Available online at www.sciencedirect.com

SCIENCE DIRECT

Developmental Biology 278 (2005) 587–606

DEVELOPMENTAL BIOLOGY

www.elsevier.com/locate/ydbio

Genomes & Developmental Control

Comparative functional genomics revealed conservation and diversification of three enhancers of the *isl1* gene for motor and sensory neuron-specific expression

Osamu Uemura[a,b], Yohei Okada[c,d,e], Hideki Ando[a], Mickael Guedj[f], Shin-ichi Higashijima[g], Takuya Shimazaki[c,e], Naoichi Chino[b], Hideyuki Okano[c,e], Hitoshi Okamoto[a,e,*]

[a]*Laboratory for Developmental Gene Regulation, Brain Science Institute, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan*
[b]*Department of Rehabilitation Medicine, Keio University, School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-0016, Japan*
[c]*Department of Physiology, Keio University, School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-0016, Japan*
[d]*Department of Neurology, Graduate School of Medicine, Nagoya University, 65 Tsurumai-cho, Showa-ku, Nagoya, Aichi 466-8550, Japan*
[e]*Core Research for Evolutional Science and Technology (CREST), Japan Science and Technology Corporation (JST), 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan*
[f]*Département Biosciences, Filaire Bioinformatique et Modélisation, Institut National des Sciences Appliquées (INSA) de Lyon, France*
[g]*Department of Developmental Neurophysiology, Center for Integrative Bioscience, Okazaki National Research Institutes, 38 Nishigonaka, Myodaiji, Okazaki, Aichi 444-8585, Japan*

Received for publication 10 October 2004, received 22 November 2004, accepted 23 November 2004
Available online 23 December 2004

Abstract

Islet-1 (Isl1) is a member of the Isl1 family of LIM-homeodomain transcription factors (LIM-HD) that is expressed in a defined subset of motor and sensory neurons during vertebrate embryogenesis. To investigate how this specific expression of *isl1* is regulated, we searched for enhancers of the *isl1* gene that are conserved in vertebrate evolution. Initially, two enhancer elements, CREST1 and CREST2, were identified downstream of the *isl1* locus in the genomes of fugu, chick, mouse, and human by BLAST searching for highly similar elements to those originally identified as motor and sensory neuron-specific enhancers in the zebrafish genome. The combined action of these elements is sufficient for completely recapitulating the subtype-specific expression of the *isl1* gene in motor neurons of the mouse spinal cord. Furthermore, by direct comparison of the upstream flanking regions of the zebrafish and human *isl1* genes, we identified another highly conserved noncoding element, CREST3, and subsequently C3R, a similar element to CREST3 with two CDP CR1 recognition motifs, in the upstream regions of all other *isl1* family members. In mouse and human, CRESTs are located as far as more than 300 kb away from the *isl1* locus, while they are much closer to the *isl1* locus in zebrafish. Although all of zebrafish CREST2, CREST3, and C3R activate gene expression in the sensory neurons of zebrafish, CREST2 of mouse and human does not have the sequence necessary for sensory neuron-specific expression. Our results revealed both a remarkable conservation of the regulatory elements regulating subtype-specific gene expression in motor and sensory neurons and the dynamic process of reorganization of these elements whereby each element increases the level of cell-type specificity by losing redundant functions with the other elements during vertebrate evolution.

Keywords: Enhancer; Islet-1; Islet-2; Islet-3; Motor neuron; Sensory neuron; Subtype specification; Comparative functional genomics; Zebrafish; Mouse

Introduction

In the development of the nervous system, many diverse subpopulations of neurons are generated, and the elucidation of genetic regulatory elements that control these events will greatly facilitate our understanding of how developing brain

* Corresponding author. Laboratory for Developmental Gene Regulation, Brain Science Institute, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan. Fax: +81 48 467 9714.
E-mail address: hitoshi@brain.riken.jp (H. Okamoto).

0012-1606/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.ydbio.2004.11.031

can self-generate such enormous complexity. As a first step to address this issue, we identified the *cis*-acting regulatory elements of the *islet-1* (*isl1*) gene that control expression in the motor and sensory neurons that are conserved throughout vertebrate evolution.

Islet-1 (Isl1) is among the LIM-homeodomain (LIM-HD) proteins expressed at the earliest stage of neural differentiation, i.e., motor and sensory neuron development, and is highly conserved during evolution (Ericson et al., 1992; Inoue et al., 1994; Jackman et al., 2000; Thor and Thomas, 1997). Isl1 and other LIM-HD transcription factors function together, and their combinatorial expression patterns (LIM codes) define the subtype specificities of motor neurons with respect to differences in target muscles to which they project (Kania et al., 2000; Segawa et al., 2001; Sharma et al., 2000; Thaler et al., 2002; Tsuchida et al., 1994). Isl1 is initially expressed in all motor neurons immediately after they exit the mitotic cycle, but is later restricted to specific subsets (Inoue et al., 1994; Tsuchida et al., 1994). In the embryonic chick, rat, and mouse spinal cord, expression of the *isl1* gene is restricted to spinal motor neurons that innervate the dorsal trunk muscles, the ventral trunk muscles, and the ventral limb bud muscles, respectively (Tsuchida et al., 1994). However, the precise mechanism regulating this fine combinatorial expression of LIM-HD proteins in each motor column remains unclear.

We previously identified two 15-kb fragments, termed CM and SS, which are located 10 and 54 kb, respectively, downstream of the zebrafish *isl1* gene (Higashijima et al., 2000). CM functions as an enhancer to activate expression of the immediately downstream *GFP* gene in cranial motor neurons and subset of the secondary spinal motor neurons innervating the dorsal part of the trunk muscles, and SS as an enhancer for expression in the primary sensory neurons. Here, we further localized the regulatory elements within CM and SS to identify the evolutionarily conserved *isl1* enhancers for specific expression of genes in motor and sensory neurons.

The regulatory elements were identified within CM and SS, and termed zCREST1 (zebrafish conserved regulatory element for *islet-1* 1) and zCREST2 using the newly optimized condition for the highly efficient transient expression assay in zebrafish embryos, and we further confirmed their activities by establishment of transgenic zebrafish. Quite intriguingly, contrary to our original expectation that zCREST2 should act as the enhancer specific for primary sensory neurons, we found in the stable transgenic zebrafish lines that zCREST2 also activates gene expression in subsets of spinal motor neurons innervating the abductor muscle of the pectoral fin bud and the ventral trunk muscles in addition to the expression in primary sensory neurons.

Two elements in the chick, mouse, and human genome were further identified that show very high levels of similarity to the zebrafish CREST1 and CREST2 but were located significantly further away from the *isl1* gene than the corresponding elements in the zebrafish genome. We further confirmed that combined action of zebrafish or human CREST1 and CREST2 is sufficient to completely recapitulate the motor neuron subtype-specific expression of *isl1* in the mouse spinal cord.

In contrast to the strong conservation of the roles for CREST1 and CREST2 in motor neuron subtype-specific gene activation, the sensory neuron-specific activity of zebrafish CREST2 depends on the sequence that is not conserved in evolution, and human CREST2 does not have this activity. This result leads us to the identification of another sensory neuron-specific regulatory element, CREST3, upstream of the *isl1* gene, by searching for the highly conserved sequence upstream of the zebrafish and human *isl1* gene. Furthermore, we found similar sequences to CREST3, termed C3R (CREST3 related), in the upstream flanking regions of other *isl1* gene family members in vertebrates and confirmed its activity as a sensory neuron-specific-enhancer in zebrafish. CREST3 activity is abrogated by mutations in two CDP CR1 recognition motifs.

Our results revealed both a remarkable conservation of the regulatory elements regulating subtype-specific gene expression in motor and sensory neurons and the dynamic process of reorganization of these elements during vertebrate evolution.

Materials and methods

Animals

Zebrafish were maintained as described elsewhere (Westerfield, 1995). The mice were maintained by the Research Resource Center (RRC) at RIKEN BSI. Embryonic day (E) 0.5 was defined as the midday time of the day when a vaginal plug was observed. All animal experiments were carried out according to the guidelines for animal experimentation at RIKEN.

Genomic DNA isolation

Human genomic DNA was isolated as described elsewhere (Sambrook et al., 1989) from fresh blood collected from a volunteer. All experiments involving human subjects were carried out according to guidelines for ethical research involving human subjects at RIKEN.

Plasmid construction

The 4.1-kb fragment (termed ICP for the *isl1* core promoter) upstream of the zebrafish *isl1* gene has been reported to contain the basal promoter and the enhancer for expression in hatching gland cells (Higashijima et al., 2000). To avoid a possible interference from this hatching gland-specific enhancer activity in the transient assay, we used the promoter of the zebrafish *heat shock protein 70* (*hsp70*) gene.

This promoter functions as a basal promoter at ambient temperatures (data not shown; Halloran et al., 2000). For the isl1:mRFP1 plasmid construction, *mRFP1* was subcloned into the *Bam*HI and *Eco*RI sites of pCS2 (Campbell et al., 2002). Then *mRFP1* was amplified using the PCR method and replaced with the *GFP* gene of the isl1:GFP plasmid (Fig. 2B; Higashijima et al., 2000). Primers used to amplify *mRFP1* were F, 5′GGATCCACCATGGCCTCCTCCGAGGAC3′ and R, 5′CTCGAGAGGCCTGAATTCTTAGG-C3′. This reverse primer for *mRFP1* was designed to target for multiple cloning sites of the pCS2. For isl1-PLAP plasmid construction, the *GFP* gene of the isl1:GFP plasmid (Fig. 2B; Higashijima et al., 2000) was replaced with the PCR-amplified *PLAP* gene (Lobe et al., 1999). The primers designed to amplify the coding region of the *PLAP* gene were F, 5′GGGGTACCCTGCCTCGCCACTGTCCTGC3′ and R, 5′GGAGATCTCAGGGAGCAGTGGCCGTCTCC3′.

Construction of deletion series

Deletion fragments of CM were generated by restriction digestion. Restriction enzymes used in this experiment are shown in Fig. 2C. Each fragment was blunt ended and ligated with *Eco*RI linker sequences to be subcloned into appropriate vectors. The enhancer activity of SS was further localized within the 4-kb subfragment (termed SSX) (Figs. 2A and D, and data not shown). It was cloned into *Xho*I and *Eco*RI sites of pBlueScript II SK+ in order to generate deletion fragments. The deletion fragments were prepared by using a "Deletion kit" (Takara) according to the manufacturer's instructions.

DNA injection into one-cell-stage zebrafish embryos

Plasmid DNA was prepared using the QIAprep spin miniprep kit (Qiagen). Plasmid injection was performed as described elsewhere (Higashijima et al., 1997).

Fluorescence microscopic examination of zebrafish embryos

The initial examination and subsequent detailed examination were performed as described elsewhere (Higashijima et al., 2000).

Comparative genomics

The genome databases from the ensembl web site (http://www.ensembl.org) and the whole-genome shotgun sequences of fugu and zebrafish (http://www.ncbi.nlm.nih.gov) were used to BLAST search for sequences similar to zCREST1 and zCREST2. For identification of CREST3, 200 kb and 1 Mb of genomic sequences, which lie immediately upstream of the zebrafish and human *isl1* gene, were retrieved from the ensemble web site (http://www.ensembl.org). Repetitive sequences were then removed from the retrieved zebrafish sequences using RepeatMasker (http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml). The sequences were then aligned using MultiPipMaker (Schwartz et al., 2000, http://bio.cse.psu.edu/pipmaker/). To distinguish whether the matched regions were noncoding or coding sequences, optional exon analysis was performed according to the instructions shown on the web page. Nine matched sequences were identified, and similar sequences to these nine regions were subsequently sought in the mouse genome using the BLAST search algorithm with default parameter set. Three of the sequences were predicted to encode genes and no similar sequences were found in five of them. One sequence, termed CREST3, was found to be conserved in zebrafish, mouse, and human genomes. To identify chick and fugu CREST3, we performed a BLAST search against the chick genome database and the whole-genome shotgun sequences of fugu using standard parameters. To find C3R, hCREST3 was divided into fragments by 100 bp, making 50 bp overlapped with following fragment. Each fragment was then searched for similar sequences using BLAST search at the ensembl web site and the whole-genome shotgun sequences of fugu.

TRANSFAC analysis

TRANSFAC analysis was performed with the program Match (http://www.gene-regulation.com/) with a parameter to minimize false negatives.

Molecular cloning

hCREST1, hCREST2, zCREST2, zCREST3, and zC3R were cloned using PCR. Zebrafish *lmx1b* was cloned by the nested degenerate PCR method using DNA purified from the cDNA library of 25–26 h postfertilization (hpf) zebrafish embryos (Hirate et al., 2001). PCR primers are as follows. Specific primers including *Eco*RI linker sequences (F-5′GGGAATTCAAACAGATGCACCTACCTC3′, R 5′GGGAATTCGGACATATGGCTAGAGTGTG3′) were designed to amplify approximately 720 nucleotides of genomic DNA (hCREST1) containing the region highly similar to zCREST1. The following primers were designed to amplify approximately 600 nucleotides of hCREST2: F, 5′GGGAATTCGAGGGTGTCTTCATCACTTCC3′; R, 5′GGGAATTCGAAAGCAGGAGAAACTCCTTG3′. To amplify zCREST2, one deletion fragment of SSX termed SSd25 in pBluescript II SK+ was used as template. Specific primers were designed as follows: zCREST2F1, 5′GGGAATTCTACCGGGCCCCCCCTCGAG3′; zCREST2R1, 5′GGGAATTCCAGCACCATAATTCAC3′. These primers were also used to generate deletion fragments of zCREST2. Specific primers for zCREST3 were as follows: F, 5′GGGGTAACAGGATGTGACACGTCGTCTGC3′; R, 5′GGGGCCTGCTGCTGGTGTCATTTACTGG3′. PCR primers for zC3R were F, 5′AACGTAATGAACACATGA-

ATAC3′; R, 5′CCAAATGTACTGTTAGGAGGATTG3′. The degenerate primers used to amplify the *lmx1b* gene were F1, 5′CTCACCACCAGCTGCTACTTCC3′; F2, 5′TGCATGGAGAAGATCGC3′; R1, 5′CGGTCRATGGGRTTYCCCAC3′; R2, 5′TCTGCATKGAGTCSAGBCGGTC3′.

Generation of mutant enhancers and deletion fragments of zCREST2

Nucleotide conversion of putative transcription factor recognition sites of zCREST1 and zCREST3 was done using the inverse PCR method. Primers used in this study were as follows. For zCREST1m: F, 5′GGTTGATTTACCAATATAAAATGCAAATG3′; R, 5′CCAGCACCCATAGCTTTTAACTTGATTGAC. For zCREST3m1: F, 5′GGTATTTTAAAACGCCGTAATGTGAC3′; R, 5′CCCATGAGAGTGCATTTAGCCAGGAG3′. For zCREST3m2: F, 5′GGTTTGCAGGGGCCTCGGCTTCAGG3′; R, 5′CCCGTATCTGTCCTCCGAGGGTTTCAC 3′. To generate deletion fragment, the PCR method was used to amplify the specific region of zCREST2 (see also Figs. 5B and C). Primers used in this study were as follows: zCREST2F2, 5′GGGCTCGAAGGAATATGTACACTTTTCAAG3′; zCREST2R2, 5′GGGAATTCACTTTACCATAATATG3′; zCREST2R3, 5′GGGAATTCGGAGGGTTAATCGTCTGAGTC3′. Primers zCREST2F1 and zCREST2R1, described above, were also used.

Generation of transgenic animals

Transgenic zebrafish and mice were generated as described elsewhere (Higashijima et al., 1997; Hogan et al., 1994). We generated one *Tg(zCREST1/isl1-PLAP)* line, two *Tg(mCREST1/isl1-PLAP)* lines, and two *Tg(hCREST1/isl1-PLAP)* lines, and all transgenic mice showed virtually the same expression pattern. For CREST2, two lines of *Tg(zCREST2/isl1-PLAP)* and three lines of *Tg(hCREST2/isl1-PLAP)* were generated. All lines of *Tg(zCREST2/isl1-PLAP)* embryos showed the same expression pattern of PLAP. One line of *Tg(hCREST2/isl1-PLAP)* embryos did not show any PLAP expression as revealed by PLAP staining, while the other two lines of *Tg(hCREST2/isl1-PLAP)* embryos showed the same pattern of PLAP expression with each other.

In situ hybridization, PLAP staining, immunofluorescence, and rhodamine phalloidin staining

In situ hybridization was performed as described elsewhere (Westerfield, 1995). Sections of in situ-hybridized embryos were prepared by refixing hybridized embryos, embedding them in JB-4 resin, and cutting them into 7 μm. PLAP staining was performed as described elsewhere (Lobe et al., 1999). A mixture of monoclonal antibodies (40.2D5 and 39.4D5, DSHB) was used to detect Isl1 protein, as described elsewhere (Segawa et al., 2001). Anti-GFP (Santa Cruz), anti-PLAP (Biogenesis), anti-Lhx3 (67.4E12, DSHB), and anti-Lhx1 (4F2, DSHB) antibodies were also used. Double labeling analysis was performed with a Zeiss 510 confocal microscope using Alexa fluor 488- and 533-conjugated secondary antibodies (Molecular Probes). For rhodamine phalloidin staining, embryos were fixed in 4% PFA at room temperature for 1 h. The fixed embryos were immersed at room temperature for 30 min in a 10^{-7} M of rhodamine phalloidin solution (Sigma) in PBS with 0.7% Triton X and then washed with PBS three times. For sectioning of transgenic zebrafish, embryos were fixed with 4% PFA overnight at 4°C. The fixed embryos were embedded in OCT (Sakura) and cut at 18-μm intervals with a cryostat. For sectioning of transgenic mice, embryos were grown until the optimal stages and fixed with 4% PFA for 5–30 min at 4°C. OCT-embedded embryos from E11.5 to E14.5 were sectioned at 14- to 20-μm intervals.

Results

Injection of circular plasmid at a very low concentration dramatically improves the efficiency of transient gene expression in zebrafish embryos

Transgenesis of zebrafish is now a common research technique. Production of stable transgenic zebrafish and the faithful expression of transgenes using promoter sequences of zebrafish origin, such as the zebrafish *GATA-1* promoter (Long et al., 1997) or the zebrafish α-*actin* promoter (Higashijima et al., 1997), are well established. However, the technique is inconvenient for identifying regulatory elements, simply because it requires a great deal of effort. One solution to this problem was derived from the enhancer analysis of the zebrafish *sonic hedgehog* gene. In this study, the investigators employed transient transgene expression assays to identify tissue-specific regulatory elements (Müller et al., 1999). However, transient transgene expression is known to entail strong mosaicism, thus superimposing the transgene expression patterns obtained from a large number of injected embryos has been the only way to obtain a spatial pattern of enhancer activity (Müller et al., 1999; Westerfield et al., 1992). To overcome this drawback, we modified the technique simply by injecting a very low concentration of circular plasmid solution, in the expectation that it would result in ubiquitous diffusion of the plasmid DNA within the embryonic cytoplasm at the one-cell stage without having a toxic effect on dividing cells. Injected plasmids are also known to replicate intensively during the blastula stage, and thus each cell should receive a sufficient concentration of plasmid no matter how little was injected into the one-cell stage embryo (Stuart et al., 1988).

We therefore injected 1–2 nl of the 23-kb CM-isl1:GFP plasmid, which contains the enhancer element of the *isl1* gene driving expression of GFP in the subset of cranial motor and sensory neurons, at a concentration of 5 ng/μl,

which is approximately one order of magnitude lower than conventionally used to create transgenic fish (Higashijima et al., 1997; Long et al., 1997). In more than 80% of the injected embryos, expression of GFP was observed in most of the types of neurons that would express GFP in the stable transgenic line carrying the same plasmid as the transgene (Fig. 1A; Higashijima et al., 2000).

This method was further assessed using α-action:GFP, which is capable of driving GFP expression in muscle cells (Higashijima et al., 1997). The same effect was observed with the injection of CM-isl1:GFP; the lower the concentration of plasmid DNA we injected, the higher the contribution of GFP-positive cells (Figs. 1B and C). However, the contribution of GFP-positive cells in craniofacial and fin muscles was much lower than the trunk muscles (Fig. 1B and data not shown). Since the former muscles begin to differentiate at later stages than the latter muscles (Weinberg et al., 1996), the efficiency of gene expression by plasmid injected using this method may decrease by 40 hpf.

Despite the limitation of the effective period, this improved transgene expression method enables a rapid and precise assay of the activity of *cis*-acting regulatory elements of a zebrafish gene by transient expression.

Identification of the minimal enhancers of the zebrafish isl1 gene for expression in cranial motor neurons and primary sensory neurons

We previously showed two 15-kb fragments, termed CM and SS, which are located 10 and 54 kb, respectively, downstream of the zebrafish *isl1* gene, function as enhancers to activate expression of the immediately downstream *GFP* gene in cranial motor neurons and in primary sensory neurons (Fig. 2A; Higashijima et al., 2000). To identify the minimal enhancers required for such specific gene expression, we generated a series of deletion fragments of CM and SS and inserted each fragment immediately upstream of the zebrafish *heat shock protein 70* (*hsp70*) promoter and the *GFP* gene (Figs. 2B, C, and D; Halloran et

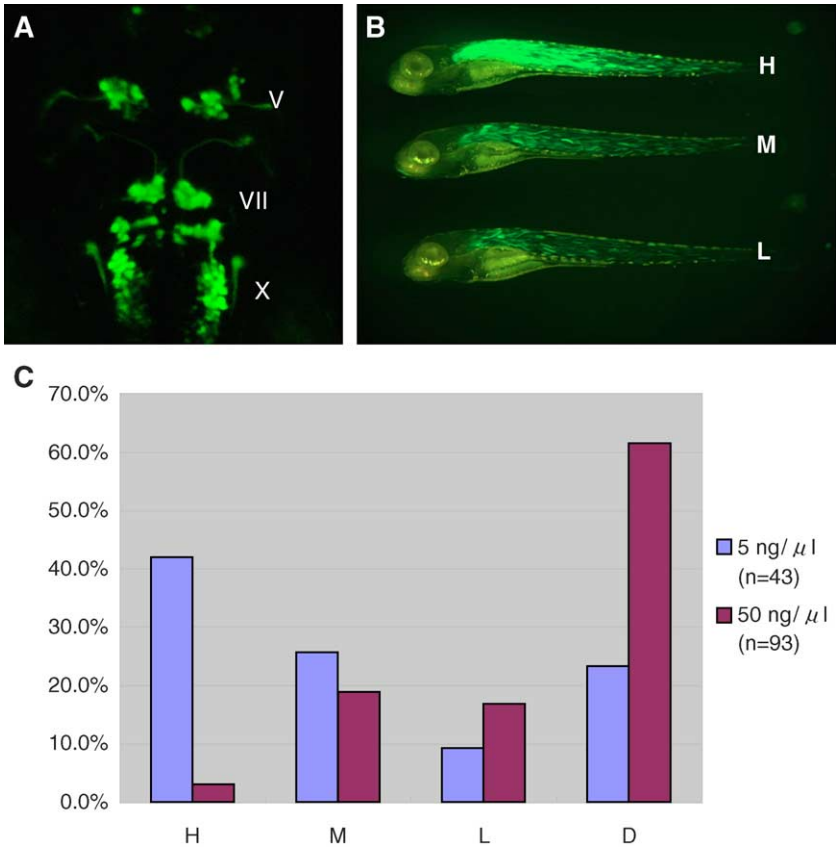


Fig. 1. Effectiveness of transient transgene expression by injection of circular plasmid at a very low concentration. (A) Dorsal view of the 48-hpf embryo injected with the CM-isl1:GFP plasmid at a very low concentration. Note that most of the cranial motor neurons transiently expressed GFP at a level comparable to that of the embryo stably transgenic for the same DNA construct (Higashijima et al., 2000). V, trigeminal motor neurons; VII, facial motor neurons; X, vagus motor neurons. (B and C) The embryos injected with the α-actin:GFP plasmid were subdivided into four groups according to the degree of contribution of GFP-positive cells to the whole embryo, i.e., to groups H, M, and L, which represent embryos that express GFP in most muscles, scattered muscle cells plus some clusters of muscle cells, and in scattered separate muscle cells alone, respectively, and group D for dead embryos. (B) Representative embryos from each group are shown. Note that even in the embryo from group H, which expressed high levels of GFP in trunk muscles, craniofacial and fin muscles rarely expressed GFP. (C) The bar graph shows the percentage of embryos belonging to each group among all the embryos injected with the α-actin:GFP plasmid. Injection was performed at a concentration of 5 ng/μl (blue bar) or 50 ng/μl (red bar).

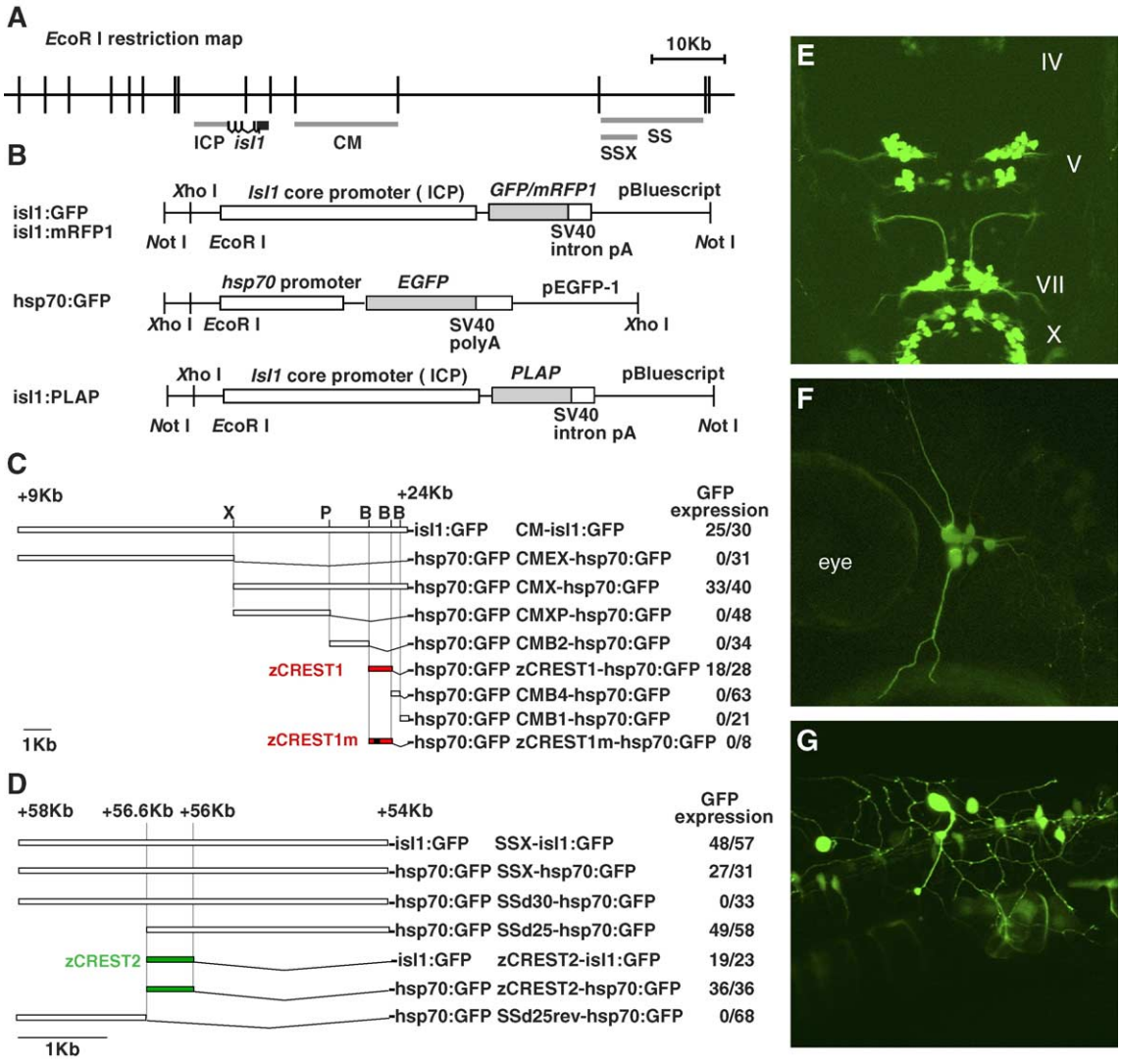


Fig. 2. Genomic structure, plasmid constructs, and results of the transient transgene expression assay. (A) *Eco*RI restriction map of the genomic region flanking the zebrafish *isl1* gene. The lines under the map indicate the regulatory regions of the *isl1* gene. ICP, the *isl1* gene core promoter; CM, an enhancer for the cranial motor neurons; SS and SSX, enhancers for the primary sensory neurons and the fin motor neurons. The thick line, *isl1*, indicates the transcribed unit of the *isl1* gene. Note that the *isl1* gene is composed of six exons divided by five introns. (B) Maps of plasmids used in this study. Each enhancer fragment was inserted immediately upstream of the promoter sequences. *PLAP*, *human placental alkaline phosphatase gene*. (C and D) Each plasmid construct shown here was injected into one-cell stage zebrafish embryos to examine its enhancer activity. The number on the right side of the name of each construct indicates the number of embryos exhibiting expression of GFP in cranial motor neurons (C) or primary sensory neurons (D) among all injected embryos. X, *Xho*I; P, *Pst*I; B, *Bst*XI. (C) Note that nucleotide conversion in TAAT to GGGG (termed zCREST1m, asterisk in Fig. 5A) completely inactivated its enhancer activity. (E) Dorsal view of a 48-hpf embryo injected with the zCREST1-hsp70:GFP plasmid and expressing GFP in the majority of cranial motor neurons IV, trochear motor neurons. (F and G) Lateral views of 32-hpf embryos injected with the zCREST2-hsp70:GFP plasmid and expressing GFP in the trigeminal ganglion neurons (F) and Rohon-Beard neurons (G).

al., 2000). All plasmid constructs shown in Figs. 2C and D were injected into one-cell stage embryos as described above. The injected embryos were tested for their enhancer activity.

Two fragments, the 800-bp zCREST1 (the zebrafish conserved regulatory element of the *islet-1* gene 1) and the 600-bp zCREST2, which lie approximately 23 and 56 kb, respectively, downstream of the *isl1* gene locus, were sufficient to drive GFP expression in cranial motor neurons (Figs. 2C and E) and primary sensory neurons (Figs. 2D, F, and G), respectively. No fragments derived from CM or SS that did not contain zCREST1 or zCREST2 sequences showed any such enhancer activity.

In transgenic zebrafish, zCREST2 drives expression of GFP in the subset of secondary motor neurons that innervates abductor muscles in the pectoral fin

Using the zCREST2-hsp70:GFP plasmid construct, we produced three different stable transgenic zebrafish lines, termed *Tg(zCREST2-hsp70:GFP)*$^{rw011a,\ b,\ and\ c}$, all of which showed the same pattern of expression of GFP. GFP

expression was observed in trigeminal ganglion neurons (Fig. 3A) and in Rohon-Beard neurons (asterisks in Fig. 3B) of these transgenic lines by 32-h postfertilization (hpf). After 36 hpf, GFP expression became prominent in the ventral spinal cord, and confocal microscopic observation revealed that the GFP-positive cells were the secondary

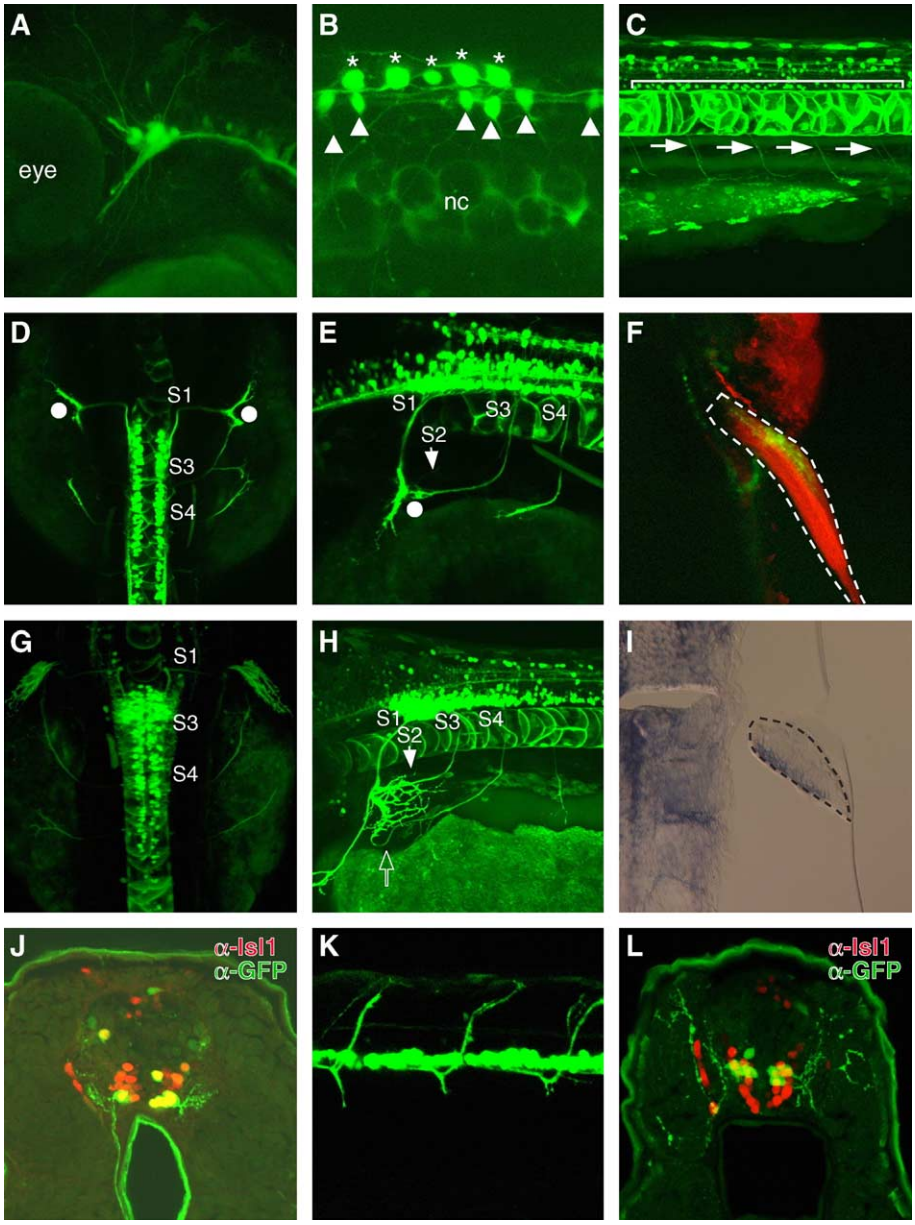


Fig. 3. Sensory and motor neuron-specific GFP expression driven by zCREST2 in stable transgenic zebrafish. zCREST2 drives GFP expression both in the primary sensory neurons and in subsets of motor neurons innervating the ventral trunk muscles and the abductor pectoral fin muscle in transgenic zebrafish embryos. (A–E, G, and H) Lateral (A–C, E, and H) and dorsal (D and G) views of *Tg(zCREST2-hsp70:GFP)*^{rw011} embryos. At 36 hpf, GFP expression was observed in the trigeminal ganglion neurons (A) and in Rohon-Beard neurons (asterisks in B). GFP expression was detected in other tissues, such as the notochord and some commissural interneurons in the spinal cord (arrowheads). nc, notochord. (C) GFP was expressed in the spinal motor neurons (bracket) that project their axons ventrally (arrows) by 72 hpf. (D, E, G, and H) The GFP-positive fin motor neurons were distributed from the spinal segment 1 (S1) through 4 (S4), as described for Medaka fish embryos (Okamoto and Kuwada, 1991). (D and E) At 54 hpf, the GFP-positive fin motor nerves from S1 to S3 had reached the base of the fin bud without entering (white dots). Meanwhile, the nerve from S4 was still extending toward the fin bud. (G and H) By 72 hpf, the GFP-positive nerves from S1 to S3 had completely converged and entered the pectoral fin at the dorsal end of its base, while the nerve from S4 had entered at the opposite end (open arrow). (F) Ventral view of the 120-hpf *Tg(zCREST2-hsp70:GFP)*^{rw011} embryo stained with rhodamine phalloidin. The GFP-positive fin motor nerves terminated in the abductor pectoral fin muscle. The pectoral fin is enclosed in a broken line. (I) Expression of the *lmx1b* gene in the adductor pectoral fin muscle of the wild-type 50-hpf embryo. The developing fin bud is enclosed with a broken line. (J) GFP-positive spinal motor neurons in 72-hpf *Tg(zCREST2-hsp70:GFP)*^{rw011} embryo were the ventral subgroup of the Isl1-positive spinal motor neurons. (K and L) In 72-hpf Isl1-GFP-transgenic zebrafish, GFP-positive spinal motor neurons projected their axons dorsally (K, lateral view). (L) GFP-positive spinal motor neurons in the 72-hpf Isl1-GFP-transgenic zebrafish embryo were the dorsal subgroup of Isl1-positive spinal motor neurons.

motor neurons that extend their axons into the ventral trunk muscles (bracket and arrows in Fig. 3C) and the pectoral fin muscles (Figs. 3D–H). The GFP-positive fin motor neurons were distributed from the spinal segment 1 (S1) through 4 (S4), as described for Medaka fish embryos (Okamoto and Kuwada, 1991). At 54 hpf, the GFP-positive fin motor nerves from S1 to S3 had reached the base of the fin bud without entering (white dots in Figs. 3D and E). Meanwhile, the nerve from S4 was still extending toward the fin bud. By 72 hpf (Figs. 3G and H), the GFP-positive nerves from S1 to S3 had completely converged and entered the pectoral fin at the dorsal end of its base, while the nerve from S4 had entered at the opposite end (open arrow in Fig. 3H).

The pectoral fin contains two fin muscles, an abductor muscle (outside) and an adductor muscle (inside), as previously reported in Medaka fish larvae (Okamoto and Kuwada, 1991). Staining of muscles of the *Tg(zCREST2-hsp70:GFP)*^{rw011} embryos at 120 hpf with rhodamine phalloidin demonstrated that the GFP-positive pectoral fin nerves terminated in the abductor muscle (Fig. 3F). The LIM-homeobox gene *Lmx1b* is expressed in the dorsal muscle precursors of the mouse limb bud (Kania et al., 2000). To determine whether the abductor and adductor muscles in the pectoral fin of zebrafish embryo correspond to the ventral and dorsal muscle precursors in mouse limb buds, we investigated the expression of zebrafish *lmx1b* in the pectoral fin bud at 50 hpf. In situ hybridization revealed expression of *lmx1b* in the adductor (inside) pectoral fin muscle (Fig. 3I), demonstrating that the GFP-positive secondary motor neurons innervate only the *lmx1b*-negative abductor pectoral fin muscle, which is equivalent to the ventral muscle precursors in the limb bud of the mouse.

Conservation of the isl1 gene enhancer sequences in vertebrate genomes

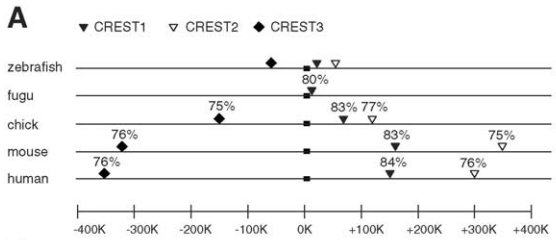
At the thoracic level of the embryonic chick spinal cord, motor neurons located in the MMCm (the medial subdivision of the medial motor column), which innervate the dorsal axial muscles, express Isl1, Isl2, and Lim3, while those in the MMCl (the lateral subdivision of the medial motor column), which innervate the ventral axial muscles, express only Isl1 and Isl2 (Tsuchida et al., 1994). At the brachial and lumbar levels, those in the LMCm (the medial subdivision of the lateral motor column), which innervate the ventral muscle mass of the limb bud, express Isl1 and Isl2, while those in the LMCl (the lateral subdivision of the lateral motor column), which innervate the dorsal muscle mass, ultimately express Lim1 and Isl2 (Tsuchida et al., 1994). Therefore, expression of the *isl1* gene is restricted to spinal motor neurons in the MMCm, MMCl, and LMCm, which innervate the dorsal trunk muscles, the ventral trunk muscles, and the ventral limb bud muscles, respectively (Tsuchida et al., 1994) (Fig. 4C).

As described above, the *Tg(zCREST2-hsp70:GFP)*^{rw011} larvae were found to express GFP in secondary motor neurons innervating the ventral trunk muscles and the abductor pectoral fin muscle (Figs. 3C–H and 4B and C). And the Isl1-GFP transgenic embryos expressed GFP under the control of the CM fragment in secondary motor neurons innervating the dorsal trunk muscles (Figs. 3K and 4B and C; Higashijima et al., 2000; Segawa et al., 2001). Therefore, taken together with the fact that *isl1* is expressed in the subset of motor neurons innervating the dorsal trunk muscles, the ventral trunk muscles, and the ventral limb bud muscles in mouse and chick spinal cord as mentioned above, the muscles innervated by GFP-positive secondary motor neurons in these two transgenic zebrafish embryos correspond exactly to the group of muscles innervated by Isl1-positive motor neurons in the mouse and chick. Based on these findings, we hypothesized that expression of the *isl1* gene is regulated by highly conserved mechanisms throughout evolution.

To address this hypothesis, we determined the nucleotide sequences of zCREST1 and zCREST2, and performed BLAST searches to identify similar sequences in other vertebrate genomes. We identified the noncoding genomic sequences in the fugu, chick, mouse, and human genome, which showed a high degree of similarity (approximately 80%) with zCREST1, and these sequences were termed fCREST1, cCREST1, mCREST1, and hCREST1, respectively (Figs. 4A and 5A). Sequences that were highly similar to zCREST2 were also found in the chick, mouse, and human genomic databases, and they were termed cCREST2, mCREST2, and hCREST2, respectively (Figs. 4A and 5B). No similar sequences to zCREST2 were identified in the fugu genome, probably because of the gaps in fugu genomic sequences in the database. We also searched elements that shared partial homology with CREST1 or 2 in all available genomic databases, but we could not find such elements. Despite a high degree of similarity, mouse and human CREST1 and CREST2 were more than five times farther downstream from the transcribed unit of the *isl1* gene than their counterparts in zebrafish and fugu (Fig. 4A). Likewise, cCREST1 and cCREST2 were approximately two to three times farther downstream (Fig. 4A).

Functional conservation of CREST1 in the regulation of gene expression in motor neurons

To verify the functional conservation of the putative enhancers described above between different vertebrate species, we first generated transgenic mice carrying the same transgene (CM/isl1-GFP) that was used to establish the Isl1-GFP zebrafish (Higashijima et al., 2000); these were referred to as Isl1-GFP mice. Expression patterns were the same as in Isl1-GFP zebrafish embryos, in that GFP was expressed in the cranial and spinal motor neurons of embryonic day 11.5 (E11.5) embryos (Fig. 6A). Double labeling of sections of the spinal cord at



B

	DM		VM		SN	
	z	m	z	m	z	m
zCREST1 (CM)	(+)	(+)	(-)	(-)	(-)	(-)
hCREST1	n.e.	(+)	n.e.	(-)	(-)*	(-)
zCREST2	(-)	(-)	(+)	(+)	(+)	(+)
hCREST2	n.e.	(-)	n.e.	(+)	(-)*	(-)
zCREST3	n.e.	n.e.	n.e.	n.e.	(+)*	n.e.
hCREST3	n.e.	n.e.	n.e.	n.e.	(+)*	n.e.
zC3R	n.e.	n.e.	n.e.	n.e.	(+)*	n.e.

DM, dorsally projecting motor neurons; VM, ventrally projecting motor neurons
SN, sensory neurons; z, zebrafish; m, mouse; n.e., not examined.
* Confirmed only by the transient transgene expression assay.

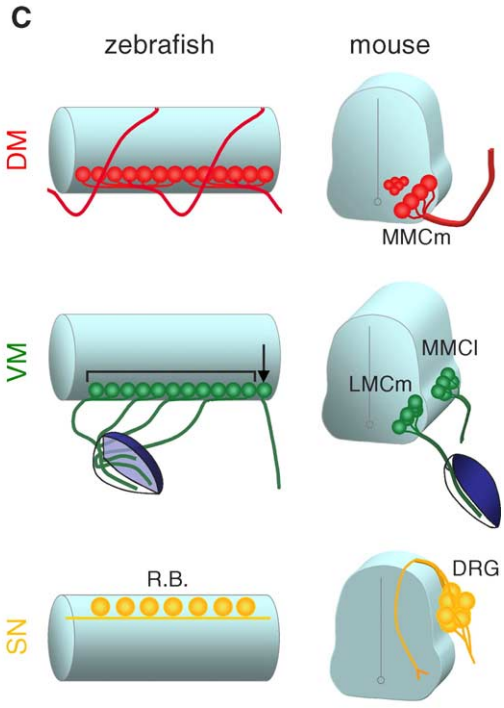


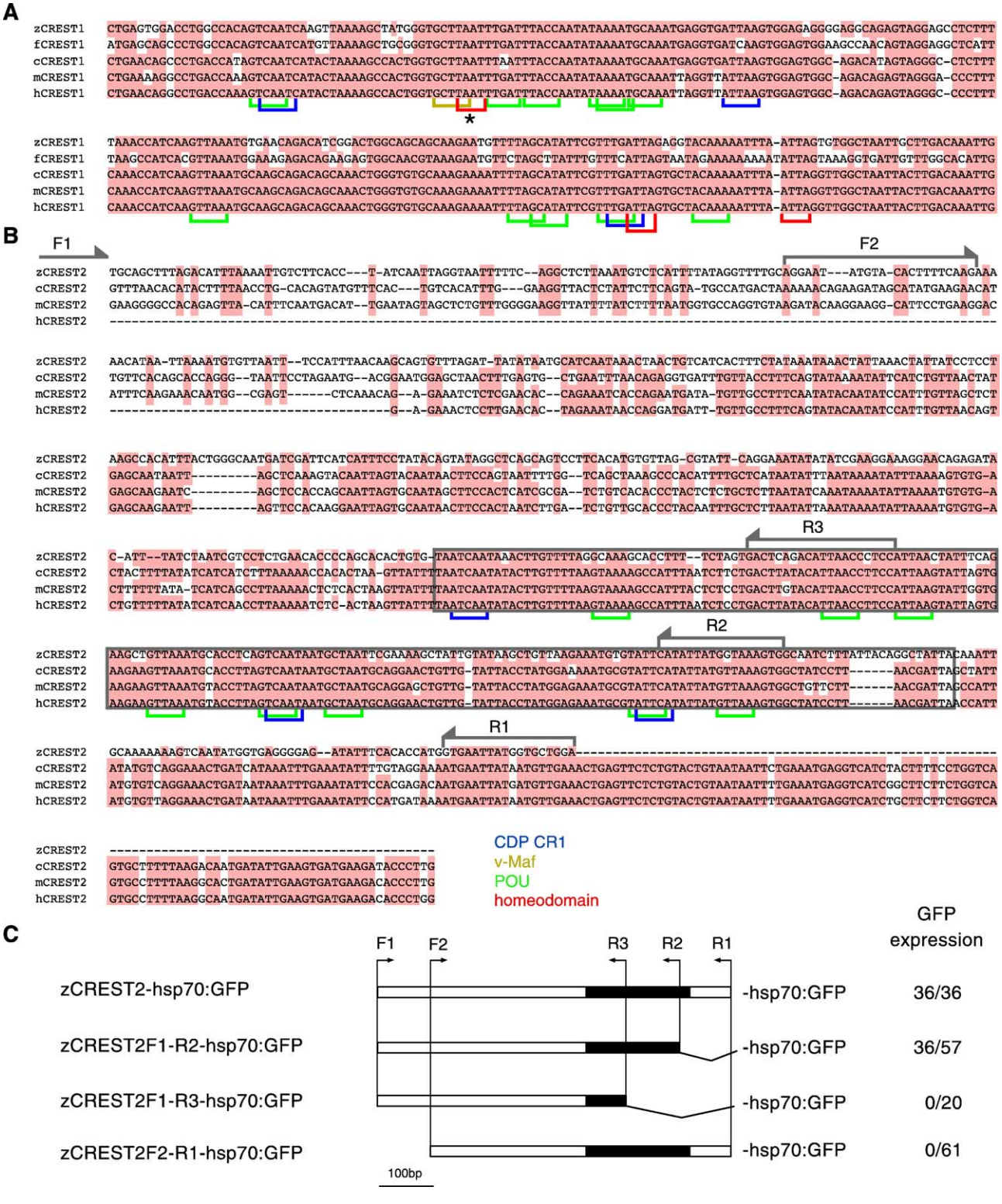
Fig. 4. Structural and functional conservation of CREST1, CREST2, and CREST3 during evolution. (A) Comparison of the distribution of CREST1, CREST2, and CREST3 in various species. The thick bars indicate the transcriptional unit of the *isl1* gene. Number on each locus indicates the percentage of sequence identity in the highly conserved core sequences of each species to corresponding element of zebrafish. zCREST1 lies 23 kb downstream of the *isl1* gene transcription initiation site, whereas fCREST1, cCREST1, mCREST1, and hCREST1 reside 9, 70, 160, and 150 kb, respectively, from the *isl1* gene. zCREST2, cCREST2, mCREST2, and hCREST2, on the other hand, lie 56, 120, 350, and 300 kb, respectively, downstream from the *isl1* gene locus. zCREST3 lies 60 kb upstream of the *isl1* gene locus, whereas cCREST3, mCREST3, and hCREST3 reside 150, 320, and 350 kb, respectively, upstream of the *isl1* gene. (B and C) Summary table (B) and figure (C) of this study. Comparisons of the function of CREST1, CREST2, CREST3, and C3R in motor and sensory neurons of each species are shown. zCREST1 (CM) drives *isl1* gene expression in dorsally projecting motor neurons (DM) in zebrafish and mice (red circles in C). DM consists of motor neurons innervating dorsal trunk muscles in zebrafish and MMCm neurons in mice. hCREST1 also has the same activity as zCREST1 in mice. In addition, zCREST1 and hCREST1 can drive gene expression in postmitotic migrating motor neurons (smaller red circles in C) only in mice. In contrast, zCREST2 drives *isl1* gene expression in ventrally projecting motor neurons (VM, green circles in C) and sensory neurons (SN, yellow circles in C) in zebrafish and mice. VM consists of motor neurons innervating abductor pectoral fin muscles (bracket in C) and ventral trunk muscles (arrow in C) in zebrafish and MMCl and LMCm neurons in mice. While zCREST2 drives gene expression in sensory neurons such as Rohon-Beard neurons (R.B.) in zebrafish and DRG in mice, hCREST2 cannot. Neither motor neurons innervating abductor pectoral fin muscles (bracket in C) nor LMCm neurons ever terminate in the *lmx1b*-positive region (blue in C). zCREST3, hCREST3, and zC3R drive gene expression specifically in sensory neurons in zebrafish. All results were confirmed using stable transgenic animals except for those observations indicated by asterisk in B. The functions of CREST3 and C3R have not been examined in mice.

E11.5 for GFP and Isl1 revealed that GFP expression is restricted to the ventromedial subset of the spinal motor neurons (arrowheads in Fig. 6B) as well as more dorsomedial cells (arrows in Fig. 6B), which are likely to be postmitotic migrating motor neurons, since GFP expression in these cells becomes undetectable by the time when motor neuron generation ends (data not shown, see also Figs. 6E, F, and H).

To further confirm the activities of the putative enhancers, we generated transgenic mice, termed *Tg(zCREST1/isl1-PLAP)*, *Tg(mCREST1/isl1-PLAP)*, and *Tg(hCREST1/isl1-PLAP)*, in which zCREST1, mCREST1, and hCREST1, respectively, are inserted immediately upstream of the zebrafish *isl1* core promoter (ICP; Higashijima et al., 2000) and the *human placental alkaline phosphatase* (*PLAP*) gene (Fig. 2B). PLAP is a glycosyl phosphatidylinositol (GPI)-linked cell surface protein that labels axons along their entire length when expressed in neurons (Leighton et al., 2001). All transgenic mice showed virtually the same expression pattern as the Isl1-GFP mice, in that reporter genes were expressed in cranial and spinal motor neurons (Figs. 6A and C). At the thoracic level, PLAP was initially expressed in all the spinal motor neurons at E11.5, including the putative postmitotic migrating motor neurons (arrows in Figs. 6D and G), but by E12.5 PLAP expression had become restricted to the most medial subtype of spinal motor neurons (arrowheads in Figs. 6E and H). These motor neurons selectively innervate the dorsal trunk muscles (open arrow in Fig. 6H), suggesting that they were MMCm neurons (Fig. 4C). We confirmed this by coexpression of PLAP and Lhx3 in these motor neurons; Lhx3 is a mouse ortholog of chick Lim3 and is a specific marker for MMCm neurons (Figs. 6F and I; Tsuchida et al., 1994).

Functional conservation and diversification of CREST2 in the regulation of gene expression in motor and sensory neurons

Next, to confirm the functional conservation of the zCREST2 and hCREST2 sequences, we generated *Tg(zCREST2/isl1-PLAP)* and *Tg(hCREST2/isl1-PLAP)* mice in which zCREST2 and hCREST2, respectively, were inserted upstream of the ICP and *PLAP* gene (Fig. 2B). At E11.5, PLAP was expressed exclusively in the sensory neurons of *Tg(zCREST2/isl1-PLAP)* mice, including the trigeminal and other cranial sensory ganglia, as well as the dorsal root ganglia (DRG), and in the spinal motor neurons (Figs. 7A and C). At E14.5, the PLAP-positive axons of the Ia



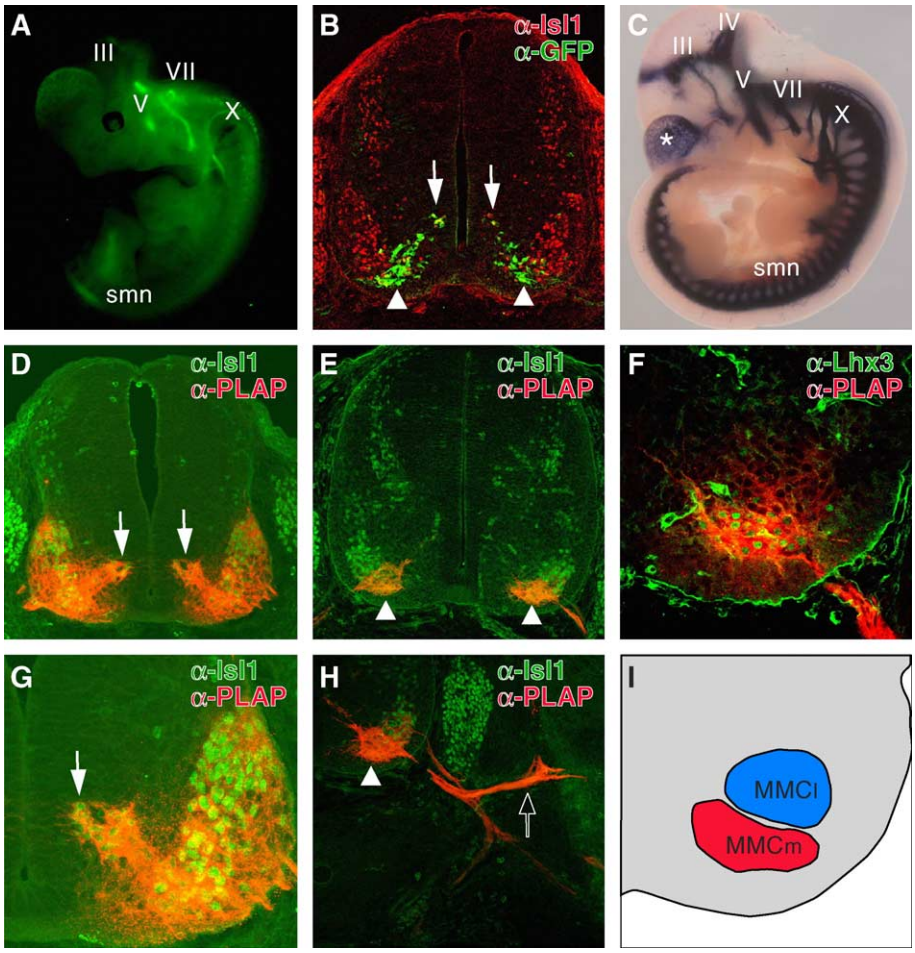


Fig. 6. Functional conservation of CREST1 for motor neuron-specific gene expression in mouse. (A) Lateral view of E11.5 Isl1-GFP transgenic mice as seen through a dissecting fluorescence microscope. III, oculomotor nerve; V, trigeminal nerve; VII, facial nerve; X, vagus nerve. This transgenic embryo carries exactly the same transgene as Isl1-GFP transgenic zebrafish (Higashijima et al., 2000). (B) Cross-section of the E11.5 Isl1-GFP embryo at the thoracic level. GFP was expressed in migrating postmitotic motor neurons (arrows) and the ventromedial subset of motor neurons (arrowheads). (C) Lateral view of the E10.5 *Tg(hCREST1/isl1-PLAP)* embryo stained for alkaline phosphatase activity. Both somatic and visceral motor neurons expressed PLAP at this stage. The asterisk indicates PLAP expression in the telencephalon that was commonly observed in lines examined (data not shown). IV, trochlear nerve. (D–H) Cross-sections at the thoracic level of E11.5 (D and G) and E12.5 (E, F, and H) *Tg(hCREST1/isl1-PLAP)* embryos double labeled with anti-Isl1 (D, E, G, and H) anti-Lhx3 (F) antibody (green) and anti-PLAP antibody (red). At E11.5 (D and G), all Isl1-positive spinal motor neurons expressed PLAP. The dorsomedially located postmitotic-migrating motor neurons (arrows) also expressed PLAP at this stage. The weak expression of PLAP in MMCl cells (F and H) may reflect this early widespread activity of CREST1. By E12.5 (E and H), PLAP expression became restricted to the ventromedial subset of Isl1-positive motor neurons (arrowheads). Note that at this stage, PLAP-positive axons projected into dorsal trunk muscles (open arrow in H). (F) All Lhx3-positive spinal motor neurons also expressed PLAP at this stage, demonstrating that the PLAP-positive neurons were the MMCm motor neurons. (I) The schematic illustration of F. The regions containing PLAP-positive MMCm neurons and PLAP-negative MMCl neurons are indicated in red and blue, respectively.

sensory neurons in the DRG were observed to extend into the ventral half of the spinal cord, which were also occupied by PLAP-positive motor neurons (arrows in Fig. 7H). In contrast, in *Tg(hCREST2/isl1-PLAP)* mice, PLAP was expressed only in spinal motor neurons, and none was expressed in sensory neurons (Fig. 7B).

Fig. 5. Sequence comparisons between CREST1s and between CREST2s and the deletion analysis of zCREST2. (A and B) Sequence comparisons between CREST1s (B) and between CREST2s (C). Putative transcription factor binding sites in evolutionarily conserved regions are indicated by brackets with different colors. The entire nucleotide sequences of zCREST1, fCREST1, cCREST1, mCREST1, hCREST1, zCREST2, cCREST2, mCREST2, and hCREST2 were deposited in DDBJ under accession numbers AB0158303, AB158312, AB167735, AB158306, AB158309, AB158304, AB167736, AB158307, and AB158310, respectively. The nucleotides conserved in more than three of all species are shaded by the red color. (A) Only the most conserved regions among species are shown. zCREST1m, nucleotide conversion from TAAT (asterisk, Fig. 2C) to GGGG in zCREST1, completely lost its enhancer activity. (B) Entire sequences of zCREST2 and hCREST2 used in this study and sequences of cCREST2 and mCREST2 retrieved from databases are shown. Highly conserved regions are enclosed with black boxes. Primers used to generate the deletion fragments of zCREST2 are indicated by arrows, except that primer F1 was designed to target to a multiple cloning sites of pBluescript II. (C) The result of deletion analysis of zCREST2. The number on the right side of each plasmid construct indicates the number of embryos positive for expression of GFP in primary sensory neurons among all injected embryos. Solid lines indicate the highly conserved region among species.

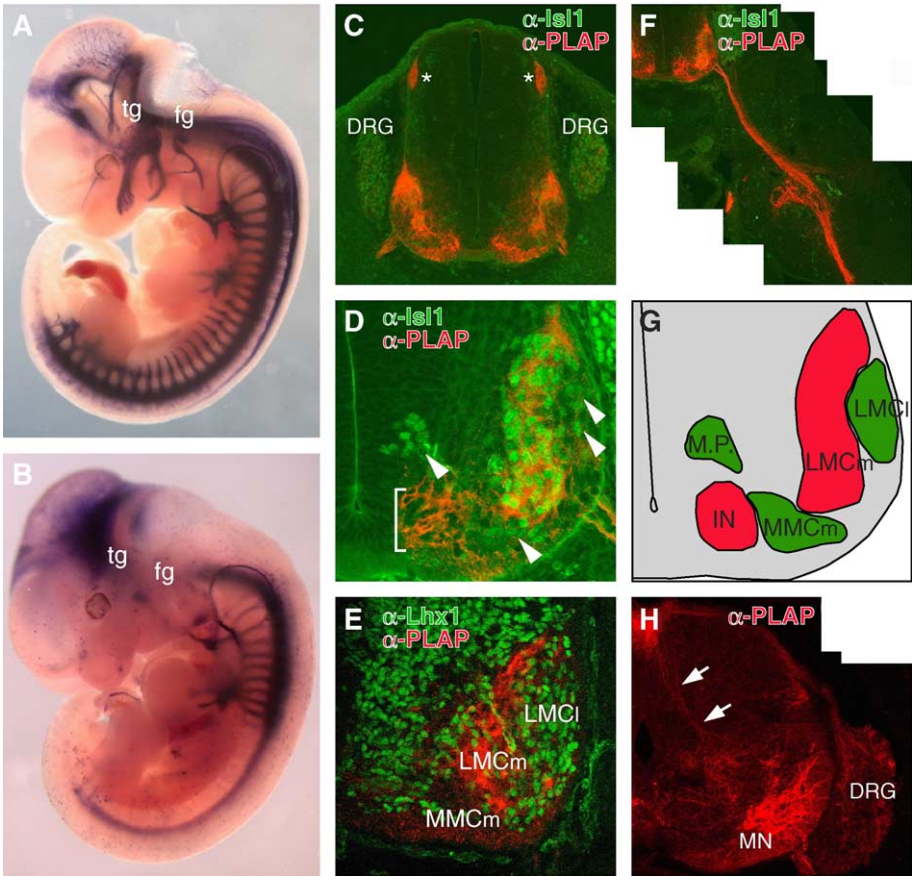


Fig. 7. Functional conservation and diversification of CREST2 for motor and sensory neuron-specific gene expression in mouse. (A and B) Lateral views of the E11.5 *Tg(zCREST2/isl1-PLAP)* embryo (A) and *Tg(hCREST2/isl1-PLAP)* embryo (B). In both embryos, PLAP was expressed in spinal motor neurons that extend their axons ventrally, while PLAP expression in the sensory neurons, such as trigeminal ganglion neurons (tg), facial ganglion neurons (fg), and dorsal root ganglion neurons (DRG), was observed only in the *Tg(zCREST2/isl1-PLAP)* embryo (see also C and H). (C, D, and F) Cross-sections at the brachial level of E11.5 *Tg(zCREST2/isl1-PLAP)* embryos doubly labeled with anti-Isl1 antibody (green) and anti-PLAP antibody (red) showing PLAP expression in the subset of spinal motor neurons innervating the ventral muscles of the limb and in DRG neurons. Asterisks in C indicate the dorsal root entry zone (DREZ) where the central axons of the DRG neurons enter the spinal cord. Close-up view of the ventral spinal cord (D) showing expression of PLAP in the specific subset of spinal motor neurons, excluding the dorsomedially, ventromedially, and dorsolaterally (arrowheads) located Isl1-positive motor neurons. In this particular line, ectopic PLAP expression was also observed in the ventralmost region of the spinal cord (bracket). The specific projection of PLAP-positive motor axons into the ventral limb bud is shown in F. (E) Double labeling with anti-Lhx1 antibody (green) and anti-PLAP antibody (red) at the brachial level of the E11.5 *Tg(zCREST2/isl1-PLAP)* embryo revealed that PLAP-positive motor neurons were negative for Lhx1, a marker for LMCl motor neurons, supporting their identity as LMCm motor neurons. Lhx1 was also expressed in other interneurons (Tsuchida et al., 1994). (G) Schematic illustration of D. Regions containing PLAP-positive and PLAP-negative neurons are indicated in red and green, respectively. M.P., motor neuron precursors; IN, interneurons. (H) Brachial level of the E14.5 *Tg(zCREST2/isl1-PLAP)* embryo labeled with anti-PLAP antibody (red). The PLAP-positive Ia afferent fibers (arrows) were observed penetrating the spinal cord and projecting ventrally to the motor neurons (MN) in the ventral horn.

Next, we examined the subtypes of motor neurons expressing PLAP in *Tg(zCREST2/isl1-PLAP)* mice at E11.5 by double labeling with anti-Isl1 antibody and anti-PLAP antibody. At the brachial level, only the dorsolaterally located Isl1-positive motor neurons coexpressed PLAP (Figs. 7D and G). The ventromedially and laterally located Isl1-positive neurons were negative for PLAP (arrowheads in Fig. 7D). The PLAP-positive motor neurons projected their axons to the ventral limb muscles (Fig. 7F), supporting that they were LMCm neurons (Fig. 4C). We further confirmed this by comparing the expression of PLAP with the expression of Lhx1 (a mouse ortholog of chick Lim1). Lhx1 is expressed specifically in LMCl neurons, and not in LMCm neurons (Tsuchida et al., 1994). The PLAP-positive motor neurons in *Tg(zCREST2/isl1-PLAP)* mice were never positive for Lhx1 (Fig. 7E). We also confirmed that PLAP was selectively expressed in LMCm neurons in *Tg(hCREST2/isl1-PLAP)* mice (data not shown). At the thoracic level, as shown in trunk level of *Tg(zCREST2-hsp70:GFP)*^{rw011} zebrafish embryos, PLAP was expressed in motor neurons projecting ventral trunk muscles in *Tg(zCREST2/isl1-PLAP)* mice (Figs. 4B and C, and data not shown).

Segregation of different subtypes of the secondary motor neurons in the zebrafish spinal cord

In contrast to primary motor neurons, there is a paucity of literature on the subtype specification of secondary motor

neurons in zebrafish embryos (Myers, 1985; Pike et al., 1992; Segawa et al., 2001). In our study, the ventrally projecting secondary motor neurons occupied the ventral half of the motor column at the trunk level (Fig. 3J). In contrast, the dorsally projecting secondary motor neurons occupied the dorsal part of the motor column (Fig. 3L). Therefore, the positional relationship between these secondary motor neurons subtypes in the zebrafish spinal cord is different from that of the corresponding subtypes of motor neurons in the chick and mouse spinal cord, where the ventrally projecting motor neurons are located within the MMCl, and the dorsally projecting motor neurons are more medially located within the MMCm at the thoracic level.

Both evolutionarily conserved and nonconserved regions are necessary for the sensory neuron-specific enhancer activity of zCREST2

While zCREST2 and hCREST2 share high sequence similarity in their evolutionarily conserved region, hCREST2 does not contain the sensory neuron-specific enhancer activity. We next examined the deletion analysis of zCREST2 to confirm which regions were necessary for sensory neuron-specific expression in zebrafish. Deletion constructs of zCREST2 were generated using the PCR method (Figs. 5B and C). Each deletion fragment was inserted in the immediate upstream of *hsp70* promoter and *GFP* gene (Fig. 2B) and injected into one cell-stage zebrafish embryos. One fragment, termed zCREST2F1-R2, which still contains the evolutionarily conserved region, could drive GFP expression in primary sensory neurons (Figs. 5B and C). In contrast, zCREST2F1-R3, which lacks most of the conserved region, could not drive GFP expression in primary sensory neurons any more (Figs. 5B and C). These data implicate the necessity of the evolutionarily conserved region for the sensory neuron-specific activity of zCREST2. However, the injection of zCREST2F2-R1-hsp70:GFP, which lacks 85 bp of nonconserved region of zCREST2, could not drive GFP expression in primary sensory neurons (Figs. 5B and C). These data and the observation that hCREST2 does not have the sensory neuron-specific enhancer activity (Fig. 4B) show that both evolutionarily conserved and nonconserved regions of zCREST2 are necessary for the sensory neuron-specific activity in zebrafish.

Prediction of a sensory neuron-specific novel enhancer of the isl1 gene in zebrafish using comparative genome informatics

Comparing the sequences and functions of CREST2s, only zCREST2 exhibits sensory neuron-specific enhancer activity. This activity depends on evolutionarily nonconserved regions of zCREST2, suggesting that there may be other conserved enhancers responsible for sensory neuron-specific expression of *isl1* that are active in other vertebrate than zebrafish. The intronic sequence of zebrafish *isl1* has an activity to enhance gene expression in a wide variety of neurons, including Rohon-Beard neurons and trigeminal ganglion neurons (Reyes et al., 2004). But this sequence is not conserved among different vertebrate species.

Comparison of the 200-kb upstream region of the zebrafish *isl1* gene with a 1-Mb region upstream of the human *isl1* gene was performed using MultiPipMaker (Schwartz et al., 2000, also see Materials and methods). Nine conserved regions were found, but only one was conserved in the chick and mouse genomes, according to a genome BLAST search. This region, termed CREST3, contains 419-bp conserved sequence with 76% identity between zebrafish and human. Zebrafish, chick, mouse, and human CREST3 reside 60, 150, 320, and 350 kb upstream of the *isl1* gene locus, respectively (Figs. 4A and 8A). CREST3 did not show significant sequence similarity to CREST2 (data not shown).

Next, we examined the enhancer activity of this region by transient transgene expression in zebrafish embryos. Approximately 1 kb of the zebrafish genomic region encompassing this conserved block was cloned in the immediately upstream of the ICP and *GFP* gene (Fig. 2B). And this plasmid DNA was injected into one-cell stage zebrafish embryos. GFP expression was observed in primary sensory neurons such as trigeminal ganglion neurons (Fig. 8B) and Rohon-Beard neurons (Fig. 8C) at 36 hpf. We also examined the function of hCREST3 by transient transgene expression in zebrafish and confirmed the functional conservation of CREST3 to drive GFP expression in primary sensory neurons (data not shown).

To examine whether zCREST2 and zCREST3 are activated in different types of sensory neurons, we injected zCREST3-isl1:mRFP1 plasmid solution into one-cell stage *Tg(zCREST2-hsp70:GFP)*^{rw011} zebrafish embryos. mRFP1 is a genetically modified red fluorescent protein that acts as a monomer (Campbell et al., 2002). A significant number of trigeminal ganglion neurons and Rohon-Beard neurons expressed both GFP and mRFP1, but there were also cells that expressed only either of them (Figs. 8D and E).

Conservation of a sensory neuron-specific enhancer of the isl1 gene family

As previously reported, all *isl1* family members (*isl1*, *isl2*, and *isl3*) are expressed in primary sensory neurons from the early stage of embryonic development of zebrafish (Inoue et al., 1994; Kikuchi et al., 1997; Segawa et al., 2001; Tokumoto et al., 1995). Moreover, mouse *isl1* and *isl2* are also expressed in DRG neurons (Thaler et al., 2004; Tsuchida et al., 1994). These data suggest a possibility that expression of all *isl1* family members in sensory neurons is regulated by an evolutionarily conserved enhancer. Then we performed a BLAST search against whole genomic sequences for sequences similar to any of 100-bp subfragments of hCREST3 with 50-bp overlaps with the adjacent subfrag-

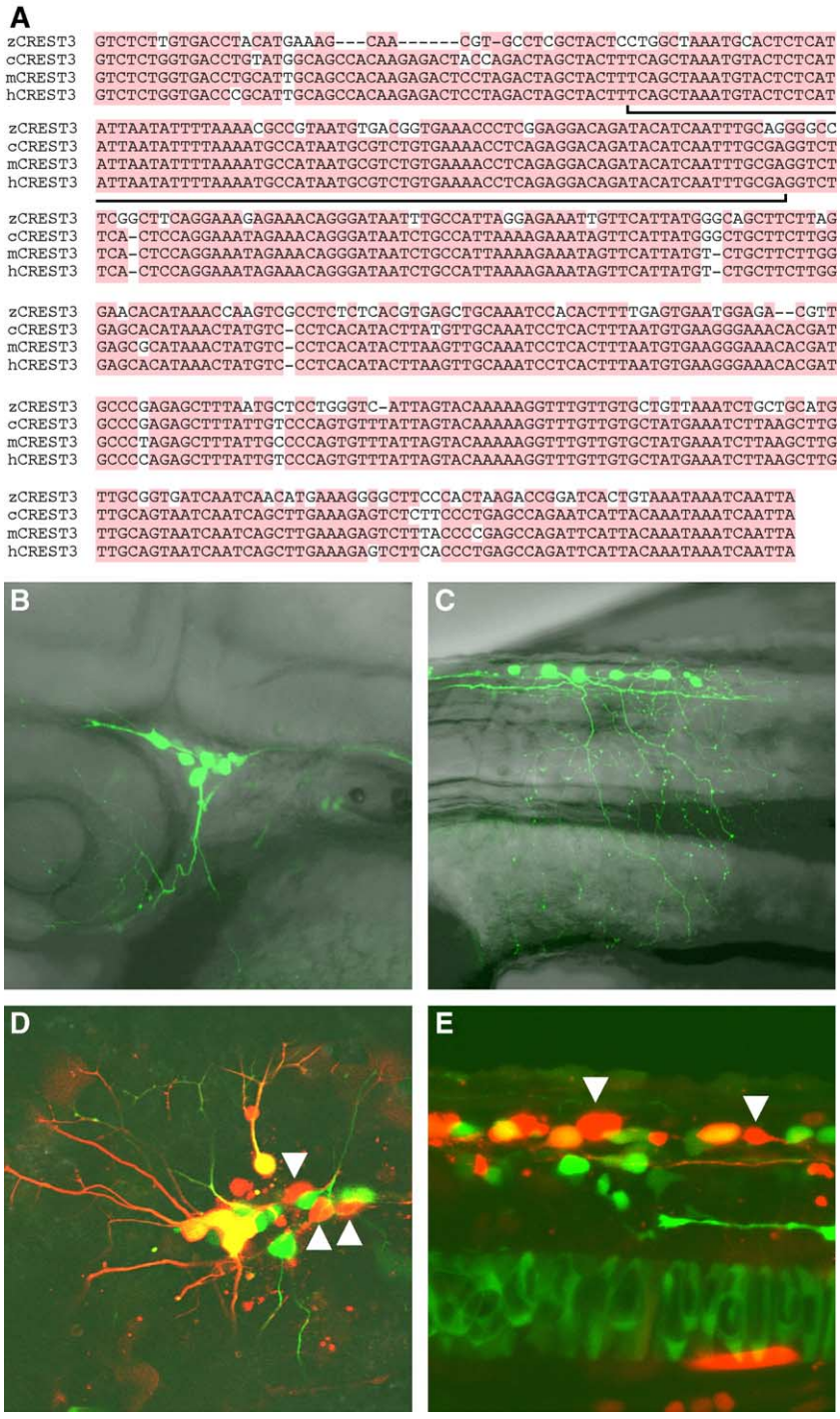


Fig. 8. Structural and functional conservation of CREST3 during evolution. (A) Sequence comparisons between CREST3s. No similar sequences to CREST3 were found in the fugu genome. Only the most conserved regions among species are shown. The entire nucleotide sequences of zCREST3, cCREST3, mCREST3, and hCREST3 were deposited in DDBJ under accession numbers AB0158305, AB167734, AB158308, and AB158311, respectively. The nucleotides conserved in more than three of all species are shaded by the red color. The bracket indicates the conserved region between CREST3s and C3Rs (see also Fig. 9D). (B and C) Lateral views of 36-hpf embryos injected with the zCREST3-isl1:GFP plasmid and exhibiting expression of GFP in trigeminal ganglion neurons (B) and Rohon-Beard neurons (C). (D and E) Lateral views of 24-hpf *Tg(zCREST2-hsp70:GFP)*^{rw011} embryos injected with the zCREST3-isl1:mRFP1 plasmid and exhibiting mRFP1 expression in trigeminal ganglion neurons (D) and Rohon-Beard neurons (E). Note that there are some mRFP1 single-positive neurons in both trigeminal ganglion neurons and Rohon-Beard neurons (arrowheads).

ments. One such subfragment (bracket in Fig. 8A) was found to show weak similarity to the 5′ flanking regions of the mouse and human *isl2* gene (Figs. 9A and D). Both regions reside in the intronic sequence of *electron transfer flavoprotein alpha-subunit* (*etfa*) gene and are 70 and 90 kb away from mouse and human *isl2* gene, respectively (Fig.

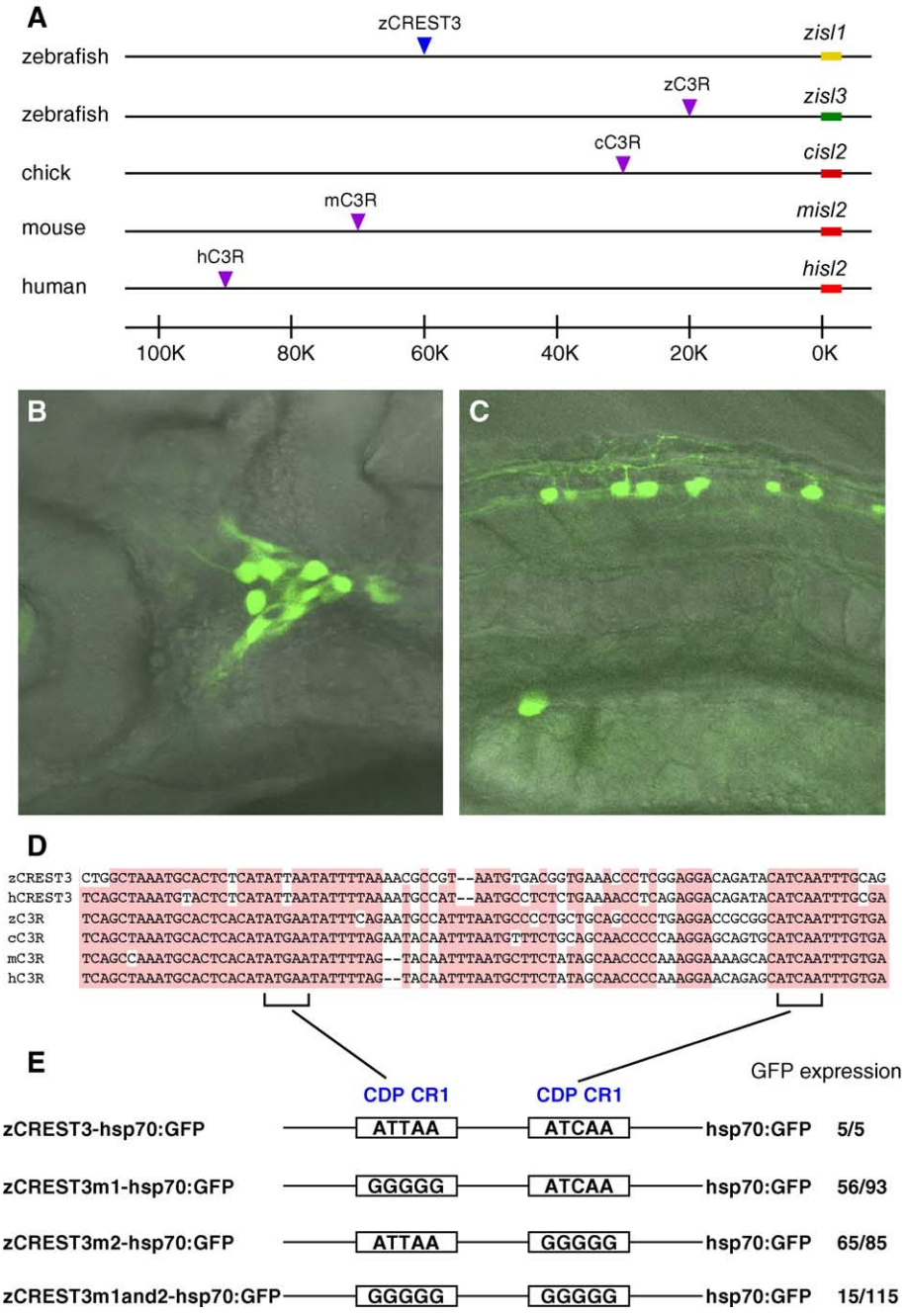


Fig. 9. Structural and functional conservation of sensory neuron-specific enhancers CREST3 and C3R. (A) Comparison of the distribution of CREST3 and C3R in each species. Yellow and green bars indicate the coding regions of zebrafish *isl1* and *isl3*, respectively. Red bars indicate the coding regions of *isl2* of chick, mouse, and human. The blue arrowhead indicates zCREST3 locus. In contrast, purple arrowheads indicate the C3R loci. No similar sequences to C3R were found in the fugu genome. (B and C) Lateral views of 32-hpf embryos injected with the zC3R-hsp70:GFP plasmid and GFP expression in the trigeminal ganglion neurons (B) and Rohon-Beard neurons (C). (D) Sequence comparisons between CREST3 and C3R. Only the most conserved regions between CREST3 and C3R are shown. Nucleotides conserved in more than four of all are shaded by the red color. The entire nucleotide sequences of zC3R, cC3R mC3R, and hC3R were deposited in DDBJ under accession numbers AB179552, AB179553, AB179554, and AB179555, respectively. The highly conserved region between CREST3 and C3R contains two CDP CR1 recognition motifs (brackets). (E) Mutational analysis of zCREST3. Both or either of the CDP CR1 sites was converted to GGGGG. The number on the right side of the name of each plasmid construct indicates the number of embryos exhibiting expression of GFP in primary sensory neurons among all injected embryos.

9A). We termed this region as C3R (CREST3 related). C3R was also found in the chick genome and showed 70% similarity with CREST3 (Figs. 9A and D). Although, in other vertebrates, C3R was found in the flanking region of the *isl2* gene, it is not clear whether the zebrafish *isl2* gene has a similar sequence in its neighborhood due to lack of sequence information around the *isl2* locus in database. Instead, we found C3R in the noncoding genomic region

20-kb upstream of the zebrafish *isl3* gene, and termed zC3R hereafter. zC3R showed 73% similarity with the subfragment of hCREST3 (Figs. 9A and D).

To confirm the enhancer activity of C3R, we next cloned the 1-kb genomic fragment encompassing zC3R and inserted them in the immediate upstream of the promoter sequence of the hsp70:GFP plasmid construct (Fig. 2B). The plasmid construct zC3R-hsp70:GFP was injected into one-cell stage zebrafish embryo. At 32 hpf, GFP expression was observed in trigeminal ganglion neurons (Fig. 9B) and Rohon-Beard neurons (Fig. 9C).

Putative homeodomain recognition site is important for the activity of zCREST1

Sequence comparison and TRANSFAC analysis revealed that there are putative recognition motifs for multiple transcription factors in evolutionarily conserved region of CREST1 (Fig. 5A). Among them, homeodomain proteins such as MNR2, Hb9, and Phox2b are known to be expressed in premature motor neurons and act as key regulators for motor neuron differentiation (Pattyn et al., 2000; Tanabe et al., 1998). Moreover, motor neurons are generated from the progenitor domain for motor neurons (pMN) delineated by the specific expression of homeodomain proteins Nkx6.1 and Pax6 (Briscoe et al., 2000). Thus, we focused on the three homeodomain recognition motifs ATTA or TAAT of zCREST1 (red brackets in Fig. 5A). We replaced each one of them with GGGG. One mutated zCREST1 termed zCREST1m, which had such nucleotide conversion in one of these motifs (asterisk in Fig. 5A), was completely inactive as assessed by transient transgene expression assay in zebrafish (Fig. 2C). This region was predicted as an Nkx protein recognition motif by our TRANSFAC analysis. This region partly overlapped with a putative v-Maf recognition motif; however, nucleotide conversion of this motif to GGGGG never reduced the expression level of GFP in our transient transgene expression assay (data not shown). Taken together, the homeodomain protein might target CREST1 via TAAT sequence shown in Fig. 5A.

Putative Cut domain recognition sites are important for the activity of CREST3 and C3R

Sequence alignment and TRANSFAC analysis revealed that the highly conserved region between CREST3 and C3R contains one POU recognition motif and two CDP CR1 (Cut Repeat1 of CDP) recognition motifs (Fig. 9D and data not shown). The CDP family constitutes a characteristic group of homeodomain proteins, containing one to three repeats of Cut repeat(s). Cut repeat itself is known to bind to specific DNA sequence (Harada et al., 1994). One such gene *D-onecut*, a member of *Drosophila* Cut domain proteins, is thought to be a transcriptional activator and implicated in the regulation of neural differentiation in the eye (Nguyen et al., 2000). Moreover, *zonecut*, a zebrafish ortholog of *D-onecut*, is expressed in primary neurons including trigeminal ganglion neurons and Rohon-Beard neurons in zebrafish embryos at the stage just prior to the expression of *isl1* begins (Hong et al., 2002; Inoue et al., 1994). Thus, the expression of *isl1* in primary sensory neurons might be directly regulated by Cut domain protein(s). To confirm this hypothesis, we replaced either or both of the core recognition motifs of CDP CR1 with GGGGG and examined their enhancer activities (Fig. 9E). Double mutations in zCREST3 alone caused remarkable reduction in the enhancer activity of zCREST3 as assessed by the number of embryos expressing GFP in primary sensory neurons among all injected embryos. These data indicate that Cut domain proteins may play important roles for the regulation of *isl1* gene expression in primary sensory neurons.

Discussion

A combination of comparative genomics and an improved transient expression assay in zebrafish embryos provides a simple and quick approach for the identification of cis-acting regulatory elements

Now that the complete genomes of a number of species have been sequenced, attention must increasingly focus on the elucidation of the mechanisms that regulate the intricate genetic networks. A rapid and simple method for the reliable identification of the *cis*-acting regulatory elements of genes is therefore an essential first step toward this goal. The accumulation of genomic sequence data from various species and a computational genome-wide sequence comparison of different species have enabled the identification of highly conserved *cis*-acting regulatory elements in the neighborhood of several genes, including *shh* (Müller et al., 1999), *otx2* (Kimura et al., 1997), *dlx5*/*dlx6* (Zerucha et al., 2000), *sox2* (Uchikawa et al., 2003), and *scl* (Göttgens et al., 2000). Furthermore, putative regulatory elements have been identified by the simple alignment of large genomic sequences from different species to find conserved sequences in the noncoding regions (Frazer et al., 2001; Loots et al., 2000). Some of these putative regulatory elements, such as the ones for the *IL-4*, *5*, and *13* genes, have been shown to be essential for controlling expression of these genes in cultured T cells. Since alignment of two large noncoding genomic sequences from two relatively close species such as human and mouse often results in identification of too many similar sequences without any functional significance, comparisons between more distant species, such as human and zebrafish, are more effective for identification of functionally significant noncoding sequences (Müller et al., 2002). Our novel and simple method for the highly efficient transient expression will provide the potent method to

confirm the activity of the putative enhancer sequences predicted by the comparative genome informatics approach.

Conservation and diversification of CREST2

We observed a strict conservation of the roles of CREST1 and CRSET2 for activation of motor neuron subtype-specific gene expression during vertebrate evolution (Figs. 4B and C). In contrast, only zebrafish CREST2 retains activity for sensory neuron-specific gene expression (Figs. 4B and C), and the nonconserved sequence within zCREST2 is essential for this activity (Figs. 5B and C).

Our data show that the mechanisms for *isl1* gene activation by CREST1 and CREST2 in nonoverlapping subsets of spinal motor neurons were established before phylogenic diversification into the teleost and the mammals in vertebrate evolution (Figs. 4B and C). And because of the nonredundant roles of these two elements, any mutation to either of these elements may have deleteriously impaired locomotion and survivability of individual animals. In contrast, zCREST2 and zCREST3 are active redundantly in an overlapping population of primary sensory neurons in zebrafish. This may have allowed for diversification of CREST2 sequence and ultimate loss of its activity to regulate gene expression in sensory neurons during evolution.

Gene expression is often controlled under summation of the influences by multiple enhancers each of which acts as an independent modular unit to activate gene expression in a highly specific subset of cells (Kirchhamer et al., 1996; Stanojevic et al., 1991). It is intriguing to see whether our results could be generalized to a hypothetical principle that multiple enhancers of one gene that originally functioned redundantly in overlapping populations of cells may individually increase the levels of cell-type specificity during evolution by losing redundant roles among themselves.

Why has only CREST3 but not CREST2 been conserved for sensory neuron-specific gene activation during evolution remains to be answered, while mouse still retains a capacity to activate zCREST2 in the sensory neurons (Figs. 4B,C and 7A,C,H). The zebrafish *isl2* and *isl3* genes are structurally very similar and located, respectively, in the regions of chromosomes 25 and 7 with highly conserved synteny to each other (http://www.zfin.org). Therefore, they are the paralogs that were generated by duplication from the putative ancestral *isl1* paralog accompanying the whole genome duplication that took place in the teleost evolution (Force et al., 1999). Although zebrafish *isl2* and *isl3* show subfunctionalization with *isl2* regulating the motor neuron identity and *isl3* regulating development of the optic vesicle and the midbrain–hindbrain boundary region, they are both expressed in the primary sensory neurons of zebrafish (Appel et al., 1995; Inoue et al., 1994; Kikuchi et al., 1997; Segawa et al., 2001; Tokumoto et al., 1995). CREST3 and its related enhancer sequence C3R are highly conserved in evolution and appear to be tightly linked to all members of the *isl1* gene family, including the mouse *isl2* gene that probably shares its origin with the putative ancestral *isl1* paralog of zebrafish *isl2* and *isl3*. Therefore, the linkage between these enhancers and the *isl1* family members has been maintained even after two rounds of whole genome duplication in vertebrate evolution. This may indicate that *isl1* family members may have evolved nonredundant and essential roles for differentiation and maintenance of sensory neurons, although their distinction has not yet been thoroughly studied.

Distinct regulation of the isl1 gene family in sensory neurons

CREST3 and its related C3R contain two CDP CR1 recognition motifs in highly conserved regions (Fig. 9D). Disruption of both motifs caused remarkable loss of the enhancer activity of zCREST3 (Fig. 9E). Thus, Cut domain proteins may act as direct regulators of the expression of the *isl1* gene family. One such gene termed *zonecut* is expressed in primary neurons including trigeminal ganglion neurons and Rohon-Beard neurons in zebrafish embryos at the stage just prior to initiation of the *isl1* gene expression (Hong et al., 2002; Inoue et al., 1994). Thus, *zonecut* seems to be a strong candidate of an upstream factor of *isl1*. However, coinjection of antisense morpholino oligonucleotide (AMO) against *zonecut* and zCREST3-isl1:GFP did not reduce the expression level of GFP in primary sensory neurons (O.U. and H.O., unpublished data). Other Cut domain genes such as *cux2*, which is also expressed in DRG neurons in mice, rather than or together with *zonecut* might be involved in the regulation of sensory neuron-specific expression of *isl1* (Iulianella et al., 2003). It is noted, however, a small portion of embryos still expressed GFP in primary sensory neurons when injected with zCREST3m1andm2-hsp70:GFP plasmid (Fig. 9E), suggesting involvement of other factors.

Distinct regulation of the isl1 gene in subgroups of motor neurons

Individual motor neuron columns are delineated by combinatorial expression (LIM code) of LIM-HD proteins (Tsuchida et al., 1994). The precise mechanism that ensures this fine combinatorial expression of LIM-HD genes in each motor column remains unclear.

In the LMC of chick spinal cord, LMCm and LMCl cells express Isl1 and Lim1, respectively. Ectopic expression of Lim1 in LMC cells represses Isl1 expression in situ and vice versa, implicating cross-repressive regulation for expression of *isl1* and *lim1* genes in the LMC (Kania and Jessell, 2003). Exposure to different levels of retinoid receptor signaling or expression of different members of the Hox-c protein family determines the rostrocaudal identity of postmitotic motor neurons, i.e., whether to take the LMC-like identity of the brachial level or that specific to the thoracic level, such as in MMCl neurons and the CT

(column of Terni) preganglionic autonomic motor neurons (Dasen et al., 2003; Sockanathan et al., 2003). Similarly, in zebrafish embryos, the environmental cue emanating from the somite overlying the spinal cord is implicated in regulating *isl1* gene expression, since transplantation experiments revealed that an immature primary motor neuron is amenable to changing its identity and expression pattern of Isl1 and Isl2 depending on its position relative to the somite (Appel et al., 1995, Lewis and Eisen, 2004).

In this study, we showed that *isl1* gene expression in the later stages of motor neuron differentiation is regulated in a subtype-specific manner by two evolutionarily conserved enhancers termed CREST1 and CREST2. They may function to enhance distinction of the *isl1* expression levels in different subtypes of motor neurons in autoregulatory response to the combinatorially expressed LIM-HD proteins. Or, distinction in capabilities of activating different enhancer elements of the *isl1* gene in different subtypes of motor neurons may implicate some unknown mechanism that prefigures motor neuron subdivision even before the establishment of LIM codes. Since injection of the AMOs against *isl1* and/or *isl2* did not prevent subtype-specific activation of these enhancers (O. U. and H O., unpublished data, and Segawa et al., 2001), autoregulatory activation of these elements is rather unlikely, although the injected AMOs may have degraded before the stages when these enhancers were normally activated. Therefore, subtype specification of motor neurons may be redundantly ensured by multiple steps, i.e., initially by activation of different subtype-specific transcriptional enhancers of LIM-HD genes and later by combinatorial expression of LIM-HD proteins.

Regulation of the isl1 gene at early stages of motor neuron differentiation

The *isl1* gene is expressed initially in all types of motor neurons soon after they leave the cell cycle. In our study using transgenic mice, we showed that this initial expression of the *isl1* gene is regulated by a single regulatory element, termed CREST1. Since CREST1 can drive gene expression specifically in all motor neurons in mouse when it is combined with ICP, which alone has no basal transcriptional activity (data not shown), CREST1 should act as the transcriptional enhancer early in motor neuron differentiation rather than as a specific silencer that represses transcription in cells other than motor neurons.

Motor neuron differentiation occurs via multiple steps. Graded Shh signaling along the dorsoventral axis of the spinal cord establishes an initial progenitor domain (pMN) for motor neurons, which is delineated by the specific expression of transcription factors, Nkx6.1 and Pax6, and by the absence of Nkx2.2 and Irx3 expression (Briscoe et al., 2000). There, the motor neuron determinants such as MNR2 and Olig2 are induced and activate the expression of other transcription factors such as NGN2, a proneuronal bHLH-type transcription factor, and the LIM-HD transcription factors, which are directly involved in the initiation of differentiation and subtype-specification of motor neurons (Mizuguchi et al., 2001; Novitch et al., 2001; Tanabe et al., 1998; William et al., 2003).

MNR2, a member of the Mnx class of HD proteins, was first found to be pivotal in the specification of motor neuron identity in chick embryo. MNR2 and its related gene product HB9 can induce Isl1 ectopically even in the dorsal spinal cord (Tanabe et al., 1998). However, a recent study suggests that Mnx class HD proteins act as transcriptional repressors (William et al., 2003). Therefore, although MNR2 and HB9 have the ability to induce ectopic Isl1 expression, they are unlikely to be the direct activators of CREST1.

The retinoid receptors have drawn recent attention as transcriptional activators that may function to link inductive signaling pathways involving both transcriptional repressors and activators that are directly involved in motor neuron specification (Novitch et al., 2003). Both retinoic acid and retinoid receptor activity are required for the progression of Pax6- and Nkx6-positive ventral progenitor cells to an Olig2-positive motor neuron progenitor and to activate the expression of transcription factors such as Mnx, Lhx3, and Isl1/2, which direct later steps in motor neuron differentiation. However, in the evolutionarily conserved block of CREST1, we could not find the consensus recognition motif sequence (AGGTCA n_{1-5} AGGTCA) for retinoid receptor binding (Rastinejad, 2001). Instead, we identified three homeodomain recognition sites, ATTA or TAAT (Fig. 5A). Disruption of one of them, which was predicted to be recognized by Nkx protein according to TRANSFAC analysis, completely inactivated CREST1 (Figs. 2C and 5A). Moreover, Nkx6.1 is expressed in both cranial and spinal motor neurons (Müller et al., 2003; Sander et al., 2000). Therefore, Nkx protein or other homeodomain protein(s) may be directly responsible for the activation of *isl1* gene expression. Further study for the CREST1-binding factor may give us better understanding on the initial stage of motor neuron differentiation.

Acknowledgments

We thank members of the Okamoto lab for their kind support and fruitful discussions, and members of the RCC, RIKEN, BSI for generation of the PLAP transgenic mice and maintenance of the zebrafish and mice. We thank Dr. John Y. Kuwada for the HSP70-EGFP (HA) vector, Dr. Connie G. Lobe for the Z/AP expression vector, Dr. Yoshikazu Hirate for the zebrafish cDNA library, Dr. Roger Y. Tsien for the *mRFP1* gene, Dr. Youichi Iwai for his valuable advice, and Dr. Angus Thomson for correction of manuscript. M.G. was supported as an internship student of the RIKEN BSI Summer Program 2003. This research was supported by grants from RIKEN BSI, CREST (JST), and a Special Coordination Fund of the Ministry of Education, Science, Technology, Sports and Culture of Japan.

References

Appel, B., Korzh, V., Glasgow, E., Thor, S., Edlund, T., Dawid, I.B., Eisen, J.S., 1995. Motoneuron fate specification revealed by patterned LIM homeobox gene expression in embryonic zebrafish. Development 121, 4117–4125.

Briscoe, J., Pierani, A., Jessell, T.M., Ericson, J., 2000. A homeodomain protein code specifies progenitor cell identity and neuronal fate in the ventral neural tube. Cell 101, 435–445.

Campbell, R.E., Tour, O., Palmer, A.E., Steinbach, P.A., Baird, G.S., Zacharias, D.A., Tsien, R.Y., 2002. A monomeric red fluorescent protein. Proc. Natl. Acad. Sci. U. S. A. 99, 7877–7882.

Dasen, J.S., Liu, J.P., Jessell, T.M., 2003. Motor neuron columnar fate imposed by sequential phases of Hox-c activity. Nature 425, 926–933.

Ericson, J., Thor, S., Edlund, T., Jessell, T.M., Yamada, T., 1992. Early stages of motor neuron differentiation revealed by expression of homeobox gene Islet-1. Science 256, 1555–1560.

Force, A., Lynch, M., Pickett, F.B., Amores, A., Yan, Y.L., Postlethwait, J., 1999. Preservation of duplicate genes by complementary, degenerative mutations. Genetics 151, 1531–1545.

Frazer, K.A., Sheehan, J.B., Stokowski, R.P., Chen, X., Hosseini, R., Cheng, J.F., Fodor, S.P., Cox, D.R., Patil, N., 2001. Evolutionarily conserved sequences on human chromosome 21. Genome Res. 11, 1651–1659.

Göttgens, B., Barton, L.M., Gilbert, J.G., Bench, A.J., Sanchez, M.J., Bahn, S., Mistry, S., Grafham, D., McMurray, A., Vaudin, M., Amaya, E., Bentley, D.R., Green, A.R., Sinclair, A.M., 2000. Analysis of vertebrate SCL loci identifies conserved enhancers. Nat. Biotechnol. 18, 181–186.

Halloran, M.C., Sato-Maeda, M., Warren, J.T., Su, F., Lele, Z., Krone, P.H., Kuwada, J.Y., Shoji, W., 2000. Laser-induced gene expression in specific cells of transgenic zebrafish. Development 127, 1953–1960.

Harada, R., Dufort, D., Denis-Larose, C., Nepveu, A., 1994. Conserved cut repeats in the human cut homeodomain protein function as DNA binding domains. J. Biol. Chem. 269, 2062–2067.

Higashijima, S., Okamoto, H., Ueno, N., Hotta, Y., Eguchi, G., 1997. High-frequency generation of transgenic zebrafish which reliably express GFP in whole muscles or the whole body by using promoters of zebrafish origin. Dev. Biol. 192, 289–299.

Higashijima, S., Hotta, Y., Okamoto, H., 2000. Visualization of cranial motor neurons in live transgenic zebrafish expressing green fluorescent protein under the control of the islet-1 promoter/enhancer. J. Neurosci. 20, 206–218.

Hirate, Y., Mieda, M., Harada, T., Yamasu, K., Okamoto, H., 2001. Identification of ephrin-A3 and novel genes specific to the midbrain-MHB in embryonic zebrafish by ordered differential display. Mech. Dev. 107, 83–96.

Hogan, B., Beddington, R., Constantini, F., Lacy, E., 1994. Manipulating the Mouse Embryo: A Laboratory Manual. Cold Spring Harbor Laboratory, New York.

Hong, S.K., Kim, C.H., Yoo, K.W., Kim, H.S., Kudoh, T., Dawid, I.B., Huh, T.L., 2002. Isolation and expression of a novel neuron-specific onecut homeobox gene in zebrafish. Mech. Dev. 112, 199–202.

Inoue, A., Takahashi, M., Hatta, K., Hotta, Y., Okamoto, H., 1994. Developmental regulation of islet-1 mRNA expression during neuronal differentiation in embryonic zebrafish. Dev. Dyn. 199, 1–11.

Iulianella, A., Vanden Heuvel, G., Trainor, P., 2003. Dynamic expression of murine Cux2 in craniofacial, limb, urogenital and neuronal primordia. Gene Expression Patterns 3, 571–577.

Jackman, W.R., Langeland, J.A., Kimmel, C.B., 2000. Islet reveals segmentation in the Amphioxus hindbrain homolog. Dev. Biol. 220, 16–26.

Kania, A., Jessell, T.M., 2003. Topographic motor projections in the limb imposed by LIM homeodomain protein regulation of ephrin-A:EphA interactions. Neuron 38, 581–596.

Kania, A., Johnson, R.L., Jessell, T.M., 2000. Coordinate roles for LIM homeobox genes in directing the dorsoventral trajectory of motor axons in the vertebrate limb. Cell 102, 161–173.

Kikuchi, Y., Segawa, H., Tokumoto, M., Tsubokawa, T., Hotta, Y., Uyemura, K., Okamoto, H., 1997. Ocular and cerebellar defects in zebrafish induced by overexpression of the LIM domains of the islet-3 LIM/homeodomain protein. Neuron 18, 369–382.

Kimura, C., Takeda, N., Suzuki, M., Oshimura, M., Aizawa, S., Matsuo, I., 1997. Cis-acting elements conserved between mouse and pufferfish Otx2 genes govern the expression in mesencephalic neural crest cells. Development 124, 3929–3941.

Kirchhamer, C.V., Bogarad, L.D., Davidson, E.H., 1996. Developmental expression of synthetic cis-regulatory systems composed of spatial control elements from two different genes. Proc. Natl. Acad. Sci. U. S. A. 93, 13849–13854.

Leighton, P.A., Mitchell, K.J., Goodrich, L.V., Lu, X., Pinson, K., Scherz, P., Skarnes, W.C., Tessier-Lavigne, M., 2001. Defining brain wiring patterns and mechanisms through gene trapping in mice. Nature 410, 174–179.

Lewis, K.E., Eisen, J.S., 2004. Paraxial mesoderm specifies zebrafish primary motoneuron subtype identity. Development 131, 891–902.

Lobe, C.G., Koop, K.E., Kreppner, W., Lomeli, H., Gertsenstein, M., Nagy, A., 1999. Z/AP, a double reporter for cre-mediated recombination. Dev. Biol. 208, 281–292.

Long, Q., Meng, A., Wang, H., Jessen, J.R., Farrell, M.J., Lin, S., 1997. GATA-1 expression pattern can be recapitulated in living transgenic zebrafish using GFP reporter gene. Development 124, 4105–4111.

Loots, G.G., Locksley, R.M., Blankespoor, C.M., Wang, Z.E., Miller, W., Rubin, E.M., Frazer, K.A., 2000. Identification of a coordinate regulator of interleukins 4, 13, and 5 by cross-species sequence comparisons. Science 288, 136–140.

Mizuguchi, R., Sugimori, M., Takebayashi, H., Kosako, H., Nagao, M., Yoshida, S., Nabeshima, Y., Shimamura, K., Nakafuku, M., 2001. Combinatorial roles of olig2 and neurogenin2 in the coordinated induction of pan-neuronal and subtype-specific properties of motoneurons. Neuron 31, 757–771.

Müller, F., Chang, B., Albert, S., Fischer, N., Tora, L., Strahle, U., 1999. Intronic enhancers control expression of zebrafish sonic hedgehog in floor plate and notochord. Development 126, 2103–2116.

Müller, F., Blader, P., Strahle, U., 2002. Search for enhancers: teleost models in comparative genomic and transgenic analysis of cis regulatory elements. Bioessays 24, 564–572.

Müller, M., Jabs, N., Lorke, D.E., Fritzsch, B., Sander, M., 2003. Nkx6.1 controls migration and axon pathfinding of cranial branchio-motoneurons. Development 130, 5815–5826.

Myers, P.Z., 1985. Spinal motoneurons of the larval zebrafish. J. Comp. Neurol. 236, 555–561.

Nguyen, D.N., Rohrbaugh, M., Lai, Z., 2000. The *Drosophila* homolog of Onecut homeodomain proteins is a neural-specific transcriptional activator with a potential role in regulating neural differentiation. Mech. Dev. 97, 57–72.

Novitch, B.G., Chen, A.I., Jessell, T.M., 2001. Coordinate regulation of motor neuron subtype identity and pan-neuronal properties by the bHLH repressor Olig2. Neuron 31, 773–789.

Novitch, B.G., Wichterle, H., Jessell, T.M., Sockanathan, S., 2003. A requirement for retinoic acid-mediated transcriptional activation in ventral neural patterning and motor neuron specification. Neuron 40, 81–95.

Okamoto, H., Kuwada, J.Y., 1991. Outgrowth by fin motor axons in wildtype and a finless mutant of the Japanese medaka fish. Dev. Biol. 146, 49–61.

Pattyn, A., Hirsch, M., Goridis, C., Brunet, J.F., 2000. Control of hindbrain motor neuron differentiation by the homeobox gene Phox2b. Development 127, 1349–1358.

Pike, S.H., Melancon, E.F., Eisen, J.S., 1992. Pathfinding by zebrafish motoneurons in the absence of normal pioneer axons. Development 114, 825–831.

Rastinejad, F., 2001. Retinoid X receptor and its partners in the nuclear receptor family. Curr. Opin. Struct. Biol. 11, 33–38.

Reyes, R., Haendel, M., Grant, D., Melancon, E., Eisen, J.S., 2004. Slow

degeneration of zebrafish Rohon-Beard neurons during programmed cell death. Dev. Dyn. 229, 30–41.
Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, New York.
Sander, M., Paydar, S., Ericson, J., Briscoe, J., Berber, E., German, M., Jessell, T.M., Rubenstein, J.L., 2000. Ventral neural patterning by Nkx homeobox genes: Nkx6.1 controls somatic motor neuron and ventral interneuron fates. Genes Dev. 14, 2134–2139.
Schwartz, S., Zhang, Z., Frazer, K.A., Smit, A., Riemer, C., Bouck, J., Gibbs, R., Hardison, R., Miller, W., 2000. PipMaker—A web server for aligning two genomic DNA sequences. Genome Res. 10, 577–586.
Segawa, H., Miyashita, T., Hirate, Y., Higashijima, S., Chino, N., Uyemura, K., Kikuchi, Y., Okamoto, H., 2001. Functional repression of Islet-2 by disruption of complex with Ldb impairs peripheral axonal outgrowth in embryonic zebrafish. Neuron 30, 423–436.
Sharma, K., Leonard, A.E., Lettieri, K., Pfaff, S.L., 2000. Genetic and epigenetic mechanisms contribute to motor neuron pathfinding. Nature 406, 515–519.
Sockanathan, S., Perlmann, T., Jessell, T.M., 2003. Retinoid receptor signaling in postmitotic motor neurons regulates rostrocaudal positional identity and axonal projection pattern. Neuron 40, 97–111.
Stanojevic, D., Small, S., Levine, M., 1991. Regulation of a segmental stripe by overlapping activators and repressors in the *Drosophila* embryo. Science 254, 1385–1387.
Stuart, G.W., McMurray, J.V., Westerfield, M., 1988. Replication, integration and stable germ-line transmission of foreign sequences injected into early zebrafish embryos. Development 103, 403–412.
Tanabe, Y., William, C., Jessell, T.M., 1998. Specification of motor neuron identity by the MNR2 homeodomain protein. Cell 95, 67–80.
Thaler, J.P., Lee, S.K., Jurata, L.W., Gill, G.N., Pfaff, S.L., 2002. LIM factor Lhx3 contributes to the specification of motor neuron and interneuron identity through cell-type-specific protein–protein interactions. Cell 110, 237–249.
Thaler, J.P., Koo, S.J., Kania, A., Lettieri, K., Andrews, S., Cox, C., Jessell, T.M., Pfaff, S.L., 2004. A postmitotic role for Isl-class LIM homeodomain proteins in the assignment of visceral spinal motor neuron identity. Neuron 41, 337–350.
Thor, S., Thomas, J.B., 1997. The *Drosophila* islet gene governs axon pathfinding and neurotransmitter identity. Neuron 18, 397–409.
Tokumoto, M., Gong, Z., Tsubokawa, T., Hew, C.L., Uyemura, K., Hotta, Y., Okamoto, H., 1995. Molecular heterogeneity among primary motoneurons and within myotomes revealed by the differential mRNA expression of novel islet-1 homologs in embryonic zebrafish. Dev. Biol. 171, 578–589.
Tsuchida, T., Ensini, M., Morton, S.B., Baldassare, M., Edlund, T., Jessell, T.M., Pfaff, S.L., 1994. Topographic organization of embryonic motor neurons defined by expression of LIM homeobox genes. Cell 79, 957–970.
Uchikawa, M., Ishida, Y., Takemoto, T., Kamachi, Y., Kondoh, H., 2003. Functional analysis of chicken Sox2 enhancers highlights an array of diverse regulatory elements that are conserved in mammals. Dev. Cell 4, 509–519.
Weinberg, E.S., Allende, M.L., Kelly, C.S., Abdelhamid, A., Murakami, T., Andermann, P., Doerre, O.G., Grunwald, D.J., Riggleman, B., 1996. Developmental regulation of zebrafish MyoD in wild-type, no tail and spadetail embryos. Development 122, 271–280.
Westerfield, M., 1995. The Zebrafish Book. University of Oregon, Eugene.
Westerfield, M., Wegner, J., Jegalian, B.G., DeRobertis, E.M., Puschel, A.W., 1992. Specific activation of mammalian Hox promoters in mosaic transgenic zebrafish. Genes Dev. 6, 591–598.
William, C.M., Tanabe, Y., Jessell, T.M., 2003. Regulation of motor neuron subtype identity by repressor activity of Mnx class homeodomain proteins. Development 130, 1523–1536.
Zerucha, T., Stuhmer, T., Hatch, G., Park, B.K., Long, Q., Yu, G., Gambarotta, A., Schultz, J.R., Rubenstein, J.L., Ekker, M., 2000. A highly conserved enhancer in the Dlx5/Dlx6 intergenic region is the site of cross-regulatory interactions between Dlx genes in the embryonic forebrain. J. Neurosci. 20, 709–721.