GCTTTCTGCTGCCTTCT-3′) under the following conditions: 94°C for 5 min; 94°C for 1 min, 52°C for 1 min, and 72°C for 1 min; and 72°C for 7 min. PCR products were analyzed by agarose gel electrophoresis after serial cycles in the linear range of amplification. RNA loading was controlled for by amplification of the housekeeping gene, *glyceraldehyde 3-phosphate dehydrogenase* (*G3PDH*). Negative controls were performed for each sample using non–reverse-transcribed RNA. A Southern blot of the samples was hybridized for *UFD1L* or *G3PDH*.

17. Human metaphase chromosome slides were prepared from Epstein-Barr virus–transformed lymphoblastoid cell lines or peripheral blood. Slides were hybridized with a genomic clone encompassing *UFD1L* that had been obtained by screening a human P1 artificial chromosome library with *UFD1L*-specific primers. FISH analysis was performed according to R. Matsuoka *et al.* [*Hum. Genet.* **103**, 70 (1998)]. Patients with 22q11 deletion had been characterized by FISH using the N25 (Oncor) marker that does not encompass *UFD1L*.
18. Whole-mount in situ hybridizations were performed with digoxygenin-labeled antisense riboprobes synthesized from the 700-bp open reading frame and the 250-bp 3′ untranslated region of mouse *UFD1L* cDNA (*9*). Embryos were photographed with and without clearing in benzyl benzoate:methyl salicylate (1:1). For histologic analysis, stained embryos were embedded in paraffin after fixation. Transverse and sagittal sections were made at 8-μm intervals throughout the embryos.
19. M. L. Kirby, T. F. Gale, D. E. Stewart, *Science* **220**, 1059 (1983).
20. E. N. Olson and D. Srivastava, *ibid.* **272**, 1069 (1996); M. L. Kirby and K. L. Waldo, *Circ. Res.* **77**, 211 (1995).
21. A research protocol for human studies was approved by the Institutional Review Board of the University of Texas Southwestern Medical Center. Informed consent was obtained from the parents for collection of blood and thymic tissue. Genomic DNA was obtained from 10 individuals at Children's Hospital, Medical Center of Dallas, with normal karyotype and negative FISH for 22q11 deletion, who had cardiac or craniofacial defects, or both, typical of 22q11 deletion; DNA from 100 control individuals with normal karyotype was also obtained. Genomic DNA extracted from human blood or thymic tissue (QIAamp blood kit, QIAGEN) was digested with Eco RI or Hind III, transferred to nylon membranes, and hybridized to ^{32}P-labeled UFD1L or CDC45 cDNA.
22. Thymic tissue was collected from patients who were undergoing surgical repair of CHD, and total RNA was extracted and used for cDNA synthesis. Northern analysis was performed on 20 μg of total thymic RNA and hybridized to a *UFD1L* cDNA probe. *HIRA*- and *CDC45*-specific amplimers were used to amplify the respective cDNAs. *HIRA* amplimers were: upper, 5′-GACGGCTCTGTGGCATTCCT-3′; lower, 5′-GCCATCTGCTGTCCGAGTCT-3′. *CDC45* amplimers were: upper, 5′-GCCTTGTTCCAGTGTGACCA-3′; lower, 5′-GTTCTCCTCATCCTCGTTCC-3′.
23. S. C. Daw *et al.*, *Nature Genet.* **13**, 458 (1996).
24. L. G. Wilming *et al.*, *Hum. Mol. Genet.* **6**, 247 (1997); P. Magnaghi *et al.*, *Nature Genet.* **20**, 74 (1998).
25. H. Yamagishi, V. Garg, D. Srivastava, unpublished observations.
26. J. McKie *et al.*, *Genome Res.* **8**, 834 (1998).
27. T. H. Shaikh *et al.*, *Am. J. Hum. Genet.* **63**, A193 (1998).
28. M. Scheffner, J. M. Huibregtse, R. D. Vierstra, P. M. Howley, *Cell* **75**, 495 (1993); C. Sears, J. Olesen, D. Rubin, D. Finley, T. Maniatis, *J. Biol. Chem.* **273**, 1409 (1998).
29. L. Monney *et al.*, *J. Biol. Chem.* **273**, 6121 (1998).
30. We thank R. Schultz and A. Bowcock for helpful discussions and assistance with genetic analyses, members of the pediatric cardiology and cardiothoracic surgery divisions for assistance with tissue collection, other members of the Srivastava laboratory for critical discussion, J. L. Goldstein, M. S. Brown, E. N. Olson, and H. H. Hobbs for critical review of this manuscript, and J. Page for manuscript preparation. Supported by grants to D.S. from NIH (R01HL57181-01) and March of Dimes.

2 November 1998; accepted 19 January 1999

Regulation of Chamber-Specific Gene Expression in the Developing Heart by *Irx4*

Zheng-Zheng Bao,[1] Benoit G. Bruneau,[1] J. G. Seidman,[1] Christine E. Seidman,[2] Constance L. Cepko[1]*

The vertebrate heart consists of two types of chambers, the atria and the ventricles, which differ in their contractile and electrophysiological properties. Little is known of the molecular mechanisms by which these chambers are specified during embryogenesis. Here a chicken *iroquois*-related homeobox gene, *Irx4*, was identified that has a ventricle-restricted expression pattern at all stages of heart development. Irx4 protein was shown to regulate the chamber-specific expression of myosin isoforms by activating the expression of the ventricle myosin heavy chain–1 (*VMHC1*) and suppressing the expression of the atrial myosin heavy chain–1 (*AMHC1*) in the ventricles. Thus, Irx4 may play a critical role in establishing chamber-specific gene expression in the developing heart.

During embryonic development, the ventricles and atria of the heart arise from a single tubular structure (*1, 2*). Mature atria and ventricles differ in their contractile and electrophysiological characteristics and express distinct sets of genes (*3, 4*). Most of the known chamber-specific genes encode isoforms of contractile proteins, including the myosin heavy chains and light chains (*4*). The correct expression of these myosin isoforms is essential for embryonic survival and proper function of the mature heart (*5*). The mechanisms involved in regulation of these chamber-specific patterns in the developing heart tube are unknown. Here we show that in chick hearts, this process requires the proper regulation of the *iroquois*-related homeobox gene *Irx4*.

The *Irx4* gene was identified by a low-stringency hybridization screening of a chick embryonic day 6 to 8 (E6–E8) retinal cDNA library, with probes derived from mouse and human EST clones that span the *Iroquois* homeodomains (*6*). The predicted open reading frame contains a homeodomain highly homologous to those in the *Drosophila* Iroquois proteins (*7, 8*), and it is most closely related to the human IRX4 (83% amino acid identity overall and 70% amino acid identity outside of the homeodomain) (Fig. 1, and supplementary data available at www.sciencemag.org/feature/data/985642.shl). We have also identified a mouse IRX4 with 71% overall amino acid homology to chick Irx4 (*9*).

By in situ hybridization (*10*), we detected *Irx4* expression in the retina, a subset of nuclei in the hindbrain, the developing feather buds, and the heart (Fig. 1, B to D) (*11*). *Irx4* was highly expressed in the ventricular myocardium, but expression was absent from the atria or the distal outflow tract in the developing heart. Low levels of expression were observed in the proximal outflow tract. The ventricle-specific expression pattern was observed as early as stage 10 in the prospective ventricular region and persisted in all developmental stages examined (Fig. 1, B to D, and Fig. 2). The expression of the mouse *Irx4* gene was similarly restricted to the ventricles in all stages of developing heart and adult heart (*9*).

We next compared *Irx4* with other genes known to have a chamber-restricted expression pattern in the developing chick heart. *Irx4* expression was first observed at Hamburger-Hamilton (HH) stage 10, when the developing heart is a linear tube. At stage 10, *Irx4* expression was already restricted to the middle portion of the heart tube, which corresponds to the prospective ventricles (Fig. 2B). The ventricle-restricted pattern persisted to later stages (Fig. 2, E and H). In contrast to *Irx4*, ventricle myosin heavy chain–1 (*VMHC1*) gene expression was detected earlier and in regions of the heart tube destined to become both atria and ventricles (Fig. 2, A and D) (*12*). As with *VMHC1*, early atrial myosin heavy chain–1 (*AMHC1*) gene expression also was not restricted (Fig.

[1]Howard Hughes Medical Institute and Department of Genetics, Harvard Medical School, 200 Longwood Avenue, Boston, MA 02115, USA. [2]Howard Hughes Medical Institute and Cardiovascular Division, Brigham and Women's Hospital, Boston, MA 02115, USA.

*To whom correspondence should be addressed. E-mail: cepko@rascal.med.harvard.edu

2F). Although previous studies suggested that *AMHC1* is expressed only in the prospective atrial region (posterior) at stage 9 (*13*), we observed posterior and anterior expression of *AMHC1* during its onset at late stage 10 to stage 12 (an ~10-hour window). As development proceeded, AMHC1 RNA became more abundant in the posterior and less abundant in the anterior. The restriction of *AMHC1* and *VMHC1* expression to different regions of the developing heart was obvious by stage 13 (Fig. 2, G and I) when the morphological constriction between atria and ventricles became visible. A similar pattern of expression has been described for the quail homolog of *AMHC1*, *slow MyHC 3* (*14*).

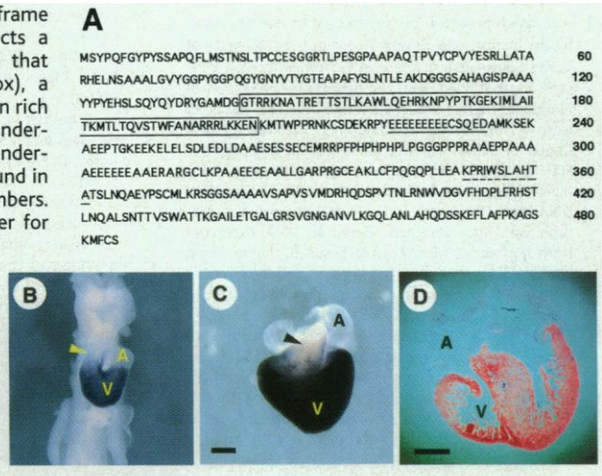

Fig. 1. (**A**) The open reading frame of the chick *Irx4* gene predicts a 485–amino acid protein (*28*) that includes a homeodomain (box), a putative transactivation domain rich in acidic amino acids (solid underline), and an iro box (dotted underline) (*8*). These features are found in all known *Iroquois* family members. The GenBank accession number for chick *Irx4* is AF091504. Expression patterns of *Irx4* in the developing heart at stage 19 (**B**), E4 (stage 24) (**C**), and E9 (stage 35) (**D**). In situ hybridizations on whole mount (B and C) and on a paraffin section of heart (D) are shown. Note that *Irx4* is only expressed in the ventricles (V) of the heart, not in the atria (A) or in the distal outflow tract (arrowheads). The section plane in (C) does not include the outflow tract. Scale bar in (C) is common to (B), 100 μm; bar in (D), 0.5 mm.

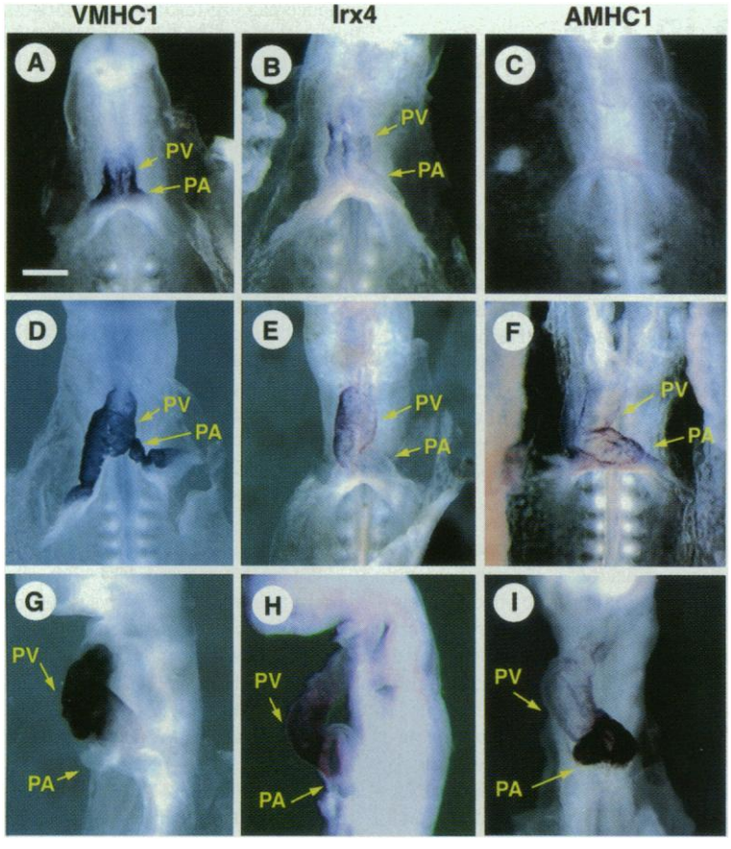

Fig. 2. Expression patterns of *Irx4*, *AMHC1*, and *VMHC1* mRNAs in the early chicken heart. All panels show whole-mount in situ hybridizations. Early expression of *Irx4* (**B**, **E**, and **H**), *VMHC1* (**A**, **D**, and **G**), *AMHC1* (**C**, **F**, and **I**) at stage 9+ (A and C), stage 10 (B), stage 10+ (D, E, and F), and stage 13 (G, H, and I). The prospective ventricular (PV) region is anterior to the prospective atrial (PA) region. Note that expression of the *Irx4* gene is restricted to the prospective ventricular region as early as stage 10, whereas both *AMHC1* and *VMHC1* are expressed throughout the heart tube initially. Chamber-specific expression of *AMHC1* and *VMHC1* began after stage 12. Scale bar, 100 μm.

Two features of the spatiotemporal expression patterns of *Irx4* and myosin heavy chain (MHC) isoform genes are noteworthy. First, *Irx4* expression was confined to the prospective ventricular region by stage 10, suggesting that the molecular specification of atria and ventricles occurs by this stage. Second, because the initial expression of *AMHC1* and *VMHC1* overlaps in the entire heart tube, chamber-specific expression observed later in development is probably achieved by down-regulation of *AMHC1* in the ventricles and down-regulation of *VMHC1* in the atria. Consistent with this hypothesis, *slow MyHC 3* contains a negative regulatory element, mutation of which results in ectopic expression of *slow MyHC 3* in ventricular cells (*15*).

To define the role of *Irx4* in cardiac development, we used both gain-of-function and loss-of-function approaches. A replication-competent avian retrovirus (RCAS-Irx4) (*16*), encoding full-length mouse *Irx4* cDNA, was injected into the precardiac regions of stage 7-8 embryos (*17*). Injected embryos were incubated for another 5 days (E6) and analyzed by in situ hybridization. At E6 (stage 29), expression of *VMHC1* and *AMHC1* is normally completely restricted to ventricles and atria, respectively. However, in hearts infected with RCAS-Irx4 virus, substantial amounts of *VMHC1* mRNA were observed in the atria (Fig. 3A), and the normally high level of atrial *AMHC1* mRNA was markedly reduced (Fig. 3B). These phenotypes were observed in more than 90% of the injected hearts (n = 28). Hearts injected with a control virus encoding alkaline phosphatase did not show altered expression of *VMHC1* or *AMHC1*, demonstrating that the effect of RCAS-Irx4 is due to *Irx4* expression, and not to viral infection alone. In addition, *VMHC1* was never induced in the distal outflow tract in RCAS-Irx4–infected hearts, even though the infection was throughout the heart, suggesting that Irx4 function is limited to a predetermined field. Thus, misexpression of *Irx4* in the atria up-regulates *VMHC1* expression and down-regulates *AMHC1* expression.

To perturb the function of Irx4 in the ventricles, we introduced a putative dominant negative construct, RCAS-H+en^r, by retroviral infection. This construct encodes a fusion protein composed of the chick Irx4 homeodomain and the repressor domain of the *Drosophila* Engrailed protein. Several studies (*18*) have shown that fusion of a DNA bind-

ing domain such as a homeodomain with the repressor domain of Engrailed can create a protein that interferes with transcriptional activation by the wild-type protein. This creates a dominant negative effect, as well as a potential gain-of-function effect due to active repression. At E6 (stage 29), *AMHC1* expression was restricted to atria in uninjected embryos as well as in embryos injected with control viruses encoding either an irrelevant homeodomain fused with the Engrailed repressor domain or the Engrailed repressor domain alone (Fig. 3). In contrast, *AMHC1* was abundantly expressed in the ventricular myocardium of the hearts injected with RCAS-H+en^r (Fig. 3, C to E). Ventricular expression of *VMHC1* was suppressed in >80% of the hearts ($n > 40$) injected with RCAS-H+en^r (Fig. 3, F to H). These results demonstrate that Irx4 function is also required to maintain a ventricular profile of myosin heavy chain gene expression.

Both RCAS-Irx4–injected hearts and RCAS-H+en^r–injected hearts displayed grossly normal morphology in both atria and ventricles. There were no detectable abnormalities in trabeculation or in the thickness of the myocardium layer in either chamber (*11*). This suggests that Irx4 does not control morphological aspects of atrial or ventricular identity. Alternatively, the misexpression experiments may have missed a critical time window, produced insufficient amounts of proteins, or the injected embryos may have died too early for a morphological defect to be manifested. We also examined the expression of endogenous *Irx4* in the misexpression experiments by using a probe specific for the 5′ untranslated region. No change of expression was detected in either RCAS-Irx4– or RCAS-H+en^r–injected hearts, suggesting that *Irx4* does not regulate its own expression (*11*).

Because the initial expression of *AMHC1* and *VMHC1* is not chamber-specific, the ectopic expression of *AMHC1* and *VMHC1* in the misexpression experiments could have resulted from early derepression of these genes, rather than later reinduction. We therefore examined the infected embryos at earlier stages. At 24 hours after infection (stage 13), when restricted expression was first detected, neither RCAS-Irx4 nor RCAS-H+en^r affected the expression of *AMHC1* or *VMHC1*. This was expected because ~24 hours is required for viral propagation and viral gene expression. At 48 hours after infection, ~20% of the hearts showed some ectopic induction of *AMHC1* or *VMHC1*. Thus, *Irx4* can affect MHC gene expression after the initial segmental restriction is completed. A minimum delay of 48 hours may also explain the observation that only a subset of cells responded to the viral misexpression (Fig. 3, E and H) even though the hearts were completely infected.

Several transcription factors, Nkx2-5, GATAs, MEF2s, dHAND, and eHAND, have been shown to play critical roles in heart development (*2*), but none is expressed in the heart in a manner that suggests a role in establishment or maintenance of atrial versus ventricular characteristics. *Irx4* is expressed only in the ventricles at all stages of heart development. Moreover, chick, mice, and zebrafish all show ventricle-specific expression of *Irx4* (*9, 19*), suggesting that Irx4 has an evolutionarily conserved role in heart development. Although fate mapping studies indicate that atrial and ventricular lineages in the precardiac mesoderm are separated during or shortly after gastrulation (stage 5 in chick) (*20*), cells are not committed to the atrial or ventricular fates until much later (*21*). Once committed, the cardiac cells maintain the expression of the proper contractile isoforms even when they are grafted into the opposite cardiac chamber, suggesting that an intrinsic factor (or factors) is responsible for maintaining this chamber-specific expression pattern (*22*). Our results indicate that Irx4 can alter the MHC expression profile after the chamber-specific pattern is established. Thus, Irx4 may be a maintenance as well as determining factor in the development of the cardiac chambers.

The roles that Irx4 plays in regional specification within the heart tube is reminiscent of the functions of other *Iroquois* family members in regionalization of the *Drosophila* wing discs (*7*), eye disc (*8*), and neural precursor domain in *Xenopus* (*23*). As noted for other family members (*7, 24*), additional factors may act in concert with Irx4 to specify multiple aspects of the chamber identities. It is intriguing that a homeodomain protein expressed only in the ventricle establishes the myosin isoform profiles in both atrial and ventricular chambers. This implies that the contractile characteristics of the cardiac chamber may be determined by imposition of a ventricular phenotype over a default atrial fate, by the action of Irx4 in the ventricles.

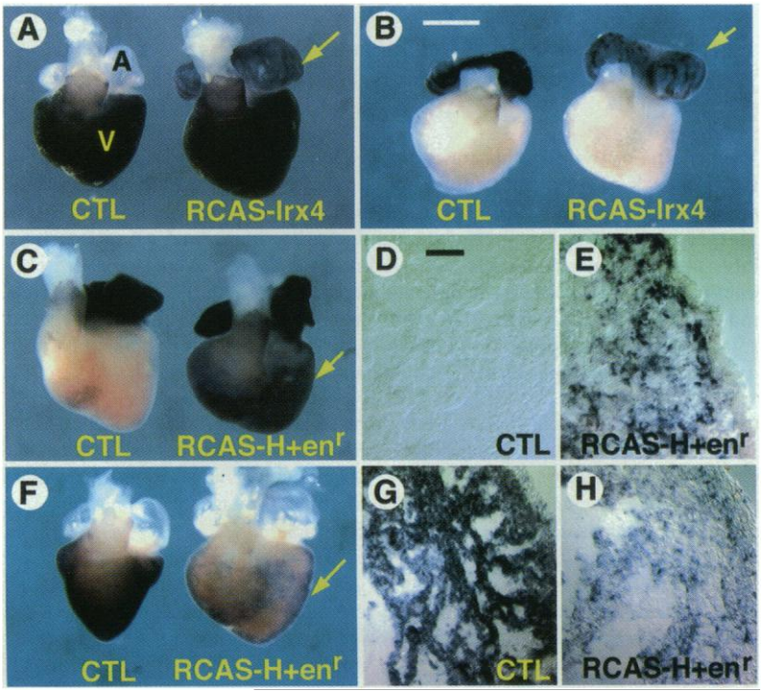

Fig. 3. Functional analyses of *Irx4* in heart development. Recombinant retroviruses were injected into the cardiogenic regions at stage 7–8 in ovo. The injected hearts were harvested at E6 for further analyses. In situ hybridization on whole mount or cryosections of hearts were performed with *VMHC1* (**A**, **F** to **H**) or *AMHC1* (**B**, **C** to **E**) probes. Results of misexpression of full-length Irx4 are shown in (A) and (B). Compared with the uninjected controls (CTL) (left), hearts infected with RCAS-Irx4 virus (right) showed up-regulation of *VMHC1* mRNA [arrow in (A)] and down-regulation of *AMHC1* mRNA in the atria [arrow in (B)]. Results of misexpression of the dominant negative Irx4 construct are shown in (C) to (H). The control hearts were injected with the virus encoding the Engrailed repressor domain only (RCAS-en^r) [CTL in (C) and (F)]. Hearts infected with RCAS-H+en^r showed ectopic induction of *AMHC1* mRNA [arrow in (C)] and down-regulation of *VMHC1* mRNA in the ventricles [arrow in (F)]. The results were confirmed by in situ hybridization on cryosections of hearts injected with RCAS-en^r (D and G) or RCAS-H+en^r (E and H), probed with *AMHC1* (D and E), or *VMHC1* (G and H). Cross sections through ventricles are shown. Induction of *AMHC1* in the ventricles of RCAS-H+en^r–injected hearts was within the myocardium. Scale bar in (B) is common to (A), (C), and (F), 0.5 mm; bar in (D) is common to (E), (G), and (H), 25 μm.

References and Notes

1. R. L. DeHaan, in *Organogenesis*, R. L. DeHaan and H. Ursprung, Eds. (Holt, Rinehart and Winston, New York, 1965), pp. 377–419.

2. M. C. Fishman and E. N. Olson, *Cell* **91**, 153 (1997); M. C. Fishman and K. R. Chien, *Development* **124**, 2099 (1997).
3. R. Zeller, K. D. Bloch, B. S. Williams, R. J. Arceci, C. E. Seidman, *Genes Dev.* **1**, 693 (1987).
4. K. R. Chien *et al.*, *Annu. Rev. Physiol.* **55**, 77 (1993); G. E. Lyons, *Trends Cardiovasc. Med.* **3**, 184 (1994).
5. J. Chen *et al.*, *J. Biol. Chem.* **273**, 1252 (1998); C. M. Pawloski-Dahm *et al.*, *Circulation* **97**, 1508 (1998).
6. Chick *Irx4* cDNA was isolated by low-stringency hybridization of an E6-E8 random-primed retinal cDNA library with probes derived from a mouse expressed sequence tag (EST) clone (GenBank accession number W41873) and a human EST clone (GenBank accession number R46202). A full-length cDNA was obtained by screening an oligo(dT)-primed whole-embryo cDNA library (stages 12 to 15), with the original *Irx4* clone as a probe. Mouse *Irx4* cDNA was isolated similarly by low-stringency screening with chick *Irx4* as a probe from an E10 mouse heart library (Stratagene). The DNA sequence was determined on both strands by an automated DNA sequencer (Applied Biosystems).
7. J. L. Gomez-Skarmeta, R. D. del Corral, E. de la Calle-Mustienes, D. Ferre-Marco, J. Modolell, *Cell* **85**, 95 (1996).
8. H. McNeill, C. H. Yang, M. Brodsky, J. Ungos, M. A. Simon, *Genes Dev.* **11**, 1073 (1997).
9. B. G. Bruneau *et al.*, unpublished data.
10. Probe synthesis and in situ hybridization procedures were performed as in (*25, 26*). Proteinase K was used at 1 μg/ml, 2.5 μg/ml, and 10 μg/ml for embryos of stages 5 to 9, stages 10 to 14, or stage 15 and older, respectively. To avoid cross-hybridization to other *Iroquois* family members, we generated a template by Bbs I digestion of the *Irx4* cDNA cloned in pBluescript SK(−) for synthesis of an antisense riboprobe (1.2 kb) without the homeobox sequence. In the misexpression experiments, a probe specific for the 5′ untranslated region (UTR) of chick *Irx4* was used. Probes specific for *VMHC1* and *AMHC1* were created on the basis of their least conserved 3′ sequence including the 3′ UTR (*12, 13*). To prepare templates for making *VMHC1* and *AMHC1* probes, we performed reverse transcription–polymerase chain reaction (RT-PCR) using total RNA from stage-17 chick embryos (primers for *VMHC1*: 5′-CTACAAACACCAAGCAGAGGAAGC, 3′-GGTTTGCACTTGATATTTATTTTTGTA; and for *AMHC1*: 5′-GCGGGTCCAGCTTCTCCACTCC, 3′-GCACCTTGACACGCCGCTCTGACTT).
11. Z.-Z. Bao, B. G. Bruneau, J. G. Seidman, C. E. Seidman, C. L. Cepko, unpublished data.
12. J. G. Bisaha and D. Bader, *Dev. Biol.* **148**, 355 (1991).
13. K. E. Yutzey, J. T. Rhee, D. Bader, *Development* **120**, 871 (1994).
14. G. F. Wang, J. Nikovits, W. M. Schleinitz, F. E. Stockdale, *Mol. Cell. Biol.* **18**, 6023 (1998).
15. ———, *J. Biol. Chem.* **271**, 19836 (1996).
16. Full-length chick and mouse *Irx4* cDNAs were fused in frame with the initiator ATG of the pSLAX-13 shuttle vector (*25*). The cDNAs were then transferred into the RCASBP(A) retroviral vector. To increase the infection rate, we made pseudotyped viruses with the vesicular stomatitis virus G coat (VSV-G) [M. Logan, S. M. Pagan-Westphal, D. M. Smith, L. Paganessi, C. J. Tabin, *Cell* **94**, 307 (1998)]. The viral titers for RCAS with chick *Irx4* were one-tenth to one-twentieth those for RCAS with mouse *Irx4*, presumably due to a high GC content in the 3′ end of the chick *Irx4*. Because a high infection rate is critical for the misexpression experiments, we used RCAS with the mouse *Irx4* insert. Because the mouse and chick Irx4 proteins are 71% homologous and the mouse *Irx4* has very similar expression patterns in the developing heart, the mouse *Irx4* is likely to have the same effect as the chick *Irx4* in the misexpression experiments. A putative dominant negative form of Irx4 (RCAS-H+en^r) was generated by fusion of a truncated chick Irx4 fragment including the homeodomain (from amino acid 109 to 220) with the repressor domain from the Engrailed protein. Control virus RCAS-en^r contains the Engrailed repressor domain only. RCAS-H+en^r and RCAS-en^r viruses were made as in (*25*), and both titers were about 5×10^8 plaque-forming units per milliliter.
17. Fertile White Leghorn chicken eggs (SPAFAS, Norwich, CT) were incubated and staged (*27*). To visualize the embryo, we introduced 1 to 10 μl of India ink, diluted 1:50 in Ringer's buffer, beneath the blastoderm. Concentrated retroviral stocks were pressure injected into the cardiogenic mesoderm at stage 7–8 with a glass micropipette. The injected embryos were returned to the incubator until further analyses.
18. P. Badiani, P. Corbella, D. Kioussis, J. Marvel, K. Weston, *Genes Dev.* **8**, 770 (1994); F. L. Conlon, S. G. Sedgwick, K. M. Weston, J. C. Smith, *Development* **122**, 2427 (1996); T. Furukawa, E. M. Morrow, C. L. Cepko, *Cell* **91**, 531 (1997).
19. D. Yelon and D. Stainier, unpublished data.
20. K. Yutzey, M. Gannon, D. Bader, *Dev. Biol.* **170**, 531 (1995); T. Mikawa and D. A. Fischman, *Annu. Rev. Physiol.* **58**, 509 (1996); D. Y. Stainier, R. K. Lee, M. C. Fishman, *Development* **119**, 31 (1993).
21. J. Satin, S. Fujii, R. DeHaan, *Dev. Biol.* **129**, 103 (1988).
22. P. J. Gruber, S. W. Kubalak, K. R. Chien, *Development* **125**, 4427 (1998).
23. E. J. Bellefroid *et al.*, *EMBO J.* **17**, 191 (1998); J. L. Gomez-Skarmeta, A. Glavic, E. de la Calle-Mustienes, J. Modolell, R. Mayor, *ibid.*, p. 181.
24. J. L. Gomez-Skarmeta and J. Modolell, *Genes Dev.* **10**, 2935 (1996).
25. R. D. Riddle, R. L. Johnson, E. Laufer, C. Tabin, *Cell* **75**, 1401 (1993).
26. Z. Z. Bao and C. L. Cepko, *J. Neurosci.* **17**, 1425 (1997); L. Tessarollo, L. Nagarajan, L. F. Parada, *Development* **115**, 11 (1992).
27. V. Hamburger and H. L. Hamilton, *J. Morphol.* **88**, 49 (1951).
28. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
29. We thank K.-H. Lee for his help with histology; D. Schulte and D. Smith for control viruses; D. Schulte, J. Lin, A. Lassar, K.-H. Lee, M. Marvin, and M. Logan for critical review of the manuscript; and D. Yelon and D. Stainier for sharing unpublished results. Supported by the Howard Hughes Medical Institute, NIH (Z.-Z.B.), the American Heart Association, and the Medical Research Council of Canada (B.G.B.).

23 November 1998; accepted 22 January 1999

Nonmethylated Transposable Elements and Methylated Genes in a Chordate Genome

Martin W. Simmen,[1]* Sabine Leitgeb,[1]* Jillian Charlton,[1] Steven J. M. Jones,[2] Barbara R. Harris,[2] Victoria H. Clark,[1] Adrian Bird[1]†

The genome of the invertebrate chordate *Ciona intestinalis* was found to be a stable mosaic of methylated and nonmethylated domains. Multiple copies of an apparently active long terminal repeat retrotransposon and a long interspersed element are nonmethylated and a large fraction of abundant short interspersed elements are also methylation free. Genes, by contrast, are predominantly methylated. These data are incompatible with the genome defense model, which proposes that DNA methylation in animals is primarily targeted to endogenous transposable elements. Cytosine methylation in this urochordate may be preferentially directed to genes.

DNA methylation in the dinucleotide sequence 5′-CpG can silence transcription. The genome defense model (*1*) posits that the primary role of methylation in animal genomes is to repress potentially damaging transposition of endogenous elements. By analogy with fungal systems (*2*), the elements are hypothesized to be targets for methylation because of their repetition in the genome (*3*). The hypothesis is difficult to test through analysis of the globally methylated mammalian genome. We therefore studied the specificity of methylation in a fractionally methylated genome belonging to the sea squirt *Ciona intestinalis*, an invertebrate member of the chordate phylum. Like most invertebrate genomes, that of *C. intestinalis* contains comparable amounts of methylated and nonmethylated DNA (*4*). We reasoned that any bias in the distribution of transposons or genes between the two fractions should therefore be readily detectable.

Three cosmids containing *C. intestinalis* genomic DNA were studied in detail (*5*). The cosmids were sequenced, and the locations of likely protein-coding regions were determined by GENEFINDER exon-prediction software (*6*) and database homology searches (Fig. 1, A through C). Putative proteins encoded by 10 of 13 potential genes showed similarity to known proteins. A systematic search for repetitive elements among the cosmid sequences and 1486 short random genomic sequences (*5*) identified four transposable elements belonging to recognizable families (*7*): a gypsy/Ty3–like long terminal repeat (LTR) retrotransposon (*Cigr-1*), a long interspersed element (LINE)–like element (*Cili-1*), a miniature inverted repeat (*Cimi-1*), and a composite short interspersed element (SINE) (*Cics-1*) with consensus RNA po-

[1]Institute of Cell and Molecular Biology, University of Edinburgh, The King's Buildings, Edinburgh EH9 3JR, UK. [2]The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: A.Bird@ed.ac.uk