Supporting Information

Suga et al. 10.1073/pnas.1008389107

SI Materials and Methods

Animal Culture. *Cladomena radiatum* Dujardin (Cnidaria, Hydrozoa, Anthomedusae) polyp colonies were kept in the laboratory at 14–19 °C. Two growth stages of medusae, "just-detached" and, "full-grown", were defined for experimental purposes; a medusa younger than 4 d after detachment from the colony is called just-detached, and a medusa with a bell of more than 2 mm in diameter is called full-grown. A medusa usually reaches the full-grown stage and develops gonads in approximately 2 wk after detachment.

Fly Strains. Flies were reared on standard medium at 25 °C. The following lines were used (all were generated in this study unless specified): (*i*) UAS-*CrPax-A*, (*ii*) UAS-*CrPax-B*, (*iii*) UAS-*CrPax-E*, (*iv*) UAS-*D-Pax2* (1), (*v*) spa^{pol} (2), (*vi*) $ey^{J5.71}$ (3), (*vii*) *spa*-Gal4; spa^{pol} (4), (*viii*) dpp^{blink}-Gal4 (5), and (*ix*) UAS-Gal4 (kindly provided by Bassem Hassan, Katholieke Universiteit Leuven, Louvain, Belgium).

Cloning and Sequencing. Primer sequences are available in Table S1. Jellyfish cDNA was generated as described previously (6). *Microciona* cDNA was kindly provided by Dr. M. M. Burger (Friedrich Miescher Institute, Basel, Switzerland). We conducted PCR with the Expand High Fidelity PCR System (Roche) using 11 degenerate primers. Thermal schedules and primer combinations are as follows: For *CrPax-A*, 20 × (30 s at 94 °C, 45 s at 40 °C, 1 min at 72 °C), then 10 × (25 s at 94 °C, 30 s at 40 °C, 1 min at 72 °C) with pax-1s and pax-1a, followed by a nested PCR 40 × (25 s at 94 °C, 30 s at 50 °C, 1 min at 72 °C) with pax-2s and pax-2a. For *CrPax-B*, 30 × (15 s at 94 °C, 1 min at 45 °C, 1 min at 72 °C) with pax-3s and AUAP (Invitrogen), followed by a nested PCR 35 × (15 s at 94 °C, 1 min at 50 °C, 1 min at 72 °C) with pax-3s and pax-3a. For *CrPax-E*, 30 × (15 s at 94 °C, 1 min at 46 °C, 2 min at 72 °C) with pax-4s and AUAP, followed by a nested PCR 40 × (15 s at 94 °C, 1 min at 46 °C, 1 min at 72 °C) with pax-4s and pax-4a. For *Microciona Pax-2/5/8*, 30 × (15 s at 94 °C, 1 min at 40 °C, 1 min at 72 °C) with pax-5s and pax-5a, followed by a nested PCR 10 × (15 s at 94 °C, 1 min at 40 °C, 30 s at 72 °C) then 30 × (15 s at 94 °C, 1 min at 45 °C, 30 s at 72 °C) with pax-5s and pax-6a. The 5′ and 3′ ends of each cDNA were amplified by RACE and inverse PCR, and then sequenced.

The 5′ upstream regions of *Cladonema* opsin genes were cloned by inverse PCR with specific primers according to the standard protocol (7).

Real-Time PCR. Real-time RT-PCR expression analysis was carried out with the Light Cycler (Roche) as previously described (6). Primer sequences are in Table S2.

Molecular Phylogenetic Tree Analysis. We inferred molecular phylogenetic trees of the *Pax* gene family by the maximum likelihood (ML) method using the GAMT (8) program with the JTT-F model of amino acid substitution. Heterogeneity of evolutionary rates among sites was modeled by a discrete Γ distribution (9) with an optimized shape parameter. Extended local bootstrap analysis, an improved method that analyzes up to eight neighboring subtrees instead of four, was carried out and the calculated probability is shown at each branch (6).

We performed Bayesian inferences of phylogenetic tree using the MrBayes program (10) with the WAG-F model of amino acid substitution and the discrete Γ distribution of heterogeneity of evolutionary rates among sites. We determined the number of Metropolis-coupled Markov Chain Monte Carlo generations by checking the convergence of average SD of split frequencies (ASDSF) (ASDSF < 0.01). The confidence of a clustering is indicated by a Bayesian posterior probability.

***In silico* Search for *Pax* Genes.** We performed BLAST searches on the genome sequences of two cnidarians, *Hydra magnipapillata* (version 1.0, dated July 26, 2006; http://hydrazome.metazome.net/) and *Nematostella vectensis*, a sea anemone (version 1.0; http://genome.jgi-psf.org/Nemve1/), and found three and eight *Pax* genes, respectively. All these sequences, except the hydra *Pax-E*-like sequence, had been already deposited in the public database. The paired domain (PD) of hydra *Pax-E*-like gene was identified at nucleotides 187733–188000 of the contig 39355. *Pax-D4* (XP_0016-23756) and an uncharacterized *Pax* (EDO38773) of *Nematostella* were not included in the phylogenetic analyses because their PD sequences are significantly corrupted.

The genome sequences of *Amphimedon queenslandica*, a marine sponge, and *Monosiga brevicollis*, a choanoflagellate, were obtained at ftp://ftp.ncbi.nih.gov/pub/TraceDB/reniera_sp__jgi-2005/ and http://genome.jgi-psf.org/Monbr1/, respectively.

In Situ Hybridization. Whole-mount in situ hybridization was performed according to the previously published protocol (6). Note that the natural eye pigment disappears during the procedure.

Protein Expression and DNA Binding Assays. Recombinant proteins containing the PDs of CrPax-A and CrPax-B tagged to histidine tails were produced in *Escherichia coli* strain Rosetta2(DE3)pLysS (Novagen) and purified with Ni-NTA Agarose (Qiagen), according to the manufacturer's instructions. The expression constructs were generated by cloning of the cDNA fragments that correspond to amino acids 1–175 of CrPax-A and 1–230 of CrPax-B into pET-21b (+) (Novagen). From these constructs, 21.5-kDa (CrPax-A PD) and 27.7-kDa (CrPax-B PD) proteins were obtained. Probes for EMSA were generated by PCR with specific primers (sequences in Table S3). The purified probes were digested with EcoRI or HindIII and subjected to end filling with Klenow fragment in the presence of [α-^{32}P]dATP. The activity of CrPax-B PD was confirmed by a positive control probe, which was generated from a region of the *Drosophila* genome (nucleotides 4087217–4087441 of GenBank entry AE014298) that carries two putative D-Pax2 binding sites (courtesy of Patrick Callaerts, Katholieke Universiteit Leuven, Louvain, Belgium, and Jorge Blanco, Institute of Medical Biology, Singapore, Singapore). Binding reaction was carried out for 20 min at 19 °C in a buffer containing 10 mM Tris-HCl (pH 7.5), 5% glycerol, 75 mM NaCl, 1 mM DTT, 50 μg/mL poly(dI-dC), and 100μg/mL BSA.

1. Kozmik Z, et al. (2003) Role of *Pax* genes in eye evolution: A cnidarian *PaxB* gene uniting *Pax2* and *Pax6* functions. *Dev Cell* 5:773–785.
2. Fu W, Noll M (1997) The *Pax2* homolog *sparkling* is required for development of cone and pigment cells in the *Drosophila* eye. *Genes Dev* 11: 2066–2078.
3. Punzo C, et al. (2004) Functional divergence between *eyeless* and *twin of eyeless* in *Drosophila melanogaster*. *Development* 131:3943–3953.
4. Jiao R, et al. (2001) Headless flies generated by developmental pathway interference. *Development* 128:3307–3319.
5. Staehling-Hampton K, Hoffmann FM (1994) Ectopic *decapentaplegic* in the *Drosophila* midgut alters the expression of five homeotic genes, *dpp*, and *wingless*, causing specific morphological defects. *Dev Biol* 164:502–512.
6. Suga H, Schmid V, Gehring WJ (2008) Evolution and functional diversity of jellyfish opsins. *Curr Biol* 18:51–55.
7. Sambrook J, Russell DW (2001) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab Press, Cold Spring Harbor, NY)3rd Ed.
8. Katoh K, Kuma K, Miyata T (2001) Genetic algorithm-based maximum-likelihood analysis for molecular phylogeny. *J Mol Evol* 53:477–484.

9. Yang Z (1994) Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods. *J Mol Evol* 39:306–314.

10. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.

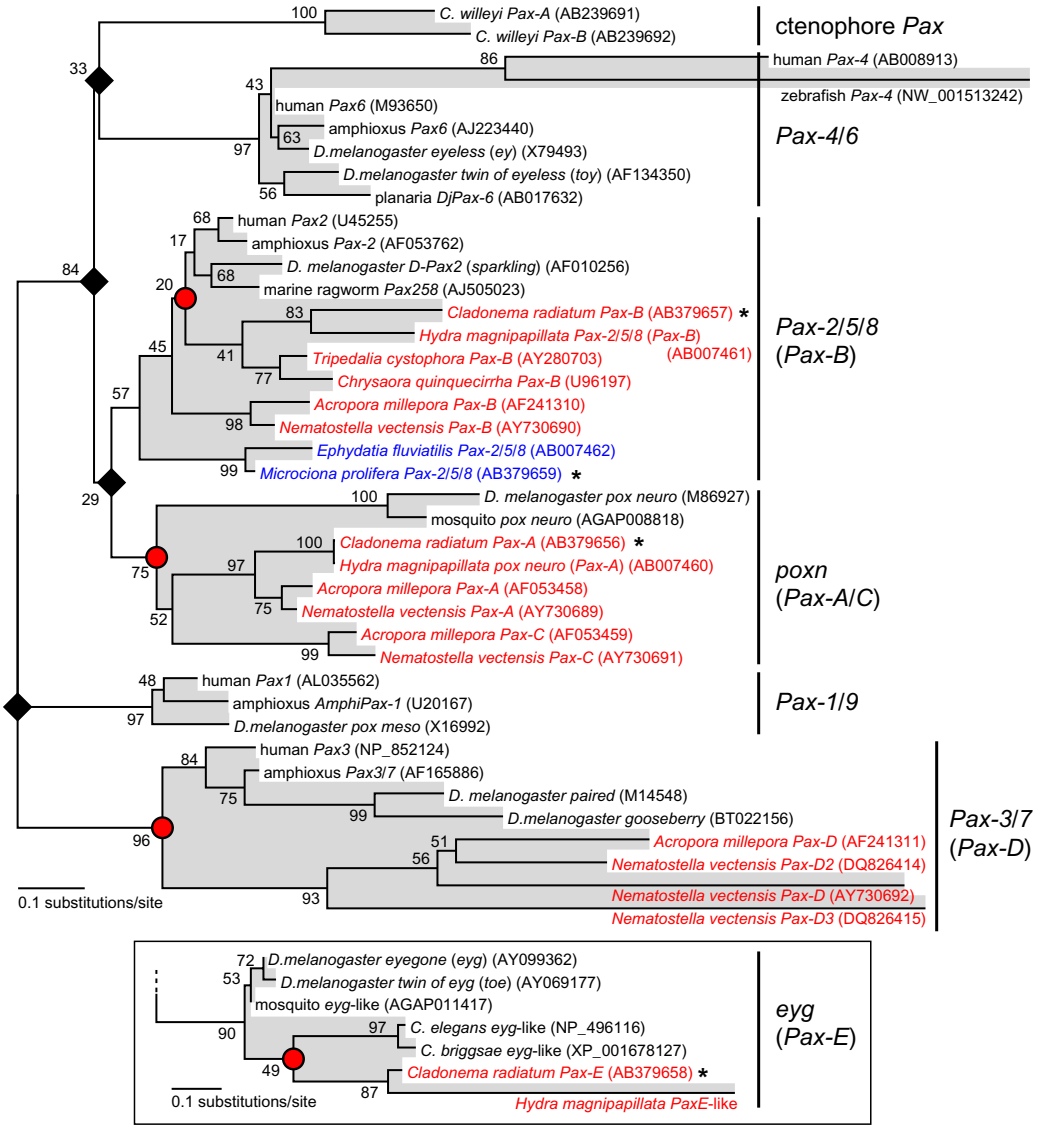

Fig. S1. Phylogenetic tree of *Pax* gene family inferred from the whole and the latter halves of (in box) PD sequences by the ML method. The tree topologies are the same as shown in Fig. 1*C*. The accession number to each sequence is given in parentheses. The extended bootstrap probability is shown comprehensively at each branch.

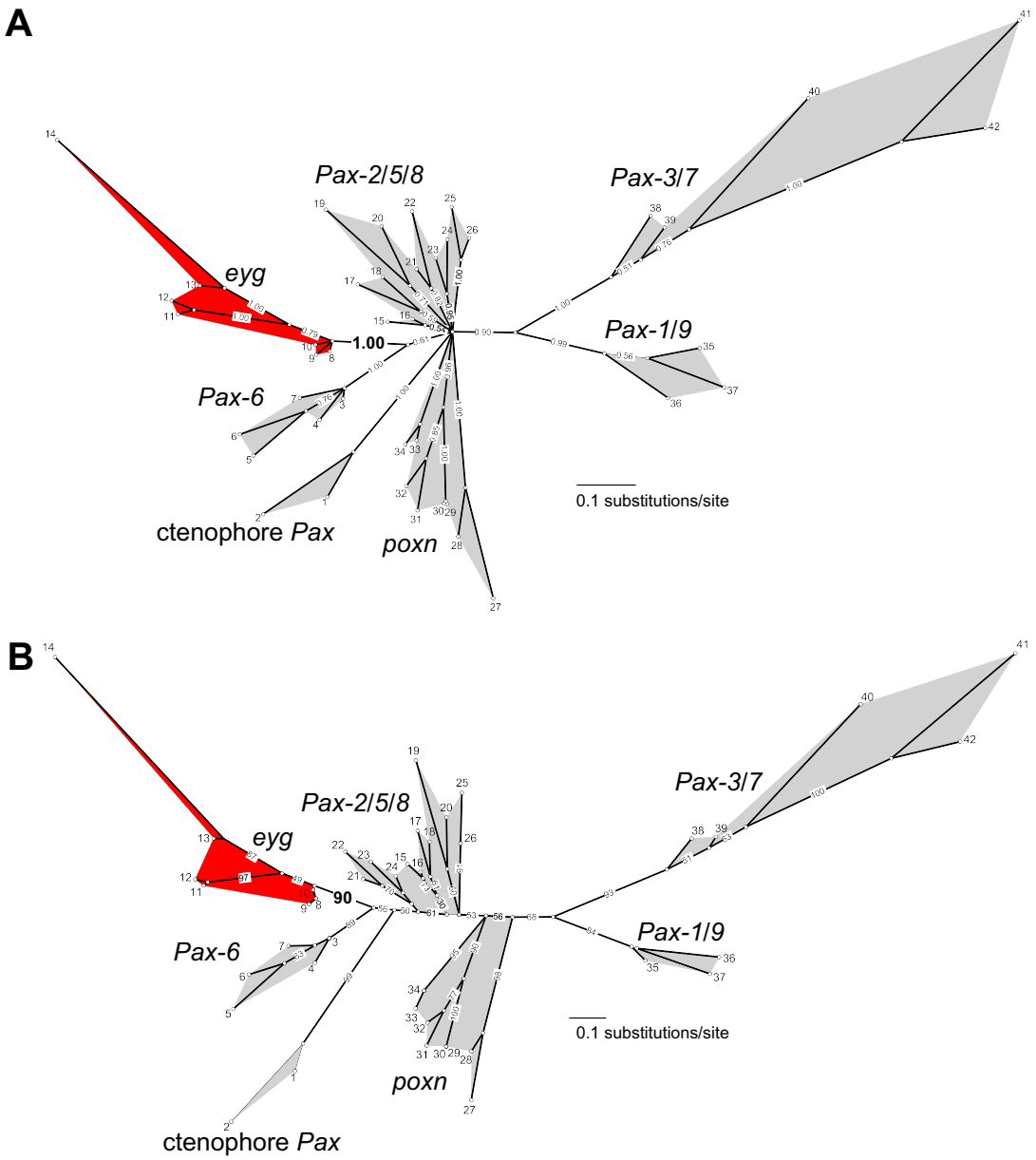

Fig. S2. Unrooted phylogenetic trees of *Pax* family including *eyg*-related genes. An alignment of the RED subdomain sequences of the PD, 57 aa sites in length, was used for inferring the Bayesian tree (*A*) and the ML tree (*B*). Five highly divergent sequences, human and zebrafish *Pax-4*, *Drosophila gooseberry*, and *Nematostella Pax-D* and *Pax-D3* were excluded from the analyses to avoid the long branch attraction artifact. The values on the branches indicate Bayesian posterior probabilities in *A* and extended local bootstrap probabilities in *B*. The values representing the support for the monophyly of *eyg* subfamily are in bold letters. The names of sequences and their GenBank accession numbers/National Center for Biotechnology Information RefSeq accession numbers/Ensembl Gene IDs are as follows: (*1*) *C. willeyi Pax-A* (AB239691); (*2*) *C. willeyi Pax-B* (AB239692); (*3*) human *Pax6* (M93650); (*4*) amphioxus *Pax6* (AJ223440); (*5*) *D. melanogaster twin of eyeless* (AF134350); (*6*) planaria *DjPax-6* (AB017632); (*7*) *D. melanogaster eyeless* (×79493); (*8*) *D. melanogaster eyegone* (AY099362); (*9*) *D. melanogaster twin of eyg* (AY069177); (*10*) mosquito *eyg*-like (AGAP011417); (*11*) *C. elegans eyg*-like (NP_496116); (*12*) *C. briggsae eyg*-like (XP_001678127); (*13*) *Cladonema radiatum Pax-E* (AB379658); (*14*) *Hydra magnipapillata Pax-E*-like; (*15*) human *Pax2* (U45255); (*16*) amphioxus *Pax-2* (AF053762); (*17*) *D. melanogaster D-Pax2* (AF010256); (*18*) marine ragworm *Pax258* (AJ505023); (*19*) *C. radiatum Pax-B* (AB379657); (*20*) *Hydra magnipapillata Pax-2/5/8* (AB007461); (*21*) *Tripedalia cystophora Pax-B* (AY280703); (*22*) *Chrysaora quinquecirrha Pax-B* (CQU96197); (*23*) *Acropora millepora Pax-B* (AF241310); (*24*) *Nematostella vectensis Pax-B* (AY730690); (*25*) *Ephydatia fluviatilis Pax-2/5/8* (AB007462); (*26*) *Microciona prolifera Pax-2/5/8* (AB379659); (*27*) *D. melanogaster pox neuro* (M86927); (*28*) mosquito *pox neuro* (AGAP008818); (*29*) *C. radiatum Pax-A* (AB379656); (*30*) *Hydra magnipapillata pox neuro* (AB007460); (*31*) *Acropora millepora Pax-A* (AF053458); (*32*) *Nematostella vectensis Pax-A* (AY730689); (*33*) *Acropora millepora Pax-C* (AF053459); (*34*) *Nematostella vectensis Pax-C* (AY730691); (*35*) human *Pax1* (AL035562); (*36*) amphioxus *AmphiPax-1* (U20167); (*37*) *D. melanogaster pox meso* (×16992); (*38*) human *Pax3* (NP_852124); (*39*) amphioxus *Pax3/7* (AF165886); (*40*) *D. melanogaster paired* (M14548); (*41*) *Acropora millepora Pax-D* (AF241311); and (*42*) *Nematostella vectensis Pax-D2* (DQ826414).

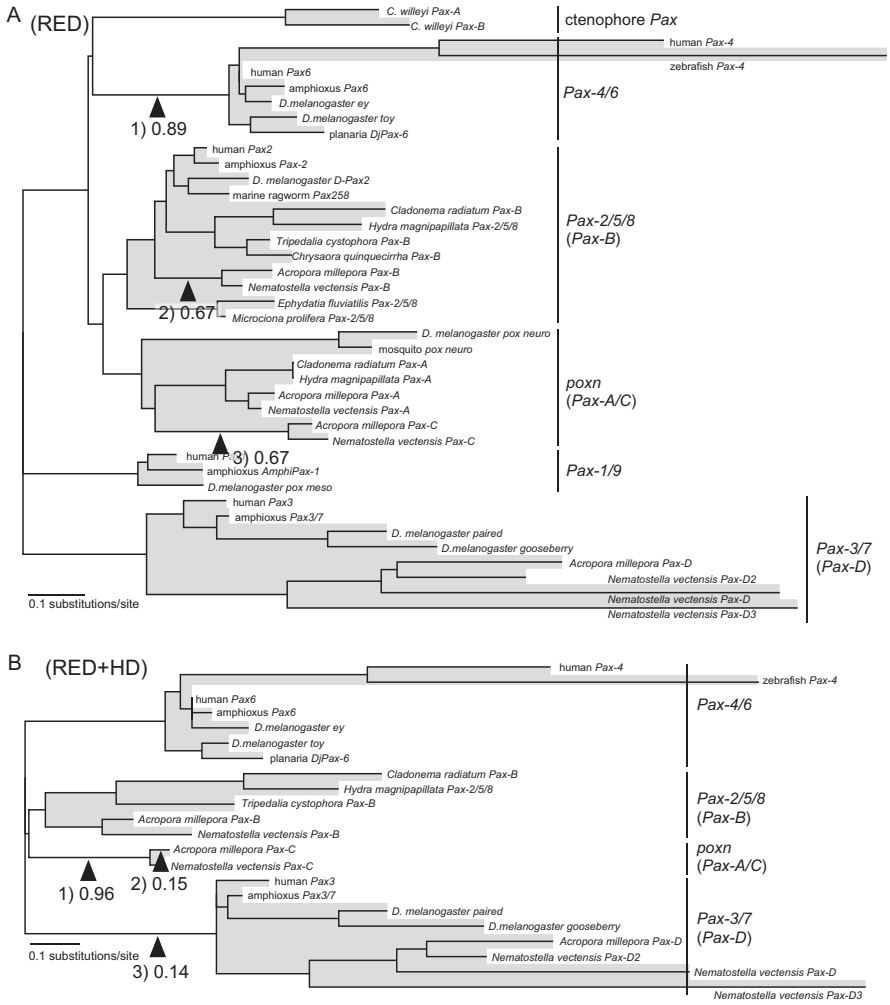

Fig. S3. Phylogenetic position of the *eyg* subfamily. All of the possible positions of *eyg* subfamily on the base topology, which is taken from Fig. 1*C* (main text), were analyzed by the ML method using the RED subdomain sequence of PD (*A*) and the RED subdomain plus the whole HD sequence (*B*). The three first topologies with the highest *P* values of the approximately unbiased test (1) are shown by arrowheads, with the respective *P* values.

1. Shimodaira H, Hasegawa M (2001) CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* 17:1246–1247.

Fig. S4. The alignment of the three CrPax-A PD binding site sequences identified by EMSA. The altered nucleotides in the mutated probes are underlined. The consensus sequence of chordate Pax-6 PD binding sites (1) is shown as sequence logos (http://weblogo.berkeley.edu/).

1. Czerny T, Busslinger M (1995) DNA-binding and transactivation properties of Pax-6: three amino acids in the paired domain are responsible for the different sequence recognition of Pax-6 and BSAP (Pax-5). *Mol Cell Biol* 15:2858–2871.

Table S1. Primers used for cloning *Cladonema Pax* genes

Name		Sequence	Corresponding amino acid sequence
Adapter Primer*†		5′-GGCCACGCGTCGACTAGTAC(T)$_{17}$-3′	-
AUAP*		5′-GGCCACGCGTCGACTAGTAC-3′	-
pax-1s	sense	5′-GGCGGTGTTAAYCARYTNGGNGG-3′	GGVNQLGG
pax-1a	anti-sense	5′-AACAAACGTTCNCKDATYTCCCA-3′	WEIRDRLL
pax-2s	sense	5′-TTCGTTAATGGNMGNCCNYTNCC-3′	FVNGRPLP
pax-2a	anti-sense	5′-GCAACTTTTGGYTTNSWNCCNCC-3′	GGSKPKVA
pax-3s	sense	5′-AATCCNACNATGTTYGCNTGGGA-3′	NPTMFAWE
pax-3a	anti-sense	5′-CGCTTATTNGARAACCANACYTG-3′	QVWFSNKR
pax-4s	sense	5′-AATCCIACIATITTYGCNTGG-3′	NPTIFAW
pax-4a	anti-sense	5′-CKATTISWRAACCANACYTG-3′	QVWFSNR
pax-5s	sense	5′-GGAGGAITIAAYCARYTNGGNGG-3′	GGVNQLGG
pax-5a	anti-sense	5′-AGCARIYKITBIYKIAYYTCCCA-3′	WEIRDRLL
pax-6a	anti-sense	5′-CGACCIARIATYTTISWIACRCANCC-3′	GCVSKILGR

*Invitrogen.
†Primer used for reverse transcription.

Table S2. Primer pairs used for real-time PCR

Gene		Sequence
CrPax-A	sense	5′-GGGAGGCCGCTACCGGATTACATGCGCC-3′
	anti-sense	5′-CGATCCGCCAATAGCACCAGGGCGTACAG-3′
CrPax-B	sense	5′-AGGCCTGGAGTGATAG-3′
	anti-sense	5′-GTCTGTTCCAAGTGTTAGCTC-3′
CrPax-E	sense	5′-GACATTCGCTTGTTGTCA-3′
	anti-sense	5′-GCACTCTCCAACAGATATCAATG-3′

Table S3. Primers used for generating EMSA probes

Gene	Probe	Sense	Antisense
CropG1	6	5′-GCGAATTCTGGGTCGTTTTGGCTGG-3′	5′-TAGAATTCTGTTCTGGAGATTGAGGG-3′
	5	5′-TTGAATTCAAGCGGTTTTCCCTCAA-3′	5′-CCGAATTCTAAAAGTAATTCAAATTTTAAAGATAAT-3′
	5.1	5′-ATTTTGTCTCAGGATGCTTGACAGGG-3′	Same as the antisense primer for probe 5
	5.1m*	5′-ATTTTGTCTCAGGAGTGTTGACAGGG-3′	Same as the antisense primer for probe 5
	4	5′-AAGAATTCTTTAAAATTTGAATTACTTTTAGTAAA-3′	5′-AAGAATTCATATTATTTGCAGAGCTCAAGAA-3′
	3	5′-TTGAATTCTTGAGCTCTGCAAATAATATTT-3′	5′-CTGAATTCGTGTCAACTTTCATGTATTATAAC-3′
	2	5′-AAGAATTCAAGTTATAATACATGAAAGTTGACAC-3′	5′-CGGAATTCCTTTATTACATCCTACTAAAATCG-3′
	1	5′-GGGAATTCTAAAGGAACCCCGCCTATAG-3′	5′-TTGAATTCATGTTTTTTTTTCGATTTCAGG-3′
CropN1	6	5′-AAGAATTCTTGTCCGCTAAATAGAGGTGT-3′	5′-AAGAATTCTATATAAGTATAATCATATACTTGGGTTATT-3′
	5	5′-CCGAATTCTATGATTATACTTATATAGTTCCATTTCAA-3′	5′-ATGAATTCTATTGACATAAATTTGCAATGA-3′
	4	5′-TTGAATTCCAAATTTATGTCAATAGAATAAATACAA-3′	0.5′-ATGAATTCATTATATACAATTACATGTTGTTATTC-3′
	4.1	Same as the sense primer for probe 4	5′-GGTGTGGCAAGTGGTGAAGTAG-3′
	4.2	5′-GCCACACCTCAGCAGCTTCC-3′	Same as the antisense primer for probe 4
	4.3	5′-CACTTGCCACACCTCAGC-3′	5′-ACGAATTCATGTTGTTATTCTATGCTCTTGC-3′
	4.3m*	5′-CACTAGCAAGACCTCAGCAGC-3′	Same as the antisense primer for probe 4.3
	3	5′-GAGAATTCACATGTAATTGTATATAATGCATGTA-3′	5′-GAGAATTCCTAAGAAACGATCTGTCAC-3′
	3.1	5′-CCATGTGTTCATTCATGATTTTTATAT-3′	5′-TTGAATTCCTATGAAGTTCTTTAATCAAAATTGTA-3′
	3.2	Same as the sense primer for probe 3.1	5′-AAGAATTCTGTCACACCTTTTGATTCTATG-3′
	3.2m*	Same as the sense primer for probe 3.1	5′-AAGAATTCTGCCATTGCATTTGATTCTATGAAG-3′
	2	5′-GAGAATTCAGGTGTGACAGATCGTTTCTTAG-3′	5′-AAGAATTCTATTAAGGTGGACACGCCT-3′
	1	5′-CGGAATTCTTGAAGGCGTGTCCACC-3′	5′-TCGAATTCTTAGAAGTTAGAAATGATGAAACAACA-3′

*Putative Pax binding sites mutated.
Restriction recognition sites underlined.