Supporting Information

Gómez-Marín et al. 10.1073/pnas.1505463112

SI Materials and Methods

Generation of BAC DK74B2-*six2a::GFP-six3a::mCherry-iTol2*. BAC clone (number 74B2) from DanioKey zebrafish BAC library containing the *six2a and six3a* genes was purchased from Source BioScience. BAC DK74B2-*six2a::GFP-six3a::mCherry-iTol2* was generated as previously described (36, 37). Briefly, we used recombineering in the *Escherichia coli* SW105 strain to introduce the iTol2-A cassette from the *piTol2-A* plasmid (37) in the BAC plasmid backbone, which contains the inverted minimal *cis*-sequences required for *Tol2* transposition. The insertion of the reporter gene *GFP* into the *six2a* locus of *DK74B2-iTol2* BAC clone was carried out using homologous recombination. To that end, the *GFP-pA-FRT-kan-FRT* reporter gene cassette from the pBSK-*GFP-pA-FRT-kan-FRT* (37) plasmid was amplified by PCR, together with 50-bp homologies to the *six2a* translation start site (Fwd_*six2a*HA1-Gfp: TTAGATAGACATACAAGTACAAAGAGGGACGTTTATTTTTGAGACAAACCgccaccatgGTGAGCAAGGGCGAGGAGCTGTTC, ReV_*six2a*HA2-frt-kan-frt: CAGACGCACGCCACTTGCTCTTGCGTAAAGCCGAATGTTGGAAGCATAGACCGCGTGTAGGCTGGAGCTGCTTC). Cells harboring the *DK74B2-iTol2* BAC were transformed with the PCR product and clones in which homologous recombination have occurred were identified by PCR (*Checksix2aHA1*-Fwd: TGTGCCTCTCTTCACCCGGTG, *Checksix2aHA1*-ReV: CAGCTCCTCGCCCTTGCTCAC, *Checksix2aHA2*-Fwd: GAAGCAGCTCCAGCCTACACG, *Checksix2aHA2*-ReV: GGGAGAACTGGTGGCTTTCGAG). To excise Kan resistance, a step of flp induction using L-arabinose cultures was then included followed by the identification of Kan-sensitive clones using plate replicas with Amp + Cam + LB (Luria-Bertani medium) and Amp + Cam + Kan + LB plates. Sensitive clones were subjected to PCR confirmation using specific primers (flp_induction_Fwd: ACGAGCTGTACAAGTAAAGCGGC and flp_induction_ReV: CCGCGTGTAGGCTGGAGCTGC). The mCherry_KanR cassette from *pCS2+_mCherry_KanR* (36) was amplified by PCR including 50-bp homologies to the six3a translation start site (Fwd_*six3aHA-mCherry*: TCGTCGTTCTTTTTTCCTTCGCAAATTTCACTCTCTCTCAGGTCATTTCCACCATGGTGAGCAAGGGCGAGGAG, ReV_*six3aHA_KanR*: TTTGGCAGGAAGAAATGAGAGGGATAAAGCTCTAAAGGCGATCTGAAAACTCAGAAGAACTCGTCAAGAAGGCG). This PCR product was used to introduce by homologous recombination the mCherry reporter gene under control of the six3a promoter in the previous BAC and to generate the *DK74B2-six2a::GFP-six3a::mCherry-iTol2* construct. Insertion was confirmed by PCR (*Checksix3aHA1*-Fwd: GATTGGCAGGGCTGCCATGAC, *Checksix3aHA1*-ReV: CCTCGCCCTTGCTCACCATGG, *Checksix3aHA2*-Fwd: GCCTTCTTGACGAGTTCTTCTG, *Checksix3aHA2-Rev*: GAAGCGACCATAGGAAGCG).

Generation of Δ18kb *DK74B2-six2a::GFP-six3a::mCherry-iTol2* BAC. Δ18kb *DK74B2-six2a::GFP-six3a::mCherry-iTol2* was generated as follows. Two fragments of 250 bp flanking the target region were amplified using the following primers: *Del_EcoRI_HA1*-Fwd: TTGAATTCGCGTTCTTAGCAAGCAGGAT and *Del_PstI_HA1*-ReV: TTCTGCAGCGTTTAGCATCCACTCGTAGC (HA1 fragment) for one side; and *Del_PstI_HA2*-Fwd: TTCTGCAGATGCGTCTGAACTCGGGACT and *Del_XhoI_HA2*-ReV: TTCTCGAGCCCCAGCCAACCCTTATTTCG (HA2) and subsequently cloned into pCR8-GW-TOPO vectors. Spectinomycine (Spe) resistance was amplified using the following primers: *PstI_SpeR*_Fwd: TTCTGCAGCTGAAGCCAGTTA and *PstI_ SpeR*_ReV: TTCTGCAGTAGCTGTTTCCTG and also cloned into the pCR8-GW-TOPO vector. The HA1 fragment, HA2 fragment, and speR were excluded from their vectors by using EcoRI + PstI, PstI + XhoI, and PstI digestions, respectively. All excised fragments were purified and ligated together including a pCS2+ plasmid previously digested with EcoRI and XhoI restriction enzymes. *pCS2+_HA1_speR_HA2* clones were selected by EcoRI + XhoI digestion. This clone was used as a PCR template of the *Del_EcoRI_HA1*-Fwd and *Del_XhoI_HA2*-ReV primers. This PCR product (deletion cassette) was used to recombine with *DK74B2-six2a::GFP-six3a::mCherry* BAC. Clones carrying the deletion cassette (Δ18kb *DK74B2-six2a::GFP-six3a::mCherry-iTol2*) were selected by growing bacteria in LB + Cam + Spe plates and subsequently confirmed by PCRs targeting both flanks of the insertion using internal and external primers [*DelCheck_six2a*_Fwd(Genom): CCAACCCTTATTTCAGGTCATGC, *DelCheck_six2a*_ReV(Spectin): GTCCACTGGGTTCGTGCCTTC; *DelCheck_six3a*_Fwd(Spectin)]. In addition, two independent digestions with XhoI and NotI were performed in parallel in full and Δ18kb (*DK74B2-six2a::GFP-six3a::mCherry-iTol2*) BAC versions to verify the presence of the deletion.

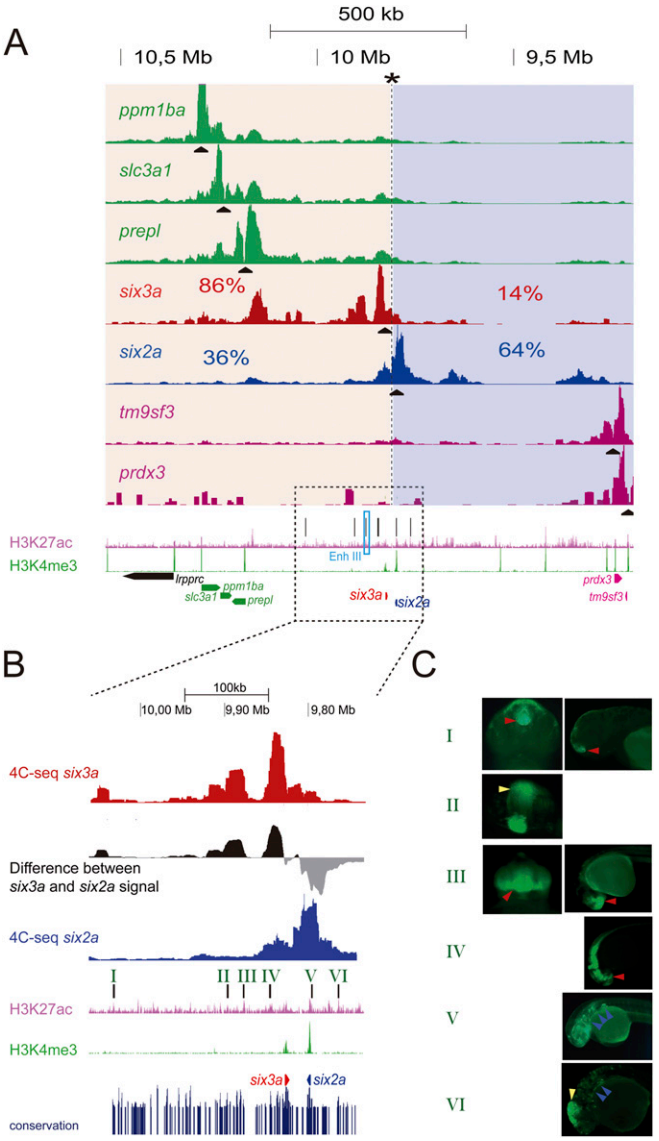

Fig. S1. The *six3a/six2a* locus is partitioned in two regulatory landscapes harboring several tissue specific enhancers. (*A*) High-resolution circular chromosome conformation capture (4C-seq) in whole zebrafish embryos at 24 hpf from different viewpoints (black triangles) along the *six3a/six2a* genomic region. The genomic position region in which there is a maximum difference between the accumulative contacts from the *six3a* and *six2a* viewpoints is shown with an asterisk, and the two 3D compartments are shaded in red and blue, respectively. The bottom tracks show the distribution of the H3K27ac (pink) and H3K4me3 (green) histone marks and the genes along this genomic region. Above the H3K27ac track, the black lines show the different regions tested for enhancer activity. The Enhancer III region used as a 4C-seq viewpoint in the fourth track is highlighted with a light blue rectangle. (*B*) Close up of the region around the *six3a* and *six2a* genes marked by a dotted rectangle in *A*. The different identified enhancers are numbered from I to VI. The last blue track shows evolutionary conserved regions. (*C*) Transgenic embryos from anterior (*Left*) or lateral (*Right*) views showing GFP expression driven by the different H3K27ac positive regions tested. The vector used in the transgenic assays to test the activity of regions II and VI contains a midbrain enhancer used as a positive control of transgenesis. This expression domain (red arrow) is therefore independent of the region under evaluation. Regions I to IV drive the reporter in *six3a*-expressing territories, whereas regions V and VI activate GFP in *six2a* domains.

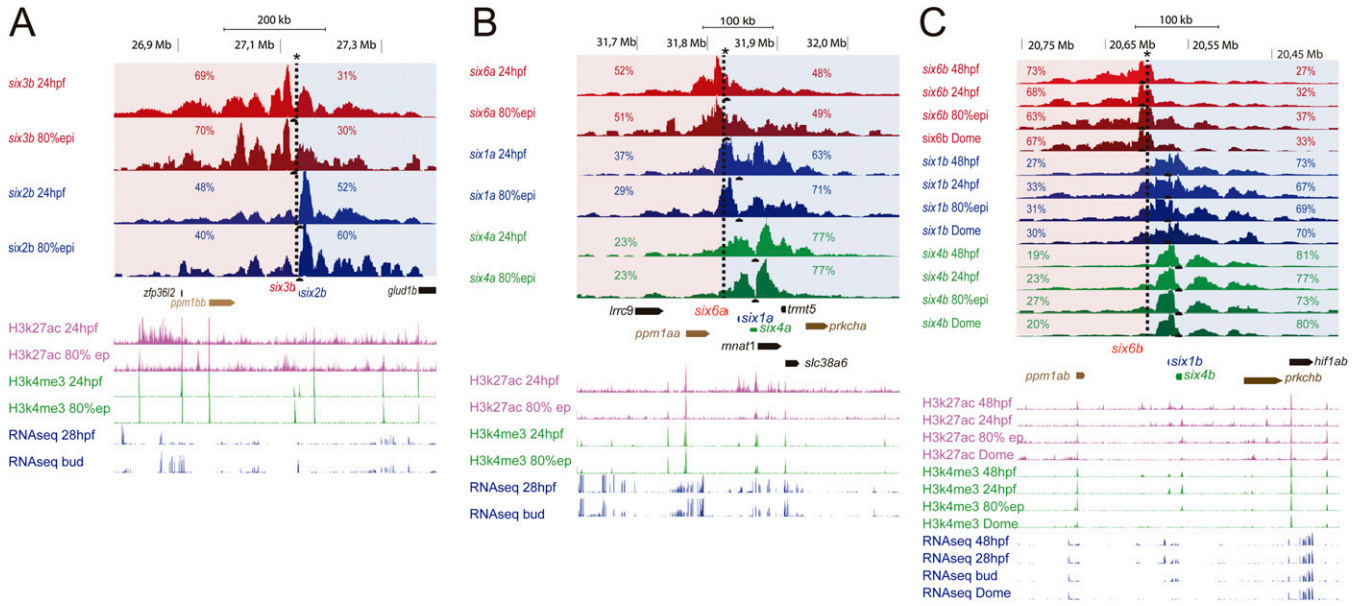

Fig. S2. Developmental dynamics of chromatin contacts, H3K27ac and H3K4me3 histone marks, and RNA-seq at the different zebrafish *six* clusters. (*A*–*C*) *six3b*/*six2b* (*A*), *six6a*/*six1a*/*six4a* (*B*), and *six6b*/*six1b*/*six4b* (*C*) clusters. In all clusters, the frequency of contacts at each side of the boundary (dashed lines and asterisks) is indicated for each promoter viewpoint.

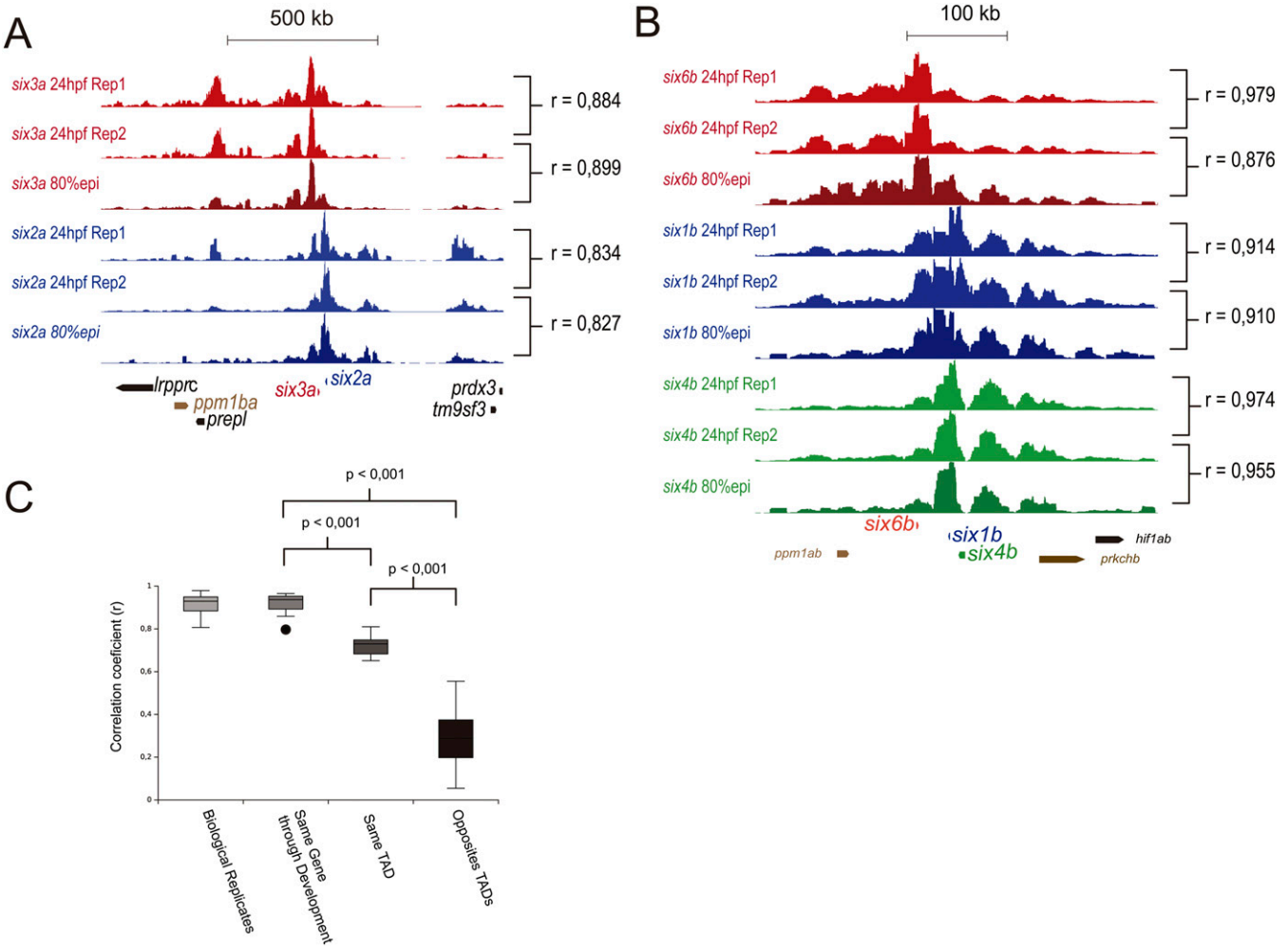

Fig. S3. 4C-seq replicas at the same stage are similar to 4C-seq at different developmental stages. (*A* and *B*) Comparison between 4C-seq replicas at a single stage and 4C-seq at different developmental stages in the *six3a/six2a* (*A*) and *six6b/six1b/six4b* (*B*) clusters. (*Right*) The correlation coefficients between 4C-seq datasets. (*C*) Correlation coefficient between 4C-seq dataset from replicas of the same gene at the same stage, same gene at different stages, genes within the same TAD, or genes at different TADs.

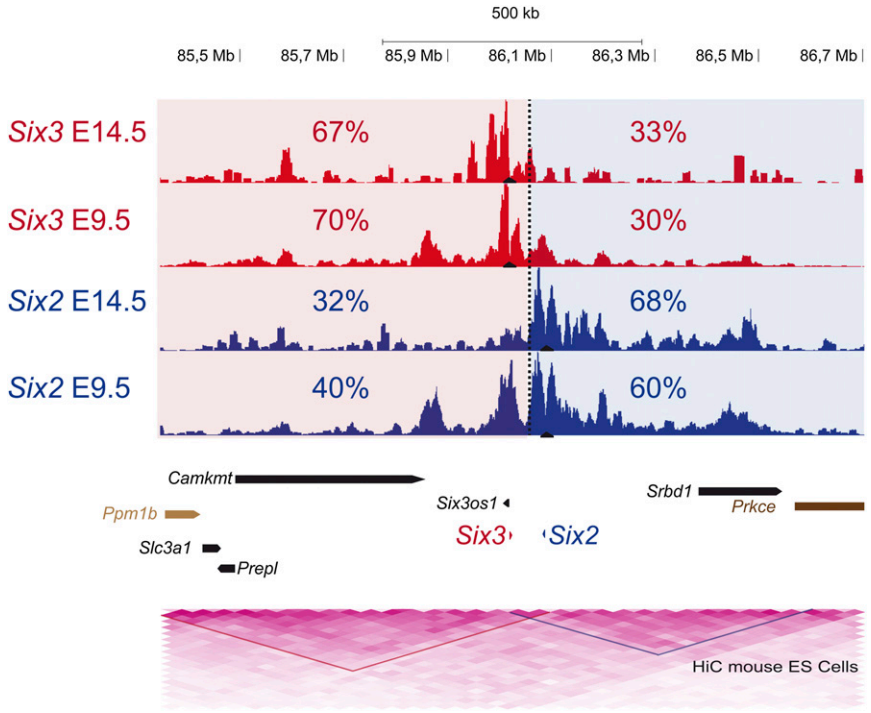

Fig. S4. 3D chromatin configuration of the mouse *Six3*/*Six2* cluster at two different developmental stages. 4C-seq from *Six3* and *Six2* viewpoints (black triangles) in whole embryos at stages E14.5 and E9.5. The genomic region in which there is maximal difference between the accumulative contacts from the *Six3* and *Six2* viewpoints is shown with a dashed line, and the two 3D compartments are shaded in red and blue, respectively. The percentage of contacts for each gene on the two 3D compartments is indicated. Below are shown the two TADs detected by HiC data from mouse ES cells (red and blue triangles).

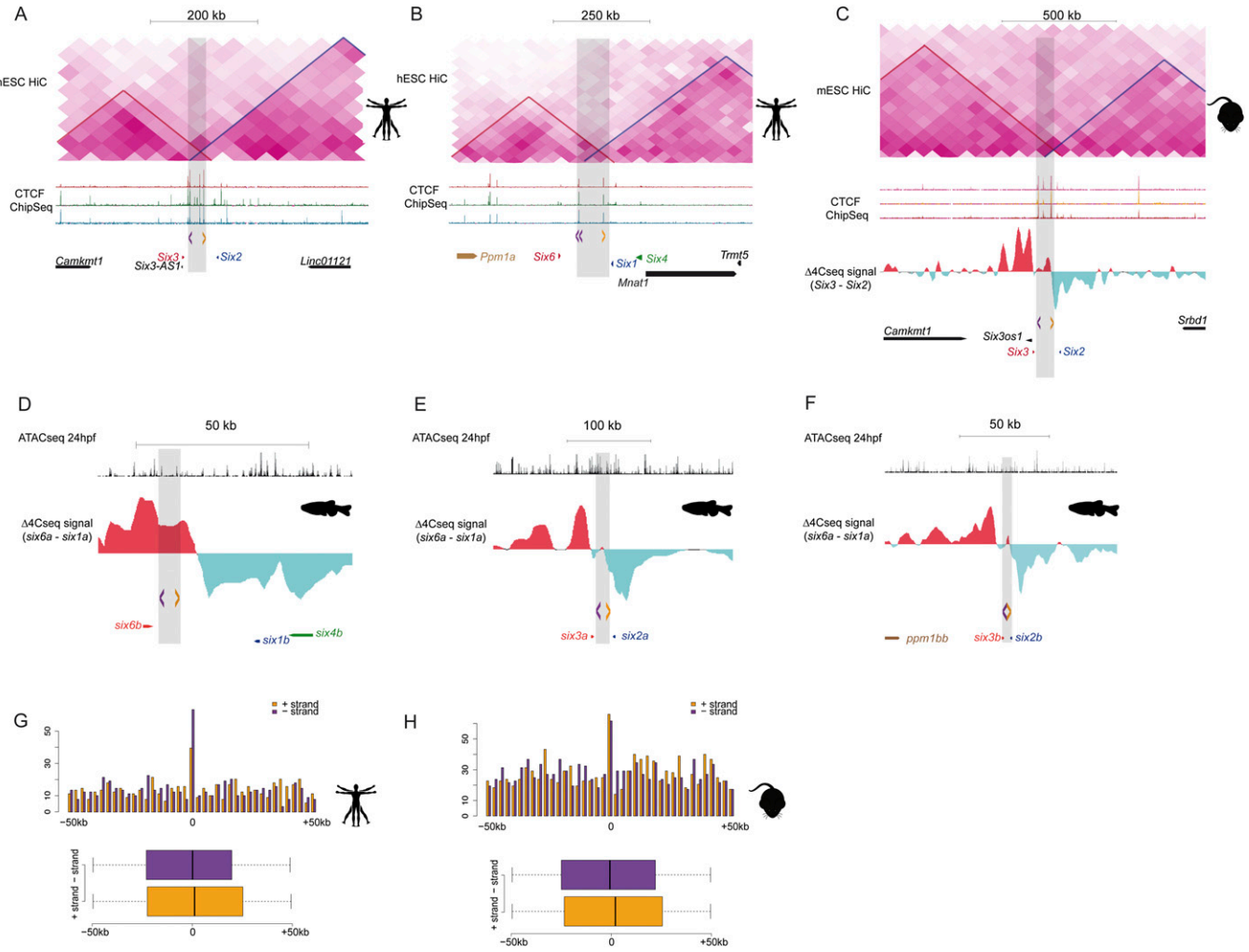

Fig. S5. Diverging CTCF sites are signature of TAD borders and are not associated with promoters. (*A* and *B*) These panels show, from top to bottom, HiC data from human ES cells, the genomic distribution of CTCF in three different cell types, the orientation of CTCF sites represented by arrowheads (purple and yellow correspond to sites in minus or plus strands, respectively) at the boundary regions, and the genes at the *Six2*/*Six3* (*A*) and *Six6*/*Six1*/*Six4* (*B*) clusters. (*C*) Mouse *Six3*/*Six2* cluster showing HiC data from mouse ES cells, the genomic distribution of CTCF in three different cell types, the difference between *Six3* and *Six2* 4C-seq signals, the orientation of CTCF sites, and the genes around this genomic region. (*D*–*F*) From top to bottom, ATAC-seq peaks from 24-hpf zebrafish embryos, difference between *Six* genes 4C-seq signals, orientation of CTCF sites represented by arrowheads, and genes at the *six6b*/*six1b*/*six4b* (*D*), *six3a*/*six2a* (*E*), and *six3b*/*six2b* (*F*) clusters. (*G* and *H*) (*Upper*) Number (*y* axis) and orientation (purple and yellow bars correspond to CTCF sites at the plus or minus strands, respectively) of CTCF sites along 50 kb (*x* axis) at each side of human (*D*) and mouse (*E*) randomly selected promoters (1,000). (*Lower*) Boxplot shows the enrichment of CTCF in diverging orientations at each side of the boundaries. The differences observed between the mean relative position of the motifs in both strands were statistically significant in the boundary-centered windows ($P = 3.27E{-}113$ in human, $P = 1.75E{-}118$ in mouse; Fig. 4), but not in the promoter-centered windows ($P = 0.107$ in human, $P = 0.066$ in mouse).

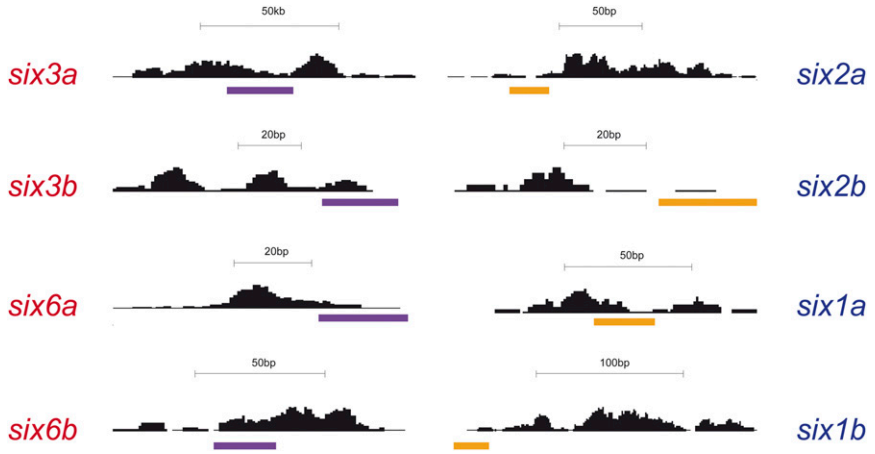

Fig. S6. ATAC-seq signal at CTCF sites in zebrafish. The different panels show the ATAC-seq peaks at each CTCF sites in the different zebrafish *Six* cluster, as indicated.

Table S1. Genomic coordinates of CTCF sites at *Six* boundaries

			CTCF matrix coordinates				
CTCF name	Specie	Pointed gene	Chromosome	Start	End	Matrix	Value
hCTCFsix6 (1)	Human	*Six6*	chr14	61026278	61026302	TGGACAGCAGGGGGCTCTC	16
hCTCFsix6 (2)	Human	*Six6*	chr14	61029012	61029036	CCGCCACTTGGAGGCAGTG	13
hCTCFsix1	Human	*Six1*	chr14	61092245	61092270	TGGCCAGCAGGTGGCACTC	23
hCTCFsix2	Human	*Six2*	chr2	45182932	45182961	TCGCCGCGAGGTGGCAGCA	16
hCTCFix3	Human	*Six3*	chr2	45208915	45208940	TCTCCAGCAGGTGGCGCCA	21
mCTCFsix6 (1)	Mouse	*Six6*	chr12	74074290	74074314	ATGCCACCAGGGGGCTCTC	17
mCTCFsix6 (2)	Mouse	*Six6*	chr12	74077887	74077911	TGTACGCCAGGTGGTGCTG	12
mCTCFsix1	Mouse	*Six1*	chr12	74122104	74122128	TATCCAGCAGGGGGCACTC	20
mCTCFsix2	Mouse	*Six2*	chr17	86033160	86033184	CGGCCGCGAGGTGGCAGCA	16
mCTCFsix3	Mouse	*Six3*	chr17	86064479	86064503	TCTCCAGCAGGTGGAGCCA	18
zCTCFsix6a	Zebrafish	*six6a*	chr13	31832136	31832159	TGACCACCAGAGGTCGCAA	19
zCTCFsix1a	Zebrafish	*six1a*	chr13	31835564	31835588	TTGCCACCAGAGGGCAGAA	11
zCTCFsix6b	Zebrafish	*six6b*	chr20	20600656	20600680	CTAACTGAAGAGGGCGCTC	12
zCTCFsix1b	Zebrafish	*six1b*	chr20	20596398	20596422	TGTCCACTAGAGGGAACCA	19
zCTCFsix2a	Zebrafish	*six2a*	chr13	9819167	9819191	TTCCCACAAGATGGCGTAA	13
zCTCFsix3a	Zebrafish	*six3a*	chr13	9804005	9804029	CGTTCAGGAGAGGGCGCCA	14
zCTCFsix2b	Zebrafish	*six2b*	chr12	27133491	27133515	TCACCACATGATGGCGACA	12
zCTCFsix3b	Zebrafish	*six3b*	chr12	27134370	27134394	TGGCCAGCAGAGGGTGCTT	18
SpCTCFsix3/6 (1)	Sea Urchin	*Sp_six3/6*	Scaffold143	896688	896712	TGACCAACAGACAGCGGCC	8
SpCTCFsix3/6 (2)	Sea Urchin	*Sp_six3/6*	Scaffold143	892778	892812	CGGCCAGGAGATGGAGAAG	13
SpCTCFsix1/2 (1)	Sea Urchin	*Sp_six1/2*	Scaffold143	888274	888298	CGGCCAGGAGATGAAGCAG	11
SpCTCFsix1/2 (2)	Sea Urchin	*Sp_six1/2*	Scaffold143	883271	883297	ATGACGACAGAGGGCGGCA	10

Table S2. Primers used for 4C-seq experiments

Organism	Primer	Sequence	Read primer (DpnII) position
Zebrafish	*six2a*_DpnII	GAAGAGAGGCACAAAACTTTAGATC	chr13:9797970
	*six2a*_Csp6I	GTCATCGCTTAGATAGACATACAAGTAC	
	*six3a*_DpnII	AGTGGGTGGAATATTATTGTGATC	chr13:9826837
	*six3a*_Csp6I	AAGTGGTGAAAGCCTCTACGTAC	
	*six2b*_DpnII	GCTTACTACTAAGTAAAGTTTTTGA	chr12:27140158
	*six2b*_Csp6I	TCATCAAATGCGGTTAAAGC	
	*six3b*_DpnII	CCAGAAGCAGAGGGCGA	chr12:27128597
	*six3b*_Csp6I	TCCAAGTACCATTGCCGAATA	
	*six4a*_DpnII	GCGCTGCAGAGTGCATTGATC	chr13:31871941
	*six4a*_Csp6I	TGTTGTCAGATTGGGAATAGGGACC	
	*six1a*_DpnII	ACTTCCGTGAGCTCTACAAGATC	chr13:31845355
	*six1a*_Csp6I	GGCTCCCCTATGCAATCCAC	
	*six6a*_DpnII	CAATGTTGCCAAACACACGAAGATC	chr13:31826506
	*six6a*_Csp6I	GCCCTATACGCCAACTTCAAGTC	
	*six4b*_DpnII	CCCACGACTCTCCCTCTTGATC	chr20:20557180
	*six4b*_Csp6I	TCGTCTCTGCAGGATATGTGGTAC	
	*six1b*_DpnII	GTCTCTCCCGGCTTGCGATC	chr20:20572332
	*six1b*_Csp6I	ACTGTCGACTCATGTCGCGC	
	*six6b*_DpnII	TGAGCTGTCAGATGTCTACGAGATC	chr20:20606589
	*six6b*_Csp6I	GCCCTATACGCCAACTTCAAGTC	
	*ppm1ba*_DpnII	GCTGATATGCATGCAAGA	chr13:10292253
	*ppm1ba*_Csp6I	CAATATTCAGAAATGAGCGAGT	
	*slc3a1*_DpnII	CCGAATTCAACCGAAAGA	chr13:10186394
	*slc3a1*_Csp6I	TGCGTTTGAAGGAATCTAGCGT	
	*prepl*_DpnII	GAGAGTGAATACACGCAGA	chr13:10237450
	*prepl*_Csp6I	CACAAGTGGAGGTGGTGT	
	*tm9sf3*_DpnII	CGTAAAATTGTTCAGCGAGTGATC	chr13:9240919
	*tm9sf3*_Csp6I	CCTCTCGATGTGTTGGTGTAC	
	*prdx3*_DpnII	GAAGGTGAAGTCTGTTCAAAGATC	chr13:9208662
	*prdx3*_Csp6I	CATAAAGGTGCTGCTGACGTAC	
	EnhIII(R5)_DpnII	TACTCTCAGAGCTGTTAAAGGATC	chr13:9877427
	EnhIII(R5)_Csp6I	GTTGCTGCATCTTCTGGAC	
	EnhVI(2.2)_DpnII	GGAGCCACGCGGATTAAGAGATC	chr13:9763732
	EnhVI(2.2)_Csp6I	GATAGCCCACAGGAATTCGGCTG	
Mouse	Six2_DpnII	CGACTCCTGAGTCACAACGATC	chr17:86087424
	Six2_Csp6I	CACTACATCGAGGCGGAGAAGC	
	Six3_DpnII	GCGCCCTCTGCGTAGAGATC	chr17:86022535
	Six3_Csp6I	CCAGCAACTGTCAGCAGCCG	
	Six4_DpnII	GTCCCTGCCCCAGAGCgatc	chr12:74213772
	Six4_Csp6I	TGCCTGCCCAGAAGTTCCGAG	
	Six1_DpnII	GGAATCCCTTCTCTCACTTGGATC	chr12:74149274
	Six1_Csp6I	GGGGACTTATACGGGCTCTC	
	Six6_DpnII	ACAGGGAGGGGAAGTGGATC	chr12:74040426
	Six6_Csp6I	CCAGGAGGCAGAGAAGCTGC	
Sea urchin	Sp_six4/5_DpnII	CGCCACTTGAAACGGTGGGATC	Scaffold143:692728
	Sp_six4/5_Csp6I	AAAGTTCGCAGGGCTTTACATC	
	Sp_six1/2_DpnII	CAAAAATAGGCGACAGCGAGATC	Scaffold143:780065
	Sp_six1/2_Csp6I	AGGATAGGGGTTGTGGGAGTAC	
	Sp_six3/6_DpnII	GCAATGGCAACTCCTCTCTTGAtc	Scaffold143:1022703
	Sp_six3/6_Csp6I	TTGGCCCCGTCGACAAGTAC	