## The *SOX9* upstream region prone to chromosomal aberrations causing campomelic dysplasia contains multiple cartilage enhancers

Baojin Yao, Qiuqing Wang, Chia-Feng Liu, Pallavi Bhattaram, Wei Li, Timothy J. Mead, James F. Crish, and Véronique Lefebvre

## SUPPLEMENTARY DATA

**Table S1.** Primers and restriction enzymes used to amplify and clone mouse *Sox9* enhancer elements. Restriction enzyme sites are italicized in the primer sequences.

| Enhancer & Strand | | Sequence | Enzyme |
|---|---|---|---|
| E20 | Forward | CACGGATCCTTAGTCTTATCAGGAGCTGGT | BamHI |
| | Reverse | CCTAGATCTCAAGCTCTCAGTTTGCAGTGG | BglII |
| E65 | Forward | TGCGGATCCTATCGGTCTCTGTAGTAACCAATAGC | BamHI |
| | Reverse | CCAGATCTGCTATTTTGTCGTTGTTATTATTTCATGATG | BglII |
| E70 | Forward | ATAGATCTAGTGGCCGACGCGTGCTCCAGACATTGAAT | BamHI |
| | Reverse | CGGGATCCAGTAGGCTCTGTCATCGTGC | BglII |
| E84 | Forward | CAGGATCCAAAGCAGTTAGTGTCAGCCTC | BamHI |
| | Reverse | ACAGATCTCTACCTCAGATGGGCTTTCGG | BglII |
| E96 | Forward | CCGGGATCCTTCTTTGGCTACTGGAAGAAAG | BamHI |
| | Reverse | CCTAGATCTGCTCAAAGGATGTGTGGAGGTG | BglII |
| E125 | Forward | TTGGATCCCATGTCTGAGCAGGGAAGAGA | BamHI |
| | Reverse | GCCAGATCTGCCTTTTCCTCTACACACTCAC | BglII |
| E161 | Forward | GACTAGTCCACAGAGTGTGAAAATGTTCAG | SpeI |
| | Reverse | GTCTAGACTGAGTGAGCAGAAGTTGCT | XbaI |
| E195 | Forward | GAGATCTCGCATTTGATCTTCTCCTTGGCTC | BglII |
| | Reverse | AGGATCCCTGGACCCCAATATCAGACAGACT | BamHI |
| E239 | Forward | CGGCTCGAGCTCCTATGTAGAGAACAAATGGATG | XhoI |
| | Reverse | TAGTCGACGCATGTCTGACCTCCACAAGAG | SalI |
| E243 | Forward | GACTAGTCAGGACTCTTTGCTTGCTGTTG | SpeI |
| | Reverse | CTCTAGAAGTAAGGAGGATACCACAGCGT | XbaI |
| E250 | Forward | TATGGATCCCTGTAGGTGTACAGGCAG | BamHI |
| | Reverse | CAGAGATCTTCTGTCTACTGTGCCTACC | BglII |

**Table S2.** Primers used for site-directed mutagenesis. Nucleotide mutations are underlined.

| Mutation | Primer strand | Sequence |
|---|---|---|
| E195 M1 | Forward | AGAAAGAGTTTCGTCCCCTTCGGAGAG |
| | Reverse | CGAAGGGGACGAAACTCTTTCTGTTTC |
| E195 M2 | Forward | GATCTTCCCTCGACGGGCGCTTATGTTCTC |
| | Reverse | CATAAGCGCCCGTCGAGGGAAGATCGCAGA |
| E195 M3 | Forward | CGCAAGGATGGGCCACCGAGGGAACTCAATTAAGGCCCCG |
| | Reverse | CGGGGCCTTAATTGAGTTCCCTCGGTGGCCCATCCTTGCG |
| E195 M4 | Forward | TTTATGAGAGGATGACCAGTTAAGGATTTGCTCAGTTTGA |
| | Reverse | TCAAACTGAGCAAATCCTTAACTGGTCATCCTCTCATAAA |
| E84 M1 | Forward | ACCTCGGATTTTTGAAAAACCAGTA |
| | Reverse | CCGAGGTAAGCCTTGTATCCCAGAT |
| E84 M2 | Forward | AAGGGGAAGTAAAGGTGGGCAAAGA |
| | Reverse | TTTACTTCCCCTTTATCGTCAGTGG |

**Table S3.** EMSA Probes. Sequences are shown for the upper strand only. Two Gs added for probe labeling are italicized and underlined. WT, wild-type; MUT, mutant

| Probe name | Sequence |
|---|---|
| E195 probe 1 WT | *GG*CTCTAGCATAGGAAACAGAAAGAGTTTTGTCCCCTTC |
| E195 probe 1 MUT | *GG*CTCTAGCATAGGAAACAGAAAGAGTTTCGTCCCCTTC |
| E195 probe 2 WT | *GG*TGCGATCTTCCCTCAATGGGCGCTTATGTTCTCGCCA |
| E195 probe 2 MUT | *GG*TGCGATCTTCCCTCGACGGGCGCTTATGTTCTCGCCA |
| E195 probe 3 WT | *GG*GCGCAAGGATGGAAAACCGAGTTAATTCAATTAAGGC |
| E195 probe 3 MUT | *GG*GCGCAAGGATGGGCAACCGAGTGAACTCAATTAAGGC |
| E195 probe 4 WT | *GG*TTATGAGAGGATGAAAAGTTAATTATTTGCTCAGTTT |
| E195 probe 4 MUT | *GG*TTATGAGAGGATGACCAGTTAAGGATTTGCTCAGTTT |
| E84 probe 1 WT | *GG*GATACAAGGCTTCCATAGTATTTTTGAAA |
| E84 probe 1 MUT | *GG*GATACAAGGCTTACCTCGGATTTTTGAAA |
| E84 probe 2 WT | *GG*CTGACGATAAAGGGGCATTCACGGTG |
| E84 probe 2 MUT | *GG*CTGACGATAAAGGGGAAGTAAAGGTG |

**Table S4.** Genomic features of *Sox9* enhancers. Distance (dist.) was calculated from the middle of the enhancers to the transcription start site of the *Sox9/SOX9* gene.

| Enh-ancer | Mouse (mm9) | | | Human (hg19) | | | Rat (rn5) | | |
|---|---|---|---|---|---|---|---|---|---|
| | Coordinates on chr11: | Size (bp) | Dist. (kb) | Coordinates on chr.17 | Size (bp) | Dist. (kb) | Coordinates on chr. 10: | Size (bp) | Dist. (kb) |
| E20 | 112,623,542-112,624,237 | 696 | 20 | 70,096,653-70,097,365 | 712 | 21 | 100,944,949-100,945,638 | 690 | 20 |
| E65 | 112,578,666-112,579,289 | 624 | 65 | 70,060,876-70,061,498 | 623 | 56 | 100,898,122-100,898,731 | 610 | 67 |
| E70 | 112,572,641-112,573,424 | 784 | 70 | 70,053,842-70,054,652 | 810 | 63 | 100,891,271-100,892,059 | 789 | 73 |
| E84 | 112,559,105-112,559,750 | 646 | 84 | 70,042,609-70,043,257 | 648 | 75 | 100,871,144-100,871,776 | 633 | 94 |
| E96 | 112,546,971-112,547,335 | 365 | 96 | 70,030,466-70,030,844 | 378 | 87 | 100,860,402-100,860,757 | 356 | 104 |
| E125 | 112,518,744-112,519,299 | 556 | 125 | 70,004,793-70,005,376 | 583 | 112 | 100,832,534-100,833,089 | 556 | 132 |
| E161 | 112,481,770-112,482,929 | 1160 | 161 | 69,971,188-69,972,311 | 1123 | 146 | 100,792,746-100,793,913 | 1168 | 172 |
| E195 | 112,447,134-112,450,155 | 3022 | 195 | 69,922,750-69,927,819 | 5069 | 194 | 100,759,205-100,761,891 | 2687 | 206 |
| E239 | 112,403,840-112,404,821 | 982 | 239 | 69,866,609-69,867,480 | 871 | 251 | 100,681,310-100,682,282 | 973 | 283 |
| E243 | 112,399,345-112,401,408 | 2064 | 243 | 69,861,669-69,863,717 | 2048 | 255 | 100,676,074-100,678,151 | 2078 | 289 |
| E250 | 112,393,446-112,393,851 | 406 | 250 | 69,856,081-69,856,452 | 371 | 261 | 100,670,678-100,671,082 | 405 | 294 |

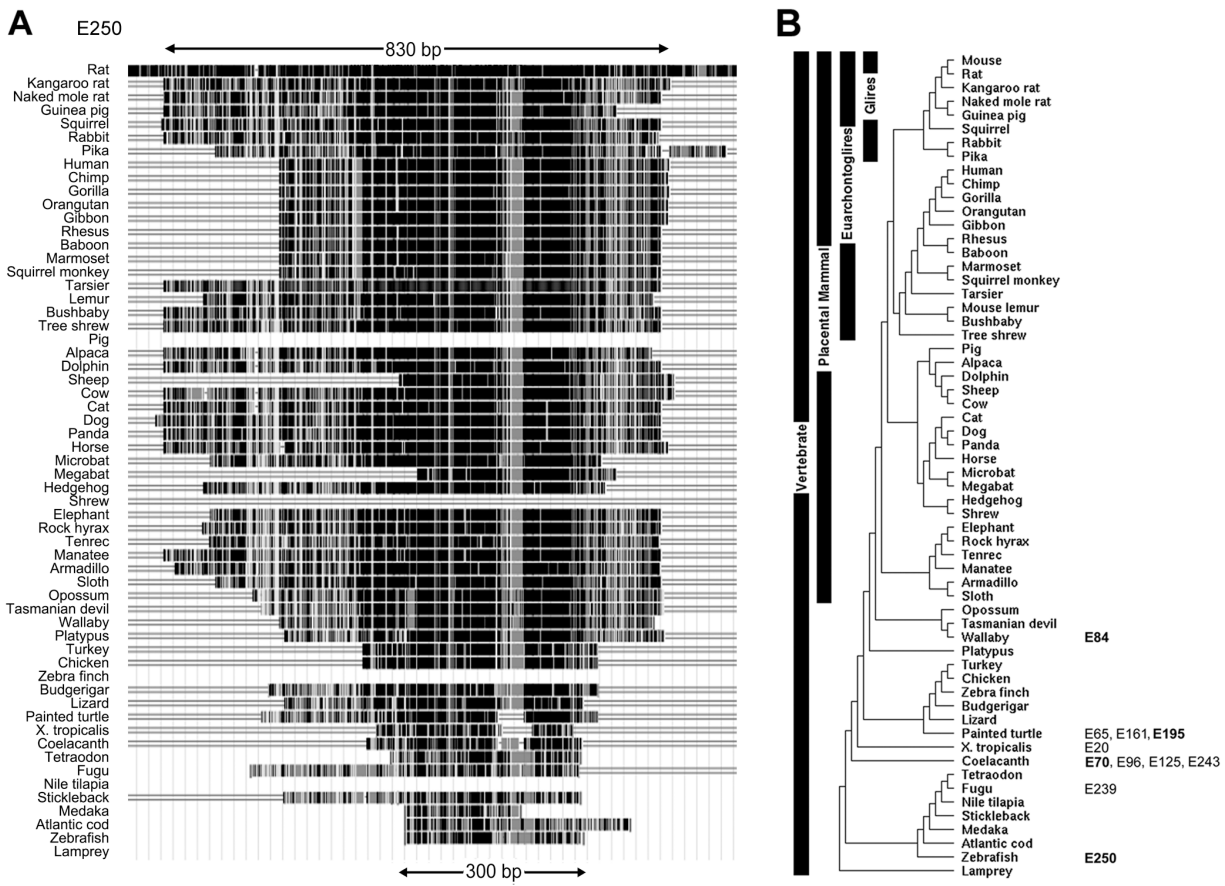


**Figure S1.** Vertebrate conservation of *Sox9* enhancers. (A) UCSC genome browser multiz alignment of the E250 region in 60 vertebrate species. (B) Schematic of phylogeny with branch lengths and clade groupings showing the most distant species from mouse that show conservation in *Sox9* enhancers.

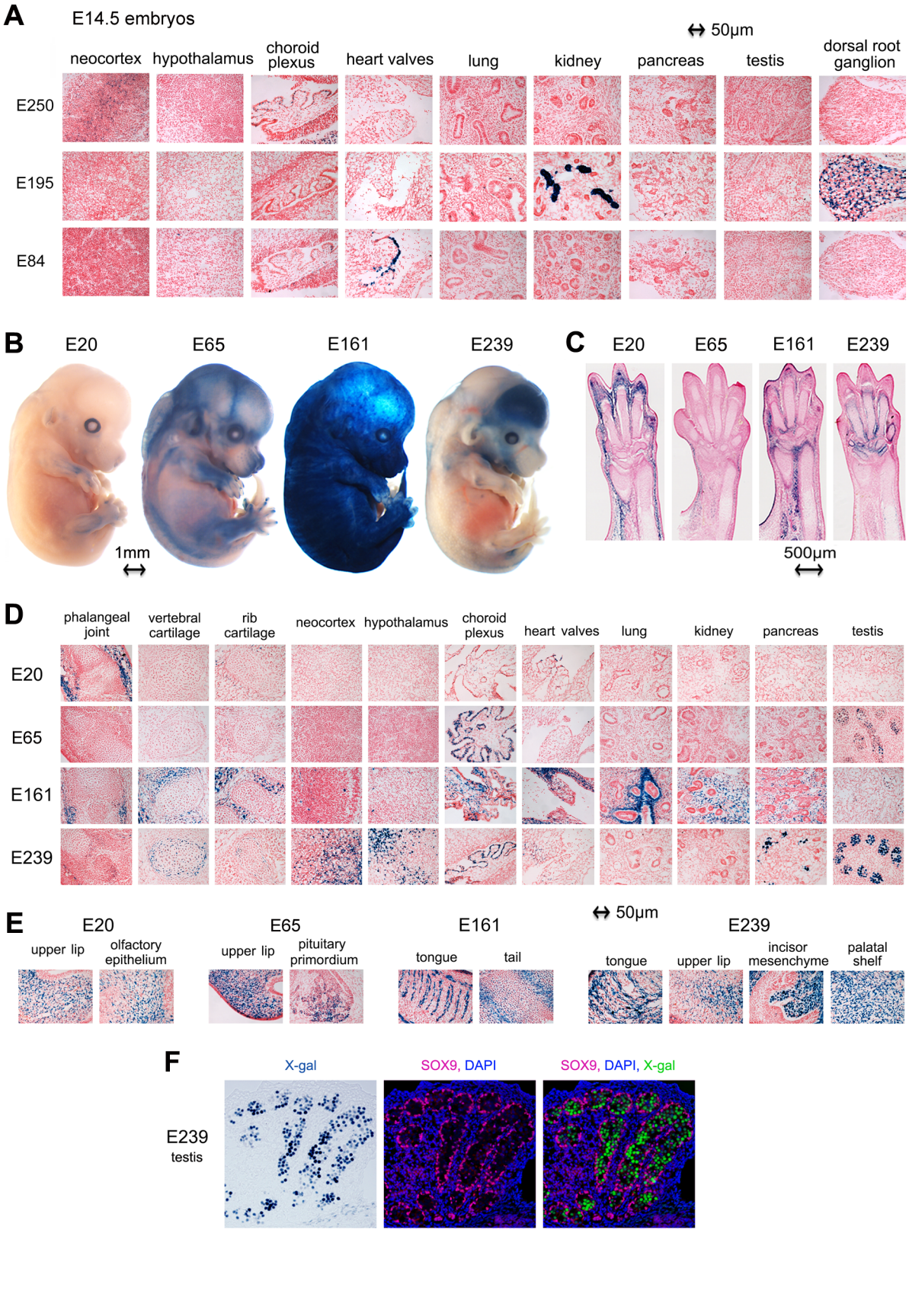
A
E14.5 embryos
↔ 50μm
neocortex
hypothalamus
choroid plexus
heart valves
lung
kidney
pancreas
testis
dorsal root ganglion
E250
E195
E84
B
E20
E65
E161
E239
1mm
C
E20
E65
E161
E239
500μm
D
phalangeal joint
vertebral cartilage
rib cartilage
neocortex
hypothalamus
choroid plexus
heart valves
lung
kidney
pancreas
testis
E20
E65
E161
E239
E
E20
upper lip
olfactory epithelium
E65
upper lip
pituitary primordium
E161
tongue
tail
↔ 50μm
E239
tongue
upper lip
incisor mesenchyme
palatal shelf
F
X-gal
SOX9, DAPI
SOX9, DAPI, X-gal
E239
testis

**Figure S2.** Activity of *Sox9* enhancers in E14.5 transgenic embryos. **(A)** High-magnification pictures of X-gal-stained sections through non-cartilaginous tissues from the same embryos as in Figure 3. **(B)** X-gal-stained E14.5 transgenic embryos harboring a *Sox9* promoter-pWHERE reporter carrying four copies of E20, E64, E161 or E239. Tissues expressing the reporters are seen in blue. **(C)** X-gal-stained sections through forelimbs of similar embryos as in panel B. Transgene activity is seen in mesenchyme surrounding skeletal elements in E20, E64 and E161 embryos, and in perichondrium in the E239 embryo. These structures no longer express *Sox9* at this stage. **(D** and **E)** X-gal-stained sections through various tissues from the same embryos as in panel C. Most sites of transgene activity are ectopic with regards to *Sox9* expression in panel D, but match *Sox9*-expressing sites in panel E. **(F)** X-gal staining (left) followed by SOX9 immunostaining and DNA staining with DAPI (middle) in a section through an E14.5 embryo testis expressing E239. The two pictures were merged in Adobe photoshop by inserting the X-gal staining picture into the green channel of the SOX9/DAPI picture (right). Note that E239 is active in cells located within seminiferous tubules (likely to be germ cells), but not in SOX9-expressing Sertoli cells lining the tubules.

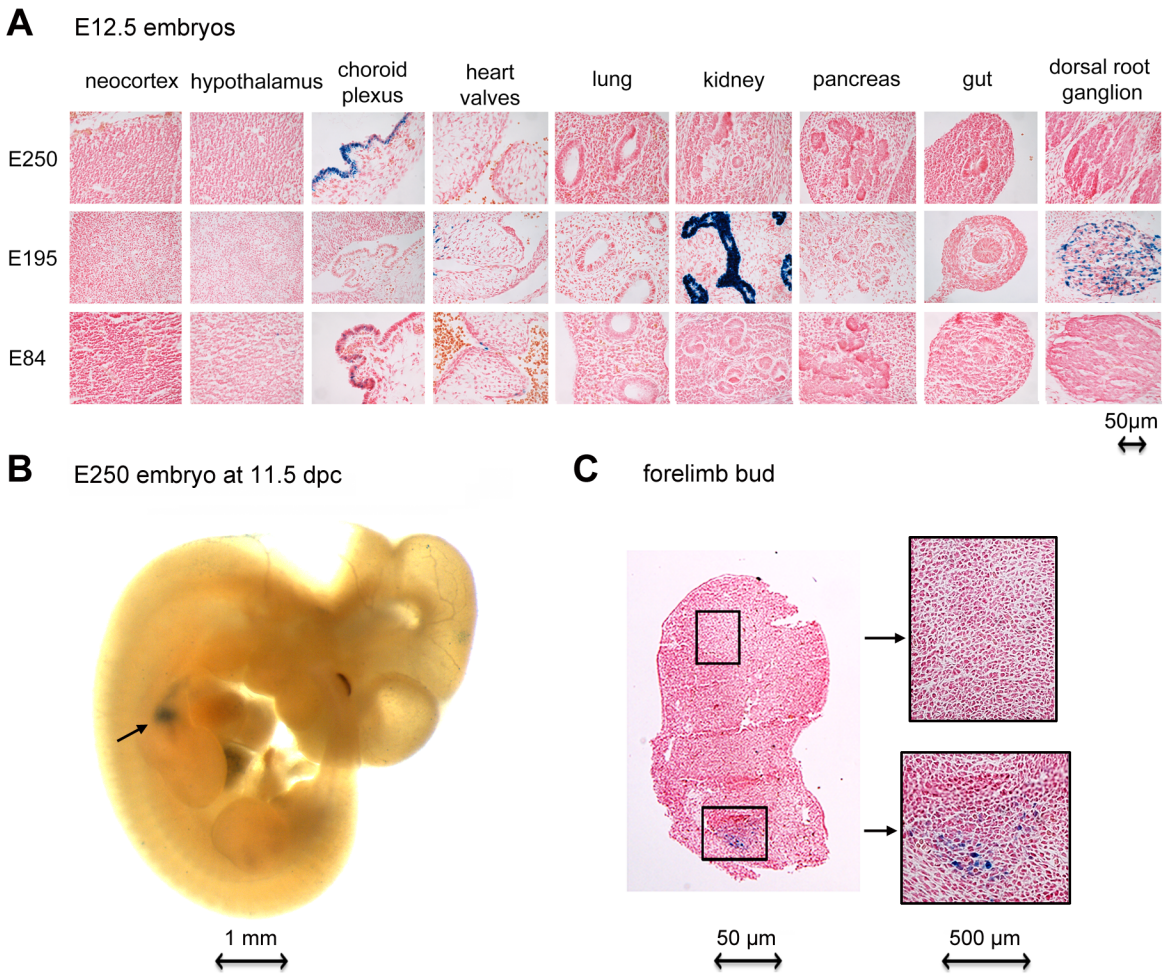


**Figure S3.** Activity of *Sox9* enhancers in E12.5 and E11.5 transgenic embryos. **(A)** X-gal-stained sections through various tissues from the same E12.5 embryos as in Figure 3E. Most sites of transgene activity match *Sox9*- or *Sox10*-expressing sites. **(B)** Whole-mount X-gal-stained E11.5 embryo carrying the E250 transgene. Arrow, transgene activity in the humerus and scapula primordia. **(C)** X-gal-stained section through a forelimb of the embryo shown in panel B. Left, low-magnification picture of the entire limb bud section. Right, high-magnification pictures of boxed areas highlighting mesenchymal cells that will soon give rise to digital precartilaginous condensations (top) and the condensing humerus primordium (bottom).

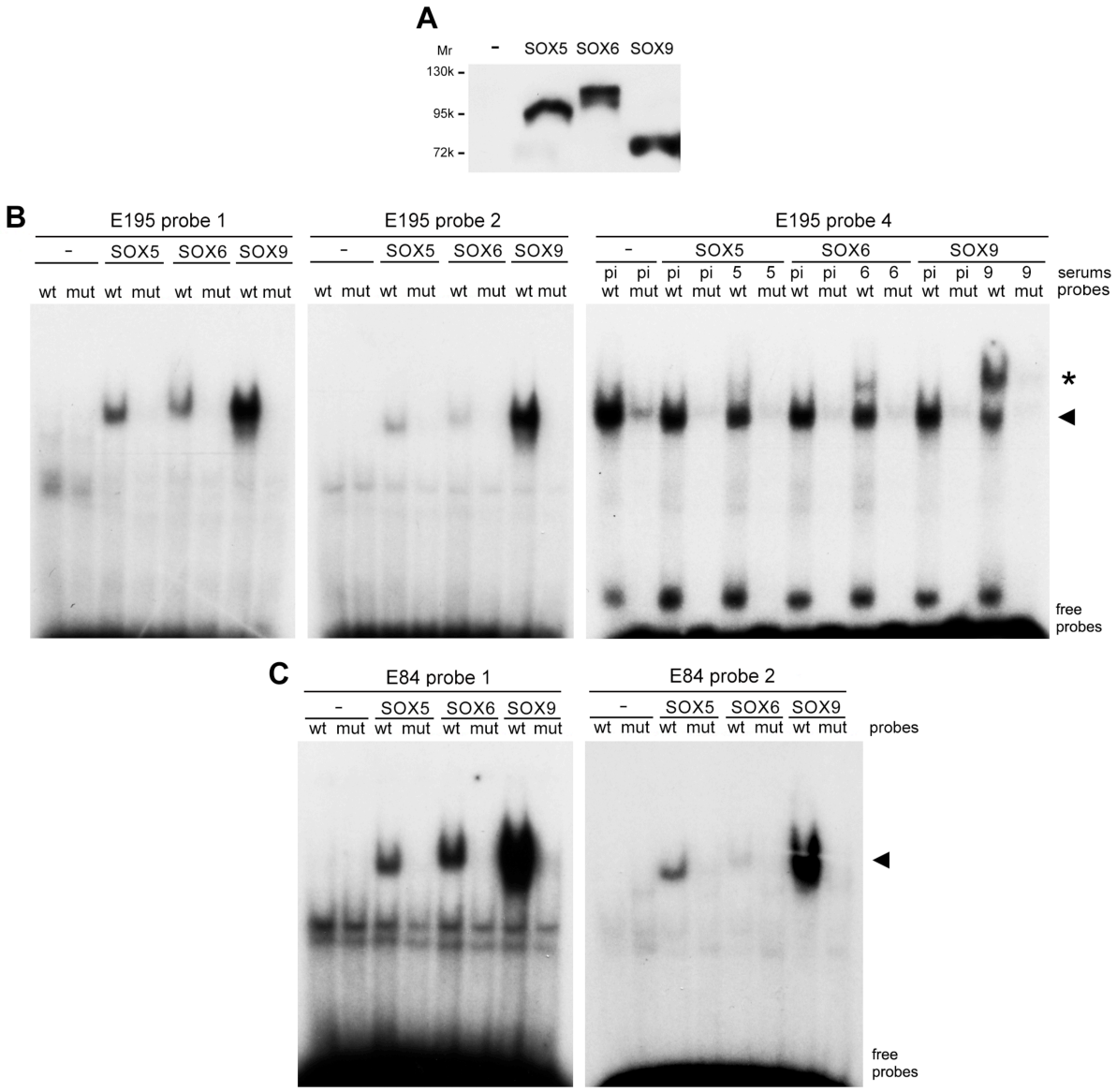


**Figure S4.** EMSAs with wild-type (wt) and mutant (mut) E195 probes 1, 2 and 4 and E84 probes 1 and 2. **(A)** Comparison of relative amounts of SOX proteins present in nuclear extracts. Nuclear extracts were obtained from COS-7 cells following cell transfection for 24 h with empty (-) or FLAG-tagged SOX-expression plasmids (SOX5, SOX6 or SOX9). The extracts were tested in western blot with anti-FLAG antibody to demonstrate that similar amounts of SOX proteins were present in all extracts. These extracts were also used in the EMSA shown in Figures 5F, 5G, and 6E. **(B)** EMSA with E195 probes and nuclear extracts from COS-7 cells forced to express no protein, SOX5, SOX6, or SOX9. Note that probe 4 forms a complex with a non-specific protein migrating at the same level as SOX/DNA complexes (arrowhead). SOX5,

SOX6 and SOX9 supershifts were obtained by EMSA of extracts incubated with probe 4 and preimmune serum (pi) or antiserum against SOX5 (5), SOX6 (6), or SOX9 (9). Star, antibody supershifts. **(C)** EMSA with E84 probes and nuclear extracts from COS-7 cells forced to express no protein, SOX5, SOX6, or SOX9.