DEVELOPMENTAL DYNAMICS 228:30–40, 2003

ARTICLE

Germ-Line Transmission of a Myocardium-Specific GFP Transgene Reveals Critical Regulatory Elements in the Cardiac Myosin Light Chain 2 Promoter of Zebrafish

Chiu-Ju Huang,[1] Chi-Tang Tu,[1] Chung-Der Hsiao,[1] Fong-Jou Hsieh,[2] and Huai-Jen Tsai[1]*

In response to the lack of a transgenic line of zebrafish labeled with heart-specific fluorescence in vivo to serve as a research model, we cloned a 1.6-kb polymerase chain reaction (PCR) -product containing the upstream sequence (−870 bp), exon 1 (39 bp), intron 1 (682 bp), and exon 2 (69 bp) of the zebrafish cardiac myosin light chain 2 gene, (*cmlc*2). A germ-line transmitted zebrafish possessing a green fluorescent heart was generated by injecting this PCR product fused with the green fluorescent protein (GFP) gene with ends consisting of inverted terminal repeats of an adeno-associated virus. Green fluorescence was intensively and specifically expressed in the myocardial cells located both around the heart chambers and the atrioventricular canal. Neither the epicardium nor the endocardium showed fluorescent signals. The GFP expression in the transgenic line faithfully recapitulated with the spatial and temporal expression of the endogenous *cmlc*2. Promoter analysis showed that the fragment consisting of nucleotides from −210 to 34 (−210/34) was sufficient to drive heart-specific expression, with a −210/−73 motif as a basal promoter and a −210/−174 motif as an element involved in suppressing ectopic (nonheart) expression. Interestingly, a germ-line of zebrafish whose GFP appeared ectopically in all muscle types (heart, skeletal, and smooth) was generated by injecting the fragment including a single nucleotide mutation from G to A at −119, evidence that A at −119 combined with neighboring nucleotides to create a consensus sequence for binding myocyte-specific enhancer factor-2. *Developmental Dynamics 228:30–40, 2003.* © 2003 Wiley-Liss, Inc.

Key words: zebrafish; transgene; *cmlc*2; promoter analysis; heart-specific expression; myocardium; nucleotide mutation; muscle-specific expression

Received 20 February 2003; Accepted 13 May 2003

INTRODUCTION

Heart disease is currently one of the most common causes of human death. A large number of researchers are currently searching for a simple animal research model to assist in finding either a cure for heart disease or a novel gene that regulates heart development. Zebrafish (*Danio rerio*) are viewed as having some advantages for heart-related research (Fishman and Stainier, 1994). Immunohistochemical detection, such as that accomplished by using MF-20 monoclonal antibody against myosin heavy chain (Bader et al., 1982; Hu et al., 2000; Sehnert et al., 2002), is commonly used to identify cardiovascular cells and to screen mutated hearts. However, this antiserum is not original against fish antigen and is not highly cardiac-specific. In addition, immunochemical work is laborious, costly, and cannot be dynamically done in vivo. Therefore, heart-specific fluorescent transgenic zebrafish will be valuable as research models for tracing

[1]Institute of Molecular and Cell Biology, National Taiwan University, Taipei, Taiwan
[2]Department of Internal Medicine and Obstetrics and Gynecology, National Taiwan University Hospital, Taipei, Taiwan
Grant sponsor: the National Science Council, Taiwan; Grant number: NSC 90-2313-B-002-319; Grant sponsor: the National Taiwan University Hospital; Grant number: NTUH 91A08.
*Correspondence to: Huai-Jen Tsai, Institute of Molecular and Cell Biology, National Taiwan University, 1 Roosevelt Road, Section 4, Taipei, Taiwan 106. E-mail: hjtsai@ntu.edu.tw

DOI 10.1002/dvdy.10356

the developmental fates of heart cells, finding new heart-specific genes and functions, establishing biological indices of environmental pollutants, and studying the efficacy of therapeutic drugs. However, an appropriate zebrafish model labeled with heart-specific fluorescence in vivo has yet to be developed.

The myosin light chain (MLC) has been identified as a major contractile component of cardiac and other striated muscles; MLC-2 is a regulatory light chain and potential modulator of contractile activity in heart and skeletal muscle cells (Stull et al., 1985; Silver et al., 1986). The *mlc*2, which has been identified in *Drosophila melanogaster* (Parker et al., 1985) and *Caenorhabditis elegans* (Cummins and Anderson, 1988), is recognized as having evolved into the atrial isoform MLC-2a and ventricular isoform MLC-2v in birds (Winter et al., 1985) and mammals (Henderson et al., 1989; Kubalak et al., 1994; Doevendans et al., 2000). Mouse *mlc*-2v is considered a ventricular-specific gene, but mouse *mlc*-2a is downregulated in the ventricular chamber during ventricular maturation and specification, making it a likely candidate for service as a cardiac development marker (Kubalak et al., 1994). Both *mlc*-2a and -2v mutations have been linked to cardiac disease (Poetter et al., 1996).

Although the *mlc*2 and its orthologs have been cloned and partly characterized in several species, little is known about the molecular structure and regulatory sequence of the cardiac MLC-2 gene (*cmlc*2) in fish except for the description provided by Yelon et al. (1999) of *cmlc*2 transcripts throughout the heart tubes of zebrafish. In this study, we cloned an upstream sequence −870 bp, exon 1, intron 1, and exon 2 of the zebrafish *cmlc*2 and fused it with a green fluorescent protein (GFP) cDNA. The result was a transgenic zebrafish line possessing GFP that was specifically expressed in the myocardial cells. Meanwhile, we also determined the minimal upstream region and key *cis*-elements of zebrafish *cmlc*2 that control heart-specific expression.

MATERIALS AND METHODS

Zebrafish Breeding

A zebrafish AB strain was cultured and maintained according to procedures described by Westerfield (1995). Embryonic stages were recorded as hours postfertilization (hpf) and days postfertilization (dpf) following Kimmel et al. (1995).

Genomic DNA Extraction and Adaptor Ligation

Genomic DNA was extracted from zebrafish larvae at 48 hpf. One microgram of *Spe*I-cut DNA fragment and 100 pmol each of Pad1 (5′-TGCGAGTAAGGATCCTCACGCAAGGAATTCCGACCAGACACC-3′) and PR-*Spe*I (5′-CTAGGGTGTCTGGTCGC-3′) adaptors were added to a final volume of 20 μl of ligation buffer. After the mixture was preheated to 70°C, six units of *Avr*II and three units of T4 DNA ligase (Promega) were added and reacted at 4°C for 16 hr. Unligated adaptors were removed with Microcon-100 (Amicon).

Polymerase Chain Reaction and PCR Products

Target DNA was amplified by performing polymerase chain reaction (PCR) twice. The first PCR was conducted in 20 μl of solution containing 20 ng of DNA (ligated with adaptors serving as a DNA template), 1 pmol of P1 primer (5′-TGCGAGTAAGGATC CTCACGCA-3′), 4 pmol of CML1 primer (5′-ACTCCATCCCGGTTCTGATCT-3′), 200 pmol of each dNTP, and 1 unit of ExTaq DNA polymerase (TaKaRa). VioTaq DNA polymerase (Viogene) was used to obtain PCR products containing the mutated sequence. After the DNA was denatured, PCR was performed for 35 cycles: each cycle was conducted at 94°C for 30 sec and then at 68°C for 6 min. The second PCR was conducted by using 1 μl of the first PCR product, 4 pmol of P1 primer, 4 pmol of CML2 primer (5′-GGAGAAGACATTGGAAGAGCCT-3′), and 1 unit of ExTaq. Secondary PCR products were inserted into a pGEM-T vector (Promega) for DNA sequencing.

Plasmid Constructs

The primers CML4-*Xho*I (5′-AACAACTCGAGTGTGACCAAAGCTTAAATC-3′) and CML2-*Nco*I (5′-CTCAACCATGGAGAAGACATTGGAAGA-3′) were designed using the sequences of the 1.6-kb PCR-product described in the preceding section. The final PCR product was cut with *Xho*I and *Nco*I and ligated into *Xho*I- and *Nco*I-cut pEGFP-ITR (Chou et al., 2001). The resulting plasmid construct, pICMLE, consisted of the upstream sequence region, exon 1, intron 1, and exon 2 of the *cmlc*2 that was fused with GFP cDNA to form a cassette. This cassette was flanked on both sides by the 145-bp inverted terminal repeats derived from an adeno-associated virus (AAV-ITR).

To identify the regulatory elements in the proximal region of the zebrafish *cmlc*2, we constructed several deletion plasmids encompassing different lengths of the GFP-fused upstream regulatory sequence: pICMLCE-(−870/787), -(−656/787), -(−210/787), -(−68/787), -(−870/34), -(−870/−73), -(−210/34), -(−210/34)Δ(−120/−111), -(−210/3), -(−210/−73), -(−173/34), -(−136/34) and -(210/−69)mCMV. Plasmid pICMLCE- (−210/34)Δ(−120/−111), in which region −120/−111 was deleted from −210/34, was designed for studying (1) whether G at −119 combined with neighboring A/T rich sequences is a sole element for heart-specific expression and (2) whether TATA at −118/−113 functions as a TATA box for zebrafish *cmlc*2 transcription. Plasmid pICMLE-(−210/−69)mCMV, in which the −210/−69 of *cmlc*2 was linked with a TATA box provided by the cytomegalovirus (CMV) basal promoter, was designed for studying whether the −210/−69 motif is capable of suppressing the nonspecific expression driven by the CMV promoter. Two other plasmids, pCMLE-(−210/34) and -(−210/−73), in which AAV-ITR was not flanked, were designed for studying the effectiveness of AAV-ITR in enhancing GFP expression driven by the regulatory *cmlc*2. All of the plasmids described above were linearized with *Not*I and microinjected into the zebrafish embryos at a concentration of 25 ng/nl.

Rapid Amplification of cDNA Ends and Primer Extension

Basically, the 5′–rapid amplification of cDNA ends (RACE) procedure followed the method described by Huang et al. (1999), except that a forward primer (5′-GGCCACGCGTCGACTAGTACTCCCCCCCCC-3′) and a reverse primer (5′-GGTTTGAGATGATGCTCTACTCATAGTC-3′) were used along with an annealing temperature of 58°C.

Adult zebrafish total heart RNA was extracted with TRI reagent (Molecular Research Center, Inc.), hybridized with the (r-^{32}P)ATP-labeled primer CML2-*Nco*I at 50°C for 22 hr, and precipitated with ethanol. Pellets were resolved in reverse transcription buffer and transcribed with M-MLV reverse transcriptase (Promega). Extension reactions were performed by means of simultaneous sequencing with the subcloned DNA using the same primer on a 6% polyacrylamide/8M urea gel.

Gene Transfer and Fluorescent Signal Observations

Following the procedure described by Hsiao et al. (2001), 3 nl of the DNA sample were microinjected into the cytoplasm of single-cell fertilized eggs, which were incubated at 28°C in dishes containing low concentrations of methylene blue solution. By using a fluorescence microscope (MZ FLIII, Leica), we observed heart development and green fluorescence in embryos. To observe in vivo the GFP expression at the cellular level in the transgenic line embryos, we used a Zeiss Axioplan microscope with Nomarski optics and fluorescence. Images were captured with a Zeiss Axiocam camera using the Zeiss Axiovision software.

Germ-Line Transmission of *cmlc2*::GFP Transgenic Zebrafish

Transgenic zebrafish embryos displaying heart-specific GFP expression were raised to adulthood. Pairs of male and female founders were initially mated to screen for transgenic zebrafish. If GFP expression was found in some of the resulting embryos, the founders were separated and allowed to mate with wild strains to identify the putative germline transmitting parent. One pair each of transgenic and wild zebrafish were kept in a 22 × 14 × 13 cm tank. A minimum of 200 embryos were examined per cross.

Whole-Mount in Situ Hybridization

Whole-mount in situ hybridization with digoxigenin (DIG) -labeled riboprobes was performed according to the protocol described by Jowett (2001). PCR products amplified from pGEM-T containing full lengths of *cmlc*2 cDNA with T7 primers were used to synthesize riboprobes using a DIG RNA Labeling Kit (Roche), followed by in vitro transcription reaction with T7 RNA polymerase. Whole-mount embryos were cleaned in 100% methanol, placed in 100% glycerol, and evaluated with a differential interference contrast microscope (DMR, Leica) equipped with a COOLPIX 990 color digital camera (Nikon).

Histology Section and GFP Observations

Adult zebrafish (3 months) of transgenic line (A34) were fixed with 4% paraformaldehyde in 1× phosphate buffered saline for 16 hr. Hearts were incubated directly into OCT without washing, and were serially cryosectioned with a thickness of 10 μm (MICROM HM5000). The sections were observed and photographed under a fluorescence microscope and a Nomarski microscope.

Injections of Morpholino Oligonucleotides

For targeted knockdown of the *nkx2.5* and *gata4* genes, morpholino antisense oligonucleotides of *nkx*2.5-MO (TCATTTGGCTAGAGAACAT TGC) and *gata*-4-MO (GCCATCGTTACACCT TGATACATAT) were synthesized (GeneTools, LLC). Each MO was injected at concentrations of 2.3, 4.6, 9.2, and 11.5 ng per embryo derived from the *cmlc*2::GFP transgenic line. *nkx*2.5-MO combined with *gata*-4-MO at a concentration of 4.6 ng was also injected. As for the control groups, we injected the following: (1) *cTnT*-MO (CTTCCACTTCTTCGTTGTCTGACAT) into embryos, because Sehnert et al. (2002) reported that *cTnT*-MO produced a silent heart in zebrafish; and (2) *EGFP*-MO (ACAGCTCCTCGCCCTTGCTCACCAT) into embryos as a positive control. Approximately 20–30 surviving embryos of each group were used to measure the beating rate of heart and to observe the intensity of GFP fluorescent signal.

RESULTS

Genomic and Deduced Amino Acid Sequences of Zebrafish *cmlc2*

The results of our sequencing of the amplified 1.6-kb PCR fragment showed that it contained a 870-bp segment of the 5′ upstream regulatory region, a 39-bp segment of the exon 1, a 682-bp segment of the intron 1, and a 69-bp segment of the exon 2 of the zebrafish *cmlc*2. The nucleotide sequences of the upstream and partial downstream segments of the gene are presented in Figure 1. According to a primer extension assay, the transcription initiation site (+1) of the zebrafish *cmlc*2 was 36 bp upstream of the start codon. Based on the results of analysis of the −210/39 region conducted by using the TRSEARCH (version 1.3) program, we propose that several putative *cis*-elements might be linked to a family of DNA binding proteins such as GATA, Cdx, MZF, CREB, and Nkx either on the coding strand or on the noncoding strand (Fig. 1).

After performing the 5′-RACE and combining it with part of the cDNA taken from AF114428 (Yelon et al., 1999), we obtained a complete cDNA of the zebrafish *cmlc*2 and found that it encoded a putative polypeptide with 172 amino acid residues (GenBank accession no. AF425743). The deduced amino acid sequence of the zebrafish cardiac MLC-2 polypeptide shared 39, 46, 74, and 60% identity with the sequences of *D. melanogaster* (AAL-25408), *C. elegans* (NP510828), mouse

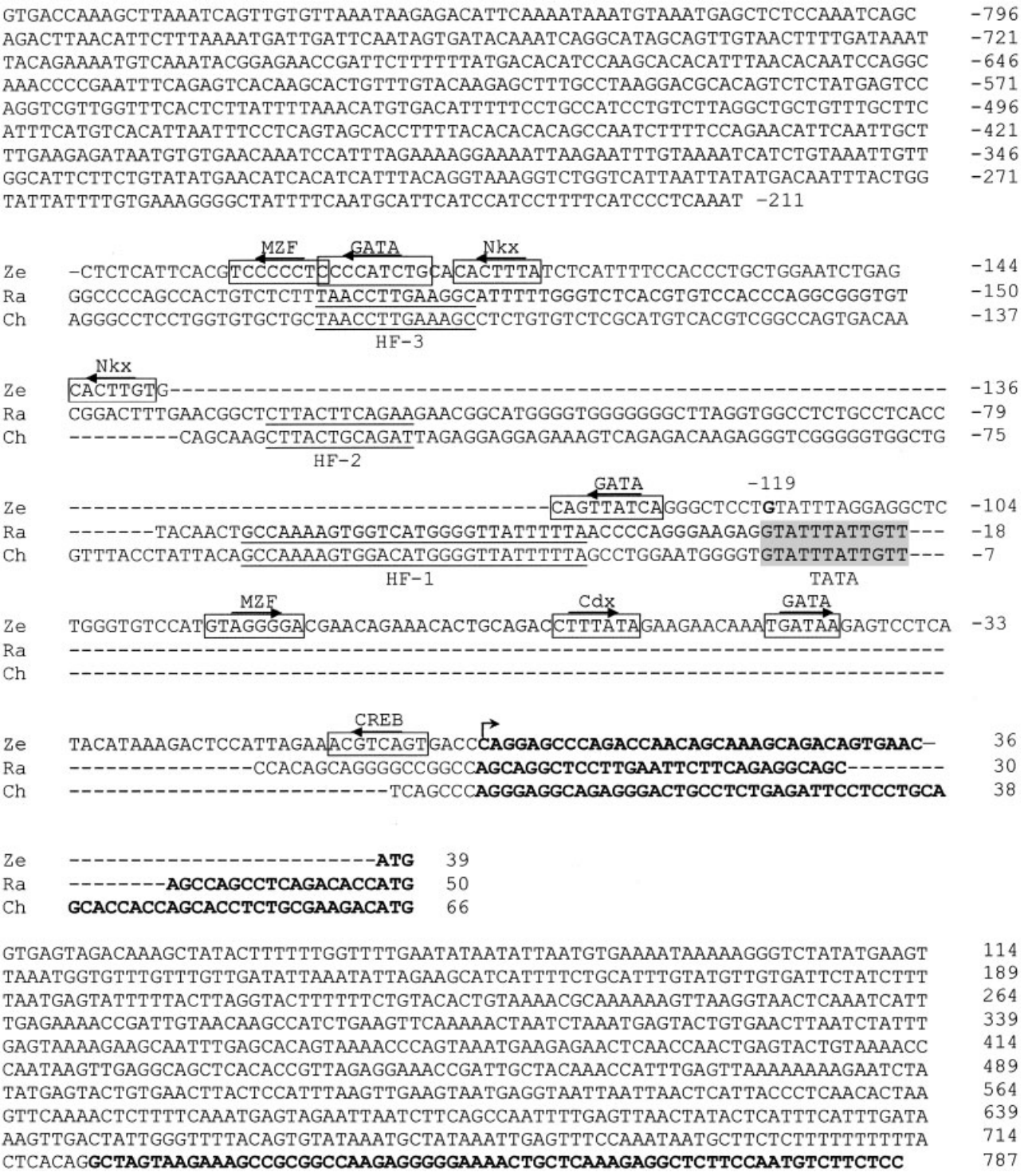

Fig. 1. The proximal regulatory sequence of the zebrafish *cmlc*2. Sequence comparisons of the upstream region from −210 to 39 in zebrafish (Ze) with their counterparts in rats (Ra; Henderson et al., 1989) and chicks (Ch; Arnold et al., 1988) are shown. Putative and known *cis*-elements for transcription factors are boxed and arrowed: forward arrows represent sites at the coding strand, whereas reverse arrows represent sites at the template strand. The single mutation of G nucleotide at −119 is marked. HF-binding sites and TATA box conserved between rats and chicken *cmlc*2 are underlined and shadowed, respectively. The transcription initiation site is marked with a curved arrow. Exons are indicated by bold type. Dashes represent gaps that were created to maximize the degree of identity among the compared sequences.

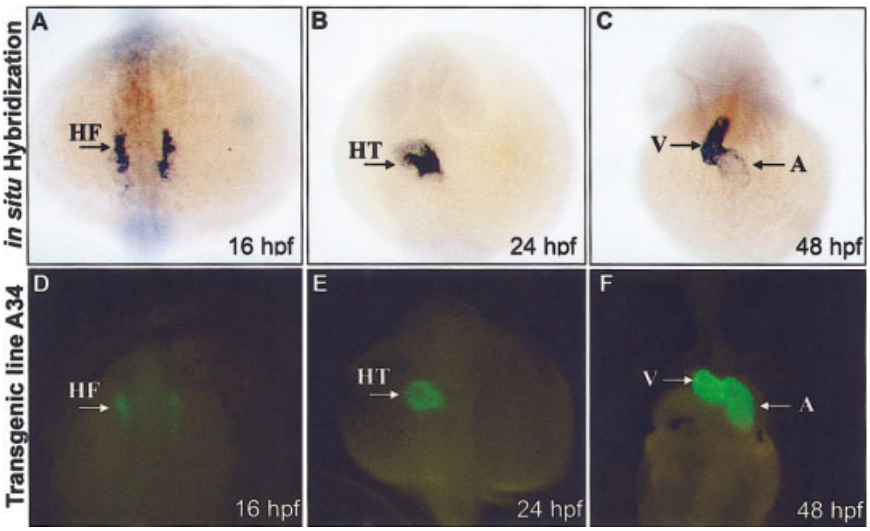

Fig. 2. Temporal and spatial expression of endogenous *cmlc*2 transcripts and green fluorescent protein (GFP) of transgenic lines. To analyze the fidelity of transgenic GFP expression driven by the regulatory region of *cmlc*2 that we cloned, we compared the endogenous *cmlc*2 expression by means of whole-mount in situ hybridization in developing embryos. **A–C:** Wild-type embryos were hybridized with a zebrafish *cmlc2* antisense riboprobe. **D–F:** Green fluorescent signals were observed in embryos derived from the transgenic line. GFP patterns at various developmental stages with a fluorescent microscope. A, atrium; V, ventricle; HF, heart field; HT, heart tube; hpf, hours postfertilization.

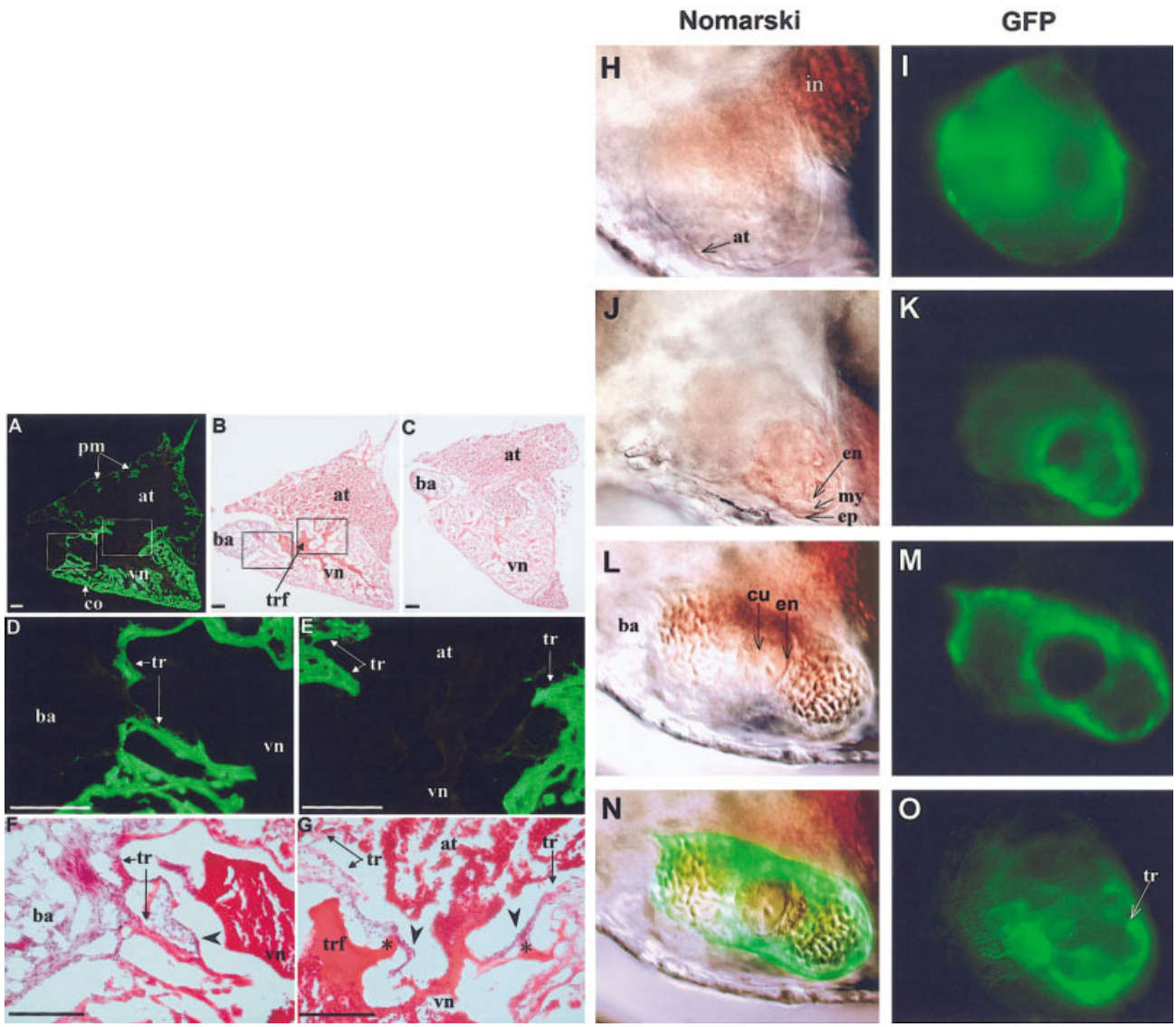

Fig. 3.

MLC-2a (NP075017), and mouse MLC-2v (AAB37470), respectively.

Generation of *cmlc*2::GFP Transgenic Zebrafish

To generate a transgenic line of zebrafish showing heart-specific fluorescence, we injected *Not*I-linearized pICMLE-(−870/787) (that is, with flanking AAV-ITR) into fertilized eggs. A 50–70% survival rate of transferred zebrafish embryos was recorded after three days of microinjection. Transient transgenic assay showed that 45–50% of the surviving embryos displayed cardiac tissue-specific GFP expression. Approximately 50% of these embryos were raised to adulthood, and heart-specific fluorescence was observed throughout their life spans.

Of the 324 founders, 37 (11.4%) individuals produced offspring that expressed GFP, including 34 showing heart-specific fluorescence. Gonad mosaic rates among the transgenic lines ranged from 1.0 to 33.7%, indicating a high degree of germ-line mosaicism. Transgene transmission rates from F1 to F2 were approximately 50%; transmission rates resulting from the inter-crossing of two F1 zebrafish showed that fluorescence ranged between 73 and 77%. These data all fit well with standard Mendelian inheritance ratios, indicating that the transgene was integrated into a single chromosomal locus.

The GFP-labeled hearts of F3 homozygotic transgenic zebrafish were observed at high resolution during the embryonic stage. Successive atrium and ventricle contractions were easily observed on both sides of individual fish. The atrium and the ventricle were highly organized into a compact structure, which could be easily observed at 7 dpf (data not shown).

Endogenous *cmlc*2 Expression and Transgenic GFP Expression Driven by the Cloned Regulatory Region of *cmlc*2

To analyze the fidelity of transgenic GFP expression driven by the regulatory region of *cmlc*2 that we cloned, the endogenous *cmlc*2 expression after whole-mount in situ hybridization in developing embryos was studied. Results showed that the zebrafish *cmlc*2 was activated in the bilateral heart field around 16 hpf (Fig. 2A) and that cardiac cells expressing *cmlc*2 were fused in a single heart tube and elongated by convergent extension by 24 hpf (Fig. 2B). These results were consistent with those reported by Yelon et al. (1999). We also noticed that *cmlc*2 transcripts were unevenly expressed in the atrium and ventricle by 48 hpf. The intensity of expression in the ventricle was much higher than that in the atrium (Fig. 2C).

F2 progeny were inter-crossed to produce a homozygotic F3 generation. The time of initial GFP detection in the transgenic lines varied: in 35 of 37 lines, GFP was activated by 20 hpf, approximately 4 hr after the detection of the endogenous *cmlc*2 transcripts. In two lines (A34 and A130), GFP was visible at approximately 16 hpf (Fig. 2D), which was exactly the stage at which the endogenous *cmlc*2 was detected by means of whole-mount in situ hybridization. At 24 hpf (after zebrafish hearts begin to beat), robust expression of GFP was noted in fish heart tubes (Fig. 2E). Although GFP expression was uniform in both the atrium and ventricle at 48 hpf (Fig. 2F), the GFP intensity in a given chamber was dependent on which chamber was contracting when the picture was taken. The cardiac-specific GFP driven by the *cmlc*2 promoter appeared throughout the life spans of all transgenic individuals.

GFP Transgene Was Expressed in All Myocardial Cells

We examined the GFP expression driven by the *cmlc*2 promoter in the embryos of a transgenic line (A34). When hearts were cryosectioned, the GFP signals were intensively observed in the ventricular trabeculae, compact layer, and arterial pectinate muscle (Fig. 3A,D,E). No green fluorescence was seen in the bulbus arteriosus, trabecular folds, or leaflet valves of the atrioventricular and bulboventricular (Fig. 3D,E). Furthermore, by using an Axioplan microscoping with Nomarski optics, we found that green fluorescent signals appeared specifically in the myocardium cells of the atrium, ventricle, and atrioventricular boundary (Fig. 3H–J). The epicardium, endocardium, and cushion all showed no green fluorescence (Fig. 3K,L). Green fluorescence could be traced in the trabeculae of 79-hpf embryos when the myocardium started to trabeculate (Fig. 3M). This signal became more prominent in the trabeculae of 5-dpf embryos, in which the ventricles showed extensive trabeculation (Fig. 3D,E). No morphologic defects (Fig. 3B,C) of hearts were observed in transgenic zebrafish lines. The transgene had been stably transmitted to the F4 generation.

Promoter Analysis of Zebrafish *cmlc*2

To determine the regulatory *cis*-elements in the proximal region of the zebrafish *cmlc*2, we constructed upstream series deletion fragments linked with GFP genes (Fig. 4A,B). Based on a transient assay, the GFP-positive signals that appeared in the 3-dpf embryos are summarized in Figure 4B. High levels of fluorescence were observed in the hearts of embryos injected with the fragments pICMLE-(−870/787), -(−656/

Fig. 3. The green fluorescent protein (GFP) transgene driven by the upstream regulatory region of the zebrafish *cmlc*2 was specifically expressed in myocardial cells. **A–G:** Histologic sections. **H–O:** In vivo observations. Frozen sections of hearts from a transgenic line observed under a fluorescence microscope (A) and hematoxylin and eosin (H & E) stain in brightfield (B); and hearts from wild-type under H & E stain under brightfield (C). The bulboventricular valve areas shown on A are magnified; D (fluorescent field) and F (brightfield) correspond to the left box; E (fluorescent field) and G (brightfield) correspond to the right box. H,J,L: Images were taken using Nomarski optics. I,K,M,O: Images were taken using a fluorescence microscope. N is an overlay figure between M and L. Arrowheads point to the leaflet of the bulboventricular valve. Asterisks indicate the leaflet supporter. At, atrium; ba, bulbus arteriosus; co, compact layer; cu, cushion; ep, epicardium; en, endocardium; in, inflow; my, myocardium; pm, pectinate muscle; tr, trabeculae; trf, trabecular fold; vn, ventricle. Scale bars = 100 μm.

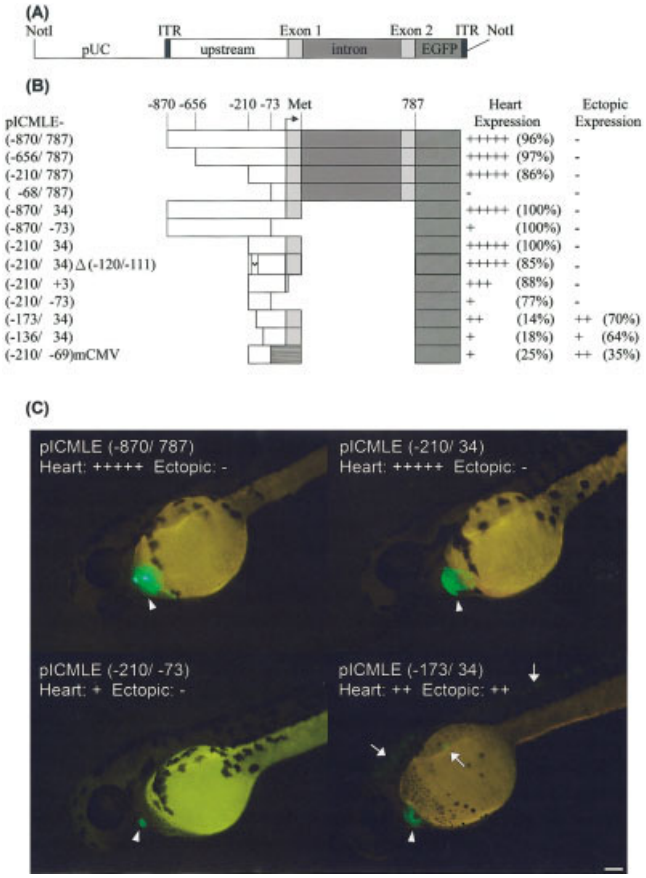

Fig. 4. Promoter analysis. Transient expression patterns of green fluorescence observed in the transgenic embryos. **A:** Schematic map of pICMLE. An enhanced green fluorescent protein (EGFP) reporter gene was fused with the upstream sequence, exon 1, intron 1, and exon 2 of the zebrafish *cmlc2*. AAV-ITR sequences (ITR) flanked both ends of the cassettes. **B:** Locations and expression levels of green fluorescent protein (GFP) in 3 days postfertilization embryos. The nucleotide positions, transcription initiation site (curved arrow), and start codon (Met) are indicated above the construct. The strongest level of fluorescence expressed in the entire heart is indicated by +++++, whereas the weakest expression level of GFP in the partial heart is indicated by +. The percentages were calculated based on the following numbers of GFP-positive transient embryos for each construct: 70, 58, 118, 0, 9, 5, 38, 53, 51, 13, 66, 11, and 20 for plasmids (−870/787), -(−656/787), -(−210/787), -(−68/787), -(−870/34), -(−870/−73), -(−210/34), -(−210/34)Δ(−120/−111), -(−210/3), -(−210/−73), -(−173/34), -(−136/34), and -(−210/−69)mCMV, respectively. **C:** Transgenic embryos derived by injecting various pICMLE plasmids as indicated, which were observed under a fluorescence microscope. Arrowheads point to the locations of heart-specific green fluorescence; arrows point to the locations of skeletal-specific green fluorescence.

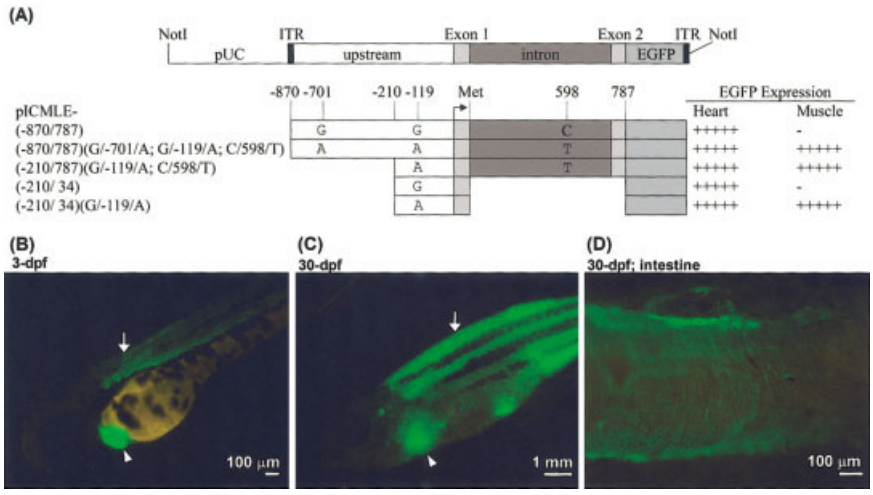

Fig. 5. Single nucleotide mutation led to the ectopic expression of green fluorescent protein (GFP) in transgenic zebrafish. The line was derived from embryos injected with mutated sequences of the zebrafish *cmlc2*. **A:** The original plasmids serving as controls were pICMLE-(−870/787) and -(−210/34), with G at −701, G at −119, and C at 598. The plasmid pICMLE-(−870/787) (G/−701/A; G/−119/A; C/598/T) contained three point mutations within the −870/787 segment: G→A at −701, G→A at −119, and C→T at 598. The plasmid pICMLE-(−210/787)(G/−119/A; C/598/T) contained two point mutations, and the plasmid pICMLE-(−210/34)(G/−119/A) contained a single nucleotide substitution. +++++ indicates a very intense green signal, whereas − indicates an undetectable signal. The percentages of GFP signals observed in both the hearts and other muscles of the GFP-positive embryos were 55% (155/283), 46% (5/11), and 70% (65/93) for injecting plasmids (−870/787)(G/−701/A; G/−119/A; C/598T), (−210/787)(G/−119/A; C/598/T), and (−210/34)(G/−119/A), respectively. **B–D:** Green fluorescent signals in the transgenic line derived from embryos injected with pICMLE-(−870/787)(G/−701/A; G/−119/A; C/598/T). Arrowheads indicate the hearts, and arrows indicate the skeletal muscles of zebrafish at 3 days postfertilization (B) and 30 days postfertilization (C). Green fluorescence was also observed in the smooth muscles of the intestine (D).

787), -(−210/787), -(−870/34), and -(−210/34) (Fig. 4B,C). These results led us to conclude that fragment −210/34 was a minimal promoter consisting of *cis*-regulatory elements required for heart specificity.

GFP expression was noted in the heart and nonheart tissues and cells (e.g., skin, skeletal muscles, eyes, blood cells, spinal cords, and yolk sacs) of embryos injected with either the pICMLE-(−173/34) or -(−136/34) fragments (Fig. 4B,C). The ectopic expression of GFP was not noted in embryos injected with either the pICMLE-(−210/34) or -(−870/34) fragments. We propose that fragment −210/−173 may play an important role in restricting the ectopic (nonheart) expression of *cmlc2*.

We found that the −72/34 sequence within the −210/34 segment was essential for expressing GFP in the heart; the level of fluorescence in hearts declined fivefold in the embryos injected with pICMLE-(210/34) versus those injected with -(−210/−73) (Fig. 4B,C). The latter construct did not contain −72/34 sequences. The tendency of the decline in fluorescence was also observed in embryos injected with pICMLE-(−870/

−73) vs. those injected with -(−870/34). The −72/34 sequence was also missing in the construct of pICMLE-(−870/−73). We also noticed that the level of fluorescence in the hearts decreased 1.7-fold in the embryos injected with pICMLE-(−210/34) vs. those injected with -(−210/3) (Fig. 4B,C). The latter construct did not include 4/34 sequences, which contained the 5′ untranslated region of the zebrafish *cmlc*2. Thus, the expression level of transgenic GFP in embryos decreased significantly when embryos were injected with a construct lacking the transcriptional start site and 5′ untranslated region of *cmlc*2. These findings suggest that there are regulatory elements within the 5′ untranslated region.

When a linearized plasmid pICMLE-(−210/34)Δ(−120/−111), in which G at -119 combined with TATTTA at −118/−113 was deleted from fragment −210/34, the GFP was still specifically and highly expressed in the hearts of transgenic embryos. Moreover, there was not a typical TATA consensus box near the transcription start site of the zebrafish *cmlc*2. Therefore, we conclude that the zebrafish *cmlc*2 that we cloned in this study is a TATA-less promoter and that G at -119 combined with neighboring A/T-rich sequences is not a sole element for controlling heart-specific expression.

Although its intensity was relatively faint, the GFP expression in all the embryos injected with fragment pICMLE-(−210/−73) was exclusively seen in hearts (Fig. 4C). In addition, we injected embryos with a −210/−69 sequence fused with the CMV minimal promoter and found that GFP was expressed in the heart as well as in other tissues. Therefore, the −210/−73 sequence acted as a basal element, allowing the exogenous GPF gene to be expressed in the heart, but this element was not capable of repressing the nonspecific expression driven by the CMV promoter.

The high-level, heart-specific expression of GFP in embryos injected with fragments pICMLE-(−870/787), -(−656/787), and -(−210/787) were similar to those observed in embryos injected with the fragments pICMLE-(−870/34) and -(−210/34) (Fig. 4B), which were missing from the intron 1 of *cmlc*2. Thus, we concluded that the intron 1 of the zebrafish *cmlc*2 was not involved in heart-specific regulation.

Compared with embryos injected with the fragment pICMLE-(−210/34) containing AAV-ITR, the level of fluorescence in hearts was remarkably reduced, 2.5-fold, in embryos injected with the fragment pCMLE-(−210/34) that did not contain AAV-ITR. No differences were noted between these two types of embryos in terms of exogenous GFP-signal tissue distribution. These results were consistent with those of embryos injected with the fragment pICMLE-(−210/−73) with or without AAV-ITR. The green fluorescent level was relatively low (++) in the hearts of embryos injected with pICMLE-(−210/−73) due to the absence of the −72/34 sequence; in comparison, no signal was detected in the hearts of embryos injected with pCMLE-(−210/−73), that is, not flanked by AAV-ITR (data not shown).

Injection of Morpholino Antisense Oligonucleotides

Several transcription factors are predicted for binding the *cis*-elements located in the upstream region of zebrafish *cmlc*2 (Fig. 1). To determine whether Nkx2.5 and GATA-4 are specifically involved in gene regulation of the zebrafish *cmlc*2, we injected *nkx*2.5- and *gata*-4-MO into the embryos of the *cmlc*2::GFP transgenic line. In control groups, we injected *EGFP*-MO and found that the intensity of GFP decreased remarkably (at concentrations of 2.3 and 4.6 ng) or that GFP disappeared (at concentrations of 9.2 and 11.5 ng). We also injected *cTnT*-MO and found that the heart stopped beating, the epicardium became enlarged, and blood circulation ceased. The phenotype of morphant was exactly the same as the silent heart mutant *sih*^{tc300b}, which was described by Sehnert et al. (2002). However, we did not notice any morphologic differences in terms of the intensity of GFP or the beating rate between the wild-type and morphants derived from embryos injected with either *nkx*2.5-MO or *gata*-4-MO at concentrations ranging from 2.3 to 11.5 ng. Only a few embryos in the highest concentration (11.5 ng) groups showed heart enlargement. Like the *nkx*2.5-MO or *gata*-4-MO injection groups, the phenotype of morphant did not change in the embryos injected with the *nkx*2.5-MO and *gata*-4-MO combination.

Transgenic Line With GFP-Expression in All Muscles Resulting From Injection of the Fragment Containing a Single Nucleotide Substitution

After screening and sequencing of the PCR products, we arbitrarily selected a pICMLE9-mutated (−870/787) clone showing three substitutions between nucleotide positions −870 and 787: G replaced by A at −701 (G/−701/A), G/−119/A, and C/598/T. When this GFP-fused mutated fragment was microinjected into zebrafish embryos, GFP was expressed not only in the heart, but also in the skeletal and smooth muscles of transgenic embryos (Fig. 5). Results from a transient assay performed using various deletions of the mutated fragment showed that fragment pICMLE-(−210/34) containing a G/−119/A was sufficient to generate embryos with the expected GFP expression in the heart, plus ectopic expression in the skeletal and smooth muscles. Furthermore, a germ-line transmitted transgenic zebrafish possessing GFP in the all three muscle types was also generated by injecting fragment pICMLE-(−870/787) containing the mutated sequences (Fig. 5B–D). Expression of GFP was observable in the heart after 16.5-hpf, whereas signals were seen in the skeletal and smooth muscles (blood vessel and intestine) after 21-hpf. This unique expression pattern, due to introducing the mutated gene construct, was stably transmitted.

DISCUSSION

By injecting a construct of mouse cardiac-myosin heavy chain promoter fused with GFP, Huang et al. (2000) generated transgenic mice lines that expressed GFP in the heart;

however, those mice suffered from severely dilated cardiomyopathy. The zebrafish is an excellent animal model for studying cardiovascular development in vertebrates (Fishman and Stainier, 1994; Fishman and Chien, 1997; Alexander et al., 1998). Due to the lack of a heart- or myocardial-specific transgene, we cloned the upstream region of the zebrafish *cmlc*2. The *cmlc2*::GFP construct made it possible to generate a transgenic line possessing myocardium-specific GFP expression. Of interest, unlike the dilated cardiomyopathy that occurred in the transgenic mice, our transgenic *cmlc*2::GFP zebrafish developed normally.

Germ-Line Transmission of Heart-Specific Fluorescent Zebrafish

Transgenic zebrafish lines showing heart-specific fluorescence were generated by injecting pICMLE-(−870/787) fragments fused with a GFP gene and flanked with AAV-ITR at both ends. The GFP transgene was expressed exclusively in the hearts of the transgenic progeny, from the embryonic stage through adulthood. Although the GFP expression level in the hearts varied among the 34 transgenic lines in this study, most of the lines showed extremely strong GFP expression. The GFP expression patterns that we observed in germ-line transmitted zebrafish recapitulated faithfully the spatial and temporal patterns of the endogenous *cmlc*2. Homozygotic F3 lines were derived through mating among the heterozygotic F2 individuals; the rate of green-tagging in the hearts of the F3 progeny was 100%, suggesting that the transgene in these lines was integrated into a single insertion site on the chromosome.

Heart Development Implications

The animal heart has three layers: the outer epicardium, the myocardium, and the inner endocardium. In mouse and avian model animals, evidence has shown that the epicardium originates from the epicardium-derived cell population located at the surface of the proepicardial serosa (Gittenberger-de Groot et al., 1998; Manner et al., 2001). T-box gene *Tbx18* (Kraus et al., 2001) and $\alpha 4\beta 1$ integrin gene (Sengbusch et al., 2002) transcripts have been found in epicardial progenitor cells. Meanwhile, the endocardium is postulated to originate from the visceral yolk sac mesoderm (DeRuiter et al., 1992). On the other hand, Laverriere et al. (1994) reported that part of the endocardium originates from the bilateral cardiogenic plates. In this report, we have clearly demonstrated that GFP driven by the zebrafish *cmlc*2 promoter/enhancer occurs in the myocardium, not in the epicardium or endocardium. This myocardium-specific marker and transgenic line may be helpful in elucidating zebrafish heart development.

Unique Proximal Upstream Sequence of the Cloned Zebrafish *cmlc*2

Transgenic assay showed that an upstream fragment as short as −210/34 in the zebrafish *cmlc*2 was sufficient to enable the reporter gene to be exclusively expressed in the myocardium of the heart. As shown in Figure 1, several putative *cis*-elements existed within this proximal upstream region. Regulatory *cis*-elements that are responsible for heart development, such as GATA- and Nkx-family binding sites, also existed in the 5′ upstream region of the zebrafish *cmlc*2. Lyons et al. (1995) reported that, in mice, Nkx2.5 is capable of activating the *mlc*2; consequently, they categorized the gene as being Nkx2.5-dependent. However, unexpectedly, when we injected both *gata*-4-morpholino and *nkx*2.5-morpholino into the embryos of the green-heart line, fluorescent signals in the hearts of the transgenic embryos were still observed. This finding reveals that the essential factors for zebrafish *cmlc*2 transcription are not totally dependent on either GATA-4 or Nkx2.5, suggesting that multiple GATA or Nkx family members may share functions in the zebrafish heart.

Zhu et al. (1991) showed that a 250-bp upstream fragment of the rat *cmlc*2 was sufficient for cardiac-specific and alpha-adrenergic–inducible expression in cultured neonatal rat myocardial cells. They also reported that three 28-bp HF regulatory elements (HF-1 (CArG, AP-2 and MEF-2 motifs), HF-2 (P element), and HF-3 (A element)) within the same upstream 250-bp region were conserved in the rat and chick *cmlc*2. However, HF regulatory elements were not present in the proximal region of the isolated zebrafish *cmlc*2. In addition, the zebrafish *cmlc*2 we isolated is a TATA-less promoter, whereas the rat (Henderson et al., 1989) and chick (Arnold et al., 1988) *cmlc*2 share a TATA box. On the other hand, we also noticed that the GFP marker was specifically expressed in the heart of another species of model fish, medaka (*Oryzias latipes*), when the fragment pICMLE-(−870/787) was injected. This finding suggests that regulatory elements controlling heart-specificity in zebrafish and medaka are conserved. Therefore, it would be worthwhile to further investigate whether zebrafish has another isoform of *cmlc*2 besides the gene we cloned in this study, and whether this fish possesses a particular genomic structure and regulation of *cmlc*2.

A Single Nucleotide Substitution Results in Additional Ectopic Expression

Green fluorescent signals were only observed in the hearts of transgenic fish that received both fragments pICMLE-(−870/787) and -(−210/34). However, in embryos injected with pICMLE-(−870/787) but having three mutated sequences (G/−701/A, G/−119/A, and C/598/T), the transgenic embryos showed GFP not only in the heart but also in skeletal and smooth muscles. This finding was consistent with that for embryos injected with pICMLE-(−210/34) but having a single mutation G/−119/A. Furthermore, green fluorescence was also consistently observed in all three muscle types of stably germline transmitted zebrafish. Thus, we conclude that the G at −119 is a critical nucleotide for controlling heart-specific expression, and that G→A at −119 is responsible for ec-

topic expression. Several possible scenarios may explain these findings: (1) A repressor binding site that normally suppresses GFP expression in skeletal and smooth muscles is disturbed or becomes unstable when G is replaced by A at −119. (2) An activator binding site that triggers the expression of GFP in skeletal and smooth muscles is generated by the same G→A replacement. However, the −210/−173 fragment plays an important role in suppressing the ectopic (nonheart) expression of GFP, meaning that the potential relationship between the −210/−173 motif and G at −119 requires further study. (3) When the original sequence of GTATTTA, located at −119 to −113 of *cmlc2*, is changed to ATATTTA through a single nucleotide replacement at −119, the resultant ATATTTA sequence is linked to the neighboring upstream CT to form a CTATATTTA motif, which is close to the consensus sequence A/T-rich *cis*-element bound by the MADS box-containing proteins, myocyte-specific enhancer factor-2 (MEF2; Gossett et al., 1989; Andres et al., 1995). MEF2 factors activate transcription in the three types of muscle. Therefore, we reason that the ectopic expressions may be due to the MEF2 binding site generated by the G→A change. It is highly likely that MEF is involved in the ectopic expression of this mutant.

Enhancement of GFP Expression in Transgenic Embryos by AAV-ITR in a Promoter-Specific Manner

Compared with embryos injected with fragments of pCMLE-(−210/34) and -(−210/−73) (which are constructs that do not contain AAV-ITR), embryos injected with fragments of pICMLE-(−210/34) and -(−210/−73) (which are constructs containing AAV-ITR) expressed green fluorescence in the heart at significantly higher intensities—a 2.5-fold increase for the -(−210/34) constructs and an increase from undetectable to a low level (++) for the -(−210/−73) constructs. Although AAV-ITR substantially enhanced the exogenous expression of GFP in the transgenic embryos, the tissue distribution of GFP signals was completely identical among the embryos injected with the fragment pICMLE-(−210/34) with and without AAV-ITR. Apparently, the AAV-ITR sequence, found at both ends of the transgene driven by a heart-specific promoter, enhanced heart-specific expression but did not cause additional nonspecific expression. This AAV-ITR function is consistent with those previously reported for transgenes driven by α- (skeletal muscle-specific) and β-actin (ubiquitous) promoters (Hsiao et al., 2001; Chou et al., 2001). The construct is capable of being transmitted to at least three subsequent generations.

ACKNOWLEDGMENTS

We thank Wei-Yuan Tsai and Wu-Hsun Lee of National Taiwan University for their help with screening germ-line transmission and medaka transgenesis and Dr. Didier Stainier and Dr. Benno Jungblut of the Department of Biochemistry and Biophysics, UCSF, for their generosity in providing and helping with the use of a Zeiss Axioplan microscope. We also thank S.W. Yih, Women's and Children's Hospital, Taipei for providing cryosection.

REFERENCES

Alexander J, Stainier DYR, Yelon D. 1998. Screening mosaic F1 females for mutations affecting zebrafish heart induction and patterning. Dev Genet 22:288–299.

Andres V, Cervera M, Mahdavi V. 1995. Determination of the consensus binding site for MEF2 expressed in muscle and brain reveals tissue-specific sequence constraints. J Biol Chem 270: 23246–23249.

Arnold HH, Tannich E, Paterson BM. 1988. The promoter of the chicken cardiac myosin light chain 2 gene shows cell-specific expression in transfected primary cultures of chicken muscle. Nucleic Acids Res 16:2411–2429.

Bader D, Masaki T, Fischman DA. 1982. Immunochemical analysis of myosin heavy chain during avian myogenesis in vivo and in vitro. J Cell Biol 95:763–770.

Chou CY, Horng LS, Tsai HJ. 2001. Uniform GFP-expression in transgenic medaka (*Oryzias latipes*) at the F0 generation. Transgenic Res 10:303–315.

Cummins C, Anderson P. 1988. Regulatory myosin light chain genes of *Caenorhabditis elegans*. Mol Cell Biol 8:5339–5349.

DeRuiter MC, Poelmann RE, Vandr-Plas-de Vries I, Mentink MMT, Gittenberger-de Groot AC. 1992. The development of the myocardium and endocardium in mouse embryos: fusion of two heart tubes? Anat Embryol (Berl) 185:461–473.

Doevendans PA, Bronsaer R, Lozano PR, Kubalak S, van Bilsen M. 2000. The murine atrial myosin light chain-2 gene: a member of an evolutionarily conserved family of contractile proteins. Cytogenet Cell Genet 90:248–252.

Fishman MC, Chien KR. 1997. Fashioning the vertebrate heart: earliest embryonic decisions. Development 124:2099–2117.

Fishman MC, Stainier DY. 1994. Cardiovascular development Prospects for a genetic approach. Circ Res 74:757–763.

Gittenberger-de Groot AC, Peeters MFMV, Mentink MMT, Gourdie RG, Poelmann RE. 1998. Epicardium-derived cells contribute a novel population to the myocardial wall and the artrioventricular cushions. Circ Res 82: 1043–1052.

Gossett LA, Kelvin DJ, Sternberg EA, Olson EN. 1989. A new myocyte-specific enhancer-binding factor that recognizes a conserved element associated with multiple muscle-specific gene. Mol Cell Biol 9:5022–5033.

Henderson SA, Spencer M, Sen A, Kumar C, Siddiqui MA, Chien KR. 1989. Structure, organization, and expression of the rat cardiac myosin light chain-2 gene. J Biol Chem 264:18142–18148.

Hsiao CD, Hsieh FJ, Tsai HJ. 2001. Enhanced expression and stable transmission of transgenes flanked by inverted terminal repeats from adeno-associated virus in zebrafish. Dev Dyn 220:323–336.

Hu N, Sedmera D, Yost HJ, Clark EB. 2000. Structure and function of the developing zebrafish heart. Anat Rec 260:148–157.

Huang CJ, Lin JY, Tsai HJ. 1999. Two distinct c-*ski* cDNAs of fish, tilapia (*Oreochromis aurea*). Mol Reprod Dev 54: 223–231.

Huang WY, Aramburu J, Douglas PS, Izumo S. 2000. Transgenic expression of green fluorescence protein can cause dilated cardiomyopathy. Nat Med 6: 482–483.

Jowett T. 2001. Double in situ hybridization techniques in zebrafish. Methods 23:345–358.

Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF. 1995. Stages of embryonic development of the zebrafish. Dev Dyn 203:253–310.

Kraus F, Haenig B, Kispert A. 2001. Cloning and expression analysis of the mouse T-box gene *Tbx18*. Mech Dev 100:83–86.

Kubalak SW, Miller-Hance WC, O'Brien TX, Dyson E, Chien KR. 1994. Chamber specification of atrial myosin light chain-2 expression precedes septation

during murine cardiogenesis. J Biol Chem 269:16961–16970.

Laverriere AC, MacNeill C, Mueller C, Poelmann RE, Burch JBE, Evans T. 1994. GATA-4/5/6 subfamily of three transcription factors transcribed in developing heart and gut. J Biol Chem 269: 23177–23184.

Lyons I, Parsons LM, Hartley L, Li R, Andrews JE, Robb L, Harvey RP. 1995. Myogenic and morphogenetic defects in the heart tubes of murine embryos lacking the homeo box Nkx25. Genes Dev 9:1654–1666.

Manner J, Perez-Pomares JM, Macias D, Munoz-Chapuli R. 2001. The origin, formation and developmental significance of the epicardium: a review. Cells Tissues Organs 169:89–103.

Parker VP, Falkenthal S, Davidson N. 1985. Characterization of the myosin light-chain-2 of *Dorsophila melanogaster*. Mol Cell Biol 5:3058–3068.

Poetter K, Jiang H, Hassanzadeh S, Master SR, Chang A, Dalakas MC, Rayment I, Sellers JR, Fananapazir L, Epstein ND. 1996. Mutations in either the essential or regulatory light chains of myosin are associated with a rare myopathy in human heart and skeletal muscle. Nat Genet 13:63–69.

Sehnert A, Huq A, Weinstein BM, Walker C, Fishman M, Stainier D. 2002. Cardiac troponin T is essential in sarcomere assembly and cardiac contractility. Nat Genet 31:106–110.

Sengbusch JK, He W, Pinco KA, Yang JT. 2002. Dual functions of $\alpha 4\beta 1$ integrin in epicardial development: initial migration and long-term attachment. J Cell Biol 157:873–882.

Silver PJ, Buja LM, Stull JT. 1986. Frequency-dependent myosin light chain phosphorylation in isolated myocardium. J Mol Cell Cardiol 18:31–37.

Stull JT, Nunnally MH, Moore RL, Blumenthal DK. 1985. Myosin light chain kinases and myosin phosphorylation in skeletal muscle. Adv Enzyme Regul 23:123–140.

Westerfield M. 1995. The zebrafish book. Eugene: University of Oregon Press.

Winter B, Klapthor H, Wiebauer K, Delius H, Arnold HH. 1985. Isolation and characterization of the chicken cardiac myosin light chain. L-2A. gene: evidence for two additional N-terminal amino acids. J Biol Chem 260:4478–4483.

Yelon D, Horne SA, Stainier DY. 1999. Restricted expression of cardiac myosin genes reveals regulated aspects of heart tube assembly in zebrafish. Dev Biol 214:23–37.

Zhu H, Garcia AV, Ross RS, Evans SM, Chien KR. 1991. A conserved 28-base-pair element (HF-1) in the rat cardiac myosin light-chain-2 gene confers cardiac-specific and alpha-adrenergic-inducible expression in cultured neonatal rat myocardial cells. Mol Cell Biol 11:2273–2281.