Homeobox transcription factor Six7 governs expression of green opsin genes in zebrafish

Yohey Ogawa, Tomoya Shiraki, Daisuke Kojima and Yoshitaka Fukada

Supplementary information

Supplementary methods

(1) Construction of TALEN targeting-vector

For construction of pTAL7DD and pTAL7RR vectors, a sequence encoding the TAL effector and the *Fok*I nuclease catalytic domain was amplified from the pTAL3 by PCR and then cloned between *BamH*I and *EcoR*I site of pCS2(+) vector. The N-terminal and C-terminal regions of the TAL effector were truncated based on pCS2TAL3DD and pCS2TAL3RR [1]. The sequence encoding the FLAG tag was inserted into the *N*-terminus of TAL effector in order to confirm the expression of TALENs by Western blot analysis. The resulting pTAL7DR was further modified by PCR for generation of pTAL7RR and pTAL7DD, the nucleotide sequence of which is shown in Supplementary material 1. A pair of pTAL7DD and pTAL7RR express a DD-RR heterodimeric form of the *Fok*I nuclease catalytic domain, and this form was previously reported to have reduced homodimeric cleavage activity, thereby decreasing toxicity of zinc finger nucleases [2].

(2) Generation of mutant zebrafish

To generate *six7* mutant zebrafish, TALEN mRNA was synthesized from the pTAL7DD and pTAL7RR by *in vitro* transcription using the SP6 mMESSAGE mMACHINE Kit (AM13140, Ambion), and purified with the RNeasy Mini kit (74104, Qiagen). Aqueous solution containing ~200 pg each of the two TALEN mRNAs and 0.05% phenol red were injected into the cytoplasm of the one cell-stage zebrafish embryos. Fish (F_0) obtained from the injected animals were crossed with wild-type zebrafish, and the resulting F_1 fish were screened for the presence of TALEN-induced mutations by a combination of PCR and subsequent digestion by endonuclease *Xho*I: first, the region surrounding the TALEN cleavage site was amplified by PCR using the primers, six7TALEN_Fw and six7TALEN_Rv (Supplementary table 1). Second, the PCR product was digested by *Xho*I, and TALEN-induced mutations were examined by sensitivities to the *Xho*I digestion (figure 3a). To sequence-verify mutations, the PCR products having no *Xho*I site were then subjected to the second round of PCR using the other primer set, DS_six7TALEN_Fw and DS_six7TALEN_Rv. The nested PCR products were digested by *Xho*I again, and then sequenced directly. After verification of the mutations, we used zebrafish having the mutant allele (*six7*^{ja51}) described in figure 3a for all the experiment except for figure S1. Genotypes were determined by PCR with two pairs of specific primers, (i) DS_six7TALEN_Fw and DS_six7TALEN_Rv, (ii) six7TALEN_Fw2WT and six7TALEN_Rv3mut (Supplementary table 1) (figure 3b). Before the experiment in figure 3, the larvae were dissected into anterior and posterior segments; the posterior ones were used for the genotyping, while the anterior segments were soaked in RNA*later*

(R0901, Sigma) for extracting RNA or soaked in 4% paraformaldehyde in Ca^{2+}- and Mg^{2+}-free Dulbecco's PBS (D-PBS) for preparing cryosections.

The *six3b* mutant zebrafish was generated as described above for the *six7* mutant. Briefly, the TAL effector repeats recognizing exon 1 of zebrafish *six3b* gene were constructed with the Golden Gate method. These repeats were cloned into pTAL7DD and pTAL7RR vectors. Aqueous solution containing ~300-pg each of the two TALEN mRNAs and 0.05% phenol red were injected into the cytoplasm of the one cell-stage zebrafish embryos. Fish (F_0) obtained from the injected animals were crossed with wild-type zebrafish, and the resultant F_1 fish were screened for the presence of TALEN-induced mutations by a combination of PCR and subsequent digestion by endonuclease *Hind*III. After the isolation of the mutant zebrafish and verifying mutations in the *six3b* gene locus, genotypes were determined by a combination of PCR and subsequent digestion by endonuclease *Hind*III (figure S1d). First, the region flanking the TALEN cleavage site was amplified by PCR using the following primers, six3bTALEN_Fw (5'-GACACCAGGAAGACCAATAG -3') and six3bTALEN_Rv (5'-TACCACTCCCTCAACAAACC -3'). Second, the PCR product was digested with *Hind*III, and TALEN-induced mutations were checked by loss of the *Hind*III digestion site. The obtained *six3b*$^{ja53/+}$ fish were crossed with *six7*$^{ja52/+}$ fish to generate *six7*$^{ja52/+}$;*six3b*$^{ja53/+}$ fish. Then, the resulting fish were crossed with each other to generate *six7* and/or *six3b* knock-out zebrafish, which were used for the experiments performed in figure S1e.

(3) Immunohistochemistry

The adult eyes were isolated from the light-adapted adult zebrafish. After removal of the lenses, the eyes were fixed in 4% paraformaldehyde in D-PBS overnight at 4°C. After several washes with D-PBS, the fixed eyes were immersed successively in 5%, 10% and 15% sucrose (w/v) in D-PBS for 30 min each at room temperature and finally immersed in 20% (w/v) sucrose in D-PBS overnight at 4°C. The embryos were fixed in 4% paraformaldehyde in D-PBS overnight at 4°C and immersed in 20% sucrose (w/v) in D-PBS overnight at 4°C. The cryoprotected samples of the adult eyes and embryos were embedded in a 2:1 solution of 20% sucrose in D-PBS and Tissue-Tek O.C.T Compound (Tissue-Tek, SAKURA), frozen, and sectioned at 10-μm thickness with cryostat (CM3000, Leica). The sections on the glass slides were rinsed with PBS (10 mM Na-phosphate buffer, 140 mM NaCl, 1 mM $MgCl_2$, pH 7.4), pretreated with a blocking solution (3% goat normal serum, 0.1% Triton X-100 in PBS) for 1 hr at room temperature, and then incubated with a primary antibody diluted in the blocking solution overnight at 4°C. After washed with PBS, the sections were treated with the blocking solution for 15 min at room temperature, and then with a secondary antibody for 4 hr at room temperature. After washed with PBS, the sections were coverslipped with VECTASHIELD Mounting Medium with DAPI (H-1200, Vector Laboratories) for staining of the cell nuclei. The primary antibodies used were mouse monoclonal antibody Zpr1 (diluted 1:400, Zebrafish International Resource Center, Eugene) directed to double cones (green and red cones), rabbit polyclonal antibody to rod transducin α-subunit, Gnat1 (diluted to 0.5 μg/ml, sc-389, Santa Cruz), mouse polyclonal antibody to rhodopsin (diluted

1:400, self-produced antibody) and rabbit polyclonal antibody to cone transducin α-subunit, Gnat2 (diluted 1:800, PM075, MBL). The secondary antibodies used were goat anti-mouse IgG antibody conjugated with Alexa-568 (A-11004, diluted to 2 μg/ml, Molecular Probes) and goat anti-rabbit IgG antibody conjugated with Alexa-488 (A-11034, diluted to 2 μg/ml, Molecular Probes). The immuno-stained sections were observed with a confocal laser scanning microscope (TCS SP8, Leica).

Whole-mount immunohistochemistry was conducted as described previously [3] with some modifications. The adult retinas were isolated from the dark-adapted adult zebrafish. The isolated retinas were fixed in 4% paraformaldehyde in D-PBS for 3 hr at 4°C. After several washes with 0.1% Triton-X100 in PBS (PBST2), the specimens were treated with 0.5% Triton-X100 in PBS for 30 min at room temperature, and subsequently washed with PBST2. The specimens were incubated in a blocking solution (2% goat normal serum, 1% Dimethyl sulfoxide in PBST2) for 1 hr at room temperature, and then with a primary antibody diluted in the blocking solution overnight at 4°C. Following several washes with PBST2, the samples were treated with the blocking solution for 1 hr at room temperature, and then with a secondary antibody for 4 hr at room temperature. After washed with PBST2, the treated retinas were mounted on a grass slide and coverslipped with VECTASHIELD Mounting Medium with DAPI (H-1200, Vector Laboratories) for staining of the cell nuclei. The primary antibodies used were mouse monoclonal antibody 1D4 (diluted to 7 μg/ml), which specifically recognized red cone opsins in zebrafish, and rabbit polyclonal antibody to cone transducin α-subunit, Gnat2 (diluted 1:400, PM075, MBL). The

secondary antibodies used were goat anti-mouse IgG antibody conjugated with Alexa-568 (A-11004, diluted to 4 μg/ml, Molecular Probes) and goat anti-rabbit IgG antibody conjugated with Alexa-488 (A-11034, diluted to 4 μg/ml, Molecular Probes). The immuno-stained retinas were observed with a confocal laser scanning microscope (TCS SP8, Leica).

(4) *In situ* hybridization

To prepare RNA probes, the DNA fragments of zebrafish *six7* and various visual opsin genes were amplified by PCR from adult zebrafish retinal cDNAs with primers listed in Supplementary table 2. To distinguish different types of duplicated green and red opsin genes, RNA probes were designed to recognize each 3' UTR region of duplicated opsin genes according to the previous study [4]. The amplified cDNA fragments were cloned into pGEM15H vector [5] or into pCR4Blunt-TOPO (K2875-20SP, Life Technologies). The resultant plasmids were linearized with appropriate restriction enzymes, and then transcribed into digoxigenin (DIG)-labelled RNA with DIG RNA labelling kit SP6/T7 (1175025, Roche) and/or with T3 RNA polymerase (1163001, Roche). The *lws* cRNA probe used in figure 3f was designed in the coding region of *lws1* to visualize *lws1* and *lws2* mRNAs.

Whole-mount *in situ* hybridization was performed as described previously [3] with some modifications. Briefly, larval zebrafish were fixed in 4% paraformaldehyde in D-PBS overnight at 4°C, dehydrated through a methanol series and stored in 100% methanol at -20°C. The samples were rehydrated and pre-treated with proteinase K (5 μg/ml in PBST, 0.1% Tween 20 in D-PBS) for 12

min [for the larvae at 1.5 days post fertilization (dpf)], for 20 min (for the 2-dpf and 2.5-dpf larvae) or for 30 min (for the 3.5-dpf larvae). The larvae were hybridized with the DIG-labelled RNA probes, and the hybridization signals were visualized by NBT/BCIP staining. Hybridization and post-hybridization washes were performed at 65°C. After staining, the specimens were washed sequentially with 30% and 75% methanol in PBST for 5 min each and washed twice in 100% methanol. The staining reaction was terminated by fixation in 4% paraformaldehyde in PBST. Then, the specimens were immersed sequentially in 30%, 50% and 70% glycerol in PBST for 5 min each, and subjected to examination with a stereoscopic microscope (MZ-FL3, Leica).

In situ hybridization using larval and adult ocular sections was carried out essentially as described previously [5,6][6,5]. In short, the 10-μm frozen ocular sections were prepared in a similar manner to immunohistochemistry. These sections were pre-treated with proteinase K and hybridized with the DIG-labelled RNA probes, and the hybridization signals were visualized by NBT/BCIP staining. Hybridization and post-hybridization washes were performed at 65°C. The treated samples were examined with a microscope, Axioplan2 (Carl Zeiss) or BZ-9000 Generation II (Keyence).

(5) Isolation of rods and cones

Retinas were dissected in ice-cold PBS from dark-adapted adult fish of *Tg*(*rho:egfp*)^{ja2} [7] or *Tg*(*gnat2:egfp*)^{ja23} transgenic lines. *Tg*(*rho:egfp*)^{ja2} fish express enhanced green fluorescent proteins (EGFP) in rods, while *Tg*(*gnat2:egfp*)^{ja23} express EGFP in all the cone subtypes. The *Tg*(*gnat2:egfp*)^{ja23}

line was generated according to the method described in the previous study [8]. The dissected retinas were digested with 0.25% trypsin, 10 U/ml DNaseI, 2 mM $MgCl_2$ and 2 mM EGTA in Ca^{2+}-free Ringer's solution (5 mM HEPES, 116 mM NaCl, 2.9 mM KCl, pH7.2) for 30 min at 37°C. The reaction was terminated by the addition of an equal volume of PBS containing 0.5% soybean trypsin inhibitor, 4 mM EDTA and 20% (v/v) fetal bovine serum. The sample was subjected to filtration through a 35-μm nylon mesh (352235, Falcon) and then to a fluorescence activating cell sorter (FACSAria, BD Biosciences) to isolate EGFP-positive rods or cones. The cell sorting was conducted with gating based on the three parameters; forward scatter, side scatter and green fluorescence. The sorted cells containing EGFP-positive rods or cones were directly collected into 1.5 ml microtubes with the TRIzol reagent (15596018, Life Technologies) for subsequent extraction of the total RNA. The extracted total RNA was purified and concentrated using the RNeasy MinElute Cleanup Kit (74204, Qiagen), and then reverse-transcribed with SuperScript II (18064, LifeTechnologies) by using anchored $(dT)_{16}$ primers. The cell purity was characterised by examining the expression levels of rod- and cone-specific genes by quantitative PCR (figure2a).

(6) Sequence alignment and phylogenetic analysis

Amino acid sequences of the *sine oculis* homeobox (Six) protein family members and the green cone opsins (Rh2) were retrieved from the Ensembl genome browser (http://www.ensembl.org) from a representative selection of the vertebrate and invertebrate clades, which were listed in Supplementary table 4 and Supplementary table 5. Accession numbers are also listed in the

Supplementary tables. The amino acid sequences of SIX protein family members of elephant shark were obtained from the database of Elephant Shark Genome Project (http://esharkgenome.imcb.a-star.edu.sg/). *SIX7* candidates were identified in the NCBI genome database of green sea turtle and Indian python by tBLASTn searches using the sequence of green anole SIX7 protein. SIX domain and homeodomain amino acid sequences (173, 143 amino acids for figure 1, S6a, respectively) were used for phylogenetic tree construction of the Six protein family. Full-length amino acid sequences without N- and C- terminal regions (about 310 amino acids) were used for phylogenetic tree construction of the teleost green opsins. The phylogenetic tree was constructed using the neighbor-joining method with 1,000 bootstrap replicates (CLC Main Workbench 6). Nematode CEH-34 was employed as an outgroup for Six family members, while spotted gar Rh2 was used as an outgroup for teleost green cone opsins.

To compare the synteny around the green cone opsin genes among vertebrate species, the adjacent genes were mapped onto the synteny regions by using the Ensembl genome browser.

References

1. Dahlem, T. J., Hoshijima, K., Jurynec, M. J., Gunther, D., Starker, C. G., Locke, A. S., Weis, A. M., Voytas, D. F. & Grunwald, D. J. 2012 Simple methods for generating and detecting locus-specific mutations induced with TALENs in the zebrafish genome. *PLoS Genet.* **8**, e1002861. (doi:10.1371/journal.pgen.1002861)

2. Szczepek, M., Brondani, V., Büchel, J., Serrano, L., Segal, D. J. & Cathomen, T. 2007 Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nat. Biotechnol.* **25**, 786–93. (doi:10.1038/nbt1317)

3. Kojima, D., Torii, M., Fukada, Y. & Dowling, J. E. 2008 Differential expression of duplicated VAL-opsin genes in the developing zebrafish. *J. Neurochem.* **104**, 1364–71. (doi:10.1111/j.1471-4159.2007.05093.x)

4. Takechi, M. & Kawamura, S. 2005 Temporal and spatial changes in the expression pattern of multiple red and green subtype opsin genes during zebrafish development. *J. Exp. Biol.* **208**, 1337–45. (doi:10.1242/jeb.01532)

5. Mano, H., Kojima, D. & Fukada, Y. 1999 Exo-rhodopsin: a novel rhodopsin expressed in the zebrafish pineal gland. *Brain Res. Mol. Brain Res.* **73**, 110–8.

6. Barthel, L. K. & Raymond, P. A. 2000 In situ hybridization studies of retinal neurons. *Methods Enzymol.* **316**, 579–90.

7. Asaoka, Y., Mano, H., Kojima, D. & Fukada, Y. 2002 Pineal expression-promoting element (PIPE), a cis-acting element, directs pineal-specific gene expression in zebrafish. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 15456–61. (doi:10.1073/pnas.232444199)

8. Kennedy, B. N., Alvarez, Y., Brockerhoff, S. E., Stearns, G. W., Sapetto-Rebow, B., Taylor, M. R. & Hurley, J. B. 2007 Identification of a zebrafish cone photoreceptor-specific promoter and genetic rescue of achromatopsia in the nof mutant. *Invest. Ophthalmol. Vis. Sci.* **48**, 522–9. (doi:10.1167/iovs.06-0975)

9. Amores, A., Catchen, J., Ferrara, A., Fontenot, Q. & Postlethwait, J. H. 2011 Genome evolution and meiotic maps by massively parallel DNA sequencing: Spotted gar, an outgroup for the teleost genome duplication. *Genetics* **188**, 799–808. (doi:10.1534/genetics.111.127324)

Supplementary table 1.

Primers used for genotyping of *six7* knock-out zebrafish.

Primer Name	Primer sequence (5' to 3')
six7TALEN_Fw	TCTGCTGCTCCTCCTTTAC
six7TALEN_Rv	CTTACGGATGCGGTACTTC
DS_six7TALEN_Fw	CCGTTGGTTGTCCGTTACTC
DS_six7TALEN_Rv	GCAGTTTAGCGTGAGATTCG
six7TALEN_Fw2WT	CCCGAGTGTGCGAGAATC
six7TALEN_Rv3mut	GCTCGATATCTCCTGTTTTCC

Supplementary table 2.

PCR primers used to amplify the target gene fragments for preparing cRNA probes.

Target gene	Forward (5' to 3')	Reverse (5' to 3')
sws1	CCTCAGAGGTTTCCAGCAAG	TTGGACAGGAGCAGACAGTG
sws2	GGGAGGTGACAAAGATGGTG	ATGTTCAGCAAGCCAAGACC
rh2-1	ATCTGCCTATTCAGTGCTCC	TGGAATGCATGCTTTTATTTAA
rh2-2	AACTAATTGCCAACTATGCTTTC	TTTGAATGCATGTATTTATTATTCTAC
rh2-3	ACTTTTTGAACTTCTTTGTGG	CACTTTTAATGCCAGTATTTATTC
rh2-4	TCTTTTGAACTTCTTTACAGGTTATG	TTTTTAATGCCGGTATTTATTTC
lws1	AGTCAGACATGGGGAAAAAG	TACATGGGCAGGCATCTAC
lws2	TGAACAAGAGGGAAGAACTG	GTACAACATATCATTTTGCTGC
six7	TTGCTTATCTTTCCTTTAAT	ACATCTAGCCTCTGTTTCCAAC

Supplementary table 3.

Primers used for RT-qPCR.

Target gene	Forward (5' to 3')	Reverse (5' to 3')
sws1	CGAGAGATATGTGGTCATCTG	TGTATCTGCTCCATCCAAAG
sws2	GGAGGAATGGTGAGTTTGTG	GGTCTTGAAGGTAAAGTTCC
rh2-1	GTAATGGAGGGATTCTTCGC	TGGTCCGCAAGAGGTTTG
rh2-2	GCGTGGGTAGATTAGTTGTG	GGCTTATGCTCAGATTTAGTGG
rh2-3	AGGAAACAAAAACAGCATTG	ATACAGTATAAATGACAGCCTTC
rh2-4	CTGAGAAATCTGGGCATCTG	TGGCGTGTAGAAAATATAAC
lws1	TACCTGGCCATCCATGCTG	GGCGAAGAACGTGTAAGGAC
lws2	ATGCCTGCCTACTTTGCC	CCAGTTCTTCCCTCTTGTTC
rho	TCCGAGACCACACAGCG	CTGCTTGTTCATGCAGATG
gnat1	AGATGCCAGGACCGTCAAAC	CAAGGGAGTAACCATCTTTGTG
gnat2	ATGCCGATAAGGAAGCCAAG	ATAACCACCTTGATGGAGAATC
arr3a	CTGATAAGCCTGTTCTTCTG	GTGGTTTGGTCAACTGTAAC
arr3b	GCAAGGCGTTGTTGTTTCC	ATGTGAGGCTTCCCAAGAGTC
nr2e3	TTCAGAGACCAGGTGATTC	CAGTTGTCCAGAGGTAGAGAC
nrl	CTACAAACGTCTCCAGCACC	AAGGCGAGGTGTTAGGATG
crx	CATAACTGGAGGGGAATCTG	AAAGCACGACACAAGAACTC
otx5	GAAGAAGTGGCTCTCAAGATC	AAACCAGACCTGGACTCTAGAC
tbx2b	CGACTCTGATGCTTCCTCAAG	GTCCATTTCGTGCTCGCTATC
six7	CTGCAGGACCCTTATCCAAAC	CATTCAGGAGAACTTCCAGAC
thrb	AACTTGGACGATTCAGAGGTG	TGAGCCACCTTGTGCTTACG
actb2	GGCAATGAGAGGTTCAGGTG	GTGGTACCACCAGACAATAC

Supplementary table 4.

Accession numbers used for alignment of the members of Six family.

Vertebrate species	Gene	Accession number
human (*Homo sapiens*)	*SIX1*	ENSP00000247182
	SIX2	ENSP00000304502
	SIX3	ENSP00000260653
	SIX4	ENSP00000216513
	SIX5	ENSP00000316842
	SIX6	ENSP00000328596
Chinese soft-shell turtle (*Pelodiscus sinensis*)	*SIX2*	ENSPSIP00000009496
	SIX3	ENSPSIP00000009008
	SIX6	ENSPSIP00000016433
	SIX7	ENSPSIP00000016631
anole lizard (*Anolis carolinensis*)	*SIX7*	ENSACAP00000021004
green sea turtle (*Chelonia mydas*)	*SIX7*	XP_007057958.1
Indian python (*Python molurus*)	*SIX7*	XP_007426202.1
coelacanth (*Latimeria chalumnae*)	*SIX1*	ENSLACP00000021104
	SIX2	ENSLACP00000015908
	SIX3	ENSLACP00000017069
	SIX4	ENSLACP00000021088
	SIX5	ENSLACP00000014916
	SIX6	ENSLACP00000021143
zebrafish (*Danio rerio*)	*six1a*	ENSDARP00000057431
	six1b	ENSDARP00000028379
	six2a	ENSDARP00000075251
	six2b	ENSDARP00000071623
	six3a	ENSDARP00000075254
	six3b	ENSDARP00000071625
	six4a	ENSDARP00000126511
	six4b	ENSDARP00000104361
	six4.3	ENSDARP00000123087
	six6a	ENSDARP00000098430
	six6b	ENSDARP00000110487
	six7	ENSDARP00000093395
medaka (*Oryzias latipes*)	*six1b*	ENSORLP00000019739
	six2a	ENSORLP00000005562
	six3a	ENSORLP00000005568
	six3b	ENSORLP00000012924
	six4b	ENSORLP00000019732
	six5	ENSORLP00000008488
	six6	ENSORLP00000019744
	six7	ENSORLP00000010004
spotted gar (*Lepisosteus oculatus*)	*six1*	ENSLOCP00000012336

	six2	ENSLOCP00000020213
	six3	ENSLOCP00000020214
	six4	ENSLOCP00000012349
	six5	ENSLOCP00000018274
	six6	ENSLOCP00000012326
	six7	ENSLOCP00000001894
elephantshark (*Callorhinchus milii*)	*SIX1*	SINCAMP00000004524
	SIX2	SINCAMP00000018264
	SIX3	SINCAMP00000012754
	SIX4	SINCAMP00000004532
	SIX6	SINCAMP00000004521
lamprey (*Petromyzon marinus*)	*SIX3*	ENSPMAP00000002840
nematode (*Caenorhabditis elegans*)	*ceh-34*	C10G8.6

Supplementary table 5.

Ensembl accession numbers used for alignment of the members of the green cone opsins in ray-finned fish.

Vertebrate species	Gene	Ensembl accession number
zebrafish (*Danio rerio*)	*rh2-1*	ENSDARP00000001158
	rh2-2	ENSDARP00000011837
	rh2-3	ENSDARP00000001943
	rh2-4	ENSDARP00000000979
cod (*Gadus morhua*)	*rh2-1*	ENSGMOP00000018628
	rh2-2	ENSGMOP00000018611
fugu (*Takifugu rubripes*)	*rh2-2*	ENSTRUP00000005472
stickleback (*Gasterosteus aculeatus*)	*rh2-1*	ENSGACP00000001857
	rh2-2	ENSGACP00000001853
tilapia (*Oreochromis niloticus*)	*rh2-1*	ENSONIP00000005294
	rh2-2	ENSONIP00000005293
	rh2-3	ENSONIP00000005291
platyfish (*Xiphophorus maculatus*)	*rh2-1*	ENSXMAP00000015304
	rh2-2	ENSXMAP00000015295
medaka (*Oryzias latipes*)	*kfh-g*	ENSORLP00000024943
	rh2-b1	ENSORLP00000024987
	rh2-b2	ENSORLP00000024993
spotted gar (*Lepisosteus oculatus*)	*rh2*	ENSLOCP00000015272

Supplementary figure legends

Supplementary figure 1. Characterization of *six7* and/or *six3b* knock-out zebrafish. (a) *Top*, exon-intron organization and partial nucleotide sequences of *six7* gene in zebrafish. The binding sites of the left and right TALENs are highlighted in *blue*. The recognition site of the endonuclease *Xho*I is highlighted in *green*. *Bottom*, the nucleotide sequences of the *six7* KO alleles (ja51, ja52, and ja54) are compared with the WT sequence. Deletions are indicated by dashes. (b) Expression profiles of phototransduction component genes in *six7*$^{ja52/ja52}$ knock-out zebrafish at ~4.5 dpf. The ocular RNA (50 ng for the larval eye) was reverse-transcribed into cDNA, and the cDNA was subjected to quantitative PCR. The data were represented by mean + SEM (n = 5). Statistical significance between WT (+/+) and *six7* KO data (ja52/ja52) is shown as the asterisks (*, p<0.05 by Student's t-test). (c) Top, exon-intron organization and partial nucleotide sequences of *six3b* gene in zebrafish. The binding sites of the left and right TALENs are highlighted in *blue*. The recognition site of the endonuclease *Hind*III is highlighted in *green*. Bottom, the nucleotide sequences of the *six3b* KO alleles (ja53 and ja55) are compared with the WT sequence. Deletions are indicated by dashes. (d) Genotyping of *six3b* knock-out zebrafish by a combination of PCR and subsequent digestion by endonuclease *Hind*III. (e) Whole body images of the larval zebrafish at 2 dpf (*top*). The larvae were obtained from crossing of *six7*$^{ja52/+}$;*six3b*$^{ja53/+}$ fish. The images were taken with a stereoscopic microscope (MZ-FL3, Leica) and subjected to measurement of the eye area with ImageJ software (*bottom*). The data shown in the *bottom* panel

were represented by mean + SEM. The number of the zebrafish used for the quantification was indicated in the bar graph. Scale bar, 200 μm.

Supplementary figure 2. Related to figure 4b. The images of *dorsal*, *central* and *ventral* retina were demagnified from those in figure 4b. The cone opsin genes were expressed in the photoreceptor layer (white arrowheads), which is adjacent to the retinal pigmented epithelium (RPE, black arrowheads). In the images of the central retina, the dorsal side (*D*) is to the left, and the ventral (*V*) is to the right. Scale bar, 100 μm.

Supplementary figure 3. Expression levels of transcription factors contributing to the photoreceptor development in mouse and/or zebrafish in the WT and *six7* KO eyes at 3 mpf. The mRNA levels were quantified by RT-qPCR. The data were represented by mean + SEM (n = 4). Statistical significance between WT and *six7* KO data is shown as the asterisks (*, $p<0.05$ by Student's *t*-test)

Supplementary figure 4. Expression patterns of *lws1* and *lws2* in WT and *six7* KO retina. (a, b) The expression of two red opsin genes, *lws1* and *lws2*, was visualized by *in situ* hybridization. (a) Cryosections were prepared from the larval zebrafish at 3.5 dpf, which was the same developmental stage as in figure 3f. The *lws1* expression was detected only in KO retina (arrowheads). Note that the photoreceptor layer is adjacent to the RPE (black arrowheads). D, dorsal side. V, ventral side. Scale bar, 50 μm. (b) Related to figure 4b. Cryosections were prepared from the eye of WT or *six7* KO at 3 mpf, which was the same

developmental stage as in figure 4a and figure S3. RPE is indicated by asterisks. Scale bar, 30 μm. (c) Fluorescent images of the flat-mounted retina which were prepared from the 6 mpf-WT and KO and immunostained with the anti-Gnat2 antibody for all the subtypes of cones (*green*; panel 3, 4) and with the 1D4 antibody for the red cones (magenta; panel 1, 2). Merged images are shown on panel 5 and penal 6, while their magnified images are on panel 5' and 6'. The images were taken from the outside of the retina with focusing on the layer corresponding to the outer segment of the double cone in the central retina. Scale bar, 30 μm.

Supplementary figure 5. Expression levels of the opsin genes in WT (+/+) and *six3b* knock-out zebrafish (ja53/ja53) at 5 dpf. The ocular RNA (120 ng for the larval eye) was reverse-transcribed into cDNA, and the cDNA was subjected to quantitative PCR. The data were represented by mean + SEM (n = 5). Statistical significance ($p < 0.05$, Student's *t*-test) was not detected between WT and *six3b* KO for any of the genes.

Supplementary figure 6. Phylogenetic tree and sequence alignment of the members in Six family including reptilian SIX7 proteins. (a) A Neighbour-Joining (NJ) tree was constructed from amino acid sequences of SIX domain and homeodomain with 1,000 bootstrapping replications. Numerical values indicating bootstrap support are shown at the base of each node. Nematode CEH-34 was employed as an outgroup. The tree was generated with reptilian SIX7 protein sequences of Chinese soft-shell turtle, green anole, green sea turtle and Indian

python. Scale bar, 0.200 substitutions per site. (b) Sequence alignment of Six protein family. Note that in the SIX domain, the reptilian SIX7 proteins are poorly aligned with all the other members of Six protein family. The asterisks represent the four residues forming the homeobox nucleic acid recognition domain (HD) recognition helix.

Supplementary figure 7. Genomic environment of the green cone opsin genes in vertebrates. Orthologous genes contributing to the conserved synteny are similarly colour-coded. See also supplementary figure 8 for genomic environment of the green cone opsin genes of teleost including additional species. The comparative analysis revealed that the chromosomal locations of the green cone opsin genes can be classified into two types, (i) the non-teleost type and (ii) the teleost type. The former synteny (i) (upper half) is characterized by the presence of a series of the genes such as *IP6K3* and *LEMD2* in close proximity to *RH2,* and is found in the non-teleost species ranging from elephant shark to tetrapod. Note that elephant shark was diverged at the most ancient period from the other vertebrate species indicated in this figure, suggesting the “non-teleost type” synteny to represent an ancient type. The latter synteny (ii) (bottom half) is characterized by an array of different genes (such as *slc6a13* and *fkbp5*) neighbouring *rh2*, and is found in the teleost species. Note that spotted gar, the closest relative to teleost retaining *six7*, has the intermediate-type synteny around the green cone opsin gene. The synteny of spotted gar suggests that the genes around *rh2* was rearranged at least twice in the teleost lineage, *i.e.*, before and after the separation from spotted gar as represented by the node “a” and “b”,

respectively. This first rearrangement may predate the third-round whole genome duplication (3R), which is hypothesized to have occurred in the teleost lineage after separation (at node a) from the lineage leading to spotted gar [9]. An arrow in grey indicates an incompletely annotated green cone opsin gene, which perhaps is a pseudogene. Elephant shark, *Callorhinchus milii*; zebrafish, *Danio rerio*; medaka, *Oryzias latipes*; fugu, *Takifugu rubripes*; spotted gar, *Lepisosteus oculatus*; coelacanth, *Latimeria chalumnae*; anole lizard, *Anolis carolinensis*; chicken, *Gallus gallus*; human, *Homo sapiens*.

Supplementary figure 8. Phylogenetic tree of the members in teleost green cone opsins (a) and genomic environment of the green cone opsin genes in teleost (b). (a) A Neighbour-Joining (NJ) tree was constructed from the amino acid sequences of the green cone opsins without N terminal and C terminal regions with 1,000 bootstrapping replications. Numerical values indicating bootstrap support are shown at the base of each node. Closed and open circles on the nodes indicate the intraspecific (*closed*) and interspecific (*open*) duplication of green opsin genes. Spotted gar Rh2 was used as outgroup. The scale bar indicates the number of amino acid substitutions per site. Teleost Rh2 was separated into three groups: Rh2-A/B, Rh2-A and Rh2-B. (b) Genomic neighbourhood of the green opsin genes are compared among teleost species, whose phylogeny is presented in the left side. Different green opsin gene groups were indicated with distinct colours, i.e., *yellow* for Rh2-A/B, *blue* for Rh2-A, and *pink* for Rh2-B. Arrows surrounded by broken line indicate incompletely annotated green opsin genes, which might be pseudogenes. Cross marks in the

species phylogeny indicate the local duplication of the green opsin genes. Zebrafish, *Danio rerio*; medaka, *Oryzias latipes*; cod, *Gadus morhua*; fugu, *Takifugu rubripes*; stickleback, *Gasterosteus aculeatus*; tilapia, *Oreochromis niloticus*; platyfish, *Xiphophorus maculatus.*

Supplementary material 1.

The nucleotide sequence of pTAL7DD vector. For generating pTAL7RR vector, two codons in FokI nuclease catalytic domain (underlined) were changed from GAT (D) to AGA (R).

>pTAL7DD

CGCCATTCTGCCTGGGGACGTCGGAGCAAGCTTGATTTAGGTGACACTATAGAATACA
AGCTACTTGTTCTTTTTGCAGGATCGCCACCATGGCTTCCTCCCCTCCAAAGAAAAAG
AGAAAGGTTGATTACAAGGATGACGACGATAAGGTGGATCTACGCACGCTCGGCTACA
GTCAGCAGCAGCAAGAGAAGATCAAACCGAAGGTGCGTTCGACAGTGGCGCAGCACCA
CGAGGCACTGGTGGGCCATGGGTTTACACACGCGCACATCGTTGCGCTCAGCCAACAC
CCGGCAGCGTTAGGGACCGTCGCTGTCACGTATCAGCACATAATCACGGCGTTGCCAG
AGGCGACACACGAAGACATCGTTGGCGTCGGCAAACAGTGGTCCGGCGCACGCGCCCT
GGAGGCCTTGCTCACGGATGCGGGGGAGTTGAGAGGTCCGCCGTTACAGTTGGACACA
GGCCAACTTGTGAAGATTGCAAAACGTGGCGGCGTGACCGCAATGGAGGCAGTGCATG
CATCGCGCAATGCACTGACGGGTGCCCCCCTGGAGACGGGCGCCGCTACAGGGCGCGT
CCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTC
GCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACG
CCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAGCGCGCGTAATACG
ACTCACTATAGGGCGAATTGGGTACCGGGCCCCCCCTCGAGGTCCTCCAGCTTTTGTT
CCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGT
GTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGT
AAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGC
CCGCTTTCCACCGGTCGTCTCCAACGACCACCTCGTCGCCTTGGCCTGCCTCGGCGGA
CGTCCTGCCATGGATGCAGTGAAAAAGGGATTGCCGCACGCGCCGGAATTGATCAGAA
GAGTCAATCGCCGTATTGGCGAACGCACGTCCCATCGCGTTGCCGGATCCCAGCTAGT
GAAATCTGAATTGGAAGAGAAGAAATCTGAACTTAGACATAAATTGAAATATGTGCCA
CATGAATATATTGAATTGATTGAAATCGCAAGAAATTCAACTCAGGATAGAATCCTTG
AAATGAAGGTGATGGAGTTCTTTATGAAGGTTTATGGTTATCGTGGTAAACATTTGGG
TGGATCAAGGAAACCAGACGGAGCAATTTATACTGTCGGATCTCCTATTGATTACGGT
GTGATCGTTGATACTAAGGCATATTCAGGAGGTTATAATCTTCCAATTGGTCAAGCA**G**
ATGAAATGCAA**GAT**TATGTCGAAGAGAATCAAACAAGAAACAAGCATATCAACCCTAA
TGAATGGTGGAAAGTCTATCCATCTTCAGTAACAGAATTTAAGTTCTTGTTTGTGAGT
GGTCATTTCAAAGGAAACTACAAAGCTCAGCTTACAAGATTGAATCATATCACTAATT
GTAATGGAGCTGTTCTTAGTGTAGAAGAGCTTTTGATTGGTGGAGAAATGATTAAAGC

TGGTACATTGACACTTGAGGAAGTGAGAAGGAAATTTAATAACGGTGAGATAAACTTT
TAATAGGAGCTCAATTCAAGGCCTCTCGAGCCTCTAGAACTATAGTGAGTCGTATTAC
GTAGATCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGC
AGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCAT
TATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTT
CAGGGGGAGGTGTGGGAGGTTTTTTAATTCGCGGCCGCGGCGCCAATGCATTGGGCCC
GGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGG
TCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAG
CCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAAT
TGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAA
TGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCT
CGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTC
AAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGA
GCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCC
ATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCG
AAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGC
TCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAA
GCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCG
CTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCC
GGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAG
CCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAA
GTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTG
AAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCG
CTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATC
TCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCA
CGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAA
ATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAG
TTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCC
ATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTG
GCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGC
AATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCC
TCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATA
GTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGG
TATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATG
TTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGG
CCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCC
ATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAG
TGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCAC
ATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTC
AAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGA
TCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAA
ATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCT
TTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTT
GAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGC
CACCTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATC
AGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAAT
AGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAA
CGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGT

GAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGA
ACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAG
AAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTC
ACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCC
ATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCT
ATTACGCCAGTCGACCATAGCCAATTCAATATGGCGTATATGGACTCATGCCAATTCA
ATATGGTGGATCTGGACCTGTGCCAATTCAATATGGCGTATATGGACTCGTGCCAATT
CAATATGGTGGATCTGGACCCCAGCCAATTCAATATGGCGGACTTGGCACCATGCCAA
TTCAATATGGCGGACTTGGCACTGTGCCAACTGGGGAGGGGTCTACTTGGCACGGTGC
CAAGTTTGAGGAGGGGTCTTGGCCCTGTGCCAAGTCCGCCATATTGAATTGGCATGGT
GCCAATAATGGCGGCCATATTGGCTATATGCCAGGATCAATATATAGGCAATATCCAA
TATGGCCCTATGCCAATATGGCTATTGGCCAGGTTCAATACTATGTATTGGCCCTATG
CCATATAGTATTCCATATATGGGTTTTCCTATTGACGTAGATAGCCCCTCCCAATGGG
CGGTCCCATATACCATATATGGGGCTTCCTAATACCGCCCATAGCCACTCCCCCATTG
ACGTCAATGGTCTCTATATATGGTCTTTCCTATTGACGTCATATGGGCGGTCCTATTG
ACGTATATGGCGCCTCCCCCATTGACGTCAATTACGGTAAATGGCCCGCCTGGCTCAA
TGCCCATTGACGTCAATAGGACCACCCACCATTGACGTCAATGGGATGGCTCATTGCC
CATTCATATCCGTTCTCACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCACTTG
GCAGTACATCAATATCTATTAATAGTAACTTGGCAAGTACATTACTATTGGAAGGACG
CCAGGGTACATTGGCAGTACTCCCATTGACGTCAATGGCGGTAAATGGCCCGCGATGG
CTGCCAAGTACATCCCCATTGACGTCAATGGGGAGGGGCAATGACGCAAATGGGCGTT
CCATTGACGTAAATGGGCGGTAGGCGTGCCTAATGGGAGGTCTATATAAGCAATGCTC
GTTTAGGGAAC

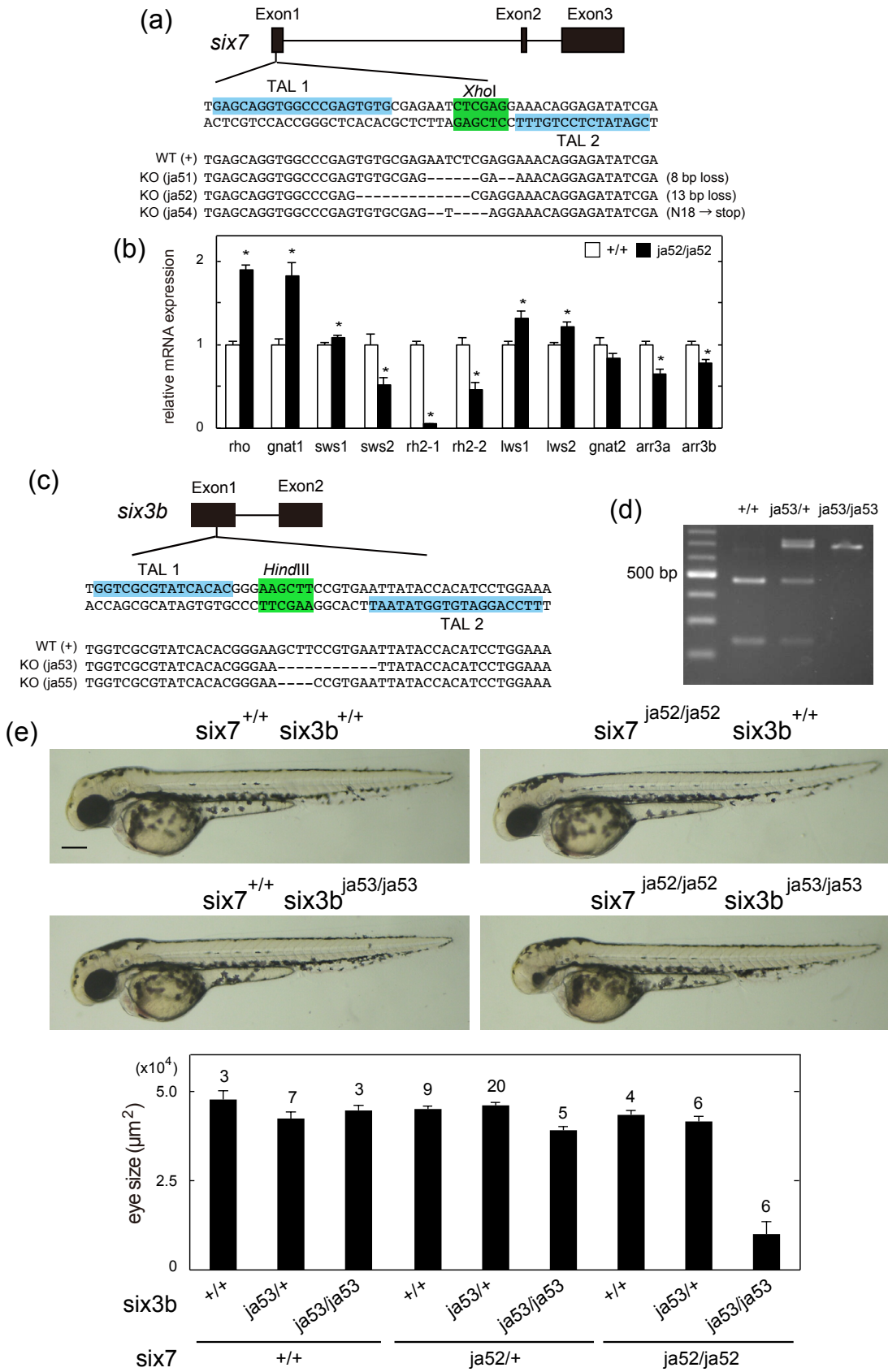
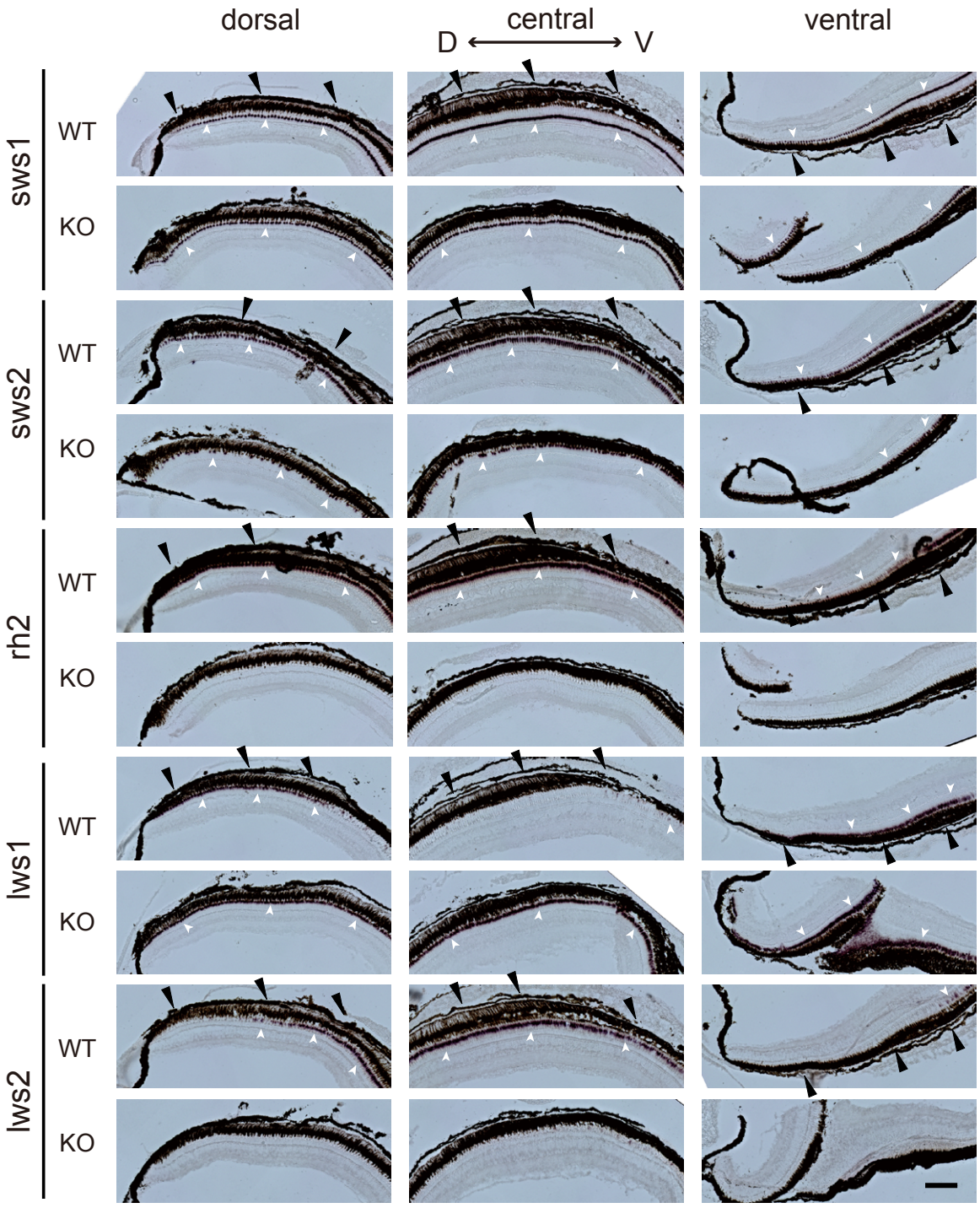

(a)
six7
Exon1
Exon2
Exon3
TAL 1
XhoI
TAL 2
WT (+) TGAGCAGGTGGCCCGAGTGTGCGAGAATCTCGAGGAAACAGGAGATATCGA
KO (ja51) TGAGCAGGTGGCCCGAGTGTGCGAG------GA--AAACAGGAGATATCGA (8 bp loss)
KO (ja52) TGAGCAGGTGGCCCGAG----------CGAGGAAACAGGAGATATCGA (13 bp loss)
KO (ja54) TGAGCAGGTGGCCCGAGTGTGCGAG--T----AGGAAACAGGAGATATCGA (N18 → stop)
(b)
+/+
ja52/ja52
relative mRNA expression
rho
gnat1
sws1
sws2
rh2-1
rh2-2
lws1
lws2
gnat2
arr3a
arr3b
(c)
six3b
Exon1
Exon2
TAL 1
HindIII
TAL 2
WT (+) TGGTCGCGTATCACACGGGAAGCTTCCGTGAATTATACCACATCCTGGAAA
KO (ja53) TGGTCGCGTATCACACGGGAA----------TTATACCACATCCTGGAAA
KO (ja55) TGGTCGCGTATCACACGGGAA----CCGTGAATTATACCACATCCTGGAAA
(d)
+/+
ja53/+
ja53/ja53
500 bp
(e)
six7+/+ six3b+/+
six7ja52/ja52 six3b+/+
six7+/+ six3bja53/ja53
six7ja52/ja52 six3bja53/ja53
(x10^4)
eye size (μm^2)
3
7
3
9
20
5
4
6
6
six3b
+/+
ja53/+
ja53/ja53
six7
+/+
ja52/+
ja52/ja52

Figure S1



4 mpf

Figure S2

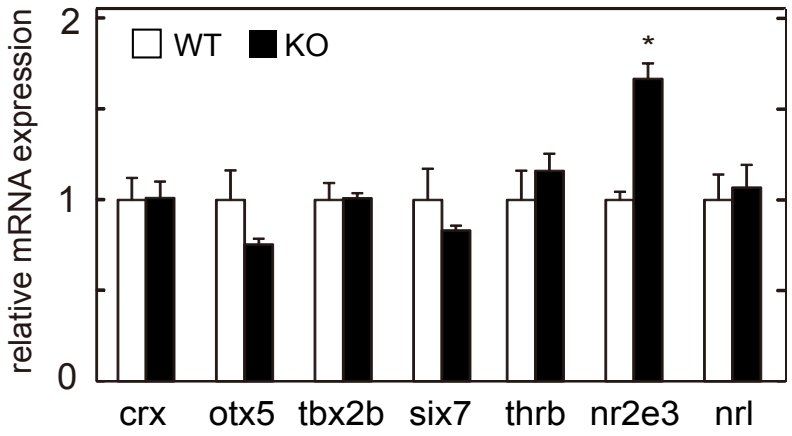

relative mRNA expression
2
1
0
WT
KO
*
crx
otx5
tbx2b
six7
thrb
nr2e3
nrl

Figure S3

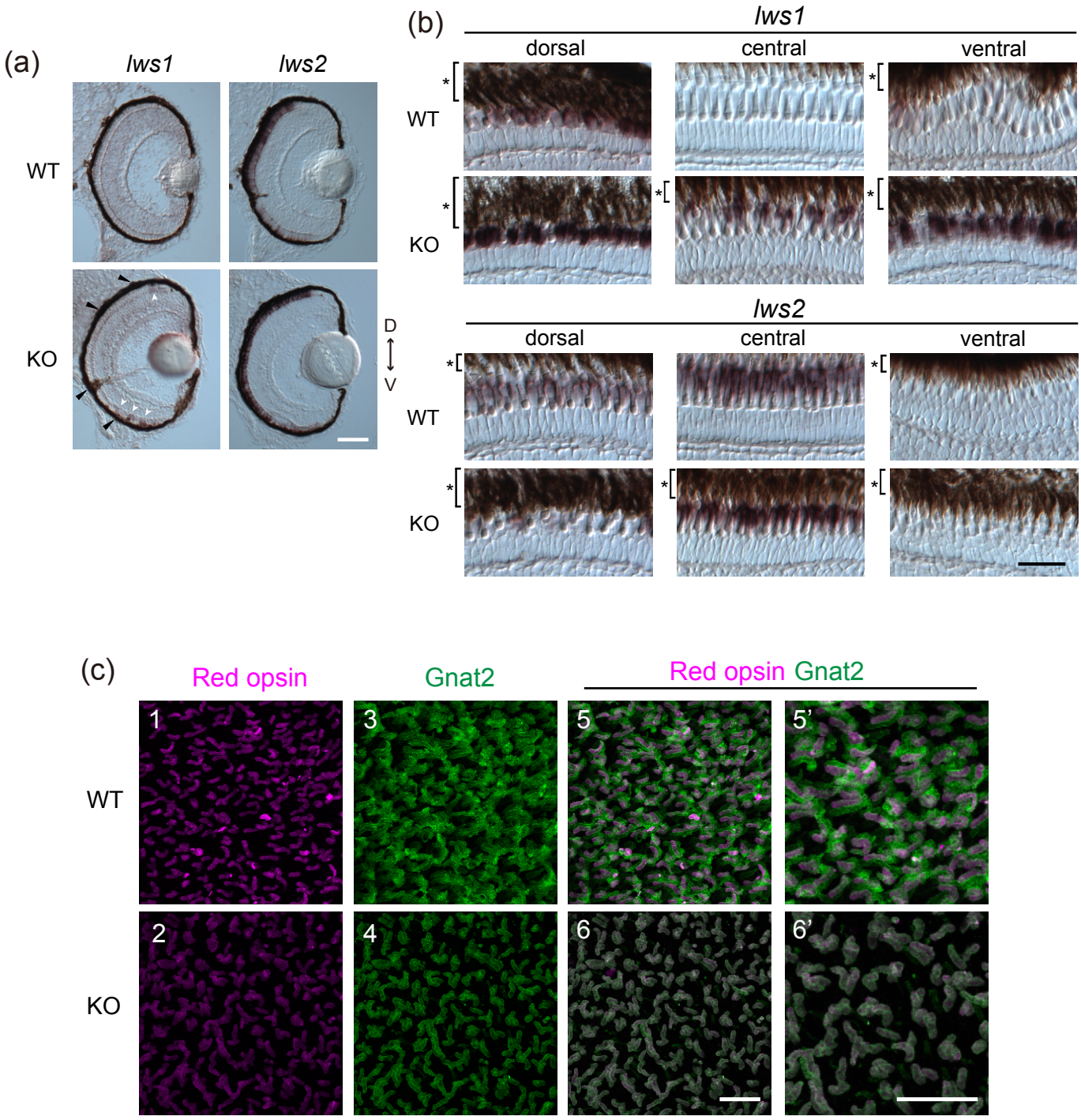

(a)
lws1
lws2
WT
KO
(b)
lws1
dorsal
central
ventral
WT
KO
D
V
lws2
dorsal
central
ventral
WT
KO
(c)
Red opsin
Gnat2
Red opsin Gnat2
WT
KO
1
2
3
4
5
5'
6
6'

Figure S4

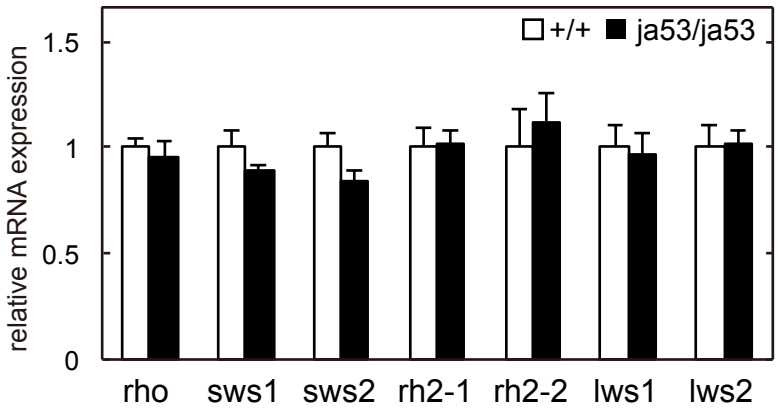

+/+
ja53/ja53
relative mRNA expression
1.5
1
0.5
0
rho
sws1
sws2
rh2-1
rh2-2
lws1
lws2

Figure S5

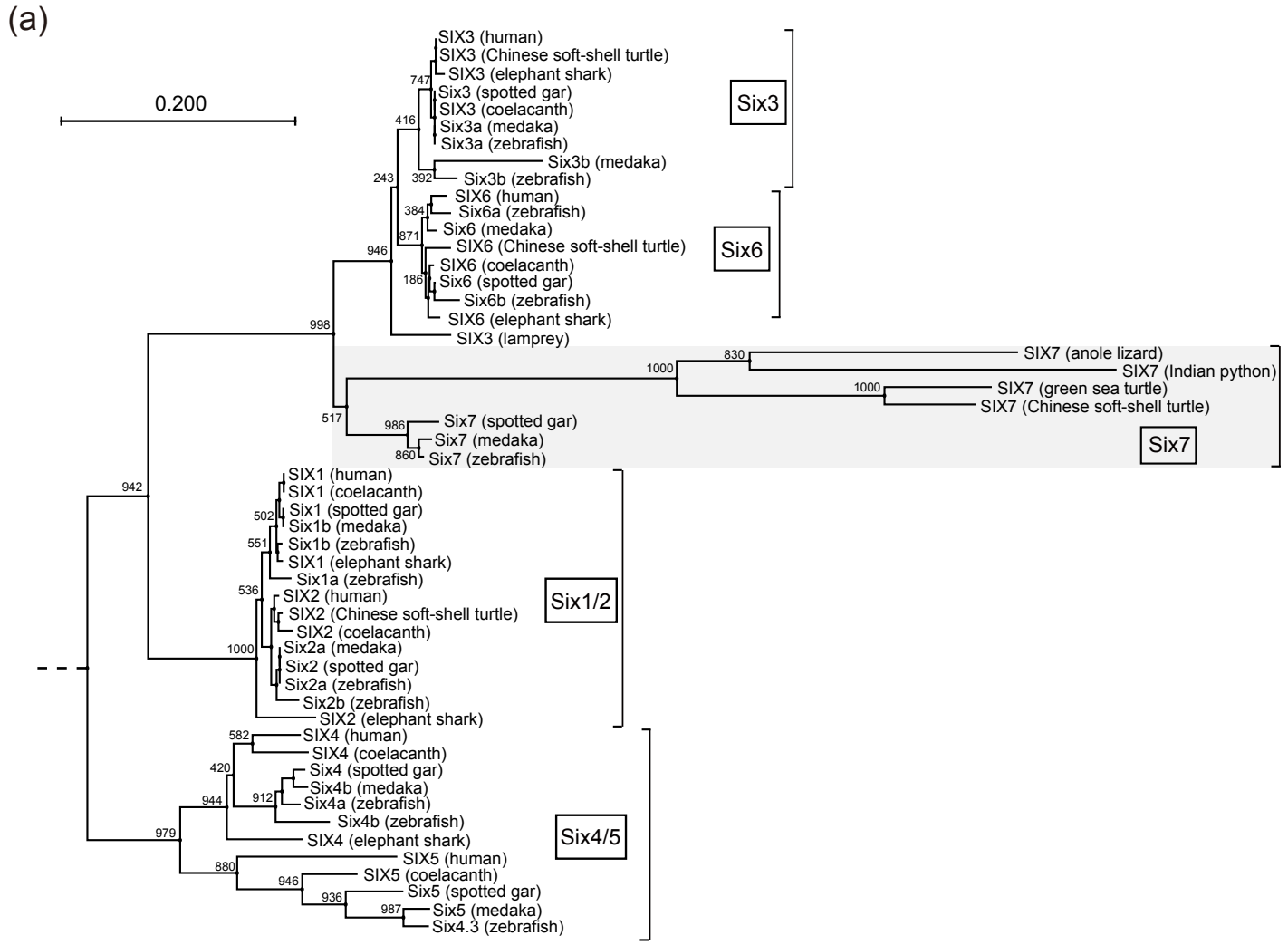
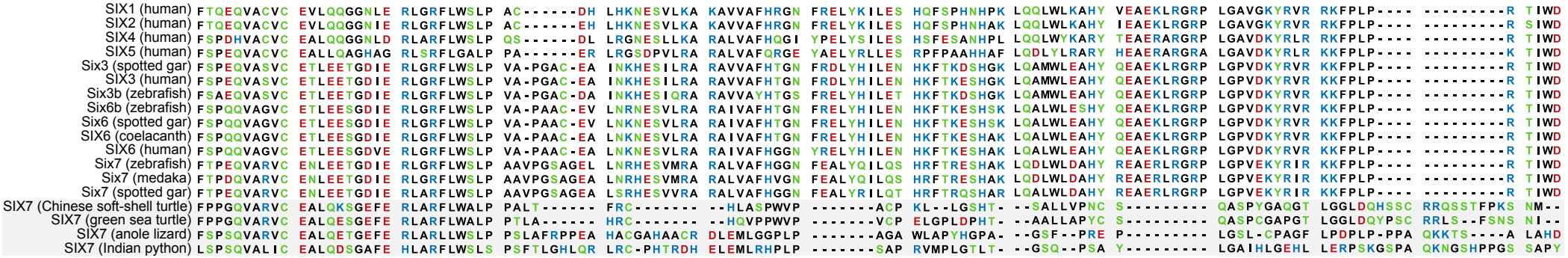

(a)
0.200
SIX3 (human)
SIX3 (Chinese soft-shell turtle)
SIX3 (elephant shark)
Six3 (spotted gar)
SIX3 (coelacanth)
Six3a (medaka)
Six3a (zebrafish)
Six3b (medaka)
Six3b (zebrafish)
Six3
SIX6 (human)
Six6a (zebrafish)
Six6 (medaka)
SIX6 (Chinese soft-shell turtle)
SIX6 (coelacanth)
Six6 (spotted gar)
Six6b (zebrafish)
SIX6 (elephant shark)
Six6
SIX3 (lamprey)
SIX7 (anole lizard)
SIX7 (Indian python)
SIX7 (green sea turtle)
SIX7 (Chinese soft-shell turtle)
Six7 (spotted gar)
Six7 (medaka)
Six7 (zebrafish)
Six7
SIX1 (human)
SIX1 (coelacanth)
Six1 (spotted gar)
Six1b (medaka)
Six1b (zebrafish)
SIX1 (elephant shark)
Six1a (zebrafish)
SIX2 (human)
SIX2 (Chinese soft-shell turtle)
SIX2 (coelacanth)
Six2a (medaka)
Six2 (spotted gar)
Six2a (zebrafish)
Six2b (zebrafish)
SIX2 (elephant shark)
Six1/2
SIX4 (human)
SIX4 (coelacanth)
Six4 (spotted gar)
Six4b (medaka)
Six4a (zebrafish)
Six4b (zebrafish)
SIX4 (elephant shark)
SIX5 (human)
SIX5 (coelacanth)
Six5 (spotted gar)
Six5 (medaka)
Six4.3 (zebrafish)
Six4/5

Figure S6

(b)

SIX domain



Homeo domain

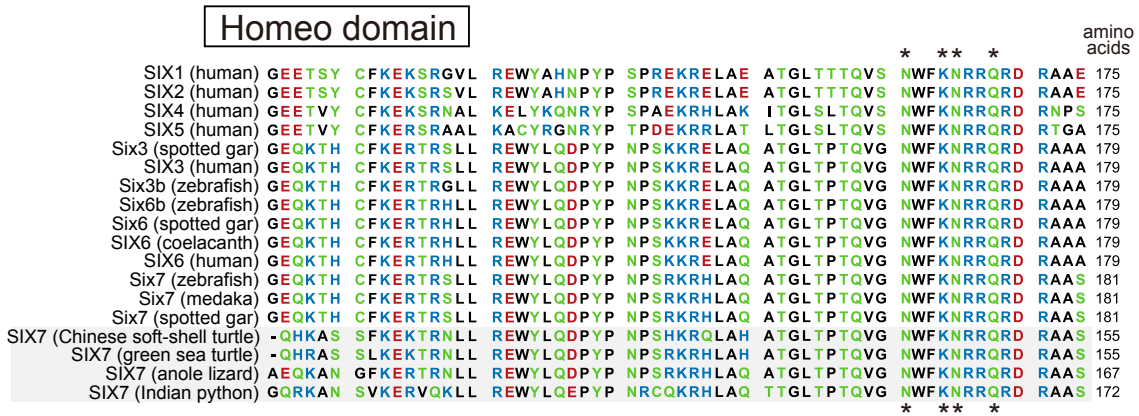


Figure S6

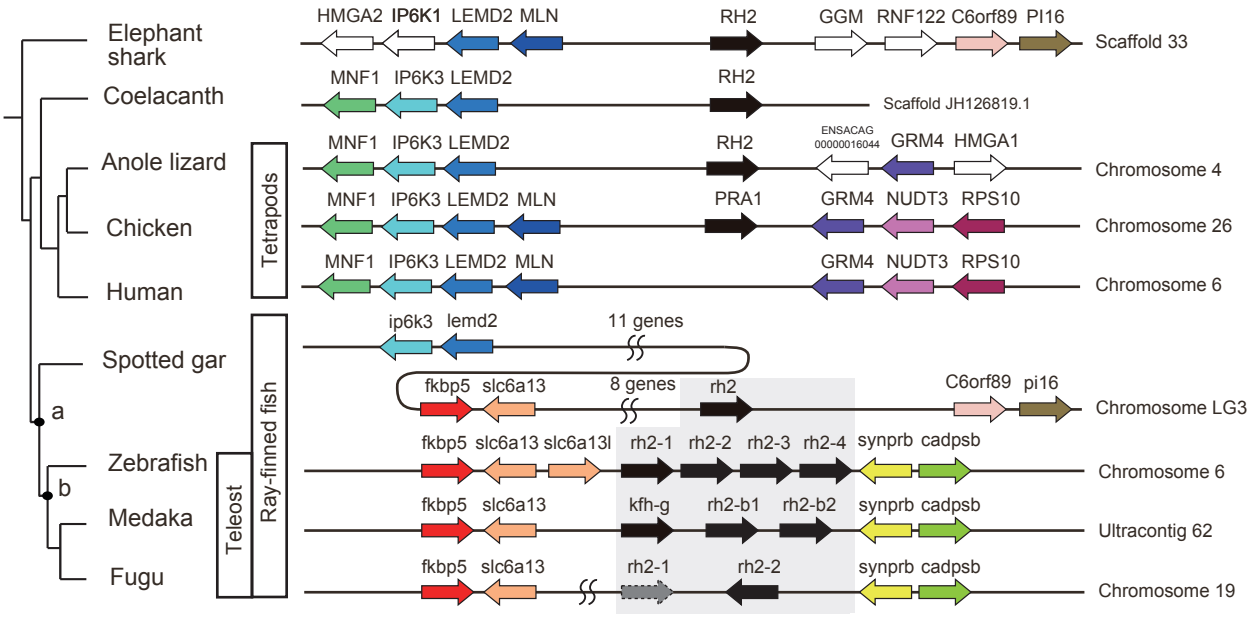

Elephant shark
Coelacanth
Anole lizard
Chicken
Human
Spotted gar
Zebrafish
Medaka
Fugu
Tetrapods
Ray-finned fish
Teleost
a
b
HMGA2 IP6K1 LEMD2 MLN
RH2
GGM RNF122 C6orf89 PI16
Scaffold 33
MNF1 IP6K3 LEMD2
RH2
Scaffold JH126819.1
MNF1 IP6K3 LEMD2
RH2
ENSACAG 00000016044
GRM4 HMGA1
Chromosome 4
MNF1 IP6K3 LEMD2 MLN
PRA1
GRM4 NUDT3 RPS10
Chromosome 26
MNF1 IP6K3 LEMD2 MLN
GRM4 NUDT3 RPS10
Chromosome 6
ip6k3 lemd2
11 genes
fkbp5 slc6a13
8 genes
rh2
C6orf89 pi16
Chromosome LG3
fkbp5 slc6a13 slc6a13l
rh2-1 rh2-2 rh2-3 rh2-4
synprb cadpsb
Chromosome 6
fkbp5 slc6a13
kfh-g
rh2-b1 rh2-b2
synprb cadpsb
Ultracontig 62
fkbp5 slc6a13
rh2-1
rh2-2
synprb cadpsb
Chromosome 19

Figure S7

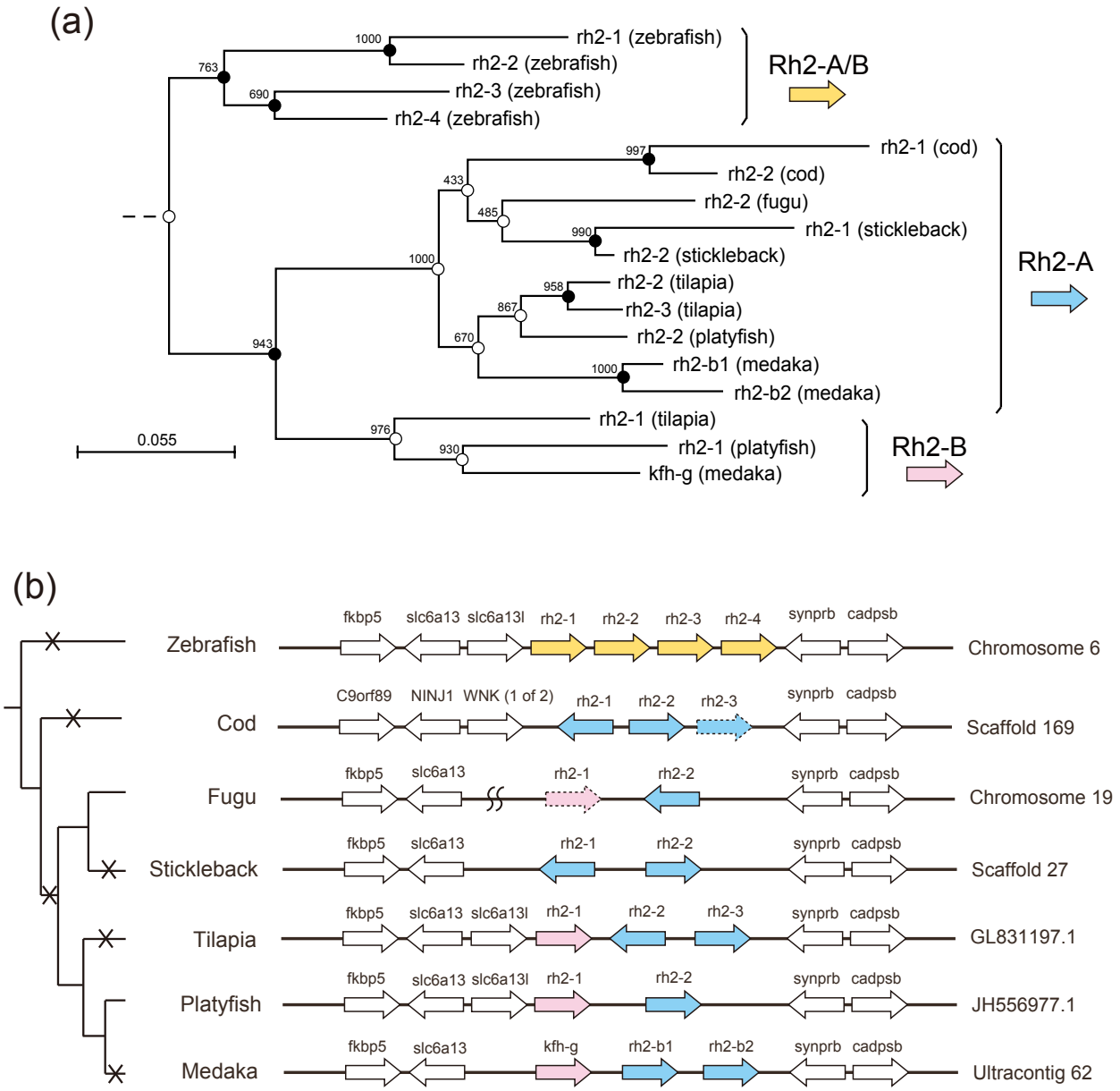

(a)
1000
763
690
rh2-1 (zebrafish)
rh2-2 (zebrafish)
rh2-3 (zebrafish)
rh2-4 (zebrafish)
Rh2-A/B
997
433
485
990
1000
958
867
670
1000
943
976
930
rh2-1 (cod)
rh2-2 (cod)
rh2-2 (fugu)
rh2-1 (stickleback)
rh2-2 (stickleback)
rh2-2 (tilapia)
rh2-3 (tilapia)
rh2-2 (platyfish)
rh2-b1 (medaka)
rh2-b2 (medaka)
Rh2-A
rh2-1 (tilapia)
rh2-1 (platyfish)
kfh-g (medaka)
Rh2-B
0.055
(b)
Zebrafish
fkbp5 slc6a13 slc6a13l rh2-1 rh2-2 rh2-3 rh2-4 synprb cadpsb
Chromosome 6
Cod
C9orf89 NINJ1 WNK (1 of 2) rh2-1 rh2-2 rh2-3 synprb cadpsb
Scaffold 169
Fugu
fkbp5 slc6a13 rh2-1 rh2-2 synprb cadpsb
Chromosome 19
Stickleback
fkbp5 slc6a13 rh2-1 rh2-2 synprb cadpsb
Scaffold 27
Tilapia
fkbp5 slc6a13 slc6a13l rh2-1 rh2-2 rh2-3 synprb cadpsb
GL831197.1
Platyfish
fkbp5 slc6a13 slc6a13l rh2-1 rh2-2 synprb cadpsb
JH556977.1
Medaka
fkbp5 slc6a13 kfh-g rh2-b1 rh2-b2 synprb cadpsb
Ultracontig 62

Figure S8