Supporting Information

Chen et al. 10.1073/pnas.1005129107

SI Materials and Methods

Generation of Mice and Genotyping. The following primers were used to clone zebrafish *Hoxa3a* protein coding exon1, exon2, and mouse *Hoxa3* intron with PCR: zfExon1f, 5′-ATGCAAAAGGCAACCTACTGCG-3′; zfExon1r, 5′-cagCTGAAATGATACTACAGG-3′; zfExon2f, 5′-tacgTAGAGAGCTGCGCTGGAGAC-3′; zfExon2r, 5′-ctaagcgtaatctggaacatcgtatgggtaTAAATGCGTCAGTTTGGGTG-3′; msIntronf, 5′-GTAAATGAATGCCTTTAGGAGG-3′; msIntronr, 5′-TGCCAGGACACAGAGAGGAAG-3′. Letters in uppercase indicate the sequence of either exon or intron, lowercase indicate introduced sequence of either enzyme sites or HA tag.

zfExon1 and zfExon2 (GenBank accession no. NM 131534) were each subcloned into the T-easy vector (Promega). The Exon2 subclone was excised from the vector and inserted into Exon1-Teasy vector to generate Exon1&2-T easy plasmid. The mouse intron was then amplified from BL6 mouse genomic DNA using Pfu DNA polymerase and subcloned into Exon1&2-Teasy, which was linearized with PvuII and SnaBI, to generate zfExon1 + msIntron + zfExon2 in one plasmid, keeping exon-intron junctions unchanged. A loxP-flanked neomycin^r (Neo^r) gene and an FRT-flanked kanamycin^r (Kan^r) gene were inserted to the 3′ of the stop codon. Then the following primers were used to add 50 base pairs of homologous arms (for BAC recombination) flanking the zfexon1-msintron-zfexon2-Neo^r-Kan^r fragment with PCR: forward, 5′-ccttacgggtgtcaagcccttgtcagagagtgtgatcacgatcgtgaaacatcgcgATGCAAAAGGCAACCTACTG-3′; reverse, 5′-ctaaccaaagaaggtcgggtgggcaactctcctggctcacagcccctgggcccatcaGAAGTTCCTATTCTCTAGAAAG-3′. Lowercase indicates the sequence of mouse *Hoxa3* 5′UTR or 3′UTR, and uppercase indicates sequence of zebrafish *Hoxa3a* exon or FRT site.

BAC DNA (RP23-253E11, Cam$^+$; GenBank) was transformed into the EL250 *Escherichia coli* strain, which expresses a inducible FLP gene. Then the 5.5-kb PCR product was transformed into EL250 containing the BAC. Cam$^+$/Kan$^+$/Amp$^-$ clones were selected and checked for the occurrence of recombination, which resulted in the replacement of mouse *Hoxa3* protein coding exons with zebrafish *Hoxa3a* protein coding exons, and the introduction of Neo and Kan genes into the BAC. The FRT-flanked Kan^r cassette was then flipped out by adding 0.1 g/mL L-arabinose into the culture, which induced the expression of FLP.

We then extracted the modified BAC DNA and digested it with NotI. The two NotI sites flanking a 12-kb mouse *Hoxa3* genomic DNA enabled us to subclone the modified Hoxa3 genomic DNA into a plasmid vector containing two TK genes to generate the final targeting vector.

The targeting vector was linearized with XhoI and electroporated into the LK-1 C57BL/6 ES cell line (1) and subjected to positive-negative selection. Southern blot with flanking probes was used to screen the targeted ES cell clones and targeted mice. The targeted ES cells were then injected into BALB/c x C3H F1 hybrid blastocysts (2) to generate chimeras. The Neo^r selection marker gene was deleted by crossing to a ubiquitous cre-expressing mouse colony (3).

The following primers were used for genotyping of this Hoxa3-zfki mouse colony with PCR (approximately 500 bp WT band and approximately 400 bps targeted band): msE2in, 5′-ctatgtggagccatgagcaa-3′; HA, 5′-ccatacgatgttcctgattacg-3′; 3′UTR, 5′-AGGAAAAGGATGCAGGGCCAG-3′.

Southern Blot Analysis. Two micrograms of genomic DNA extracted from the ES cell clones were digested with restriction endonucleases PmeI&SpeI or BamHI in 25-μL reactions and separated on 0.7% agarose gels in 1× TBE overnight at 60 V. The gels were then denatured in 0.5 M NaOH, 1.5 M NaCl for 30 min followed by neutralized in 0.5 M Tris, pH 7.5, 1.5 M NaCl for 30 min. The DNA was then blotted onto Hybond N+ membranes (Amersham) by upward capillary transfer in 10× SSPE for overnight. The nylon membranes were then UV-crosslinked and baked at 80 °C for 2 h. The membranes can then be stored at 4 °C until use. The membrane was prehybridized in 4× SSPE, 0.1% sodium pyrophosphate, 0.1% SDS, 0.2% BSA, 0.2% Ficoll 400, 0.2% polyvinylpyrrolidone, and 100 ng/mL salmon sperm DNA at 55 °C for 1 h, and hybridized with radioactive DNA probes in a fresh solution at 55° C overnight. The probes were labeled with [^{32}P]dCTP using the Klenow enzyme (New England Biolabs). The hybridized membranes were washed twice in 2× SSC, 0.1% SDS at room temperature for 15 min; once in 0.2× SSC, 0.1% SDS at 60 °C for 10 to 15 min. An x-ray film was placed over the membrane in a developing cassette for overnight at −80 °C to detect the signal.

Absolute Quantification RT-PCR. Embryonic d 10.5 mouse embryos were homogenized in trizol (Invitrogen) and total RNA was extracted according to the manual coming with TRIzol reagent. The total RNA was then incubated with DNaseI (Roche) at 37 °C for 20 min to remove any DNA from the RNA samples. First-strand cDNA was reverse transcribed with superscript III (Invitrogen) and random primers (Invitrogen), incubated at 42 °C for 90 min and then at 70 °C for 15 min. RNaseH and RNaseA (Promega) were used to remove RNA from the transcribed first-strand cDNA.

Absolute quantitative PCR was performed on an ABI 7500 real time PCR system with SYBR green PCR master mix (Applied Biosystems) and the following primers designed with Primer Express3.0 software: mouse Hoxa3 pair 1, forward, 5′-CAGCCAATGGGTTCGCTTA-3′; reverse 5′-GCGGACGGCGCGTAT-3′; mouse Hoxa3 pair 3, forward, 5′-CCCGGTGCAGGAGGCTAT-3′; reverse, 5′-GGCTCATATGGGACACTGTTGA-3′; zebrafish Hoxa3a pair 3, forward, 5′-GATTCCGCCGCCAAGAA-3′; reverse, 5′-AAATGTGTTTCCTTGTGGATGGT-3′; Zebrafish Hoxa3a pair 7, forward, 5′-GCCAAGAATCCAGTCCACGTA-3′; reverse, 5′-TCATCCAAGGGAAAATGTGTTTC-3′.

Two plasmids containing full length cDNA of either mouse Hoxa3 gene or zebrafish Hoxa3a gene were used to generate standard curves for absolute quantification. The PCR condition is as the following: 50 °C, 2 min; 95 °C, 10 min; 40 cycles of 95 °C for 15 sec and 60 °C for 1 min, followed by dissociation steps. Absolute copy number was determined using 7500 SDS software (Applied Biosystems).

Hoxa3 Relative Protein Quantification. The target peptide (SPLLNSPTVGK) was chosen from the tryptic peptides of mouse Hoxa3 using the criteria of it not being conserved in other Homeobox family proteins and not having amino acids susceptible to oxidation and alkylation during sample workup. Proteins were extracted from mouse embryos (two embryos for each genotype, $^{+/+}$ and *mz*/*mz*) and digested by trypsin (Promega) following reduction and alkylation. The resulting peptides were separated by an offline strong cation exchange chromatography. The tryptic peptides were separated by offline strong cation exchange liquid chromatography. Solvent A (5 mM KH_2PO_4/30% acetonitrile, pH 2.7), solvent B (solvent A with 350 mM KCl), and solvent C (0.1 M Tris/0.5 M KCl, pH 7.0) were used to develop a linear gradient consisting of 5 min at 100% solvent A, 48 min gradient at variable slope to 100% solvent B, 12 min at 100% solvent B, 15 min to 100% solvent C, and 10 min to 100% solvent A. Fractions were collected every 2 min, and then

combined into five fractions, desalted, and dried. Seven fractions were collected for each genotype sample, and each fraction was analyzed in SRM mode via LC-MS/MS (LTQ-Orbitrap XL;ThermoFisher). The acquired spectra were searched against a mouse protein database (Swissprot, updated on March 24, 2009) using Bioworks (version3.3.1 SP1; Thermo Fisher Scientific). The calculation of the ratio was based on the peak area of the reconstructed ion chromatogram of respective peptides following normalization by a high-scoring trpytic peptide that coeluted from titin in both samples.

In Situ Hybridization. Whole-mount in situ hybridizations were performed essentially as previously described (4). Probes for mouse *Hoxa3* and *Pax1* (4), *Foxn1*, and *Gcm2* (5) have been described. Probe for zebrafish *Hoxa3a* was a full-length cDNA fragment (GenBank accession no. NM 131534). Embryonic d 10.5 and 11.5 mouse embryos were fixed in 4% PFA overnight then dehydrated in methanol and stored at −20°C until use. The roof plate of embryonic d 10.5 embryos was opened to reduce background, and embryonic d 11.5 mouse embryos were hemisected to help penetration. Digoxigenin-labeled RNA probes were used at 0.25 to 0.5 μg/mL. Signal was detected using 1:5,000 alkaline phosphatase–conjugated anti-dig antibody Fab fragments (Roche) and color reaction was carried out in BM-purple substrate (Roche) from 2 to 7 h. Embryos stained in whole mount were processed for sectioning by standard paraffin embedding and then cut into10-μm sections and counterstained with nuclear fast red.

Histology and Immunofluorecence (Anti-HA Staining). Newborn mice or embryos were fixed in 4% paraformaldehyde, dehydrated in gradient ethanol, embedded in paraffin, and then cut into 10-μm sections. Paraffin sections were either stained with hematoxylin and eosin or used for immunohistochemistry. Paraffin sections of embryonic d 10.5 mouse embryos were dewaxed and rehydrated with gradient ethanol/PBS solution. Antigen retrieval was performed by incubating in 10 mM citrate buffer in a boiling water bath for 20 min. The sections were then cooled down and rinsed with water before treated with 3% H_2O_2 in PBST (0.3% Triton X-100 in PBS solution). The sections were blocked in TNB [0.5% blocking reagent (PerkinElmer) in TNT (0.1 M Tris-HCl, pH 7.5, 0.15 M NaCl)] and then incubated with anti-HA (Roche) 1:50 diluted in TNB overnight at 4 °C. After the wash in TNT, the sections were incubated with 1:100 antirat IgG-Biotin (Jackson Laboratory) in TNB for 1 h at room temperature followed by wash and incubation with SA-HRP (PerkinElmer) for 30 min. Biotinyl tyramide (TSA biotin system kit; PerkinElmer) was applied followed by incubation in 1:100 SAV-Texas red (Jackson Laboratory) for 30 min. The sections were then stained with DAPI and mounted with aqueous mounting medium.

Calcitonin Immunohistochemistry. Anticalcitonin staining was performed using anticalcitonin (AnaSpec) essentially as described (4). The sections were dewaxed, rehydrated, and treated with 3% H_2O_2 in PBST before blocked in 10% goat serum/0.5% Triton X-100 in PBS solution (PBST) for 30 min at room temperature. The sections were then incubated with 1:200 anticalcitonin (AnaSpec) in PBST for 1 h at room temperature or overnight at 4 °C. After washing, Goat antirabbit IgG-HRP (Santa Cruz Biotechnology) was diluted 1:1,000 in PBST and incubated for 30 min at room temperature followed by wash in PBST. Color reaction was developed by incubating in 0.6 mg/mL diaminobenzidine and 0.06% H_2O_2 in PBS solution for 10 to 15 min at room temperature. Sections were then counter stained with nuclear fast red and then dehydrated and mounted with Cytoseal 280 mounting medium (Richard-Allan Scientific).

Neurofilament Staining. Whole-mount neurofilament staining was performed as described previously (6). Embryonic d 11.5 mouse embryos were fixed in methanol:DMSO (4:1) overnight followed by incubation in methanol:DMSO:30% H_2O_2 (4:1:1) for 4 to 6 h at room temperature. The embryos were then stored in methanol at −20 °C for upwards of 1 month until use. The embryos were dehydrated with a gradient methanol in PBST, then blocked in PBSTMD (2% skim milk powder, 1% DMSO in PBST) for 2 h at room temperature. The embryos were then incubated in PBSTMD:neurofilament monoclonal antibody 2H3 cell supernatant (4:1; Developmental Studies Hybridoma Bank) overnight at 4 °C. After five 1-h washes in PBSTMD, the embryos were incubated in goat anti-mouse antibody (1:100 diluted in PBSTMD; Santa Cruz Biotechnology) for 3 d at 4 °C. Color reaction was performed by incubating with 1 mg/mL diaminobenzidine in PBS solution for 1 h, followed by incubation in 0.03% H_2O_2/PBS solution for 1 to 2 min. The embryos were then dehydrated in methanol and cleared in BABB buffer (1:2 benzyl benzoate:benzyl alcohol).

Skeleton Preparation. Alizarin red and alcian blue staining of skeleton preparations was performed as previously described for zebrafish embryos (7) and newborn mice (8). Newborn mice were killed by CO_2 asphyxiation. The skin and the internal organs up to the diaphragm were removed. The newborn mice were then fixed in 95% ethanol for 5 d, soaked in acetone for 2 d, then stained in the staining solution (0.015 g alcian blue 8GX, 0.005 g alizarin red S, 5 mL acetic acid, 75 mL 95% ethanol, 20 mL water) at 37 °C for 10 d. The staining solution was then replaced with pure water and then with the trypsin solution (30% saturated sodium borate, 1% trypsin). The samples were cleared in the trypsin solution at 37 °C for 4 to 6 h and then in gradient glycerol/1% KOH.

Zebrafish. WT zebrafish stocks were maintained at the University of Georgia using standard husbandry techniques. Collected embryos were incubated at 28 °C in egg water. For some experiments, 0.003% 1-phenyl-2-thiourea was added into the egg water from 12 h postfertilization (hpf).

In Situ Hybridization and Sections on Zebrafish Embryos. Embryos were fixed 24 h in 4% paraformaldehyde and dehydrated overnight in methanol at −20 °C. Then the embryos were rehydrated stepwise in methanol/PBS solution and finally put back in 100% PBT (1× PBS 0.1% Tween 20). Embryos were then treated with proteinase K (10 mg/mL in PBT) for 25 min (48 hpf) or 30 min (72 hpf). The digestion was stopped by rinsing in glycine (2 mg/mL in PBT). Embryos were postfixed in 4% paraformaldehyde 1′ PBS for 20 min and then washed in PBT five times for 5 min each. The embryos were prehybridized at least 1 h before hybridization at 70 °C overnight in hybridization buffer (50% formamide, 5′SSC, 50 mg/mL heparin, 500 mg/mL tRNA, 0.1% Tween 20, 9 mM citric acid). Then the embryos were washed at 70 °C for 15 min in 75% hybridization buffer, 25% 2′SSC; 15 min in 50% hybridization buffer, 50% 2′SSC; 10 min in 25% hybridization mix, 75% 2′SSC; 15 min in 2′SSC; and twice for 15 min in 0.2′SSC. Further washes were performed at room temperature for 10 min in 75% 0.2′SSC, 25% PBT; 10 min in 50% 0.2′SSC, 50% PBT; 5 min in 25% 0.2′SSC, 75% PBT; 10 min in PBT; and then 2 h in PBT with 2 mg/mL BSA, 2% sheep serum. Then the embryos were incubated overnight at 4 °C with the preabsorbed alkaline phosphatase–labeled anti-digoxigenin antiserum (Roche) at a 1/5,000 dilution in a PBT buffer containing 2 mg/mL BSA, 2% sheep serum. The embryos were then washed six times for 15 min each in PBT at room temperature, followed by washing in staining buffer (100 mM Tris HCl, pH 9.5, 50 mM $MgCl_2$, 100 mM NaCl, 0.1% Tween-20) three times for 5 min each. Color reaction was performed in BM purple AP substrate (Roche). When the color was developed, the reaction was stopped in 1′PBS solution/5 mM EDTA. The whole-mount images were performed with a Leica MZ125 dissection scope and a Q imaging

digital camera. The embryos were then embedded with 7.5% gelatin solution (7.5% gelatin, 15% sucrose in PBS solution) and sectioned into 10-μm sections with cryostat. Imaging was done on a Zeiss microscope with Optronics digital camera.

MO Injection. Splicing-suppressing MOs were designed by Gene Tools to complement the sequence around the splicing junction. The sequence of the MO is as follows, with the slash line representing the exon/intron junction. Hoxa3a: splicing donor, 5′-GTGGTTGATGTAATCAC/CTGAAATG-3′; splicing acceptor, 5′- GCGCAGCTCTCTA/CTGTATGAGTAC-3′. Hoxb3a: splicing donor, 5′-TTGGCAAACACAC/CATTAGCTGAGC-3′.

The MOs—MOhoxa3a, MOhoxb3a, and control MO—were microinjected (approximately 1 nL) at concentrations ranging from 2.5 to 10 mg/mL in phenol red buffer (0.25% phenol red, 120 mM KCl, 20 mM Hepes-NaOH, pH 7.5)

RT-PCR on Zebrafish Embryos. MO-injected or control embryos at different stages were homogenized in TRIzol (Invitrogen) and total RNA was extracted according to the manual coming with TRIzol reagent. The total RNA was then incubated with DNaseI (Roche) at 37 °C for 20 min to remove any DNA from the RNA samples. The same amount of RNA from each sample was used for reverse transcription. First-strand cDNA was reverse transcribed with superscript III (Invitrogen) and random primers (Invitrogen), incubated at 42 °C for 90 min and then at 70 °C for 15 min. RNaseH and RNaseA (Promega) were used to remove RNA from the transcribed first-strand cDNA. Primers used in semiquantitative PCR are as follows: Hoxa3a exon fP, 5′- TACTGCCCCTTCTGGACCATC-3′; Hoxa3a exon rP, 5′- CATTCCAAGCCCCTTCTGGTC-3′; Hoxa3a intron rP, 5′- CAGACAACATATGACTGCGC-3′; Hoxb3a exon fP, 5′- CTATGGCCACAAACCCAACT-3′; and Hoxb3a exon rP, 5′- ATCTTGATCTGCCGCTCACT-3′.

Sequences Used for Evolutionary Analyses. Sequences used for evolutionary analysis in Fig. S5 were obtained from publically available databases. Each sequence is listed by the taxon and common name; the database: accession number and size in amino acids of the sequence used are listed for each species.

Anolis carolinensis (anole lizard), Ensembl: ENSACAP0000001022, 420 aa.
Astatotilapia (haplochromine cichlid fish), GenBank: ABS70768.1, 417 aa.
Callorhinchus milii (elephant shark chimera), GenBank: ACU32552.1, 409 aa.
Danio rerio (zebrafish), Hoxa3a, Swissprot: Q8AWZ2.1, 410 aa.
Gallus gallus (chicken), NCBI: NP_989879.1, 413 aa.
Heterodontus francisci (horn shark), Swiss-Prot: Q9IA21.1, 410 aa.
Homo sapiens (human), NCBI: NP_109377.1, 443 aa.
Latimeria menadoensis (coelacanth fish), GenBank: ACL81435.1, 411 aa.
Leucoraja erinacea (skate), GenBank: ACT65754.1, 411 aa.
Mus musculus (mouse), NCBI: NP_034582.1, 443 aa.
Oryzia latipes (medaka fish), Hoxa3a, GenBank: BAE44251.1, 417 aa.
Polypterus senegalis (bichir fish), GenBank: AC126321.1, 415 aa.
Rhinolophus ferrumequinum (bat), GenBank: ACC68935.1, 443 aa.
Salmo salar (salmon fish), Hoxa3aa, NCBI: NP_001135143.1, 422 aa.
Salmo salar (salmon fish), Hoxa3ab, NCBI: NP_001124470.1, 422 aa.
Sorex araneus (shrew), GenBank: ACE79086.1, 442 aa.
Takifugu rubripes (puffer fish), Hoxa3a, Swiss-Prot: Q1KL12.1, 417 aa.
Xenopus tropicalis (clawed frog), GenBank: AAI66398.1, 406 aa.

1. Keskintepe L, Norris K, Pacholczyk G, Dederscheck SM, Eroglu A (2007) Derivation and comparison of C57BL/6 embryonic stem cells to a widely used 129 embryonic stem cell line. *Transgenic Res* 16:751–758.
2. Pacholczyk G, Suhag R, Mazurek M, Dederscheck SM, Koni PA (2008) Generation of C57BL/6 knockout mice using C3H x BALB/c blastocysts. *Biotechniques* 44:413–416.
3. Koni PA, et al. (2001) Conditional vascular cell adhesion molecule 1 deletion in mice: impaired lymphocyte migration to bone marrow. *J Exp Med* 193:741–754.
4. Manley NR, Capecchi MR (1995) The role of Hoxa-3 in mouse thymus and thyroid development. *Development* 121:1989–2003.
5. Gordon J, Bennett AR, Blackburn CC, Manley NR (2001) Gcm2 and Foxn1 mark early parathyroid- and thymus-specific domains in the developing third pharyngeal pouch. *Mech Dev* 103:141–143.
6. Manley NR, Capecchi MR (1997) Hox group 3 paralogous genes act synergistically in the formation of somitic and neural crest-derived structures. *Dev Biol* 192:274–288.
7. Kimmel CB, et al. (1998) The shaping of pharyngeal cartilages during early development of the zebrafish. *Dev Biol* 203:245–263.
8. Mansour SL, Goddard JM, Capecchi MR (1993) Mice homozygous for a targeted disruption of the proto-oncogene int-2 have developmental defects in the tail and inner ear. *Development* 117:13–28.

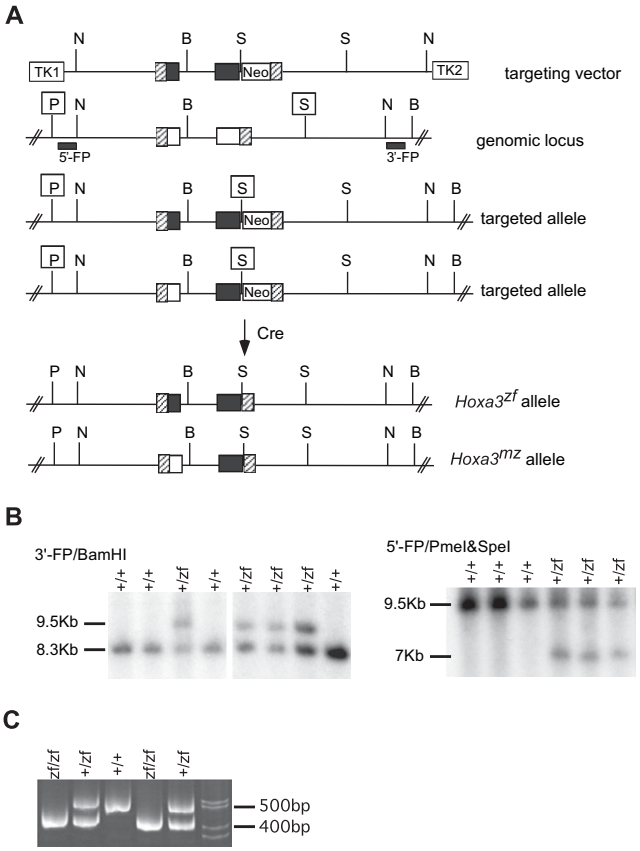

Fig. S1. Strategy for gene targeting. (*A*) Map of targeting vector, WT mouse Hoxa3 genomic locus, and targeted locus before and after the deletion of neomycin with cre recombinase. Horizontal thin lines represent noncoding genomic DNA at mouse Hoxa3 locus, boxes with heavy diagonal lines represent 5′ or 3′UTR of mouse Hoxa3, boxes in black represent coding exons of zebrafish Hoxa3a, open boxes represent coding exons of mouse Hoxa3. Black bars under the horizontal line identify the 5′ and 3′ franking probes. P, PmeI; N, NotI; B, BamHI; S, SpeI. (*B*) Southern blot analysis of the indicated mouse genotypes. Genomic DNA was digested with the PmeI and SpeI and probed with 5′-franking probe, or digested with BamHI and probed with 3′-franking probe. (*C*) PCR identification of the indicated genotypes.

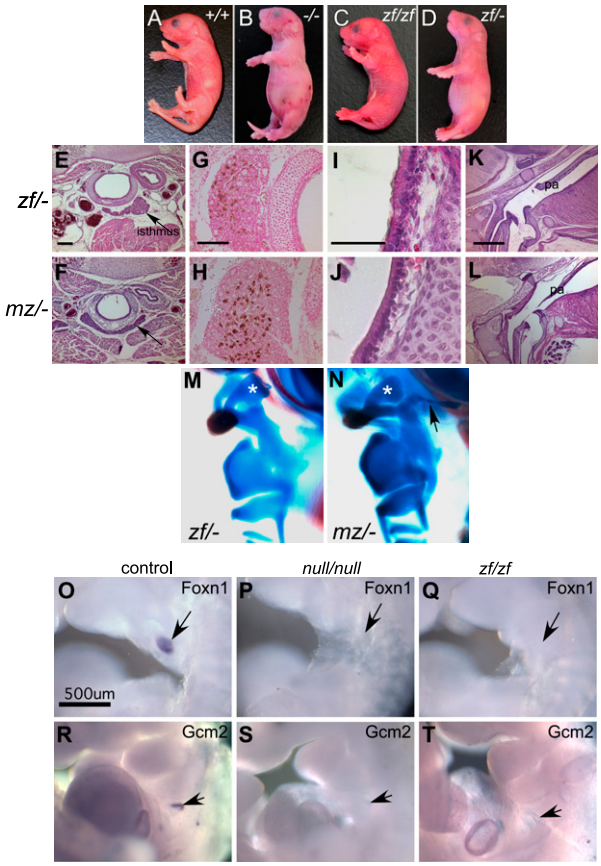

Fig. S2. Molecular phenotype and dosage effect of the *Hoxa3*^{zf} and *Hoxa3*^{mz} alleles. (*A–D*) *Hoxa3*^{zf} rescued bloated phenotype in a dosage-dependent manner. Newborn animals of indicated genotypes. *Hoxa3*$^{null/null}$ ($^{-/-}$) animals are characterized by a bloated abdomen, which is never seen in *Hoxa3*$^{zf/zf}$ (*zf/zf*) mice. However, with only one copy of *Hoxa3*^{zf} allele, the *Hoxa3*$^{zf/null}$ (*zf/–*) mutant mimics null mutant in the bloated abdomen phenotype. (*E–J*) Hematoxylin and eosin staining on transverse paraffin sections of newborn animals (dorsal is up). One copy of the *Hoxa3*^{zf} or *Hoxa3*^{mz} allele is sufficient for the normal development of thyroid isthmus (arrow) (*E* and *F*), ultimobrachial body (*G* and *H*), and tracheal epithelia (*I* and *J*). (*K* and *L*) Hematoxylin and eosin staining on newborn sagittal paraffin sections (anterior is up, dorsal is to the left). One copy of the *Hoxa3*^{mz} is sufficient to rescue the soft palate defect associated with *Hoxa3*-null mutant, but one copy of *Hoxa3*^{zf} allele is not sufficient to do so. (*M* and *N*) Lateral views of the throat cartilages in cleared newborn skeletal preparations (anterior is up, dorsal is to the right). Asterisk indicates lesser horn of hyoid bone. (Scale bars: 200 μm for *E* and *F*, 100 μm for *G* and *H*, 50 μm for *I* and *J*, 800 μm for *K* and *L*, and 500 μm for *M* and *N*.) (*O–T*) Defective organogenesis of thymus and parathyroid. Whole-mount in situ hybridization was used to analyze the expression of Foxn1 at E11.5 (*O–Q*), Gcm2 (*R–T*) at embryonic d 10.5. Foxn1 expression in the third pouch is absent in $^{-/-}$ and *zf/zf* embryos. Gcm2 expression at embryonic d 10.5 is detected at the third pouch in control embryo, but is greatly reduced in $^{-/-}$ and *zf/zf* embryos. pp, pharyngeal pouch; pa, pharyngeal arch. Cranial is up. (Scale bar: 500 μm.)

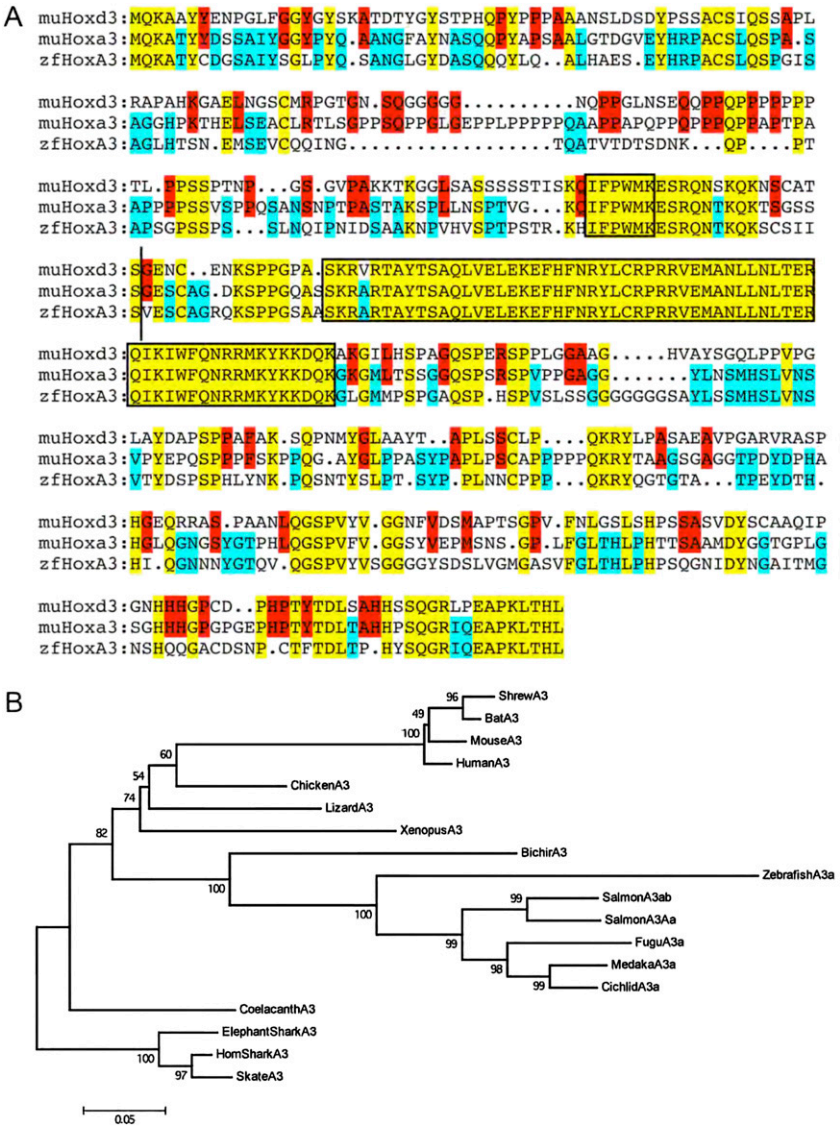

Fig. S3. Protein sequence alignment and molecular evolutionary analysis of Hoxa3 genes. (*A*) Amino acid sequence alignment of zebrafish Hoxa3a, mouse Hoxa3, and mouse Hoxd3. Yellow blocks highlight the identical amino acids among the three proteins. Red blocks highlight the identity between mouse HoxA3 and mouse HoxD3, which is 52%. Cyan blocks show the 59% identity between mouse HoxA3 and zebrafish HoxA3a. The hexapeptide motifs are shown in the small box. The two larger boxes highlight the homeodomains. The vertical line indicates the splice junctions. (*B*) Phylogenetic tree of vertebrate Hoxa3. The alignment in Fig. S6 was used for generating this phylogenetic tree. The MEGA 4 package (1) was implemented and the tree was generated using the neighbor-joining method (2) with 1,000 bootstrap replications. The tree was rooted to the cartilaginous fish lineage. Alternative methods of phylogenetic analysis gave essentially the same topology as the neighbor-joining tree. The salient point of this tree is that the zebrafish Hoxa3a sequence is evolving at a faster rate than any of the Hoxa3 sequences used in the analysis, including that of other teleost fishes. This is borne out in the relative rate tests (Table S3).

1. Kumar S, Nei M, Dudley J, Tamura K (2008) MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief Bioinform* 9:299–306.
2. Saitou N, Nei M (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425.

```
ShrewA3              MQKATYYDSSAIYGGYPYQAANGFAYNANQQPYP--ASAALGADG-EYHRPACSLQSP-- 55
BatA3                MQKATYYDSSAIYGGYPYQAANGFAYNASQQPYP--ASVALGADG-EYHRPACSLQSPA- 56
HumanA3              MQKATYYDSSAIYGGYPYQAANGFAYNANQQPYP--ASAALGADG-EYHRPACSLQSPS- 56
MouseA3              MQKATYYDSSAIYGGYPYQAANGFAYNASQQPYA--PSAALGTDGVEYHRPACSLQSPA- 57
ChickenA3            MQKATYYDSSAIYGAYPYQGANGFTYNASQQQYP--PSSSL-VET-EYHRPACSLQSPG- 55
LizardA3             MQKATYYDSSAIYGGYPYQGTNGFAYNASQQQYPPLPPSSSLVET-EYHRPACSLQSPGG 59
XenopusA3            MQKATYYDSSAIYGGYPYQGANGFTYNASQQQYP---PSSSLLET-EYHRPACSLQSPG- 55
HornSharkA3          MQKATYYDSSAIFGGYTYQGANGFNYNASQQQYP----PSSHVES-DYHRPACSLQSPG- 54
SkateA3              MQKATYYDSSAIFGGYTYQGANGFNYNANQQQYP----PSSHVEN-DYHRPACSLQSPA- 54
ElephantSharkA3      MQKATYYDSSAIFGGYTYQAANGFNYNASQQQYP----PSSHVES-DYHRPACSLQSPA- 54
CoelacanthA3         MQKATYYDSSAIYGGYPYQGANGFTYNASQQQYPQ---SSSLVES-DYHRPACSLQSPA- 55
MedakaA3a            MQKATYYDSSAIYSGYPYQSANGFSYDANQVQYP----RASHVES-EYHRPACSLQTPG- 54
CichlidA3a           MQKATYYDSSAIYSGYPYQSANGFSYDANQVQYP----RASHVES-EYHRPACSLQSPD- 54
FuguA3a              MQKATYYDSSAIYSGYPYQSANGFSYDANQIQYP----RTSHVES-EYHRPACSLQSPG- 54
SalmonA3ab           MQKATYYDSSAIYSGYPYQSANGFSYDANQVQYP----RASHVES-EYHRPACSLQSPD- 54
SalmonA3Aa           MQKATYYDSSAIYSGYPYQSANGFSYDANQVQYP----RVSHVES-EYHRPACSLQSPD- 54
ZebrafishA3a         MQKATYCDGSAIYSGLPYQSANGLGYDASQQQYL----QALHAES-EYHRPACSLQSPG- 54
BichirA3             MQKATYYDSSAIYGGYPYHSAKGFSYNVNQQQYP----PSSHVES-DYHRPACSLQSSA- 54
                     ****** *.***:.. .*:.::*: *:..*  *          :  :*********:.

ShrewA3              SAGGHPKAHELSEACLRTLSGPPSQPPGLGEPPLAPPPPQAAPPAPQQPQPPPQPPAPAA 115
BatA3                SAGGHPKAHELSEACLHTLSGPPSQPPGLGEPPLAPPPPQAAPPAPQQPQPPPQPPAPAP 116
HumanA3              SAGGHPKAHELSEACLRTLSAPPSQPPSLGEPPLHPPPPQAAPPAPQPPQPAPQPPAPTP 116
MouseA3              SAGGHPKTHELSEACLRTLSGPPSQPPGLGEPPLPPPPPQAAPPAPQPPQPPPQPPAPTP 117
ChickenA3            SAVSHHKANDISESCMRTLPSQPLQPPGLTDP-----------------QAPPQPPPAPQ 98
LizardA3             SGGGGGGATDLNESCMRTLTSQTLAPQVLPEQP-------------QQPPPPSQPGPQQP 106
XenopusA3            SAVPHHKANDINESCMRTINSQSNQAPVIPEQ-----------------QPTPQ------ 92
HornSharkA3          -TVPHHKPNDINESCMRTSASQPSHHPVIAEQ--------------------QQQKQPPP 93
SkateA3              STVPHHKPNDINESCMRTSANQTSHLPVIAEQ--------------------QKPPPPPP 94
ElephantSharkA3      STVPHHKPNDINESCMRTSASQPSHPPVIPEQ--------------------QQKQ--PP 92
CoelacanthA3         STVPHHKHNDINESCMRTSNNQPPQPPGISEH--------------------QQTQ--PP 93
MedakaA3a            GSVALQKPGEMAEGCDRTTSIQAAQSKVIPEN-------------------NQ--VPVSG- 93
CichlidA3a           GSVALQKPGEMAESCDRTTAIQAAQSKVHPES-------------------NQPQVPVSG- 95
FuguA3a              GSVALQKREMAAENCDRTTAVQAVQSKVHPES-------------------NQPQVPVSA- 95
SalmonA3ab           GLVALQKPGEMAENCDRGTAIQAAQPPVLTES-------------------NQPQMSVSV- 95
SalmonA3Aa           GSVALQKSGEIAESCDMSTAIQAAQPPVLTDS-------------------NKPQVSVSV- 95
ZebrafishA3a         ISAGLHTSNEMSEVCQQ---INGTQATVTDTS-------------------DNKQPPTAP- 92
BichirA3             VNAPLHKSNEITESCLQTNTTQQKQPTIVPDN-------------------QTQAQPLSRG 96
                           * *

ShrewA3              AVPPTPSSVSPPQNAS-NNPAPASAAK---SPLLNSP-TVAKQIFPWMKESRQNTKQKTS 170
BatA3                AAPPPPSSVSPPQNAS-SNPTPASAAK---SPLLNSP-TVAKQIFPWMKESRQNTKQKTS 171
HumanA3              AAPPPPSSASPPQNAS-NNPTPANAAK---SPLLNSP-TVAKQIFPWMKESRQNTKQKTS 171
MouseA3              AAPPPPSSVSPPQSAN-SNPTPASTAK---SPLLNSP-TVGKQIFPWMKESRQNTKQKTS 172
ChickenA3            AQPPPPSSASPSQNAS-SNPAPANSTK---SPALNSP-TVSKQIFPWMKESRQNTKQKNS 153
LizardA3             APPPPSSSVSPPPNAPNATSAPANATK---SAILNSP-TMSKQIFPWMKESRQNTKQKHS 162
XenopusA3            -GPPP--SVSPPQTTS--NAATASSNK---ATSITSP-TMSKQIFPWMKESRQNTKQQKA 143
HornSharkA3          PPPPPPPSVSPPQNTS-SNSTQSSTSK---NPTLTSQATISKQIFPWMKESRQNAKQKTS 149
SkateA3              PPPPPPPSVSPPQNTS-NNSTQSSTSK---NPTLTSQTTISKQIFPWMKESRQNTKQKTS 150
ElephantSharkA3      PPPPPPPSVSPPQNTS-SNSTQSSTSK---NPTLTSQTTISKQIFPWMKESRQNAKQKNS 148
CoelacanthA3         QAPQPTNPVSPSQTSS-NNSTPSNSNK---NPGITSPPTIAKQIFPWMKESRQNSKQKNS 149
MedakaA3a            -PPPS--SQSPGAISQNASNGSSQPG---AKNASPTSGARSKQIFPWMKECRQNTKQKPA 147
CichlidA3a           -PPPP--SQSPGAISQNTSNGSNQPVS--AKNDSPTSTARSKQIFPWMKESRQNTKQKPT 150
FuguA3a              -PPPP--PQSPGAISQTTSNGSNQPT---AKNSSPTSASRGKQIFPWMKESRQATKQKST 149
SalmonA3ab           -PPPPPQSQSPGSLNQNTSNGSCQPNSKSSVHSSPT-TTR-KHIFPWMKESRQNQKPK-T 151
SalmonA3Aa           -PPPPPQSQSPGSLNQNKSNGSSQADSNGSVHGSPTSTTR-KHIFPWMKESRQNQKPK-T 152
ZebrafishA3a         -SGPS----SPSSLNQIPNIDSAAKN---PVHVSPTPSTR-KHIFPWMKESRQNTKQK-S 142
BichirA3             TVAPPSPSISPPENTSSSSNHSCTNSN--PPHSSPSPATT-KQIYPWMKESRQNSKPK-N 152
                         .    **       .               :  :  *:*:*****.**  * :
```

Fig. S4. Amino acid alignment of selected Hoxa3 sequences (*SI Materials and Methods*) using ClustalW2 (1). Shaded region indicates the homeodomain, which is completely conserved across all taxa. Gaps were introduced by the Clustal algorithm. Asterisk indicates identical residues; colon indicates conserved substitutions; period indicates semiconserved substitutions.

1. Larkin MA, et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23:2947–2948.

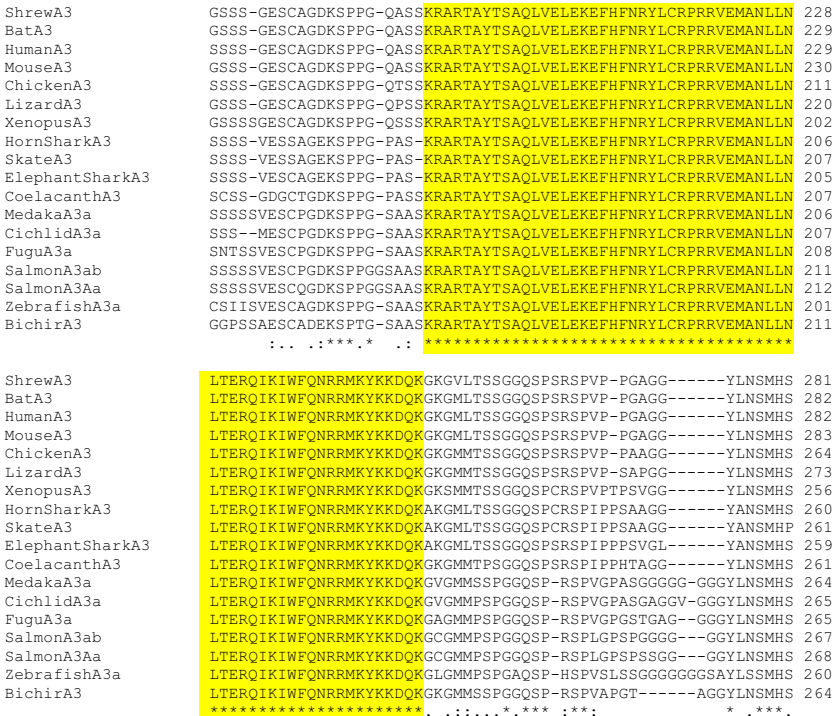

Fig. S4. (Continued)

```
ShrewA3              TPDYDPHAHGLQGNGSYGTPHLQGSPVFVGGS--YVEPMSNSGPALFGLTHLPHAASAAM 395
BatA3                TPDYDPHAHGLQGNGSYGTPHLQGSPVFVGGS--YVEPMSNSGPALFGLTHLPHAASAAM 396
HumanA3              TPDYDPHAHGLQGNGSYGTPHIQGSPVFVGGS--YVEPMSNSGPALFGLTHLPHAASGAM 396
MouseA3              TPDYDPHAHGLQGNGSYGTPHLQGSPVFVGGS--YVEPMSNSGP-LFGLTHLPHTTSAAM 396
ChickenA3            TPEYDTHP--LQGNG-YGNPHIQGSPVYVGGN--YVETMTNSGPSIFGLTHLSHPPSANM 368
LizardA3             TPEYDPHS--LQGNG-YGNPHIQGSPVYVGSN--YVDTMTNSGPSIFGLTHLPHPSAANM 377
XenopusA3            TPEYEPHP-LQQSSGAYGNPHVQGSPVYVGGN--YVETMTNSGPSMFGLSHLSHSSS-NM 361
HornSharkA3          TPEYDTHP--LQGNGGYGTPHLQGSPVYVGGN--FVESMPSSGPSLFSLTHLGHPPSGNM 365
SkateA3              TPEYDTHP--LQGNGGYGTPHLQGSPVYVGGN--FVETMPSSGPSLFSLTHLGHPPSGNM 366
ElephantSharkA3      TPEYETHP--LQGNGGYGTPHLQGSPVYVGGN--FVESMSSSGPSLFSLTHLGHPPSGNM 364
CoelacanthA3         TPEYDTHP--LQGNGTYGNPHLQASPVYVGGN--YVETMGNSGSSIFGLTHLPHPSTTNM 366
MedakaA3a            TPEYDAHP--LQGSGSYGTHMQ-GSPVYVGGG--YIDSMPSSGASVFGLTHLSHPPSANM 372
CichlidA3a           TPEYDAHP--LQGNGSYGTHMQ-GSPVYVGGG--YIDSMPSSGASVFGLAHLTHPPSANM 373
FuguA3a              TPEYDAHP--LQGNGSYGTHMQ-GSPVYVGGG--YIDSVPNSGSSVFGLTHLPHPPSANM 373
SalmonA3ab           TPEYDAHP--LQGNGGYGTHLQQVSPVYVGGG--YIDSMPNNRASVFGLTHLSHPPSTNM 377
SalmonA3Aa           TPEYDAHP--LQGNGGYGAHLQQGSPVYVSGG--YIDSMPNNGSSIFGLTHLSHPPSTHM 378
ZebrafishA3a         TPEYDTHH--IQGNNNYGTQVQ-GSPVYVSGGGGYSDSLVGMGASVFGLTHLPHPSQGNI 365
BichirA3             TPEYETHS--LQGNGNYGTAHLQTSPVY--GS--YVDSIANSGPSIFGLTHLPHPSSANM 370
                     **:*:.*    *... **      ***:  ..  : :.: .  . :*.*:** *..   :

ShrewA3              DYGGAGPLGSGHHHGPGPGEPHP-TYTDLTAHHPSQGRIQEAPKLTHL 442
BatA3                DYGGAGPLGSGHHHGPGPGEPHP-TYTDLTAHHPSQGRIQEAPKLTHL 443
HumanA3              DYGGAGPLGSGHHHGPGPGEPHP-TYTDLTGHHPSQGRIQEAPKLTHL 443
MouseA3              DYGGTGPLGSGHHHGPGPGEPHP-TYTDLTAHHPSQGRIQEAPKLTHL 443
ChickenA3            DYSGAGPMGNNHHHG--PCDPHP-TYTDLTAHHPSQGRIQEAPKLTHL 413
LizardA3             DYS--GPMGNNHPHG--PCDPHP-TYTDLTSHHASQGRIQEAPKLTHL 420
XenopusA3            DYSGAGPMNSGHHHG--PCDSHP-TYTDLSAHHNPQGRIQEAPKLTHL 406
HornSharkA3          DYNGAGPMTSNHHHG--PCDPHP-TYTDLSSHHPSQGRIQEAPKLTHL 410
SkateA3              DYNSAGPMTSNHHHG--PCDPHP-TYTDLSSHHPSQGRIQEAPKLTHL 411
ElephantSharkA3      DYSGAGPMTSNHHHG--PCDPHP-TYTDLSSHHPSQGRIQEAPKLTHL 409
CoelacanthA3         DYSGAGPMASNHHHG--PCDPHP-TYTDLTSHHPSQGRIQEAPKLTHL 411
MedakaA3a            EYNGAITMGNSQHHG--VCDPTP-TYTDL-TPHYSQGRIQEAPKLTHL 416
CichlidA3a           EYNGAITMGNSQHHG--VCDPTP-TYTDL-TPHYSQGRIQEAPKLTHL 417
FuguA3a              DYNGAITMGNSQHHG--VCDPTP-TYTDL-TSHYSQGRIQEAPKLTHL 417
SalmonA3ab           DYNGAITMGNSHHQR--VCDPNP-TYTDLNTPHYSQGRIQEAPKLTHL 422
SalmonA3Aa           DYNGAITMGNSHHHG--LCDPNP-TYTDL-TPHYSQGRIQEAPKLTHL 422
ZebrafishA3a         DYNGAITMGNSHQQG--ACDSNPCTFTDL-TPHYSQGRIQEAPKLTHL 410
BichirA3             DYNGAGSMASNHQHG--PCEPHP-TYTDLSAHHPSQGRIQEAPKLTHL 415
                     :*.   .: ..: :     :. * *:***   * .************
```

Fig. S4. (Continued)

```
HornSharkA3          MQKATYYDSSAIFGGYTYQGANGFNYNASQQQYPPSSVESDYHRPACSLQSPTVPHHKPN 60
SkateA3              MQKATYYDSSAIFGGYTYQGANGFNYNANQQQYPPSSVENDYHRPACSLQSPTVPHHKPN 60
ElephantSharkA3      MQKATYYDSSAIFGGYTYQAANGFNYNASQQQYPPSSVESDYHRPACSLQSPTVPHHKPN 60
CoelacanthA3         MQKATYYDSSAIYGGYPYQGANGFTYNASQQQYPSSSVESDYHRPACSLQSPTVPHHKHN 60
ShrewA3              MQKATYYDSSAIYGGYPYQAANGFAYNANQQPYPAALADGEYHRPACSLQSPAGGHPKAH 60
HumanA3              MQKATYYDSSAIYGGYPYQAANGFAYNANQQPYPAALADGEYHRPACSLQSPAGGHPKAH 60
BatA3                MQKATYYDSSAIYGGYPYQAANGFAYNASQQPYPVALADGEYHRPACSLQSPAGGHPKAH 60
MouseA3              MQKATYYDSSAIYGGYPYQAANGFAYNASQQPYAAALTDGEYHRPACSLQSPAGGHPKTH 60
ChickenA3            MQKATYYDSSAIYGAYPYQGANGFTYNASQQQYPSSLVETEYHRPACSLQSPAVSHHKAN 60
LizardA3             MQKATYYDSSAIYGGYPYQGTNGFAYNASQQQYPSSSVETEYHRPACSLQSPGGGGGGAT 60
XenopusA3            MQKATYYDSSAIYGGYPYQGANGFTYNASQQQYPSSSLETEYHRPACSLQSPAVPHHKAN 60
BichirA3             MQKATYYDSSAIYGGYPYHSAKGFSYNVNQQQYPPSSVESDYHRPACSLQSSNAPLHKSN 60
MedakaA3a            MQKATYYDSSAIYSGYPYQSANGFSYDANQVQYPRASVESEYHRPACSLQTPSVALQKPG 60
CichlidA3a           MQKATYYDSSAIYSGYPYQSANGFSYDANQVQYPRASVESEYHRPACSLQSPSVALQKPG 60
FuguA3a              MQKATYYDSSAIYSGYPYQSANGFSYDANQIQYPRTSVESEYHRPACSLQSPSVALQKRE 60
SalmonA3ab           MQKATYYDSSAIYSGYPYQSANGFSYDANQVQYPRASVESEYHRPACSLQSPLVALQKPG 60
SalmonA3Aa           MQKATYYDSSAIYSGYPYQSANGFSYDANQVQYPRVSVESEYHRPACSLQSPSVALQKSG 60
ZebrafishA3a         MQKATYCDGSAIYSGLPYQSANGLGYDASQQQYLQALAESEYHRPACSLQSPSAGLHTSN 60
                     ****** *.***:.. .*:.::*: *:..*  *     : :*********:.

HornSharkA3          DINESCMRSQPSHHPVIAEQPPPPSPPQNTSNSTQSSTSNPTLTSTIKQIFPWMKESRQN 120
SkateA3              DINESCMRNQTSHLPVIAEQPPPPSPPQNTSNSTQSSTSNPTLTSTIKQIFPWMKESRQN 120
ElephantSharkA3      DINESCMRSQPSHPPVIPEQPPPPSPPQNTSNSTQSSTSNPTLTSTIKQIFPWMKESRQN 120
CoelacanthA3         DINESCMRNQPPQPPGISEHAPQPSPSQTSSNSTPSNSNNPGITSTIKQIFPWMKESRQN 120
ShrewA3              ELSEACLRGPPSQPPGLGEPVPPTSPPQNASNPAPASAASPLLNSTVKQIFPWMKESRQN 120
HumanA3              ELSEACLRAPPSQPPSLGEPAPPPSPPQNASNPTPANAASPLLNSTVKQIFPWMKESRQN 120
BatA3                ELSEACLHGPPSQPPGLGEPAPPPSPPQNASNPTPASAASPLLNSTVKQIFPWMKESRQN 120
MouseA3              ELSEACLRGPPSQPPGLGEPAPPPSPPQSANNPTPASTASPLLNSTVKQIFPWMKESRQN 120
ChickenA3            DISESCMRSQPLQPPGLTDPQPPPSPSQNASNPAPANSTSPALNSTVKQIFPWMKESRQN 120
LizardA3             DLNESCMRSQTLAPQVLPEQPPPPSPPPNAPTSAPANATSAILNSTMKQIFPWMKESRQN 120
XenopusA3            DINESCMRSQSNQAPVIPEQGPPPSPPQTTSNAATASSNATSITSTMKQIFPWMKESRQN 120
BichirA3             EITESCLQTQQKQPTIVPDNVAPPSPPENTSSNHSCTNSPHSSPSTTKQIYPWMKESRQN 120
MedakaA3a            EMAEGCDRIQAAQSKVIPENPPPSSPGAISQSNGSSQPGKNASPTARKQIFPWMKECRQN 120
CichlidA3a           EMAESCDRIQAAQSKVHPESPPPPSPGAISQSNGSNQPVKNDSPTARKQIFPWMKESRQN 120
FuguA3a              MAAENCDRVQAVQSKVHPESPPPPSPGAISQSNGSNQPTKNSSPTSRKQIFPWMKESRQA 120
SalmonA3ab           EMAENCDRIQAAQPPVLTESPPPPSPGSLNQSNGSCQPNVHSSPTTRKHIFPWMKESRQN 120
SalmonA3Aa           EIAESCDMIQAAQPPVLTDSPPPPSPGSLNQSNGSSQADVHGSPTTRKHIFPWMKESRQN 120
ZebrafishA3a         EMSEVCQQINGTQATVTDTSSGPSSPSSLNQNIDSAAKNVHVSPTTRKHIFPWMKESRQN 120
                        * *                 .**     .           :: *:*:*****.**

HornSharkA3          AKQKSSSSVESSAGEKSPPGPASAKGMLTSSGGQSPRSPIPSAAGGANSMHSLATSAPYD 180
SkateA3              TKQKSSSSVESSAGEKSPPGPASAKGMLTSSGGQSPRSPIPSAAGGANSMHPLATSAPYD 180
ElephantSharkA3      AKQKSSSSVESCAGEKSPPGPASAKGMLTSSGGQSPRSPIPPSVGLANSMHSLASNAPYD 180
CoelacanthA3         SKQKSSCSGDGCTGDKSPPGPASGKGMMTPSGGQSPRSPIPHTAGGLNSMHSLVNSVPYE 180
ShrewA3              TKQKSGSSGESCAGDKSPPGQASGKGVLTSSGGQSPRSPVPPGAGGLNSMHSLVNSVPYE 180
HumanA3              TKQKSSSSGESCAGDKSPPGQASGKGMLTSSGGQSPRSPVPPGAGGLNSMHSLVNSVPYE 180
BatA3                TKQKSGSSGESCAGDKSPPGQASGKGMLTSSGGQSPRSPVPPGAGGLNSMHSLVNSVPYE 180
MouseA3              TKQKSGSSGESCAGDKSPPGQASGKGMLTSSGGQSPRSPVPPGAGGLNSMHSLVNSVPYE 180
ChickenA3            TKQKSSSSGESCAGDKSPPGQTSGKGMMTSSGGQSPRSPVPPAAGGLNSMHSLVNSVPYE 180
LizardA3             TKQKSGSSGESCAGDKSPPGQPSGKGMMTSSGGQSPRSPVPSAPGGLNSMHSLVNSVPYE 180
XenopusA3            TKQQAGSSGESCAGDKSPPGQSSGKSMMTSSGGQSPRSPVPPSVGGLNSMHSLVNSVPYE 180
BichirA3             SKPKNGGPAESCADEKSPTGSAAGKGMMSSPGGQSPRSPVAGTAGGLNSMHSLVNSVSYE 180
MedakaA3a            TKQKASSSVESCPGDKSPPGSAAGVGMMSSPGGQSPRSPVGASGGGLNSMHSLVNSVPYE 180
CichlidA3a           TKQKTSSSMESCPGDKSPPGSAAGVGMMPSPGGQSPRSPVGASGAGLNSMHSLVNSVPYE 180
FuguA3a              TKQKTSNTVESCPGDKSPPGSAAGAGMMPSPGGQSPRSPVGGSTGALNSMHSLVNSVPFE 180
SalmonA3ab           QKPKTSSSVESCPGDKSPPGSAAGCGMMPSPGGQSPRSPLGSPGGGLNSMHSLVNSVPYE 180
SalmonA3Aa           QKPKTSSSVESCQGDKSPPGSAAGCGMMPSPGGQSPRSPLGSPSSGLNSMHSLVNSVPYE 180
ZebrafishA3a         TKQKSCSIVESCAGDKSPPGSAAGLGMMPSPGAQSPHSPVSSSGGGLSSMHSLVNSVTYD 180
                      * :     :.. .:***.* .:. .::...*.***:**:    .  .***.*.....::
```

Fig. S5. Amino acid alignment that was used for generating the phylogenetic tree (Fig. S4B) and for calculating relative rates (Table S4). The alignment from Fig. S5 has been manually edited to remove the homeodomain, to eliminate gaps, and to maximize homology. Asterisk indicates identical residues; colon indicates conserved substitutions; period indicates semiconserved substitutions.

```
HornSharkA3          PHSPTSFSKPHQNAYAIPTSYPGLNCPPPQKRYAGTAVTPEYDTHPLQGNGYGTPHQGSP 240
SkateA3              PHSPTSFGKPHQNAYVIPTSYPGLNCPPLQKRYAGTAVTPEYDTHPLQGNGYGTPHQGSP 240
ElephantSharkA3      QHSPTSFNKPHQNTYGIPTSYPALNCPPPQKRYAGTAVTPEYETHPLQGNGYGTPHQGSP 240
CoelacanthA3         PQSPPSFGKPHQNAYGLSTSYPALNCPPPQKRYTGTAVTPEYDTHPLQGNGYGNPHQASP 240
ShrewA3              PQSPPPFSKPPQGAYGLPASYPALPCAPPQKRYTAAGGTPDYDPHALQGNGYGTPHQGSP 240
HumanA3              PQSPPPFSKPPQGTYGLPASYPALPCAPPQKRYTAAGGTPDYDPHALQGNGYGTPHQGSP 240
BatA3                PQSPPSFSKPPQGAYGLPASYPALPCAPPQKRYTTAGGTPDYDPHALQGNGYGTPHQGSP 240
MouseA3              PQSPPPFSKPPQGAYGLPASYPALPCAPPQKRYTAAGGTPDYDPHALQGNGYGTPHQGSP 240
ChickenA3            PQSPPPFNKPHQNTYGIPASYTALNCPPPQKRYTGTAVTPEYDTHPLQGNGYGNPHQGSP 240
LizardA3             PQSPPSFNKPHQNAYGIPTSYPALNCPPPQKRYTGTAVTPEYDPHSLQGNGYGNPHQGSP 240
XenopusA3            PQSPPAFNKHHPSAYGVPAPYPSHNCPPHQKRYSGTAVTPEYEPHPQQSSGYGNPHQGSP 240
BichirA3             PQSPTSYNKPHQNSYGMSTSYPALNCPPSQKRYSGTSVTPEYETHSLQGNGYGTAHQTSP 240
MedakaA3a            SQSPTSYSKPHQNAYAMPTSYPPLNCPPSQKRYPGTSATPEYDAHPLQGSGYGTHMQGSP 240
CichlidA3a           SQSPNSYSKPHQNAYGMPTSYPPLNCPPSQKRYPGNSATPEYDAHPLQGNGYGTHMQGSP 240
FuguA3a              SQSPTSYNKPHQNAYGMATSYPPHNCPPTQKRYAGTSATPEYDAHPLQGNGYGTHMQGSP 240
SalmonA3ab           SQSPTSYNKPAHNAYGMPTSYPPLNCPPSQKRYPGTSATPEYDAHPLQGNGYGTHLQVSP 240
SalmonA3Aa           SQSPTSYNKPSHNAYGMPTSYPPLNCPPSQKRYSGTSPTPEYDAHPLQGNGYGAHLQGSP 240
ZebrafishA3a         SPSPASYNKPQSNTYSLPTSYPPLNCPPPQKRYQGTTATPEYDTHHIQGNNYGTQVQGSP 240
                       ** .:.*   .:* :.:.*.   *.* ****     **:*:.*  *...**   * **

HornSharkA3          VYGNFVESMPSSGPLFSLHLGHPPSNMDYNGPMTSNHHHGPCDPHPTYTDLSHHPSQGRI 300
SkateA3              VYGNFVETMPSSGPLFSLHLGHPPSNMDYNGPMTSNHHHGPCDPHPTYTDLSHHPSQGRI 300
ElephantSharkA3      VYGNFVESMSSSGPLFSLHLGHPPSNMDYSGPMTSNHHHGPCDPHPTYTDLSHHPSQGRI 300
CoelacanthA3         VYGNYVETMGNSGSIFGLHLPHPSTNMDYSGPMASNHHHGPCDPHPTYTDLSHHPSQGRI 300
ShrewA3              VFGSYVEPMSNSGPLFGLHLPHAASAMDYGGPLGSGHHHGPGEPHPTYTDLAHHPSQGRI 300
HumanA3              VFGSYVEPMSNSGPLFGLHLPHAASAMDYGGPLGSGHHHGPGEPHPTYTDLGHHPSQGRI 300
BatA3                VFGSYVEPMSNSGPLFGLHLPHAASAMDYGGPLGSGHHHGPGEPHPTYTDLAHHPSQGRI 300
MouseA3              VFGSYVEPMSNSGPLFGLHLPHTTSAMDYGGPLGSGHHHGPGEPHPTYTDLAHHPSQGRI 300
ChickenA3            VYGNYVETMTNSGPIFGLHLSHPPSNMDYSGPMGNNHHHGPCDPHPTYTDLAHHPSQGRI 300
LizardA3             VYSNYVDTMTNSGPIFGLHLPHPSANMDYSGPMGNNHPHGPCDPHPTYTDLSHHASQGRI 300
XenopusA3            VYGNYVETMTNSGPMFGLHLSHSSSNMDYSGPMNSGHHHGPCDSHPTYTDLAHHNPQGRI 300
BichirA3             VYGSYVDSIANSGPIFGLHLPHPSSNMDYNGSMASNHQHGPCEPHPTYTDLAHHPSQGRI 300
MedakaA3a            VYGGYIDSMPSSGAVFGLHLSHPPSNMEYNITMGNSQHHGVCDPTPTYTDLTPHYSQGRI 300
CichlidA3a           VYGGYIDSMPSSGAVFGLHLTHPPSNMEYNITMGNSQHHGVCDPTPTYTDLTPHYSQGRI 300
FuguA3a              VYGGYIDSVPNSGSVFGLHLPHPPSNMDYNITMGNSQHHGVCDPTPTYTDLTSHYSQGRI 300
SalmonA3ab           VYGGYIDSMPNNRAVFGLHLSHPPSNMDYNITMGNSHHQRVCDPNPTYTDLTPHYSQGRI 300
SalmonA3Aa           VYGGYIDSMPNNGSIFGLHLSHPPSHMDYNITMGNSHHHGLCDPNPTYTDLTPHYSQGRI 300
ZebrafishA3a         VYGGYSDSLVGMGAVFGLHLPHPSQNIDYNITMGNSHQQGACDSNPTFTDLTPHYSQGRI 300
                     *:..: :.: .  .:*.*** *..  ::*. .: ..: :   :. **:***  * .****

HornSharkA3          QEAPKLTHL 309
SkateA3              QEAPKLTHL 309
ElephantSharkA3      QEAPKLTHL 309
CoelacanthA3         QEAPKLTHL 309
ShrewA3              QEAPKLTHL 309
HumanA3              QEAPKLTHL 309
BatA3                QEAPKLTHL 309
MouseA3              QEAPKLTHL 309
ChickenA3            QEAPKLTHL 309
LizardA3             QEAPKLTHL 309
XenopusA3            QEAPKLTHL 309
BichirA3             QEAPKLTHL 309
MedakaA3a            QEAPKLTHL 309
CichlidA3a           QEAPKLTHL 309
FuguA3a              QEAPKLTHL 309
SalmonA3ab           QEAPKLTHL 309
SalmonA3Aa           QEAPKLTHL 309
ZebrafishA3a         QEAPKLTHL 309
                     *********
```

Fig. S5. (Continued)

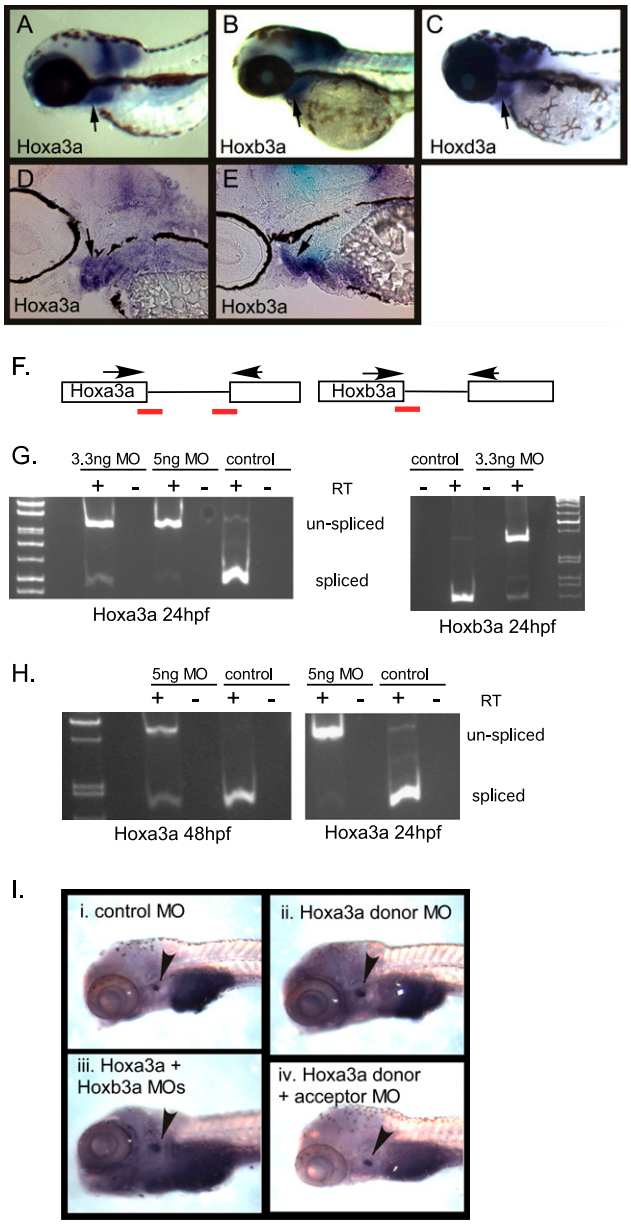

Fig. S6. Expression and functional analysis of zebrafish Hox3 genes. (*A*–*C*) Expression of PG3 Hox genes in 72 hpf zebrafish embryos. Whole-mount in situ hybridization with 72 hpf embryos showing the expression patterns of Hoxa3a (*A*), Hoxb3a (*B*), and Hoxd3a (*C*). These three PG3 Hox genes were expressed in similar patterns. Arrows indicate expression in the pharyngeal arch region. Sagittal sections of Hoxa3a (*D*) and Hoxb3a (*E*) whole-mount in situ hybridization embryos. Hoxa3a and Hoxb3a were both detected to be expressed in the pharyngeal arch mesenchyme and pouch endoderm. Arrows indicate the third pouch. Anterior is to the left, dorsal is up. (*F*–*I*) Early knockdown of Hoxa3a and Hoxb3a by the injection of MOs does not affect later thymus development in zebrafish. (*F*) Scheme of prespliced mRNA of Hoxa3a and Hoxb3a. The red bars indicate the MO specific to the splicing donor or acceptor site. The arrows represent the primers used in RT-PCR. (*G*) RT-PCR products were run on a 1.0% agarose gel; 3.3 ng/embryo or 5 ng/embryo of Hoxa3a MOs (donor and receptor) or mismatched MO were injected into one-cell stage embryos. At 24 hpf, the embryos were collected and analyzed for mRNA splicing. On the gel, the lower bands represent the RT-PCR product of postspliced mRNA and the higher bands represent the product of prespliced mRNA. In the MO-injected embryos, most of the mRNA was prespliced, indicating that the MOs suppressed splicing of target genes effectively at this time point. Hoxa3a (*Left*) and Hoxb3a (*Right*) are shown. RT+, PCR with first-strand cDNA; RT–, PCR with non-RT control. (*H*) Hoxa3a RNA splicing in 24 hpf and 48 hpf morphants. At 24 hpf, there was no postspliced mRNA product in the morphants, but at 48 hpf, the postspliced mRNA product was restored. (I) Whole-mount in situ hybridization analysis on Rag-1 expression in the Hoxa3a/Hoxb3a morphants at 1 week. (*i*) Rag-1 expression was detected in the embryos injected with mismatched MO. (*ii*) Embryos injected with 5 ng/embryo of Hoxa3 MO specific to splicing donor site showed normal Rag-1 expression. (*iii*) A total of 5 ng of MOs that were specific to Hoxa3a splicing donor site and acceptor site were coinjected into the embryos. Rag-1 expression in these embryos was not changed. (*iv*) A total of 3.3 ng/embryos of Hoxb3a MO (splicing donor site) was coinjected with Hoxa3a MOs. Rag-1 expression in these morphants was normal compared with control. Arrowheads indicate Rag-1 expressing cells.

Table S1. Rescue of bloated abdomen and truncated soft palate phenotype by *Hoxa3*^{zf} allele

Genotype	Bloated abdomen, %	Truncated soft palate, %
Hoxa3$^{+/+}$	0 (0/20)	0 (0/3)
Hoxa3$^{+/-}$	0 (0/10)	0 (0/3)
Hoxa3$^{+/zf}$	0 (0/22)	0 (0/3)
Hoxa3$^{-/-}$	75% (6/8)	100% (3/3)
Hoxa3$^{zf/zf}$	0 (0/15)	0 (0/3)
Hoxa3$^{zf/-}$	83% (5/6)	100% (3/3)
Hoxa3$^{mz/mz}$	0 (0/10)	0 (0/3)
Hoxa3$^{mz/-}$	0 (0/8)	0 (0/3)

Newborn animals of different genotypes were checked for the bloated abdomen and the soft palate phenotype. Numbers in parentheses are the number of animals showing the phenotype per total number of individuals examined.

Table S2. Summary of embryonic neurofilament analysis

Characteristic	$^{+/+}$	zf/zf	mz/mz
Unconnected*	0	14% (2)	22% (2)
Unconnected/fused†	0	7% (1)	11% (1)
Fused/fused	0	21% (3)	11% (1)
Fused/normal	20% (3)	43% (6)	44% (4)
Normal/normal	80% (12)	14% (2)	11% (1)
Total number affected	20% (3)	86% (12)	89% (8)
Total number unaffected	80% (12)	14% (2)	11% (1)
Total number analyzed	15	14	9

Numbers in parentheses are the total numbers observed.
*The IX cranial nerve was unconnected to the hindbrain, including both unilateral and bilateral phenotyeps.
†IX nerve on one side and IX nerve fused to X nerve on the other side.

Table S3. Relative rate tests of Hoxa3 among different vertebrate taxa

Outgroup	Ingroup 1	Ingroup 2	*P* Value
Elephant shark	Mouse	Chicken	0.00051†
Elephant shark	Mouse	Coelacanth	0.00029†
Elephant shark	Mouse	*Xenopus*	0.07505
Elephant shark	Mouse	Bichir	0.74773
Elephant shark	Mouse	Fugu	0.04401*
Elephant shark	Mouse	Zebrafish	0.00005†
Elephant shark	Bichir	Zebrafish	0.00005†
Elephant shark	Bichir	Fugu	0.04550*
Elephant shark	Fugu	Zebrafish	0.00753†
Elephant shark	Coelacanth	Bichir	0.00006†
Elephant shark	Coelacanth	Zebrafish	0.00000†
Elephant shark	*Xenopus*	Bichir	0.04417*
Elephant shark	*Xenopus*	Zebrafish	0.00000†

Tajima relative rate tests were carried out on aligned sequences in Fig. S6 using the MEGA 4 package (1). The elephant shark Hoxa3 sequence was used as the outgroup. Only results from salient comparisons are given in the table.
*Significant at $P < 0.05$;
†highly significant at $P < 0.01$. Note all comparisons involving zebrafish are highly significant.

1. Kumar S, Nei M, Dudley J, Tamura K (2008) MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. *Brief Bioinform* 9:299–306.