

Hintermann*, Bolt*, Hawkins* et al.

Supplementary Table 1: Counts of pseudocloacal phenotypes in fish

Genotype	n	Phenotype
wild type	10	separate openings
<i>hoxa13a</i>	7	separate openings
<i>hoxa13b</i>	7	separate openings
<i>hoxd13a</i>	6	separate openings
<i>hoxa13a</i> ^{+/-} <i>hoxa13b</i> ^{+/-} <i>hoxd13a</i> ^{+/-}	6	separate openings
<i>hoxa13a</i> ^{-/-} <i>hoxa13b</i> ^{-/-}	4	fused openings
<i>hoxa13a</i> ^{-/-} <i>hoxa13b</i> ^{-/-} <i>hoxd13a</i> ^{-/-}	4	fully fused openings

	separate openings	fused openings
wild type	10	0
<i>hoxa13a</i> ^{-/-} <i>hoxa13b</i> ^{-/-}	0	4

p-value < 0.001 (Fisher Exact Test)

	separate openings	fully fused openings
wild type	10	0
<i>hoxa13a</i> ^{-/-} <i>hoxa13b</i> ^{-/-} <i>hoxd13a</i> ^{-/-}	0	4

p-value < 0.001 (Fisher Exact Test)

Hintermann*, Bolt*, Hawkins* et al.

Supplementary Table 2: 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	CCTTCAAACTCAAGGCCCATC
3DOM_Del_r	CTCCCGGATTTGCTGTAACAC
evx2_f	CGCACTGGCATTCTCTGTTTT
evx2_r	GGAAGTGTGTCGTTGTGGTGG
zf_CsA_f	CAGCCCGCAAAGCCTCATTTTA
zf_CsA_r	GTGTCAACGAGAGGAGAAGGCT
zf_CsB_f	ACCAGGAGAAACACCACACACA
zf_CsB_r	TGACCAACTGATAACCCCAACC
zf_islandV_f	CTCATTTGCGCCGCTGTCTTTA
zf_islandV_r	GGTTAGATGTGGGGTTTGGGGA
zf_islandII_f	AGCAAAGCCCGGCTAATAGACA
zf_islandII_r	TGACGCGTGGGCTTAAATCAC
hoxa13a_8_del_f	GCCAAGGAGTTTGCCTTGTA
hoxa13a_8_del_r	TGACGACTTCCACACGTTTC
hoxa13b_14_ins_f	GATTGACCCGGTGATGTTTC
hoxa13b_14_ins_r	TACACTGGTTCGCAGCAAAA
hoxd13a_f	AAGCCGGTGATACATCAGGAG
hoxd13a_r	GTGGCCTTCCATTGTCAAAC
hoxd13a_KI_F	TGCCATTTGGAAAACAGTTACA
hoxd13a_KI_R	TGCTCTCTGAACCACAGGTAGA
LacZ_F	CGGCATCAGAGCAGATTGTA
hsp70_R	CCTCCTTGTTCAAGTCGTGGT

crRNA

Name	Sequence
hoxd13a_crRNA	CTGAGAGGATCCCATTTGCGAAACACCTGGG
atp5g3a_crRNA	AACCATATCCACTCTTCAGGAGGTCATGTG
hoxd3a_crRNA	TGATGCTGCACCCTAAATGG
hnmpa3_crRNA	ATAATCTAGTCATAGCTGGA
CsB_g1_F	CAATGGGCACAGGCTAAATACT
CsB_g1_R	GCATTATGGCCTTGAAGACTGT
CsB_g2_F	TGTGTGAGCGTCTGGTTAATCT
CsB_g2_R	CAGCCGCTCAATATTTTAGGTT

MOUSE

Genotyping primers

primer name	Sequence	Reference
<i>Inv(Itga6-nsi)d11lac</i>		
Inv(Itga6-nsi)d11lac_WT_f	GCAAGCCACTTGGAACAACCTGTTAATGG	(Tschopp and Duboule 2011)
Inv(Itga6-nsi)d11lac_WT_r	CCGTCCAATGTGCGTGTTC	
Inv(Itga6-nsi)d11lac_Inv_f	GAGTTTCTCTTGCTGAATGAAGAGCTG	
Inv(Itga6-nsi)d11lac_Inv_r	CCGTCCAATGTGCGTGTTC	
<i>Inv(Itga6-attP)</i>		
Inv(nsi-itga6)d11lac_WT_f	GCAAGCCACTTGGAACAACCTGTTAATGG	(Schep et al. 2016)
Inv(nsi-itga6)d11lac_WT_r	CCGTCCAATGTGCGTGTTC	
Inv(nsi-itga6)d11lac_Inv_f	GAGTTTCTCTTGCTGAATGAAGAGCTG	
Inv(nsi-itga6)d11lac_Inv_r	CCGTCCAATGTGCGTGTTC	
<i>Del(HoxD)</i>		
Del(HoxD)_WT_f	GAGCCCGACGCATCGAGATAGC	(Spitz et al. 2001)
Del(HoxD)_WT_r	CAAGGTCCTCAGCCTTAAGAGTGG	
Del(HoxD)_Del_f	AGGGATCCGGAGCATACCACTG	
Del(HoxD)_Del_r	CTCTCTCTACGAGGGAATGTGGAG	
<i>tgBAC(HoxD), tg(GT2), tg(islandE), tg(CsB)</i>		tgBAC(HoxD) in (Schep et al. 2016)
tgLacZ_PCRb_f	CCTGCTGATGAAGCAGAACA	
tgLacZ_PCRb_r	CAGCGACCAGATGATCACAC	
<i>Del(Atf2-SB1)</i>		
Del(Atf2-SB1)_WT_f	GACAATCGTATGCATGGCATACTCGG	(Montavon et al. 2011)
Del(Atf2-SB1)_WT_r	GATAGGAGTGACATTCAGACACGGC	
Del(Atf2-SB1)_Del_f	GTTTTCCAGTCACGACGTTG	
Del(Atf2-SB1)_Del_r	GCCACTGGCCGAATATTACCTATTTGTG	
<i>Del(SB1-Rel5)</i>		
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	
<i>Del(Rel5-Rel1)</i>		
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	AAAAGACTCCAACCTCTCGGA
Hoxd11_R	AGACGGTCCCTGTTCAATTTC
Hoxd10_F	GCTGGTCCCCGAGTCTTGCCT
Hoxd10_R	CCGGTGGCGTAGGTCTGACTCA
Hoxd9_F	CTCCACCCGGAAAAAGCGCTGT
Hoxd9_R	CGGTCCCGGGTGAGGTACATGT
Hoxd8_F_	TTCCCTGGATGAGACCACAAG
Hoxd8_R_	CTAGGGTTTGGAAGCGACTGT
Tbp_F	CCTTGTACCCTTACCAATGAC
Tbp_R	ACAGCCAAGATTACGGTAGA

Primers to clone transgenes

Name	Sequence
GT2_F	tccggtcgacTGTCACCACCATCGACAAGT
GT2_R	tccggtcgacATGCATTTACCGTCTTTC
IsE_F1	ccccccctcgagCTCAAGCCAGACAGGGATGATTA
IsE_R1	cgataccgtcgacGTGGGCTGTTTACTGGCAA
CsB_F1	ccccccctcgagAACTGCAGGGCTTAAACCGAT
CsB_R1	cgataccgtcgacTGGGCCCAAGTGCCTTAATC

Hintermann*, Bolt*, Hawkins* et al.

Supplementary Table 3: Counts of embryos used for zebrafish WISH

Line	Target	Stage	Experiment	n hom	n het	n wt	n siblings
5DOM	<i>hoxd4</i>	16-19	A	1	3	0	
5DOM	<i>hoxd4</i>	36 hpf	B	1	4	1	
5DOM	<i>hoxd4</i>	48 hpf	A	0	1	3	
5DOM	<i>hoxd4</i>	48 hpf	C	1	2	1	
5DOM	<i>hoxd4</i>	60 hpf	A	1	2	1	
5DOM	<i>hoxd4</i>	60 hpf	D	1	2	1	
3DOM	<i>hoxd4</i>	16-19	E	0	3	1	
3DOM	<i>hoxd4</i>	36 hpf	F	2	2	2	
3DOM	<i>hoxd4</i>	36 hpf	G	6	1	1	
3DOM	<i>hoxd4</i>	48 hpf	C	7	3	1	
3DOM	<i>hoxd4</i>	60 hpf	H	0	4	1	
3DOM	<i>hoxd4</i>	60 hpf	I	1	4	0	
5DOM	<i>hoxd10</i>	16-19	J	2	3	2	
5DOM	<i>hoxd10</i>	36 hpf	J	1	3	3	
5DOM	<i>hoxd10</i>	48 hpf	J	2	3	1	
5DOM	<i>hoxd10</i>	60 hpf	K	2	2	3	
3DOM	<i>hoxd10</i>	16-19	K	3	3	3	
3DOM	<i>hoxd10</i>	36 hpf	K	2	1	4	
3DOM	<i>hoxd10</i>	48 hpf	K	2	6	5	
3DOM	<i>hoxd10</i>	60 hpf	K	1	2	2	
5DOM	<i>hoxd13</i>	16-19	A	2	0	2	
5DOM	<i>hoxd13</i>	24 hpf	L	2	NA	NA	5
5DOM	<i>hoxd13</i>	28 hpf	L	4	NA	NA	4
5DOM	<i>hoxd13</i>	40 hpf	L	3	NA	NA	3
5DOM	<i>hoxd13</i>	48 hpf	M	4	2	3	
5DOM	<i>hoxd13</i>	48 hpf	A	2	1	2	
5DOM	<i>hoxd13</i>	60 hpf	E	1	1	2	
5DOM	<i>hoxd13</i>	60 hpf	N	7	3	0	
5DOM	<i>hoxd13</i>	72 hpf	O	7	1	5	
5DOM	<i>hoxd13</i>	72 hpf	P	2	NA	NA	4
3DOM	<i>hoxd13</i>	16-19	E	1	3	0	
3DOM	<i>hoxd13</i>	16-19	Q	2	5	0	
3DOM	<i>hoxd13</i>	36 hpf	R	2	1	2	
3DOM	<i>hoxd13</i>	48 hpf	S	2	5	1	
3DOM	<i>hoxd13</i>	60 hpf	G	2	5	0	
3DOM	<i>hoxd13</i>	60 hpf	I	2	3	3	

Hintermann*, Bolt*, Hawkins* et al.
Supplementary Table 4: HCR *hoxd13a* probes

oPool name	Sequence
hoxd13a_B2_opool	TAAAAGTCTGATATAAACAACTATGAAATCATCCAGTAAACCGCC
hoxd13a_B2_opool	CCTCGTAAATCCTCATCAAAACACAGTTATGCACTCAAATTTAAG
hoxd13a_B2_opool	TATAACATTTTGAGTGTATGGGAAAAAATCATCCAGTAAACCGCC
hoxd13a_B2_opool	CCTCGTAAATCCTCATCAAAGGAAATTGACGGGAGAATTCATTGA
hoxd13a_B2_opool	AACCAACTGTTTAACATTACTTCTAAAAATCATCCAGTAAACCGCC
hoxd13a_B2_opool	CCTCGTAAATCCTCATCAAATATAATTTACATTGTATCCCACTAT
hoxd13a_B2_opool	TACTTTATGCAAACCAACTCTACAAAAATCATCCAGTAAACCGCC
hoxd13a_B2_opool	CCTCGTAAATCCTCATCAAAATCTCAATGGGATTAACATTTGATG
hoxd13a_B2_opool	TCTCTTCTTATCCTTGACCCGTGCGAAATCATCCAGTAAACCGCC
hoxd13a_B2_opool	CCTCGTAAATCCTCATCAAATTGAAACCATATAGTTACTTGTCTC
hoxd13a_B2_opool	GGTAGAAGAAGCGATCCGCCGTCTGAAATCATCCAGTAAACCGCC
hoxd13a_B2_opool	CCTCGTAAATCCTCATCAAACCTCTTGTAAATGAACTTAGTGGTG
hoxd13a_B2_opool	CTCGAGCTCTTTCAGCTGAAATTTTAAATCATCCAGTAAACCGCC
hoxd13a_B2_opool	CCTCGTAAATCCTCATCAAAGTAAGGAACCCGCTTCTTCTCCCG
hoxd13a_B2_opool	TGAAGCCGCTGCTGCTTCTCTGTAAATCATCCAGTAAACCGCC
hoxd13a_B2_opool	CCTCGTAAATCCTCATCAAATGATGGTTTCCAGATATGCGGGCTC
hoxd13a_B2_opool	TTGAAAGCAATAAAGTTGACTGCTGAAATCATCCAGTAAACCGCC
hoxd13a_B2_opool	CCTCGTAAATCCTCATCAAAATTGTTTGACCAGTCCCAATGTTGG
hoxd13a_B2_opool	TGTCAAACAAGTCTCATGTCTAAGAAAATCATCCAGTAAACCGCC
hoxd13a_B2_opool	CCTCGTAAATCCTCATCAAATCCCATCATTGCTCTCTGAACCACA
hoxd13a_B2_opool	GAAGGCAGGGATTCTCGGATAGGAGAAATCATCCAGTAAACCGCC
hoxd13a_B2_opool	CCTCGTAAATCCTCATCAAACGCGCAGCCTTGATAAAAGGCGAAC
hoxd13a_B2_opool	GCTGGAGGCTTACCGTGGTCCACGAAATCATCCAGTAAACCGCC
hoxd13a_B2_opool	CCTCGTAAATCCTCATCAAATTTATTGGCGAACTGTCCGTTTTGC
hoxd13a_B2_opool	AGGCGGGAACACGGCGTTCTGTGAAAAATCATCCAGTAAACCGCC
hoxd13a_B2_opool	CCTCGTAAATCCTCATCAAATTTGTAAGTGTCTTCCAAATGGCAT
hoxd13a_B2_opool	GGAGCCAGAGCCGATATTGGATGGAAAATCATCCAGTAAACCGCC
hoxd13a_B2_opool	CCTCGTAAATCCTCATCAAAAGAAAAGTAAGGACTGATAGACTCA
hoxd13a_B2_opool	AGTCGGTCGATCAACAGCGGAGAGCAAATCATCCAGTAAACCGCC
hoxd13a_B2_opool	CCTCGTAAATCCTCATCAAACGGCGCTCCTGATGTACACCGGCTT
hoxd13a_B2_opool	GAAGGCAGATGGGTAAACATTTATAAAATCATCCAGTAAACCGCC
hoxd13a_B2_opool	CCTCGTAAATCCTCATCAAACCTTTCATCCAGTCCTCCTCCGTCC
hoxd13a_B2_opool	TGCATGGGATTTAACTCATGTCTGTAAATCATCCAGTAAACCGCC
hoxd13a_B2_opool	CCTCGTAAATCCTCATCAAATGGGCTCAAAACCAAGAAAAATGCA

Hintermann*, Bolt*, Hawkins* et al.

Supplementary Table 5: Accession numbers for re-analysed data

Raw data

Name	SRA number	Publication	DOI
CTCF_ChIP_E105_PT_rep1	SRR17750150	(Hintermann et al. 2022)	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

Processed data

Name	GSE	Publication	DOI
scATAC general fragment file	GSE243256	(Sun et al., 2024)	10.1038/s41556-024-01449-0
scRNA-seq raw matrix	GSE223922	(Sur et al., 2023)	10.1016/j.devcel.2023.11.001

Genome assemblies

Name	Assembly
Mus_musculus	GCA_000001635.9
Ornithorhynchus_anatinus	GCA_000002275.2
Gallus_gallus	GCA_000002315.5
Anolis_carolinensis	GCA_000090745.2
Xenopus_tropicalis	GCA_000004195.4
Latimeria_chalumnae	GCA_000225785.1
Danio_rerio	GCA_000002035.4
Takifugu_rubripes	GCA_901000725.3
Petromyzon_marinus	GCA_010993605.1
Lepisosteus_oculatus	GCA_000242695.1
Scyliorhinus_canicula	GCA_902713615.2
Leucoraja_erinacea	GCA_028641065.1
Amia_calva	GCA_036373705.1

Hintermann*, Bolt*, Hawkins* et al.

Supplementary Table 6: 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
<i>Hoxd4</i>	NM_010469.2	74552322	<i>hoxd4a</i>	NM_001126445.2	1951004
<i>Hoxd13</i>	NM_008275.4	74498569	<i>hoxd13a</i>	NM_131169.3	1990311
<i>Nfe2l2</i>	NM_010902.5	75534860	<i>nfe2l2a</i>	NM_182889.1	1654399
<i>Atp5g3</i>	NM_001301721.1	73741670	<i>atp5mc3a</i>	NM_201176.1	2333895

Domain	Domain_name	mm39	danRer11	mm39/danRer11
<i>Atp5g3-Hoxd13</i>	5DOM	756899	343584	2.2
<i>Hoxd4-Nfe2l2</i>	3DOM	982538	296605	3.3
<i>Hoxd13-Hoxd4</i>	cluster	53753	39307	1.4
<i>whole genome</i>	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)