Development 126, 1439-1450 (1999)
Printed in Great Britain © The Company of Biologists Limited 1999
DEV4082

Complex modular *cis*-acting elements regulate expression of the cardiac specifying homeobox gene *Csx/Nkx2.5*

Makoto Tanaka[1,*], Stephanie Burns Wechsler[3,*], Ike W. Lee[1,*], Naohito Yamasaki[1], Joel A. Lawitts[2] and Seigo Izumo[1,‡]

[1]Cardiovascular Division, and [2]Transgenic Core Facility, Beth Israel Deaconess Medical Center and Department of Medicine, Harvard Medical School, 330 Brookline Avenue, Boston, MA 02215, USA
[3]Department of Pediatrics and Communicable Diseases, University of Michigan Medical Center, 8220D MSRB III, Box 0646, 1150 West Medical Center Drive, Ann Arbor, MI 48109, USA
*These authors contributed equally to this work
‡Author for correspondence (e-mail: sizumo@caregroup.harvard.edu)

Accepted 15 January; published on WWW 3 March 1999

SUMMARY

The murine homeobox gene *Csx/Nkx2.5* is an evolutionarily highly conserved gene related to the *Drosophila tinman* gene, which specifies cardiac and visceral mesoderm. Since *Csx/Nkx2.5* plays an essential role in heart development, studying its regulation is essential for the better understanding of molecular mechanisms of cardiogenesis and the pathogenesis of congenital heart disease in humans. In this study, we characterized the murine *Csx/Nkx2.5* gene and identified two novel untranslated exons, 1a, and 1b, resulting in three different *Csx/Nkx2.5* transcripts. To examine the tissue-specific transcriptional regulation in vivo, we analyzed a total of 23 kb of *Csx/Nkx2.5* upstream and downstream sequences by generating transgenic embryos carrying *lacZ* reporter constructs containing various lengths of flanking sequence. With 14 kb of 5′ flanking sequence, *lacZ* expression was observed in the cardiac crescent at E7.5, and in the outflow tract, the interatrial groove, the atrioventricular canal and right and left ventricles, as well as in pharyngeal floor, thyroid primordia, and stomach at E10.5. In adult animals, *lacZ* expression of the transgene was limited to the atrioventricular junction and the subendocardium of the ventricular septum. Reducing the size of flanking sequence to 3.3 kb of intron 2 restricted *lacZ* expression to the outflow tract and the basal part of the right ventricle in E10.5 embryos. In contrast, the addition of 6 kb of 3′ flanking sequence caused strong expression of the reporter gene in the entire right ventricle. Interestingly, *Csx/Nkx2.5* seems to be negatively regulated by its own gene product, because when *lacZ* was "knocked-in" to replace the entire coding exons, *lacZ* expression was much higher in the heart of homozygous embryos than that in the heterozygote. These results indicate that the transcriptional regulatory elements of *Csx/Nkx 2.5* seems unexpectedly highly modular, and is temporally regulated in a dynamic manner by different enhancer regions. Since *Csx/Nkx2.5*-like genes are expressed in all species having a heart, their complex modular organization with multiple enhancers probably reflects progressive addition of regulatory elements during the evolution from a simple heart tube to a complex four-chambered organ.

Key words: Heart development, *Csx/Nkx2.5*, Transcriptional enhancers, Transgenic mice, Homeobox gene

INTRODUCTION

The molecular mechanisms which guide development of the heart, in vertebrates and many other species, have been the subject of intense investigation in recent years (Fishman and Chien, 1997; Olson and Srivastava, 1996). These studies have suggested common molecular pathways which are evolutionarily conserved from simpler species such as *Drosophila* to higher species including mammals (Bodmer and Venkatesh, 1998; Harvey, 1996). This evolutionary conservation is particularly striking, since insects and vertebrates diverged more than 500 million years ago and the heart is quite distinct anatomically in these widely divergent species.

In most vertebrates, the heart begins initially as a crescent shaped mesodermal structure located anteriorly and laterally. This precardiac mesoderm is brought ventrally and caudally, by folding of the embryo, to form a single midline heart tube with the inflow region located most caudally and the outflow region located most rostrally. This tube undergoes looping, bringing the inflow, ventricular and outflow regions of the heart into the alignment seen in the mature heart. Later, chamber septation occurs, valves develop in the atrioventricular junction as well as in the outflow tract, and the outflow tract itself is divided into two great vessels. Additional refinements occur with the development of the coronary arteries and the cardiac conduction system (Sadler, 1995).

Heart development is governed by complex signals

including inductive and positional signals from adjacent structures, as well as signals from a number of transcription factors (Fishman and Chien, 1997; Lyons, 1996; Mohun and Sparrow, 1997; Olson and Srivastava, 1996). Since transcription factors have the ability to activate multiple genes, they are important regulators of organ development. A number of cardiac transcription factors have been identified that have important influences on the early stages of specification and differentiation of the cardiac mesoderm (Tanaka et al., 1998). *Csx/Nkx2.5* (Komuro and Izumo, 1993; Lints et al., 1993), *MEF-2C* (Edmondson et al., 1994), *GATA4* (Heikinheimo et al., 1994; Kelley et al., 1993) as well as *dHAND* and *eHAND* (Srivastava et al., 1995) are members of four different classes of transcription factors all expressed in the heart at early stages of development. Targeted disruption of these genes have all yielded severe cardiac and extracardiac phenotypes, and result in death of the embryo between E9.5 and E10.5 of development (Kuo et al., 1997; Lin et al., 1997; Lyons et al., 1995; Molkentin et al., 1997; Srivastava et al., 1997).

Among these cardiac transcription factors, the homeobox gene *Csx/Nkx2.5* is the first vertebrate homolog of the *Drosophila tinman* gene to be described (Komuro and Izumo, 1993; Lints et al., 1993). *Tinman* is required for specification of the cardiac and visceral mesoderm in *Drosophila* (Azpiazu and Frasch, 1993; Bodmer, 1993). *C. elegans* does not have a heart, but a tinman-like gene, *ceh*-22, is specifically expressed and required for the development of pharyngeal muscle which contracts rhythmically like a heart (Okkema and Fire, 1994). Csx/Nkx2.5 protein is initially expressed in the precardiac mesoderm and the adjacent pharyngeal endoderm in E7.5 mouse embryos (Kasahara et al., 1998). It is homogeneously expressed in the myocardium of the atria and the ventricles by E8.5 in the mouse, and continues to have strong expression throughout the embryonic and neonatal periods and into adulthood (Kasahara et al., 1998; Komuro and Izumo, 1993; Lints et al., 1993). Targeted disruption of this gene causes the arrest of heart development at the initial stage of looping, and the mice die by E9.5 – 11.5 (Lyons et al., 1995; Tanaka et al., 1999). *Csx/Nkx2.5* mutant hearts do not form trabeculae and endocardial cushion. The human *Csx/Nkx2.5* was localized to chromosome 5q35 (Turbay et al., 1996). Very interestingly, heterozygous mutations in the human *CSX/NKX2.5* gene were found to cause familial congenital heart disease (secondum atrial septal defects and other forms) associated with atrioventricular conduction defects(Schott et al., 1998).

Because *Csx/Nkx2.5* plays such an essential role in cardiac development, including that of humans, studying the regulation of *Csx/Nkx2.5* gene is important for the better understanding of the molecular mechanism of cardiogenesis as well as the pathogenesis of congenital heart disease in humans. In this study, we describe the genomic structure and the regulation of the *Csx/Nkx2.5* gene in transgenic mice. The genomic structure of *Csx/Nkx2.5* is more complex than initially anticipated, including two 5′ untranslated exons and three transcriptional start sites coupled with alternative splicing. Functional dissection of the regulatory regions of the *Csx/Nkx2.5* gene revealed multiple distinct enhancer elements which direct highly restricted expression of the *Csx/Nkx2.5* gene to distinct regions of the heart itself and to other tissues.

MATERIALS AND METHODS

Isolation of genomic clones and sequencing

Two lambda phage genomic DNA clones containing the *Csx/Nkx2.5* coding and flanking sequences (Fig. 1) were isolated from a 129 mouse genomic library in λDASH II. Clone 1 contained 14 kb of 5′ flanking region in addition to a portion of the *Csx/Nkx2.5* coding region. Clone 2 contained 6 kb of 5′ flanking region, the *Csx/Nkx2.5* coding region, and 6 kb of 3′ flanking region. In addition, a mouse P1 clone (Genome Systems, St. Louis, MO) which contained the *Csx/Nkx2.5* coding sequence, as well as more than 10 kb 5′ and 3′ flanking sequence was isolated. Restriction enzyme mapping and serial Southern blotting were performed in a standard manner (Ausubel et al., 1993). Automated dideoxynucleotide sequencing was performed using ABI Applied Biosystems machinery.

Definition of 5′ end of mRNA

5′ RACE was performed following previously described methods with some modifications (Reecy et al., 1997). The first strand cDNA synthesis reaction was performed using 5 μg total RNA from the heart of 3-week-old mice and 100 ng of random primers (Promega, Madison, WI) in a 40 μl volume containing 50 mM Tris-HCl (pH 8.3), 75 mM KCl, 3 mM $MgCl_2$, 10 μM DTT, 40 U RNase inhibitor (Promega), 0.5 mM each deoxynucleotides, 200 U M-MLV reverse transcriptase (Life Technologies, Gaithersburg, MD) for 90 minutes at 37°C. The cDNA was purified using a QIAquick PCR purification column (QIAGEN, Santa Clarita, CA) after incubation with 5 μg RNase A (Ambion, Austin, TX). Ligation-anchored PCR was performed as described by Ali Ansari-Lari et al. (1996) with some modifications. Initially, 3 nmoles of a 3′ amino-modified, 5′ phosphorylated anchor primer (5′-TCTCTACTCCGAATTCCGT-CGTCCACACCT-3′; Integrated DNA Technologies, Coralville, IA) was end ligated to one half of the purified first strand cDNA using 20 U of T4 RNA ligase (New England Biolabs, Beverly, MA) in a 50 μl reaction volume containing 50 mM Tris-HCl (pH 8.0), 10 mM $MgCl_2$, 1 mM hexamine cobalt chloride, 20 μM ATP, 25% PEG 8000 and incubated for 24 hours at 16°C. The anchor-ligated cDNA was further purified using a QIAquick PCR purification column. The first round PCR was performed using one-tenth of the purified anchor-ligated cDNA, an anchor-specific primer (5-AGGTGTGGACGACGGAATT-CGGAGTAGAGA-3′) and a Csx-specific primer (5′-GGGGGCGG-CTGGGAAAGCAGGAGAGCACTT-3′). PCR conditions were as follows: denaturation at 94°C for 1 minute, annealing at 60°C for 1 minute, extension at 72°C for 2 minutes with an additional 5 minutes incubation at 72°C at the end of the 30 cycles. Subsequent PCR reactions were performed using 5 μl of each PCR product, a nested anchor primer (5′-CGACGGAATTCGGAGTAGAGA-3′), and one of the Csx-specific primers (5′-TTGAAGGCGGCCAGCATGC-AGGAGGCA-3′ or 5′-ACAGGAGCGACGGGCAGTTCTGCGT-3′) at the same PCR conditions as above. The PCR products were visualized on a 2% agarose gel with ethidium bromide, and subsequently subcloned and sequenced.

Detection of exons 1b and 1c by primer extension

5 μg samples of mouse heart poly(A) RNA were coprecipitated with 2×10^4 cpm end-labeled, gel-purified oligonucleotide probe (probe sequence: 5′-CGGAGCACCAGGGGCAGAAGAGGC-3′). The samples were resuspended in 10 μl annealing buffer (0.1 M NaCl, 0.01 M Tris, pH 8.0, 0.001 M EDTA), heated to 85°C for 5 minutes then incubated at 37°C for 2 hours. 20 μl of reverse transcription buffer (10 mM DTT, 16 mM $MgCl_2$, 1 mM dNTP, 1 U/μl RNasin, 0.1 M Tris, pH 8.0) was added to each sample as well as 20 U Superscript reverse transcriptase (Life Technologies, Gaithersburg, MD). Reverse transcription was allowed to proceed at 40°C for 1 hour, and 7.5 μl of each sample was mixed with an equal volume of formamide loading buffer. After denaturation at 95°C for 5 minutes, 7 μl of each sample was loaded onto an 8% denaturing

polyacrylamide gel. After electrophoresis, the gel was dried and exposed to film.

Detection of exon 1a by RT-PCR

RT-PCR was performed using rTth DNA polymerase (rTth RT-RNA PCR kit; Perkin Elmer, Branchburg, NJ) following the manufacturer's protocol. Reverse transcription was performed using an exon 1c-specific primer (5′-ACAGGAGCGACGGGCAGTT-CTGCGT-3′), and the subsequent PCR reaction was performed by adding an exon 1a-specific primer (5′-GAGTGCTCTGCCTGATG-ATC-3′) to the RT reaction according to the manufacturer's instruction. The PCR reaction consisted of 35 cycles of denaturation at 95°C for 10 second, annealing/extension at 55°C for 15 seconds, and at the end of the 35 cycles, an additional final extension at 55°C for 7 minutes. The PCR products were visualized on a 3% agarose gel with ethidium bromide, and a Southern analysis was performed using a ^{32}P end-labeled exon 1b-specific primer (5′-CCAGTCTAG-AAGCGGTGATCGCCA-3′).

Construction of reporter gene constructs, generation and analysis of transgenic mice

An *Xba*I-*Pst*I *lacZ* cassette from pnlacF (Bonnerot et al., 1987) was subcloned into the *Xba*I and *Pst*I sites of pBluescript SK– (Stratagene) and regions of the genomic *Csx/Nkx2.5* DNA were cloned 5′ or 3′ to the *lacZ* cassette. The CsxlacZ-1 construct included 3.3 kb of genomic sequence, including most of intron 2 and the beginning of exon 1c, between the *Not*I site and the ATG codon (Fig. 1). An *Xho*I-*Pst*I fragment containing part of exon 1c, intron 3 and part of exon 2 was subcloned 3′ to the *lacZ* gene in CsxlacZ-1 to make the construct CsxlacZ-2 (Fig. 1). The CsxlacZ-3 construct contains a 4 kb fragment between the *Spe*I site and the ATG codon of *Csx/Nkx2.5*, thus including a part of intron 1, exon 1b, intron 2 and part of exon 1c (Fig. 1). CsxlacZ-4 and CsxlacZ-5 contain 6 kb and 14 kb, respectively, of upstream 5′ *Csx/Nkx2.5* genomic sequence cloned upstream of the *lacZ* gene. A 6 kb long 3′ downstream genomic fragment was fused 3′ to the *lacZ* gene in CsxlacZ-1 to create the construct CsxlacZ-6 (Fig. 1).

For microinjection of CsxlacZ-1, -2, -3 and -6, *Csx* genomic sequence together with the *lacZ* cassette portion of each construct was excised from the vector using *Xho*I and *Not*I, and purified by gel electrophoresis and Geneclean III (Bio101, Vista, CA). In preparation for microinjection, CsxlacZ-4 and CsxlacZ-5 were linearized with *Xho*I, then purified as indicated above. The creation of transgenic mice was done in a standard manner (Hogan et al., 1994).

For transient transgenic analysis, F_0 embryos were dissected at embryonic day 10.5 (E10.5). Subsequent genotyping was performed using PCR on chromosomal DNA isolated from the yolk sacs. PCR primer pairs used for detection of the transgenes were as follows: 5′-CCGTCCGATGAAAAACAGGAG-3′ and 5′-TCTGCTCTTC-GTCTGGCTGATG-3′ for CsxlacZ-1, -2, and -3; 5′-CCGTCCGAT-

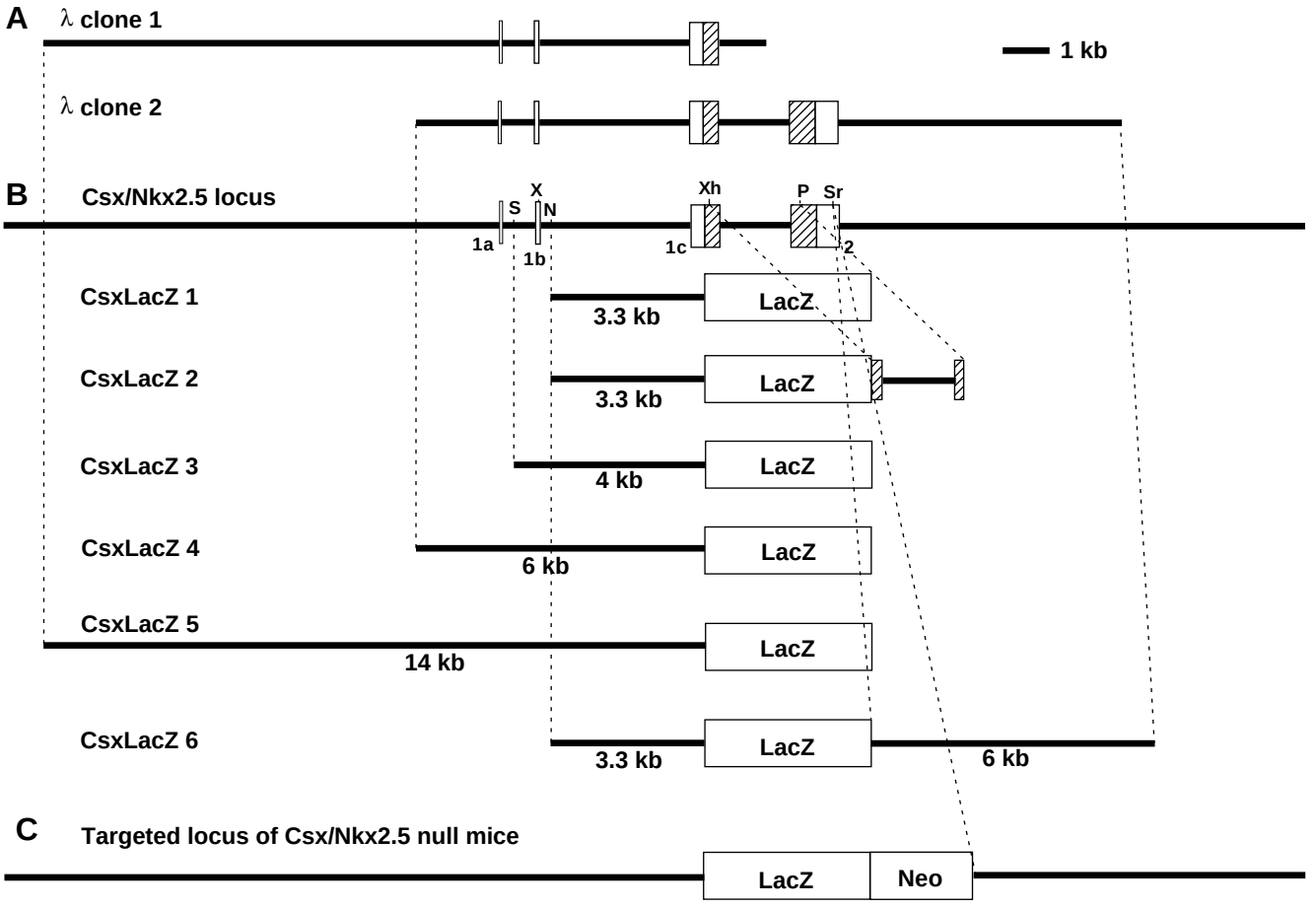


Fig. 1. Structure of the *Csx/Nkx2.5* locus (B), phage clones (A) and constructs utilized for transgenic experiments, and the targeted locus of *Csx/Nkx2.5* null mutant mice (C). Open boxes and shaded boxes represent untranslated sequence and coding sequence, respectively. (A) Two genomic λ phage clones (clones 1 and 2) were isolated from a 129 mouse genomic library. (B) Various subfragments of these two λ phage clones were utilized for construction of transgenes (CsxlacZ-1 through CsxlacZ-6). (C) The entire coding region of *Csx/Nkx2.5* was replaced by *lacZ* and the neomycin-resistance gene (Neo).

GAAAAACAGGAG-3′ and 5′-TTAAGTTGGGTAACGCCAGGG-3′ for CsxlacZ-4 and -5; 5′-AACTTGCTAGGTAGACTAGGCTGGC-3′ and 5′-TCTGCTCTTCGTCTGGCTGATG-3′ for CsxlacZ-6.

Whole-mount β-gal staining was performed according to the method of Schlaeger et al. (1995). After photographs were taken of the whole mounts, embryos were dehydrated through graded ethanol and xylene, embedded in paraffin, sectioned and counterstained with Nuclear Fast Red (Vector Laboratories, Burlingame, CA).

Generation of mouse lines carrying CsxlacZ-1 and CsxlacZ-5

For CsxlacZ-1 and CsxlacZ-5, we established transgenic lines of mice. The F_0 mice carrying each transgene were backcrossed with FVB mice and F_1 embryos were examined for *lacZ* expression. We identified one line of transgenic mice with *lacZ* expression for each construct and analyzed them at different time points, including E7.5, E8.25, E9.5, E15.5, and adult.

Generation of Csx/Nkx2.5 knock-out and *lacZ* knock-in mice

The entire coding region of *Csx/Nkx2.5* was replaced with a *lacZ–neomycin resistance gene (Neo)* cassette by homologous recombination (Fig. 1) in a standard manner (Hogan et al., 1994). Detailed phenotypic characterization of the gene targeted mouse will be reported separately (Tanaka et al., 1999).

In situ hybridization

In situ hybridization was performed as described previously (Tanaka et al., 1998). Briefly, embryos were fixed in 4% paraformaldehyde at 4°C overnight and embedded in paraffin. Tissue sections were hybridized with a S^{35}-labeled *Csx/Nkx2.5* cRNA probe at 55°C, washed and treated with RNase A. After emulsion autoradiography, sections were counter-stained with hematoxylin and eosin.

RESULTS

Csx/Nkx2.5 gene produces three different RNA transcripts by differential promoter usage and alternative splicing

As an initial step in defining the regulatory region of *Csx/Nkx2.5*, we determined the 5′ end of the *Csx/Nkx2.5* transcript by 5′ RACE with total RNA from hearts of 3-week-old mice. Three different transcripts which were identical for most of their length, except for distinct 5′ ends, were identified (transcripts I, II, and III, Fig. 2A). The sequence comparison with the genomic DNA sequence showed exon – intron boundaries in accordance with the consensus exon-intron junction sequences (Fig. 2D). Therefore, we named the most 5′ cDNA segment exon 1a, the

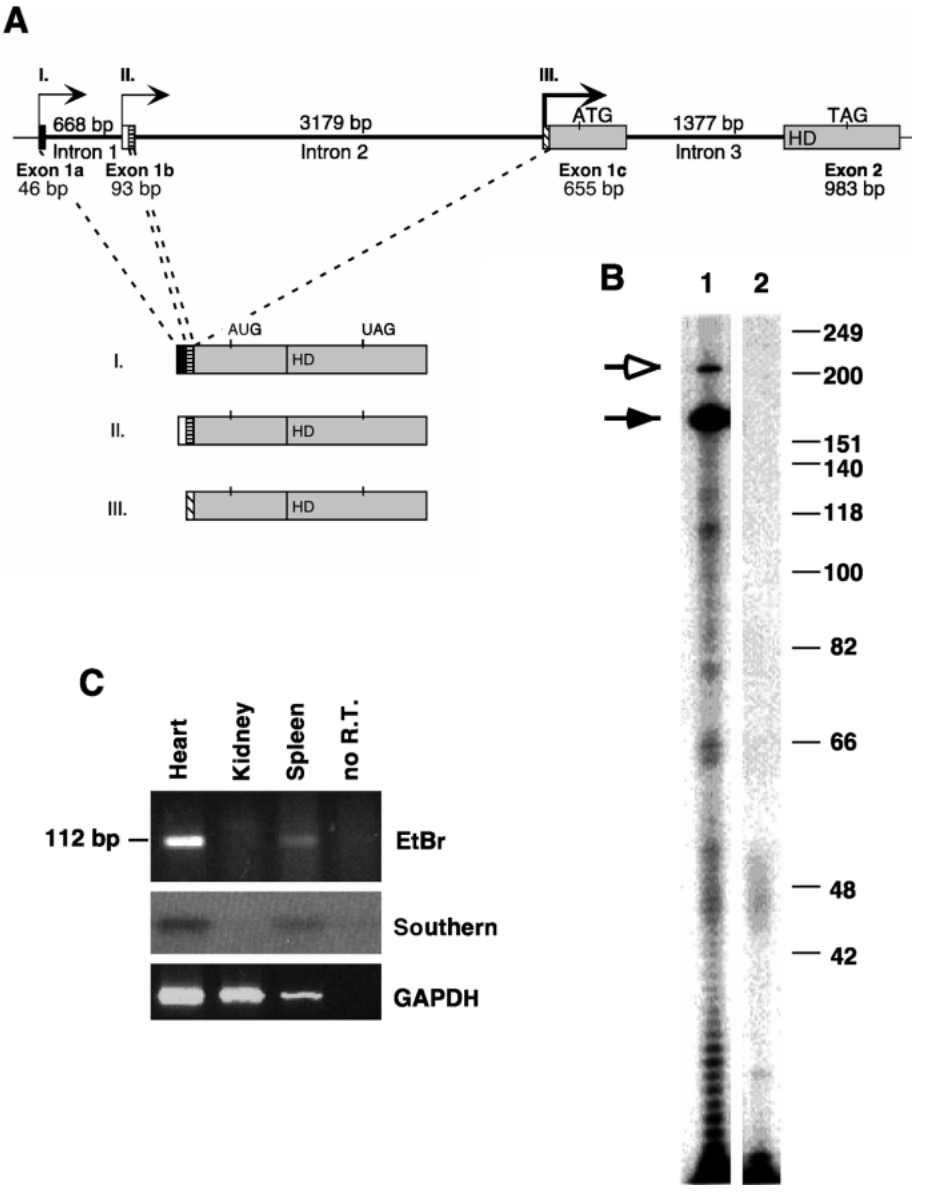


Fig. 2. (A) The complete genomic structure of the mouse *Csx/Nkx2.5* gene. Exons are denoted as boxes, and the intervening thick lines represent introns. Three transcription initiation sites and their transcripts are marked with arrows and roman numerals (I, II, and III), respectively. In transcript I, exons 1b and 1c are shorter (36 bp and 584 bp, respectively) than those in the transcripts II or III. HD, homeodomain. (B) Identification of transcripts II and III by primer extension analysis. Lane 1 contains 5 μg of mouse heart poly(A)$^+$ RNA. Lane 2 contains 20 μg tRNA. The strong band seen at 179 bp (black arrow) corresponds to transcript III. The band seen at 201 bp (open arrow) corresponds in size to transcript II. (C) Identification of transcript I. RT-PCR was carried out using an exon 1a and exon 1c-specific primer pair, and the amplified bands were hybridized with an exon 1b-specific primer probe. (D) Nucleotide sequence of the regions flanking exons of the mouse *Csx/Nkx 2.5* gene. Exons are boxed, and underlined sequences are part of the exon sequence found only in either transcript II or III. A consensus polyadenylation sequence is underlined, and translation initiation and stop codons are double underlined. Three transcription initiation sites are marked with arrows.

second cDNA segment exon 1b, and the intervening genomic sequence intron 1. Likewise, the exon that contains the start codon is termed exon 1c, flanked by introns 2 and 3, and followed by the last exon, exon 2, which includes the homeodomain (HD) and TAG translation stop codon.

Transcript I contains two segments of cDNA sequence at its most 5′ end which are discontinuous in the genomic sequence. The most 5′ segment of transcript I is 46 bp in length (exon 1a in Fig. 2A) and is 669 bp upstream from the second discontinuous segment. The second segment is 36 bp long (a part of exon 1b in Fig. 2A) and 3504 bp upstream from the translation start codon. For transcript II, exon 1b is the most 5′ sequence, however, it is extended more 5′ than in transcript I, in this case 93 bp long including the 36 bp internal fragment found in transcript I (horizontal hatch in Fig. 2A). In transcript III, the first exon at the 5′ end, exon 1c, includes 71 bp of continuous genomic sequence which is spliced out in transcripts I and II (Fig. 2A,D).

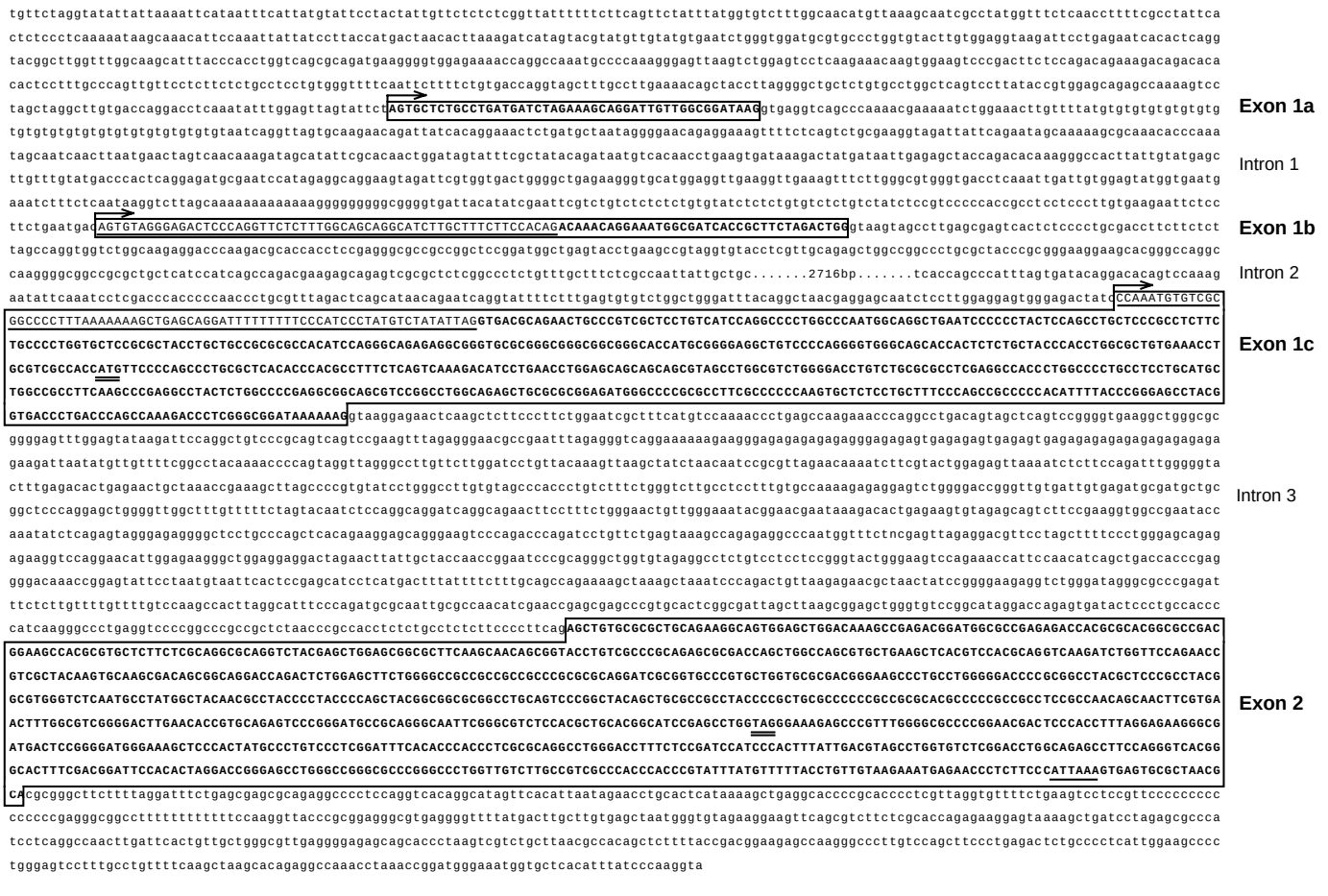
To confirm the existence of the novel exons, we performed primer extension analysis. With adult mouse heart poly(A) RNA, two bands were observed which correspond to the sizes of transcript II and transcript III (Fig. 2B). The band representing transcript III was much stronger, suggesting that it is the major site of initiation. Since we could not detect a band corresponding to transcript I by primer extension analysis, we performed RT-PCR using total RNA samples from several tissues. With exon 1a-specific, and exon 1c-specific primer pairs, RT-PCR generated predicted size (112 bp) products from the heart and spleen. No band was amplified with RNA from kidney which does not express *Csx/Nkx2.5*. By Southern blotting, this 112 bp band was shown to include exon 1b using an exon 1b-specific oligonucleotide probe (Fig. 2C).

To further characterize the genomic structure of the murine *Csx/Nkx2.5* gene, we subcloned and sequenced 8.6 kb of genomic DNA. (Fig. 2D). The 3′ end of exon 1a is located 4265 bp upstream from the ATG codon, followed by intron 1 (668 bp), exon 1b (93 bp), intron 2 (3179 bp), exon 1c (655 bp) that includes the ATG translation start site, intron 3 (1377 bp), and exon 2 (983 bp) with TAG stop codon and a poly(A) addition signal (Fig. 2A,D). The genomic sequences around the three putative transcription initiation sites show no TATA box, indicating that *Csx/Nkx2.5* is a TATA-less gene.

3.3 kb and 4 kb of 5′ flanking sequence drive *lacZ* expression in the outflow tract and basal portion of the right ventricle as well as in the pharynx, thyroid primordium and stomach

Since the major *Csx/Nkx2.5* transcript (transcript III) starts 325 bp upstream from the ATG codon, we began our search for enhancer elements in the genomic region immediately upstream to exon 1c. The 3.3 kb of 5′ flanking sequence between the *Not*I site and the ATG codon was fused to the *lacZ* reporter gene (Fig. 1; CsxlacZ-1) and injected into mouse

D



embryos to generate transgenic mice. Of 40 embryos examined, 8 embryos carried the transgene and three were noted to have β-gal staining (Table 1). When the *lacZ* expression pattern was assessed in these embryos at E10.5, we found that this 3.3 kb upstream region could drive *lacZ* expression in cardiac and extra-cardiac tissues (Fig. 3). Myocardial cells in the outflow tract and the basal part of the RV were strongly positive for *lacZ*, and there were also a few positive cells in the trabecular layer of the RV (Fig. 3B-D). In the outflow tract, *lacZ* was strongly expressed in the myocardium, but no *lacZ* expression was detected in the aortic sac or endocardial cushions (Fig. 3C).

At E10.5, extracardiac expression of *lacZ* was also observed in locations where the endogenous *Csx/Nkx2.5* was detected, as the pharyngeal floor (PF in Fig. 3C,D), thyroid primordium (TP in Fig. 3D), and in a distal part of the stomach (ST in Fig. 3E). In the stomach, *lacZ* expression was observed in mesenchymal cells. Ectopic expression of *lacZ* was observed in the surface ectoderm (SE) of pharyngeal arches and the laryngotracheal groove (LG) (data not shown). The *lacZ* expression pattern, in both cardiac and extracardiac tissues, was identical in all three *lacZ*-positive embryos, with the exception of an ectopic expression in the glossopharyngeal ganglions in one embryo (Table 2).

The CsxlacZ-2 construct included the same 3.3 kb of 5′ flanking sequence plus 1.4 kb of intron 3 sequence (Fig. 1). Of 34 embryos, six carried the CsxlacZ-2 transgene and of these, three had positive β-gal staining (Table 1). Cardiac expression was seen in the basal portion of the RV and in the outflow tract and extracardiac expression was seen in the pharyngeal floor, the thyroid primordium, and the stomach in two embryos (Table 2). In one embryo, *lacZ* expression was noted in the pharyngeal floor, but not in other sites, probably due to a negative effect of the transgene integration site.

The addition of most of intron 1 and exon 1b to the CsxlacZ-1 construct yielded CsxlacZ-3 (Fig. 1). Two embryos were positive for CsxlacZ-3 transgene, with one embryo having positive β-gal staining (Table 1). The *lacZ* expression was observed in the outflow tract, RV, pharynx, thyroid primordium and stomach, (Table 2). Thus, the cardiac β-gal staining pattern was quite similar for constructs CsxlacZ-1, CsxlacZ-2, and CsxlacZ-3. The extracardiac expression pattern was also very similar except for the sites of ectopic expression (Table 2).

A strong negative regulatory element exists between 4 kb and 6 kb upstream of the *Csx/Nkx2.5* coding sequence

CsxlacZ-4 contains 6 kb of 5′ flanking sequence (Fig. 1). Seventy-two injected embryos were analyzed with ten found to carry the CsxlacZ-4 transgene. Interestingly, none of the CsxlacZ-4 transgenic embryos showed any β-gal staining in the heart nor in any extracardiac tissues (Table 1). This suggests that a strong negative regulatory element exists between 6 kb and 4 kb 5′ to exon 1c.

Table 1. Summary of transgene integration and *lacZ* expression

Construct	No. of mice	No. of embryos	Transgene (+)	*lacZ* (+) embryos among transgene (+) embryos
CsxlacZ-1	7	40	8 (20%)	3 (38%)
CsxlacZ-2	6	34	6 (18%)	3 (50%)
CsxlacZ-3	4	25	2 (8%)	1 (50%)
CsxlacZ-4	11	72	10 (14%)	0 (0%)
CsxlacZ-5	8	83	11 (13%)	4 (36%)
CsxlacZ-6	5	38	12 (32%)	6 (50%)

A *cis*-regulatory element for expression of *Csx/Nkx2.5* in the medial wall and inner trabeculae of RV and LV exists between 14 and 6 kb of 5′ flanking sequence

With 14 kb of 5′ flanking sequence driving the reporter gene (CsxlacZ-5), *lacZ* was expressed more widely in the heart (Fig. 4A-C) than with CsxlacZ-1. β-gal staining was observed in the entire RV and LV except for the compact layer of the lateral walls (Fig. 4B,E,G-I). In the atrium, there was a cluster of *lacZ*-positive myocardial cells in the interatrial groove (arrows in Fig. 4D,H), which is above the future atrial septum. β-gal staining was also seen in the outflow tract (Fig. 4E-G) as in CsxlacZ-1. Cardiac staining was similar in all four transgenic embryos with positive β-gal staining.

Extracardiac expression of CsxlacZ-5 was observed in the pharyngeal floor (PF) and thyroid primordium (TP) in all four

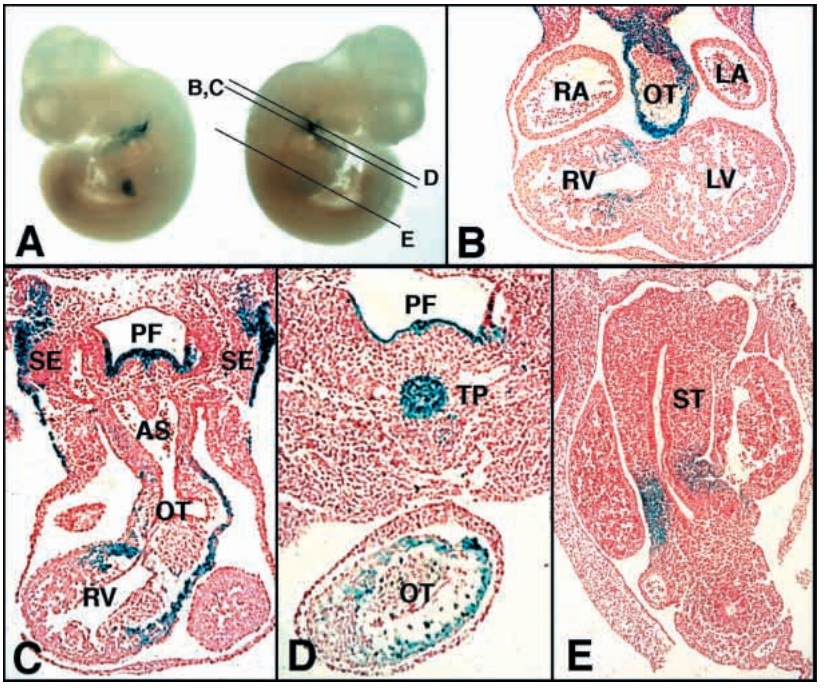


Fig. 3. Whole-mount X-gal staining (A) and transverse sections (B-E) of an E10.5 transgenic embryo carrying the CsxlacZ-1 transgene. Transgene expression was observed in the outflow tract (B-D), a basal portion of the RV (B,C), pharyngeal floor (C,D), thyroid primordium (D) and a distal part of the stomach (E). RA, right atrium; LA, left atrium; RV, right ventricle; LV, left ventricle; OT, outflow tract; PF, pharyngeal floor; AS, aortic sac; SE, surface ectoderm of pharyngeal arches; TP, thyroid primordium; ST, stomach.

CsxlacZ-5 transgenic embryos (Fig. 4F,G). Only one of the four CsxlacZ-5 embryos was noted to have weak β-gal staining in the stomach in contrast to transgenic embryos harboring CsxlacZ-1,2 and 3 constructs, all of which exhibited strong expression in the stomach. No ectopic expression was noted in all embryos, indicating the presence of repressor elements to restrict expression of Csx/Nkx2.5 in a tissue-specific manner (Table 2).

An enhancer for RV is located within 6 kb of the 3′ flanking sequence

To examine *cis*-regulatory elements downstream of the coding sequence of *Csx/Nkx2.5*, we linked 6 kb of 3′ flanking sequence to the CsxlacZ-1 construct to create CsxlacZ-6 (Fig. 1). Of 38 embryos examined, 12 carried the CsxlacZ-6 transgene, of which six were *lacZ* positive (Table 1). Three transgenic embryos had β-gal staining in the entire RV (Fig. 5B,E,F) including the compact layer of the lateral walls. The other three had no cardiac but weak extracardiac stainings (not shown). Extracardiac staining in all six embryos included the stomach, pharynx, and thyroid primordium (Fig. 5A,D,E).

Different expression patterns of the reporter gene in adult transgenic mice

In addition to the evaluation of transient transgenic embryos described above, lines of transgenic mice were created and F_1 transgenic mice were evaluated at different time points in development.

In the transgenic CsxlacZ-1 line which carried the 3.3 kb 5′ flanking sequence, no β-gal staining could be observed at E7.5 (data not shown). At E9.5, the pattern of cardiac and extracardiac *lacZ* expression was identical to that seen in the transient transgenic embryos (Fig. 6A). At E15.5, *lacZ* expression was also observed in the spleen (SP in Fig. 6B), but not in the tongue (not shown). In the E15.5 heart, *lacZ* expression was reduced, but detectable in the outflow tract region of the RV (Fig. 6D), while the transgene expression in the thyroid gland was still strong (TG in Fig. 6C). Interestingly, when adult animals of the same transgenic CsxlacZ-1 line were analyzed, β-gal staining was completely absent in the adult hearts (Fig. 6E).

In the transgenic CsxlacZ-5 line, which carries the 14 kb 5′ flanking sequence, β-gal staining was observed in the cardiac crescent at E7.5 (Fig. 7A), in the common ventricle and outflow tract at E8.25 (Fig. 7B) and in the RV, LV, septum and outflow tract at E9.5 (Fig. 7C). At E15.5, the cardiac β-gal staining was significantly down-regulated except for patchy expression in the AV junction and the interventricular septum (data not shown). The spleen was also stained, but the tongue

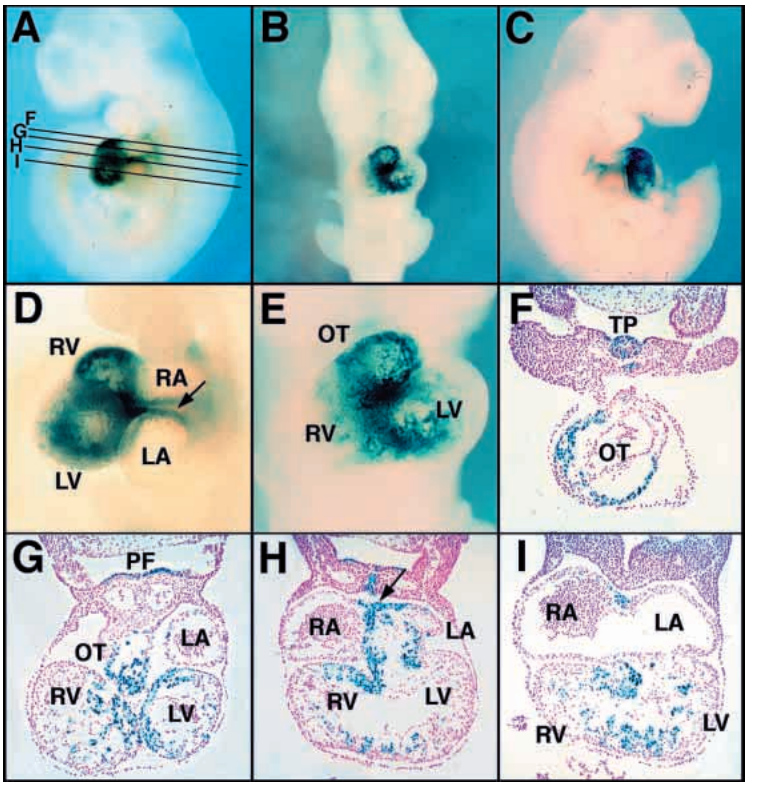


Fig. 4. Whole-mount X-gal staining (A-E) and transverse sections (F-I) of an E10.5 transgenic embryo carrying the CsxlacZ-5 transgene. Myocardial cells at the groove between the right and left atria above the future atrial septum were positive for *lacZ* expression (A, arrows in D,H). In connection to these cells, there were *lacZ* positive cardiac myocytes at the atrioventricular canal (A, D, H, I) and basal and septal sides of RV and LV (B,E,H,I). Transgene expression was also observed in the outflow tract (B,C,E-G) and thyroid primordium (F).

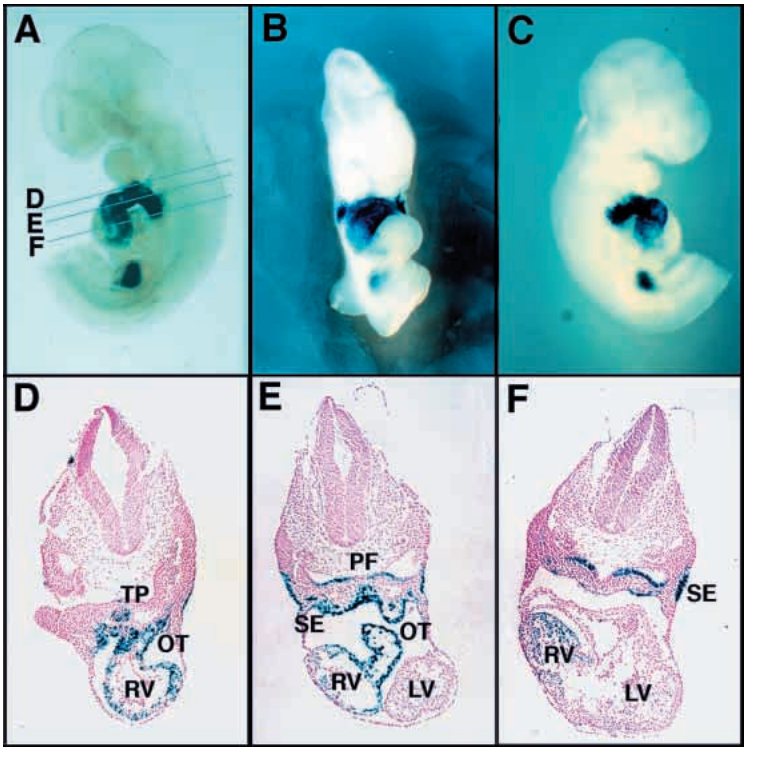


Fig. 5. Whole-mount X-gal staining (A-C) and transverse sections (D-F) of an E10.5 transgenic embryo carrying the CsxlacZ-6 transgene. *lacZ* was expressed in myocardial cells in the entire RV (B,E,F) and the outflow tract (B-E). Transgene expression was also seen in the pharyngeal floor (E,F) and thyroid primordium (D). Ectopic expression was observed in the surface ectoderm of pharyngeal arches (E,F).

Table 2. Summary of transgene expression

Construct	Regulatory sequence	Tissues examined for expression of transgene													
		OFT	RV(b)	RV(m)	RV(l)	LV(m)	LV(l)	IAG	A	PF	TP	ST	GG‡	SE‡	LG‡
CsxlacZ-1	5′ 3.3 kb	+	+	–	–	–	–	–	–	+	+	+	1*	+	+
CsxlacZ-2	5′ 3.3 kb +3rd int	+	+	–	–	–	–	–	–	+	+	+	–	–	–
CsxlacZ-3	5′ 4 kb	+	+	–	–	–	–	–	–	+	+	+	–	+	+
CsxlacZ-4	5′ 6 kb	–	–	–	–	–	–	–	–	–	–	–	–	–	–
CsxlacZ-5	5′ 14 kb	+	+	+	2*	+	2*	+	–	+	+	3*	–	–	–
CsxlacZ-6	5′ 3.3 kb +3′ 6 kb	+	+	+	+	–	–	–	–	+	+	+	–	+	–

OFT, outflow tract; RV(b), basal part of RV; RV(m), medial side of RV; RV(l), lateral side of RV; LV(m), medial side of LV; LV(l), lateral side of LV; IAG, interatrial groove (groove between right and left atria); A, atrium; PF, pharyngeal floor; TP, thyroid primordium; ST, stomach; GG, glossopharyngeal nerve ganglion; SE, surface ectoderm of pharyngeal arches; LT, laryngotracheal groove.

‡Expression in these tissues is ectopic.

1* *lacZ* expression was observed in one embryo.

2* Expression of *lacZ* was observed only in the trabecular layers.

3* *lacZ* expression was detected in one embryo.

was not at E15.5 (data not shown). Analysis of adult CsxlacZ-5 transgenic mice showed patchy b-gal staining along the luminal surface of the RV, particularly along the ventricular septal surface (Fig. 7D). β-gal staining was also seen in a small part of the LV base (Fig. 7E) and in the AV junction area (the arrow in Fig. 7E).

Autoregulation of Csx/Nkx2.5 expression

One of the mechanisms to ensure tissue-specific expression of a transcription factor would be a positive autoregulation of its own promoter, as has been shown in MyoD gene (Goldhamer et al., 1995). The *Csx/Nkx2.5* promoter contains multiple NKE elements, the binding sites for Csx/Nkx2.5 protein (Chen and Schwartz, 1995). In order to determine whether or not a positive autoregulation of *Csx/Nkx2.5* is present, we examined *lacZ* expression in mice homozygous null mutant for *Csx/Nkx2.5* created by homologous recombination. Since *lacZ* expression is under the control of 5′ and 3′ regulatory regions of *Csx/Nkx2.5* in the mutant embryos (Fig. 1C), β-gal staining will be weaker or absent in homozygous null embryos, if positive regulation by Csx/Nkx2.5 per se is important. Unexpectedly, β-gal staining was far stronger in the

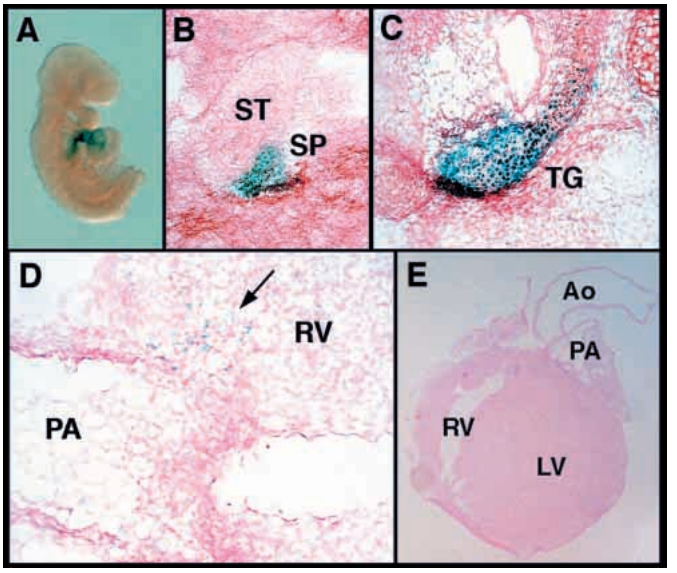


Fig. 6. Whole-mount X-gal staining (A), transverse sections (B-D) and a sagittal section (E) of mice from transgenic lines carrying the transgenes CsxlacZ-1. In a CsxlacZ-1 embryo at E9.5 (A), the cardiac expression pattern of β-gal staining is similar to that seen in the transient transgenic CsxlacZ-1 embryos. At E15.5, *lacZ* expression can be observed in the spleen (B), thyroid gland (C) and the outflow region of the RV (the arrow in D). Note no β-gal staining in the adult CsxlacZ-1 heart (E).

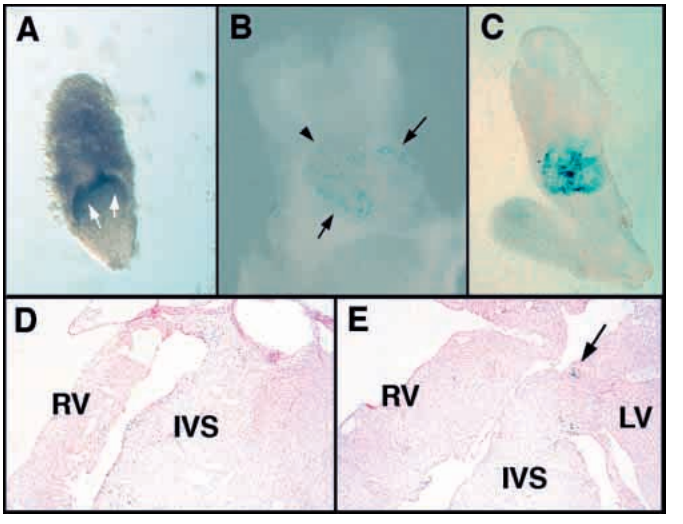


Fig. 7. X-gal staining of whole-mount (A-C), and transverse sections (D,E) of mice from transgenic lines carrying the transgenes CsxlacZ-5. At E7.5, β-gal staining was observed in the cardiac crescent (arrows in A) and at E8.25, the common ventricle (arrows in B) and the outflow tract (arrowhead in B) was positive for *lacZ* expression (B). In a CsxlacZ-5 embryo at E9.5 (C), the cardiac expression pattern of β-gal staining is also similar to that seen in the transient transgenic embryos. Sections through the heart of an adult CsxlacZ-5 mouse (D,E) showed β-gal staining along the luminal surface of the RV and a small part of the LV as well as staining in the subendocardial region of the interventricular septum and in the AV junction region (the arrow in E).

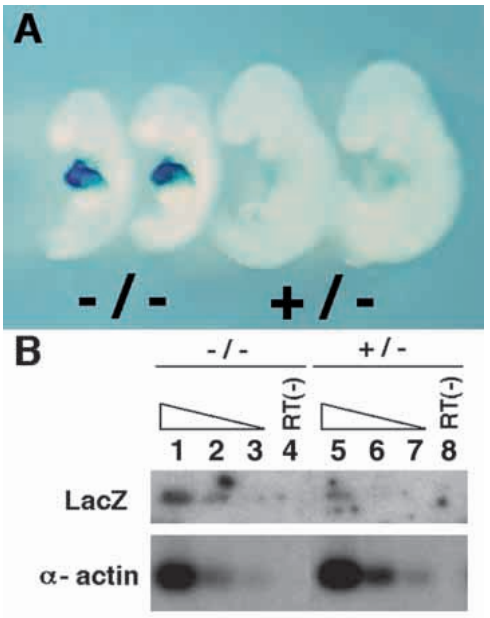


Fig. 8. (A) Whole-mount X-gal staining of homozygous (–/–) and heterozygous (+/–) null embryos for *Csx/Nkx2.5*. Note strong expression of *lacZ* in the homozygous mutant hearts. (B) Five fold serial dilutions of RT products (lanes 1-3 and 5-7) were used for PCR as template. RT(–) (lanes 4 and 8): RT reaction was carried out without reverse transcriptase.

homozygous mutant embryos than in the heterozygous mutant embryos stained simultaneously (Fig. 8A). This argues against a positive feedback regulation of *Csx/Nkx2.5*. Instead, a negative feedback, either directly or indirectly, of *Csx/Nkx2.5* seems to exist, since the intensity of β-gal staining in the homozygous mutant heart was much more than double that in the heterozygous mutant heart (Fig. 8A). In order to confirm this result, we performed semi-quantitative RT-PCR using RNA extracted from the heart of heterozygous and homozygous mutant embryos at E9.5. Transcripts of *lacZ* in homozygous mutant hearts were approximately 8 fold those in heterozygous mutant hearts, after adjustment using the relative abundance of the transcript for *α-cardiac actin* (Fig. 8B).

DISCUSSION

In this study, we have shown that *Csx/Nkx2.5* gene has three TATA-less promoters which generate at least three distinct transcripts by differential promoter usage coupled with alternative splicing. The most abundant transcript, however, is transcript III, which contains exon 1c and exon 2, and the previously described TN and NK-2 domains, and the homeodomain. The regulatory region of the *Csx/Nkx2.5* gene seems highly complex and modular, with separate enhancer elements allowing expression of the gene in anatomically distinct areas of the heart at specific developmental stages, as well as in the other tissues in which *Csx/Nkx2.5* is expressed.

Our results suggest that a major *cis*-acting element for extracardiac expression resides within intron 2, since the extracardiac β-gal staining pattern observed in mice carrying the CsxlacZ-1 construct is essentially replicated in mice transgenic for other constructs tested. The only exception is CsxlacZ-4 which had no staining in any extracardiac tissue nor

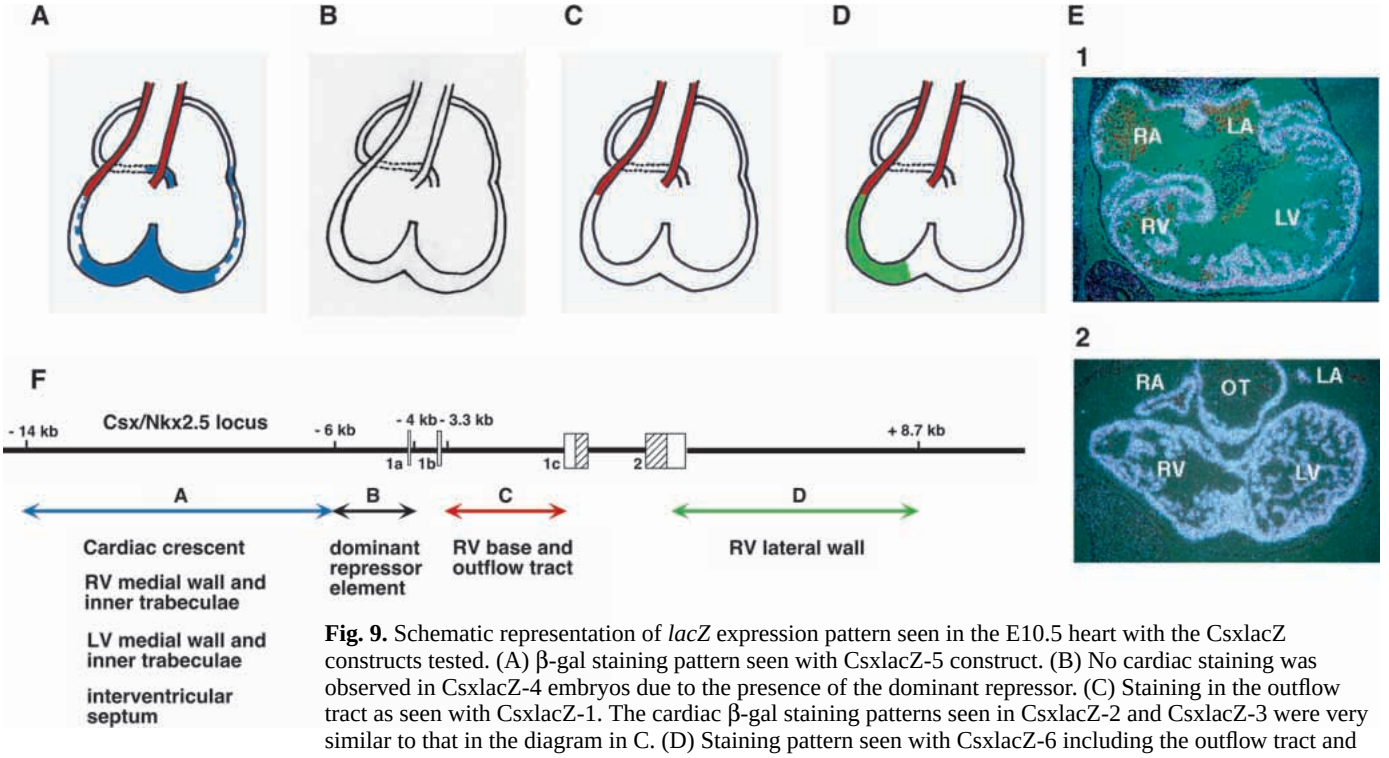


Fig. 9. Schematic representation of *lacZ* expression pattern seen in the E10.5 heart with the CsxlacZ constructs tested. (A) β-gal staining pattern seen with CsxlacZ-5 construct. (B) No cardiac staining was observed in CsxlacZ-4 embryos due to the presence of the dominant repressor. (C) Staining in the outflow tract as seen with CsxlacZ-1. The cardiac β-gal staining patterns seen in CsxlacZ-2 and CsxlacZ-3 were very similar to that in the diagram in C. (D) Staining pattern seen with CsxlacZ-6 including the outflow tract and the entire RV. **(**E) In situ hybridization of a E10.5 embryo using a *Csx/Nkx2.5* cRNA probe. Note a homogenous expression pattern of *Csx/Nkx2.5* mRNA in the embryonic heart, including the outflow tract (OT). (F) The summary of different enhancer elements of *Csx/Nkx2.5* that direct expression of the transgene at different subregions of the heart shown in A-D.

in the heart, presumably due to the presence of a strong negative *cis*-element between −4 kb and −6 kb.

The cardiac *lacZ* expression pattern observed in the heart suggests a highly modular control of *Csx/Nkx2.5* expression in the heart itself (Fig. 9). β-gal staining in the basal section of the RV and the outflow tract was driven by the intron 2 region in construct CsxlacZ-1. The expression pattern was essentially the same after addition of intron 3 (CsxlacZ-2) and also with the addition of most of intron 1, exon 1b and intron 2 (CsxlacZ-3) as the regulatory regions driving *lacZ* (Fig. 9C). The addition of 14 kb of 5′ flanking sequence (CsxlacZ-5) drove most of cardiac expression except for the atrium. The CsxlacZ-5 expression pattern included the outflow tract and the basal portion of the RV seen with other constructs, but also added *lacZ* expression to the interatrial groove, the atrioventricular canal, the ventricular septum and the medial walls and inner trabeculae of the RV and LV(Fig. 9A). Interestingly, the addition of 6 kb of 3′ flanking sequence downstream to the coding region of *Csx/Nkx2.5* along with the 3.3 kb intron 2 region drove strong *lacZ* expression in the entire RV, including the compact layer of the lateral walls (Fig. 9D).

In transgenic lines for CsxlacZ-5, *lacZ* expression was seen in the cardiac crescent at E7.5 and in the common ventricle and the outflow tract at E8.25, suggesting that −14 kb to −6 kb fragment also drives cardiac expression of *Csx/Nkx2.5* at early stages. In the adult transgenic mice, the limitation of β-gal staining along the luminal surface of the septum and in the AV junction suggests that this element may be driving reporter gene expression in a part of the cardiac conduction system in the adult heart. This is particularly intriguing in view of the recent report demonstrating that mutations in the homeodomain region of the human *CSX/NKX2.5* gene are responsible for non-syndromic atrial septal defects associated with AV conduction system defects (Schott et al., 1998).

Although we have analyzed the genomic region encompassing 22.7 kb surrounding the *Csx/Nkx 2.5* gene in this study, enhancers that guide expression of the gene in some regions have not been identified yet. In particular, only small subsections of the atrium were positive for the reporter gene expression; thus, additional atrial enhancer(s) remain to be described. Also, while strong expression of *Csx/Nkx2.5* was seen with the construct CsxlacZ-6 in both trabecular and compact layers of the RV, expression in the compact layer of the lateral wall of the LV has not been observed with these transgenic constructs, suggesting that additional enhancers are present elsewhere. Though expression of *Csx/Nkx2.5* in the tongue has been noted (Kasahara et al., 1998; Lints et al., 1993), the tongue did not stain in any of the transgenic constructs tested. Together with a progressively restricted regional pattern of expression of the constructs CsxlacZ-1 and CsxlacZ-5 during development, *Csx/nkx2.5* gene seems to require a very large number of *cis*-acting elements spread in a large area of the genome to recapitulate the homogeneous pattern of expression of this gene throughout development.

The endogenous *Csx/Nkx2.5* expression in the heart seems quite uniform and symmetrical within the heart throughout development. (Fig. 9E). Why then should *Csx/Nkx2.5* have such widely dispersed modular organization of *cis*-acting regulatory elements instead of having a compact tissue-specific enhancer, as described for the *α-myosin heavy chain* or *M creatine kinase* (Donoviel et al., 1996; Subramaniam et al., 1991)? It is possible that as the heart has increased in its complexity during evolution from a primitive tube to the four chambered structure, additional regulatory modules have been added to allow expression of the gene in new regions. Alternatively, a modular organization allows the expression of the *Csx/Nkx 2.5* gene at different sites within the heart at different developmental stages. Thus, a complex organogenesis may be achieved not only through the addition of new gene modules (Fishman and Olson, 1997) but also by adding new regulatory modules to the pre-existing genes during evolution.

There is precedent for modular regulation of other cardiac muscle genes. The proximal promoter element of the *ventricular myosin light chain-2* (*MLC 2V*) gene drives a reporter gene in a pattern restricted to the RV, while the endogenous *MLC 2V* is expressed throughout both ventricles (Ross et al., 1996). Two different regulatory sequences of the *myosin light chain 3f* gene have been shown to direct expression of a reporter gene to distinct regions of the heart (Franco et al., 1997). However, the complexity of modular enhancers seems to be much more pronounced in *Csx/Nkx2*.5 compared to the above contractile protein genes. This is probably due to the fact that *tinman*-like genes are evolutionarily much more ancient than sarcomeric contractile protein genes.

Modular regulation of expression over time has been described for *tinman*, the *Drosophila* homologue of *Csx/Nkx2.5*. *Tinman* expression has been shown to have three phases of mesodermal expression: initially, it is expressed in all cells of the trunk mesoderm; later, it is expressed in a broad band of dorsal mesoderm; and finally, *tinman* is limited in its expression to the heart precursor cells(Azpiazu and Frasch, 1993; Bodmer, 1993; Bodmer et al., 1990). It has recently been shown that these three phases of expression are each controlled by a separate enhancer element in the *tinman* promoter (Yin et al., 1997). Our results suggest that this differential temporal gene regulation by separate enhancers seems operative with *Csx/Nkx2.5* as well (Fig. 9). However, our transgenic constructs did not fully recapitulate the expression pattern of the endogenous gene. There must, therefore, be other regulatory elements which drive *Csx/Nkx2.5* expression at different developmental stages.

Mutant embryos homozygous null for *Csx/Nkx2.5* showed strong β-gal staining in the heart (Fig. 8). CSX/NKX2.5 protein binding sites were found in the intron of human *CSX/NKX2.5* gene and it was suggested that a positive autoregulation might be important for expression of *CSX/NKX2.5* based on a transient transfection assay (Oka et al., 1997). However, in a chromosomal context, there was no evidence of a positive autoregulation of *Csx/Nkx2.5* in the heart, at least in mice. Instead, a negative autoregulation of *Csx/Nkx2.5*, either directly or indirectly, seems to exist, which may serve to tightly control the levels of expression of *Csx/Nkx2.5*. It is noteworthy that the truncation mutations in human *CSX/NKX2.5* gene result in autosomal dominant forms of congenital heart disease, most likely due to haploid insufficiency (Schott et al., 1998). A precise regulation of the levels of expression of *Csx/Nkx2.5* seems to be important for normal cardiac development.

Interestingly, no β-gal staining was detectable in the stomach of either heterozygous or homozygous Csx/Nkx2.5 null embryos (Fig. 8). This raises an interesting possibility that, in the developing stomach, *Csx/Nkx2.5* gene is positively

autoregulated, either directly or indirectly, by its own gene product. Thus, autoregulation by *Csx/Nkx2.5* could be either negative or positive depending on the type of tissues.

Clearly, there are complex, time-dependent and combinatorial genetic pathways working together to form the vertebrate heart (Fishman and Chien, 1997; Lee et al., 1998; Lyons, 1996; Olson and Srivastava, 1996). Although the present study points to highly modular nature of the *Csx/Nkx2.5* enhancer elements we have not narrowed down the exact sequence requirement for each enhancer, which will require much more additional work. However, this is the first promoter analysis of a cardiac transcription factor to be reported in vertebrates. Further characterization of the *Csx/Nkx2.5* promoter/enhancer elements will not only give new insight into molecular mechanisms of cardiac organogenesis, but also allow site-directed expression of other genes in subregions of the heart to further our understanding of the intricacies of cardiac development.

We thank Sonia Bartunkova and Brett Richards for technical assistance, and Martina Shinke, Dan Wechsler and Mark Russell for critical review of the manuscript. This work was supported by grants from the National Institute of Health K08HL03507 and HD28820 to SBW, RO1HL51253 to SI, NRSA Fellowship to IWL and Paul Dudley White Fellowship from the American Heart Association Massachusetts Affiliate to MT.

REFERENCES

Ali Ansari-Lari, M., Jones, S. N., Timms, K. M. and Gibbs, R. A. (1996). Improved ligation-anchored PCR strategy for identification of 5′ ends of transcripts. *Biotechniques* **21**, 34-36, 38.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seiman, J. G., Smith, J. A. and Struhl, K. (1993). *Current Protocols in Molecular Biology.* John Wiley & Sons Inc., New York.

Azpiazu, N. and Frasch, M. (1993). tinman and bagpipe: two homeo box genes that determine cell fates in the dorsal mesoderm of Drosophila. *Genes Dev.* **7**, 1325-1340.

Bodmer, R. (1993). The gene tinman is required for specification of the heart and visceral muscles in *Drosophila*. *Development* **118**, 719-729.

Bodmer, R., Jan, L. Y. and Jan, Y. N. (1990). A new homeobox-containing gene, msh-2, is transiently expressed early during mesoderm formation of *Drosophila*. *Development* **110**, 661-669.

Bodmer, R. and Venkatesh, T. V. (1998). Heart development in Drosophila and vertebrates: conservation of molecular mechanisms. *Dev. Genet.* **22**, 181-186.

Bonnerot, C., Rocancourt, D., Briand, P., Grimber, G. and Nicolas, J. F. (1987). A beta-galactosidase hybrid protein targeted to nuclei as a marker for developmental studies. *Proc. Natl. Acad. Sci. USA* **84**, 6795-6799.

Chen, C. Y. and Schwartz, R. J. (1995). Identification of novel DNA binding targets and regulatory domains of a murine tinman homeodomain factor, nkx-2. 5. *J. Biol. Chem.* **270**, 15628-15633.

Donoviel, D. B., Shield, M. A., Buskin, J. N., Haugen, H. S., Clegg, C. H. and Hauschka, S. D. (1996). Analysis of muscle creatine kinase gene regulatory elements in skeletal and cardiac muscles of transgenic mice. *Mol. Cell. Biol.* **16**, 1649-1658.

Edmondson, D. G., Lyons, G. E., Martin, J. F. and Olson, E. N. (1994). Mef2 gene expression marks the cardiac and skeletal muscle lineages during mouse embryogenesis. *Development* **120**, 1251-1263.

Fishman, M. C. and Chien, K. R. (1997). Fashioning the vertebrate heart: earliest embryonic decisions. *Development* **124**, 2099-2117.

Fishman, M. C. and Olson, E. N. (1997). Parsing the heart: genetic modules for organ assembly. *Cell* **91**, 153-156.

Franco, D., Kelly, R., Lamers, W. H., Buckingham, M. and Moorman, A. F. (1997). Regionalized transcriptional domains of myosin light chain 3f transgenes in the embryonic mouse heart: morphogenetic implications. *Dev. Biol.* **188**, 17-33.

Goldhamer, D. J., Brunk, B. P., Faerman, A., King, A., Shani, M. and Emerson, C. P., Jr. (1995). Embryonic activation of the myoD gene is regulated by a highly conserved distal control element. *Development* **121**, 637-649.

Harvey, R. P. (1996). NK-2 homeobox genes and heart development. *Dev. Biol.* **178**, 203-216.

Heikinheimo, M., Scandrett, J. M. and Wilson, D. B. (1994). Localization of transcription factor GATA-4 to regions of the mouse embryo involved in cardiac development. *Dev. Biol.* **164**, 361-373.

Hogan, B., Beddington, R., Constantini, F. and Lacy, E. (1994). *Manipulating the Mouse Embryo.* Cold Spring Harbor Laboratory Press, New York.

Kasahara, H., Bartunkova, S., Schinke, M., Tanaka, M. and Izumo, S. (1998). Cardiac and extracardiac expression of Csx/Nkx2. 5 homeodomain protein. *Circ. Res.* **82**, 936-946.

Kelley, C., Blumberg, H., Zon, L. I. and Evans, T. (1993). GATA-4 is a novel transcription factor expressed in endocardium of the developing heart. *Development* **118**, 817-827.

Komuro, I. and Izumo, S. (1993). Csx: a murine homeobox-containing gene specifically expressed in the developing heart. *Proc. Natl. Acad. Sci. USA* **90**, 8145-8149.

Kuo, C. T., Morrisey, E. E., Anandappa, R., Sigrist, K., Lu, M. M., Parmacek, M. S., Soudais, C. and Leiden, J. M. (1997). GATA4 transcription factor is required for ventral morphogenesis and heart tube formation. *Genes Dev.* **11**, 1048-1060.

Lee, Y., Shioi, T., Kasahara, H., Jobe, S. M., Wiese, R. J., Markham, B. E. and Izumo, S. (1998). The cardiac tissue-restricted homeobox protein Csx/Nkx2. 5 physically associates with the zinc finger protein GATA4 and cooperatively activates atrial natriuretic factor gene expression. *Mol. Cell. Biol.* **18**, 3120-3129.

Lin, Q., Schwarz, J., Bucana, C. and Olson, E. N. (1997). Control of mouse cardiac morphogenesis and myogenesis by transcription factor MEF2C. *Science* **276**, 1404-1407.

Lints, T. J., Parsons, L. M., Hartley, L., Lyons, I. and Harvey, R. P. (1993). Nkx-2. 5: a novel murine homeobox gene expressed in early heart progenitor cells and their myogenic descendants *Development* **119**, 419-431.

Lyons, G. E. (1996). Vertebrate heart development. *Curr. Opin. Genet. Dev.* **6**, 454-460.

Lyons, I., Parsons, L. M., Hartley, L., Li, R., Andrews, J. E., Robb, L. and Harvey, R. P. (1995). Myogenic and morphogenetic defects in the heart tubes of murine embryos lacking the homeo box gene Nkx2-5. *Genes Dev.* **9**, 1654-1666.

Mohun, T. and Sparrow, D. (1997). Early steps in vertebrate cardiogenesis. *Curr. Opin. Genet. Dev.* **7**, 628-633.

Molkentin, J. D., Lin, Q., Duncan, S. A. and Olson, E. N. (1997). Requirement of the transcription factor GATA4 for heart tube formation and ventral morphogenesis.. *Genes Dev.* **11**, 1061-1072.

Oka, T., Komuro, I., Shiojima, I., Hiroi, Y., Mizuno, T., Aikawa, R., Akazawa, H., Yamazaki, T. and Yazaki, Y. (1997). Autoregulation of human cardiac homeobox gene CSX1: mediation by the enhancer element in the first intron. *Heart Vessels* **Suppl 12**, 10-14.

Okkema, P. and Fire, A. (1994). The Caenorhabditis elegans NK-2 class homeoprotein CEH-22 is involved in combinational activation of gene expression in pharyngeal muscle. *Development* **120**, 2175-2186.

Olson, E. N. and Srivastava, D. (1996). Molecular pathways controlling heart development. *Science* **272**, 671-676.

Reecy, J. M., Yamada, M., Cummings, K., Sosic, D., Chen, C.-Y., Eichele, G., Olson, E. N. and Schwartz, R. J. (1997). Chicken Nkx-2. 8:A novel homeobox gene expressed in early heart progenitor cells and pharyngeal pouch-2 and -3 endoderm. *Dev. Biol.* **188**, 295-311.

Ross, R. S., Navankasattusas, S., Harvey, R. P. and Chien, K. R. (1996). An HF-1a/HF-1b/MEF-2 combinatorial element confers cardiac ventricular specificity and established an anterior-posterior gradient of expression. *Development* **122**, 1799-1809.

Sadler, T. W. (1995) *Medical Embryology*. Williams & Wilkins, Baltimore.

Schlaeger, T. M., Qin, Y., Fujiwara, Y., Magram, J. and Sato, T. N. (1995). Vascular endothelial cell lineage-specific promoter in transgenic mice. *Development* **121**, 1089-1098.

Schott, J.-J., Benson, D. W., Basson, C. T., Pease, W., Silberbach, G. M., Moak, J. P., Maron, B. J., Seidman, C. E. and Seidman, J. G. (1998). Congenital heart disease caused by mutations in the transcription factor NKX2-5. *Science* **281**, 108-111.

Srivastava, D., Cserjesi, P. and Olson, E. N. (1995). A subclass of bHLH proteins required for cardiac morphogenesis. *Science* **270**, 1995-1999.

Srivastava, D., Thomas, T., Lin, Q., Kirby, M. L., Brown, D. and Olson, E. N. (1997). Regulation of cardiac mesodermal and neural crest development by the bHLH transcription factor, dHAND. *Nature Genet.* **16**, 154-160.

Subramaniam, A., Jones, W. K., Gulick, J., Wert, S., Neumann, J. and Robbins, J. (1991). Tissue-specific regulation of the alpha-myosin heavy chain gene promoter in transgenic mice. *J. Biol. Chem.* **266**, 24613-24620.

Tanaka, M., Kasahara, H., Bartunkova, S., Schinke, M., Komuro, I., Inagaki, H., Lee, Y., Lyons, G. E. and Izumo, S. (1998). Vertebrate homologs of tinman and bagpipe: roles of the homeobox genes in cardiovascular development. *Dev. Genet.* **22**, 239-249.

Tanaka, M., Chen, Z., Bartunkova, S., Yamasaki, N. and Izumo, S. (1999). The cardiac homeobox gene *Csx/Nkx2.5* lies genetically upstream of multiple genes essential for heart development. *Development* (in press).

Turbay, D., Wechsler, S. B., Blanchard, K. M. and Izumo, S. (1996). Molecular cloning, chromosomal mapping, and characterization of the human cardiac-specific homeobox gene hCsx. *Mol. Med.* **2**, 86-96.

Yin, Z., Xu, X. L. and Frasch, M. (1997). Regulation of the twist target gene tinman by modular *cis*-regulatory elements during early mesoderm development. *Development* **124**, 4971-4982.