

Name	Sequence
GT2_F	tccggtcgacTGTCACCACCATCGACAAGT
GT2_R	tccggtcgacATGCATTTCACCGTCTTTC
IsE_F1	ccccccctcgagCTCAAGCCAGACAGGGATGATTA
IsE_R1	cgataccgtcgacGTGGGCTGTTTACTGGCAA
CsB_F1	cccccctcgagAACTGCAGGGCTTAAACCGAT
CsB_R1	cgataccgtcgacTGGGCCCAAGTGCCTTAATC