GENOMICS 27, 27–32 (1995)

Genomic Structure and Chromosomal Localization of the Mouse LIM/Homeobox Gene *Lhx3*

ALEXANDER B. ZHADANOV,* NEAL G. COPELAND,† DEBRA J. GILBERT,† NANCY A. JENKINS,† AND HEINER WESTPHAL*,[1]

*Laboratory of Mammalian Genes and Development, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892; and †Mammalian Genetics Laboratory, ABL-Basic Research Program, NCI-Frederick Cancer Research and Development Center, Frederick, Maryland 21702

Received December 21, 1994; accepted February 14, 1995

We have previously cloned the murine *Lhx3* cDNA that encodes a predicted protein containing two tandemly repeated LIM domains and a homeodomain. Early expression of *Lhx3* in oral ectoderm that is committed to contribute to the anterior and intermediate lobes of the pituitary and its perseverance in the adult gland strongly suggest an involvement of the gene in mediating and maintaining the differentiation program of this important endocrine system. Additional functions are suggested by the fact that *Lhx3* is also expressed bilaterally along the spinal cord and the hindbrain at early stages of mouse development. Here we report the structural organization and chromosomal localization of the *Lhx3* gene. The gene is composed of six exons and five introns. Two different exons, Ia and Ib, appear to be alternatively spliced to exon II. The first LIM domain is encoded by exon II and the second by exon III. The homeobox is shared by exons IV and V. We have mapped *Lhx3* to the proximal region of mouse chromosome 2 in a region that shares homology with human chromosomes 9q and 10p.

INTRODUCTION

Homeodomain-containing proteins constitute a major class of transcriptional regulators (Levin and Hoey, 1988). A subgroup of LIM-type homeodomain proteins is characterized by two tandemly repeated LIM motifs. The LIM motif, named after the three homeodomain proteins *lin-11* (Freyd *et al.*, 1990), *Islet-1* (*isl-1*) (Karlsson *et al.*, 1990), and *mec-3* (Way and Chalfie, 1988), is a zinc-binding motif composed of about 50–60 amino acid residues that form a conserved pattern of cysteine and histidine residues, $CxxCx_{17\text{-}19}HxxCxxCxxCx_{16\text{-}20}Cxx(D/H/C)x$ (reviewed by Sanchez-Garcia and Rabbitts, 1993). This structure can be defined as a pair of zinc fingers separated by a linker of two amino acids (Sanchez-Garcia and Rabbitts, 1994). The LIM domain has been shown to bind Zn^{2+} (Michelsen *et al.*, 1993; Archer *et al.*, 1994). Most zinc-finger motifs function by binding to specific DNA sequences (Klug and Rhodes, 1987); however, the LIM domain has not yet been shown to have DNA-binding activity. Recent *in vitro* and cell-based studies have demonstrated a specific role of the LIM motifs in protein dimerization that appears to depend on the coordination of two zinc atoms in the finger doublet (Feuerstein *et al.*, 1994; Schmeichel and Beckerle, 1994). So far, only negative regulatory functions have been ascribed to the LIM-type homeodomain proteins (Sanchez-Garcia *et al.*, 1993; Xue *et al.*, 1993; Taira *et al.*, 1994).

We have previously cloned and sequenced the murine *Lhx3* cDNA (Zhadanov *et al.*, 1995) based on its homology to the *Xenopus laevis Xlim-3* gene (Taira *et al.*, 1993). *Lhx3* mRNA accumulates in Rathke's pouch, the primordium of the pituitary, at day 9.5 of mouse embryonic development and is detected predominantly in the anterior and intermediate lobes of the adult pituitary. This suggests that the gene product may be involved in the establishment and maintenance of the differentiated phenotype of pituitary cells. The *Lhx3* gene encodes two different exons, Ia and Ib, that are alternatively spliced to the second exon generating two mRNAs that can specify the proteins Lhx3a and Lhx3b, 400 and 402 amino acids in length, respectively. In this report we describe the genomic organization and the chromosomal location of *Lhx3* in the mouse.

MATERIALS AND METHODS

Screening of a mouse genomic DNA library. A mouse 129^{SV} genomic library (*Sau*3A-digested chromosomal fragments cloned into the EMBL 3A vector) was kindly provided by Dr. S. Tonegawa (Massachusetts Institute of Technology, Cambridge, MA). The library was screened with an *Lhx3* cDNA labeled with [α-^{32}P]dCTP using a Prime-It II kit (Stratagene). Three clones that hybridized with 5′-specific (position 151–464 according to the cDNA sequence) and 3′-specific (position 1377–1826) probes were isolated. One clone that contained the full *Lhx3* sequence was subjected to restriction endonuclease digestion, followed by Southern hybridization to appropriate probes. The whole 17-kb insert of genomic DNA was isolated by *Sal*I

[1] To whom correspondence should be addressed at Building 6B, Room 413, NIH, Bethesda, MD 20892-2790. Telephone (301) 496-1855. Fax: (301) 402-0543.

0888-7543/95 $6.00

digestion using the corresponding linker sites of the EMBL 3A vector, subcloned into the *SalI* site of Bluescript and sequenced by the dideoxynucleotide method using a Sequenase kit (U.S. Biomedical).

Establishment of the intron–exon structure of the Lhx3 gene. The intron–exon boundaries were determined by direct comparison of the nucleotide sequences of *Lhx3* cDNA, 5′ RACE (Zhadanov *et al.*, 1995), and genomic clones. Exon Ia was mapped by restriction enzyme digestion followed by Southern hybridization using a 186-bp PCR-generated fragment of plasmid pRC17 (Zhadanov *et al., 1995*) as a probe. The sizes of introns were confirmed by direct sequencing.

Interspecific backcross mapping. Interspecific backcross progeny were generated by mating (C57BL/6J × *Mus spretus*) F1 females and C57BL/6J males as described (Copeland and Jenkins, 1991). A total of 205 N2 mice were used to map the *Lhx3* locus (see text for details). DNA isolation, restriction enzyme digestion, agarose gel electrophoresis, Southern blot transfer, and hybridization were performed essentially as described (Jenkins *et al.*, 1982). All blots were prepared with Hybond-N^+ nylon membrane (Amersham). The probe, a ~1960-bp *Eco*RI fragment of mouse cDNA, was labeled with [α-^{32}P]dCTP using a nick-translation labeling kit (Boehringer Mannheim); washing was performed to a final stringency of 1.0 × SSC, 0.1% SDS, 65°C. Fragments of 10.0 and 3.6 kb were detected in *Sac*I-digested C57BL/6J DNA and fragments of 7.8 and 3.5 kb were detected in *Sac*I-digested *M. spretus* DNA. The presence or absence of the 7.8 kb *M. spretus*-specific fragment was followed in backcross mice. A description of the probes and RFLPs for the loci linked to *Lhx3* including vimentin (*Vim*), Drosophila notch gene homolog 1 (*Notch1*), and spectrin 2, alpha, brain (*Spna2*) has been reported previously (Del Amo *et al.*, 1993). Recombination distances were calculated as described (Green, 1981) using the computer program SPRETUS MADNESS. Gene order was determined by minimizing the number of recombination events required to explain the allele distribution pattern.

RESULTS

Genomic Organization of the Lhx3 Gene

A 17-kb fragment of genomic DNA that hybridized strongly with the *Lhx3* cDNA was analyzed in detail by PCR, restriction endonuclease digestion, Southern hybridization, and direct sequencing. Figure 1 shows the genomic organization of the mouse *Lhx3* gene. The transcription unit contains six exons plus exon Ib that can be alternatively spliced to exon II (Zhadanov *et al.*, 1995), spanning about 8 kb. The sequences at the intron–exon boundaries (Fig. 2) fit well with consensus sequences reported for both the donor and the acceptor splice sites (Shapiro and Senapathy, 1987). Exon Ia contains the first 188 bp of the *Lhx3* transcript. The start of the open reading frame is located 100 bp from the extreme 5′-end of the sequence obtained by 5′-RACE. The alternative exon Ib, located about 2 kb downstream from exon Ia, also includes 103 bp of the 5′-untranslated part of the transcript and 94 bp of coding region and presumably is transcribed from an alternative promoter. Exon II extends from position 198 to 369, exon III from position 370 to 572, exon IV from position 573 to 724, exon V from position 725 to 893, and exon VI from position 894 to the 3′-end of the cDNA. The exons define structural domains that we previously described for the Lhx3 protein (Zhadanov *et al.*, 1995). Exons Ia and Ib encode two different NH_2-termini thereby defining two isoforms of the protein: Lhx3a and Lhx3b. Exon Ia encodes the first 29 amino acids of Lhx3a, exon Ib 31 amino acids of Lhx3b. The first LIM domain is encoded by nucleotides 209 to 364 in the cDNA and therefore constitutes almost the entire second exon. Exon III includes the sequence encoding the second LIM domain. The fourth exon encodes part of the homeodomain including helix 1 and helix 2. The fifth exon defines the rest of the homeodomain together with 42 amino acids located downstream from the homeodomain. The sixth exon encodes the proline-rich C-terminal part of the Lhx3 protein, and nucleotides 1313 to 2179 specify the 3′-untranslated region of the gene. The *Lhx3* gene has one polyadenylation site (aataaa) at position 2166–2172. Also, an ATTTA motif, previously detected in the 3′-untranslated regions of several mRNAs with rapid turnover including those of lymphokines, cytokines, and oncogenes (Shaw and Kamen, 1986; Sachs, 1993), is located at residues 2130–2134.

As mentioned above, analysis of clones generated by 5′-RACE revealed that *Lhx3* has two different exons, Ia and Ib, that can be alternatively spliced to exon II. No 5′-RACE clones were found containing both exons spliced together. Moreover, we could not detect any product by RT-PCR (pituitary RNA was used as a template for generation of the single-strand DNA) using exon Ia-specific forward primers and exon Ib-specific reverse primers. By way of control, we were able to amplify fragments of the expected sizes using an exon III-specific reverse primer and exon Ia- or exon Ib-specific forward primers (A. Zhadanov, data not shown). These data along with the observation that transcripts of *Lhx3a* isoform are detected by RT-PCR at Day 8.5 of prenatal development, whereas *Lhx3b*-specific transcripts appear at Day 9.5 (A. Zhadanov, unpublished), led us to entertain the notion that alternative processing of the *Lhx3* gene may depend on promoter usage. Figure 2 shows the sequences of the 5′-upstream regions of exons Ia and Ib. They do not contain a TATA box or "initiator" sequence characteristic for several TATA-less promoters (Smale and Baltimore, 1989). The presence of two GGGCGG Sp1 sites (Mitchell and Tjian, 1989) close to the 5′-end of exon Ia (see Fig. 2) and two CCGCCC Sp1 sites (Letovsky and Dynan, 1989) near the 5′ end of exon Ib supports the notion of alternative promoter sites. Experiments are in progress to analyze pituitary-specific cell lines for differential *Lhx3* promoter activity.

Chromosomal Localization

The mouse chromosomal location of *Lhx3* was determined by interspecific backcross analysis using progeny derived from matings of [(C57BL/6J × *M. spretus*)F1 × C57BL/6J] mice. This interspecific backcross mapping panel has been typed for over 1700 loci that are well distributed among all of the autosomes as well as the X chromosome (Copeland and Jenkins, 1991). C57BL/6J and *M. spretus* DNAs were digested with several enzymes and analyzed by Southern blot hybridization for informative restriction fragment length polymorphisms (RFLPs) using a mouse cDNA probe.

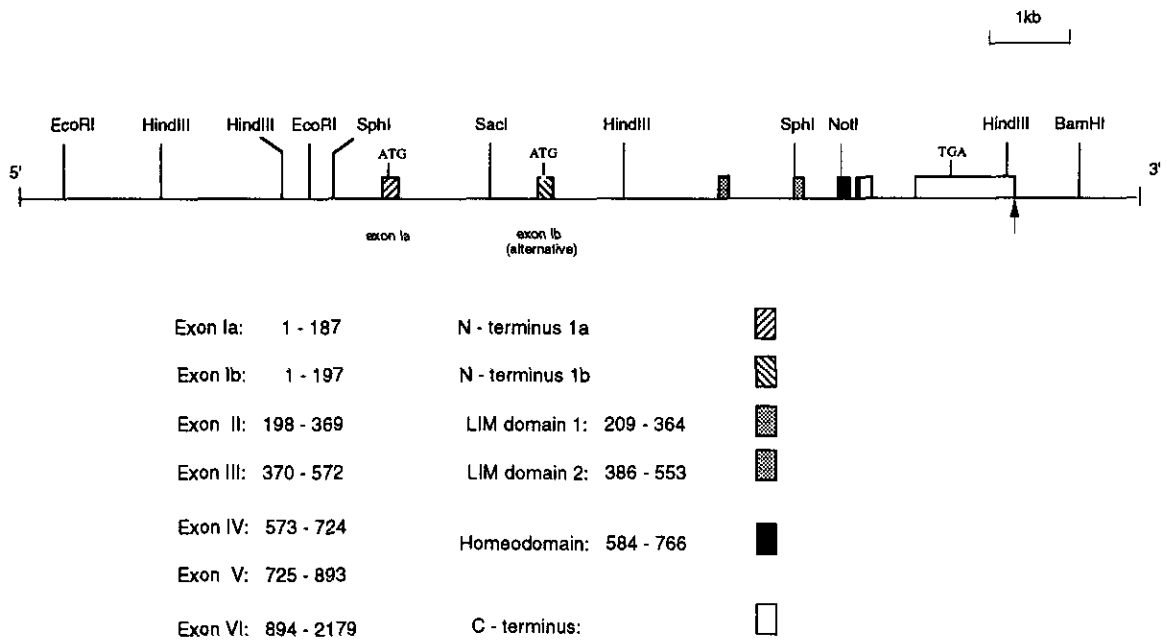

FIG. 1. Genomic structure of *Lhx3*. Exons are represented as rectangles. Initiation ATG codons and termination TGA codons are shown. The arrow points to the aataaa polyadenylation signal. Position numbers of each exon and each structural motif correspond to the numbered nucleotides of the cDNA or, in the case of exon Ia, to a sequence obtained by 5′-RACE (Zhadanov *et al.*, 1995). Each LIM domain is encoded by a separate exon. The homeobox is shared by exons IV and V.

The 7.8-kb *Sac*I *M. spretus* RFLP (see Materials and Methods) was used to follow the segregation of the *Lhx3* locus in backcross mice. The mapping results indicated that *Lhx3* is located in the proximal region of mouse chromosome 2 linked to *Vim, Notch1,* and *Spna2*. A total of 102 mice were analyzed for every marker as shown in the segregation analysis (Fig. 3), and 173 mice were typed for some pairs of markers. Each locus was analyzed in pairwise combination for recombination frequencies using the additional data. The ratios of the total number of mice exhibiting recombinant chromosomes to the total number of mice analyzed for each pair of loci and the most likely gene order are centromere–*Vim*–15/173–*Lhx3*–0/157–*Notch1*–1/116–*Spna2*. The recombination frequencies (expressed as genetic distances in centimorgans ± the standard error) are *Vim*–8.7 ± 2.1—(*Lhx3.Notch1*)–0.9 ± 0.9–*Spna2*. No recombinants were detected between *Lhx3* and *Notch1* in 157 animals typed in common, suggesting that the two loci are within 1.9 cM of each other (upper 95% confidence limit).

DISCUSSION

Comparison of the LIM Domain Encoding Exons of the Lhx3 Gene with Those of Other LIM Genes

The two LIM domains of the *Lhx3* gene are encoded by separate exons, whereas the homeodomain encoding sequence is interrupted by a short 86-bp intron. The positions of all intron–exon boundaries are exactly the same for mouse and zebrafish *Lhx3* genes (Glasgow, unpublished). The LIM-only proteins, however, have a more complex structure. The first LIM domain of the human rhombotine genes *RBTN1/Ttg1, RBTN2/Ttg2,* and *RBTN3/Ttg3* (McGuire *et al.*, 1989; Boehm *et al.*, 1990, 1991), as well as their mouse homologs, are encoded by single exons, whereas in the case of the second LIM domain this is true only for *RBTN2* (Foroni *et al.*, 1992). The *RBTN1* and *RBTN3* genes have an intron in the middle of the second LIM domain that disrupts the second putative Zn finger defined by the LIM motif. The authors suggest that the extra intron may have been inserted into a precursor *RBTN1/3* gene prior to duplication that created the present *RBTN1* and *RBTN3* genes (Foroni *et al.*, 1992).

The genomic structure of the *Lhx3* gene seems to share some features with those of RBTN genes. Interestingly, all three RBTN genes have a different genomic organization of the 5′-regions. The *RBTN3* gene contains two exons upstream of the first LIM domain (Boehm *et al.*, 1990). In the *RBTN1* gene exons I and Ia are used alternatively depending on promoter usage (Boehm *et al.*, 1990). The mouse retinoic acid X receptor (*RXR-β*) gene is another example of a gene where two isoforms of the protein designated *β*-1 and *β*-2 are generated from separate exons transcribed from different CpG island promoters and spliced to a common acceptor site in the transactivation domain (Nagata *et al.*, 1994). Most adult tissues have a similar amount of both *RxR-β1* and *RxR-β2*. The gene structure of *Lhx3* is also designed to yield two protein isoforms (Zhadanov *et al.*, 1995). The *Lhx3* gene contains Sp1 motifs that are characteristic for TATA-less promoters, immediately upstream of the transcription start site of exon Ia and close to the 5′-end of exon Ib. These motifs were shown to have an enhancing effect on this class of pro-

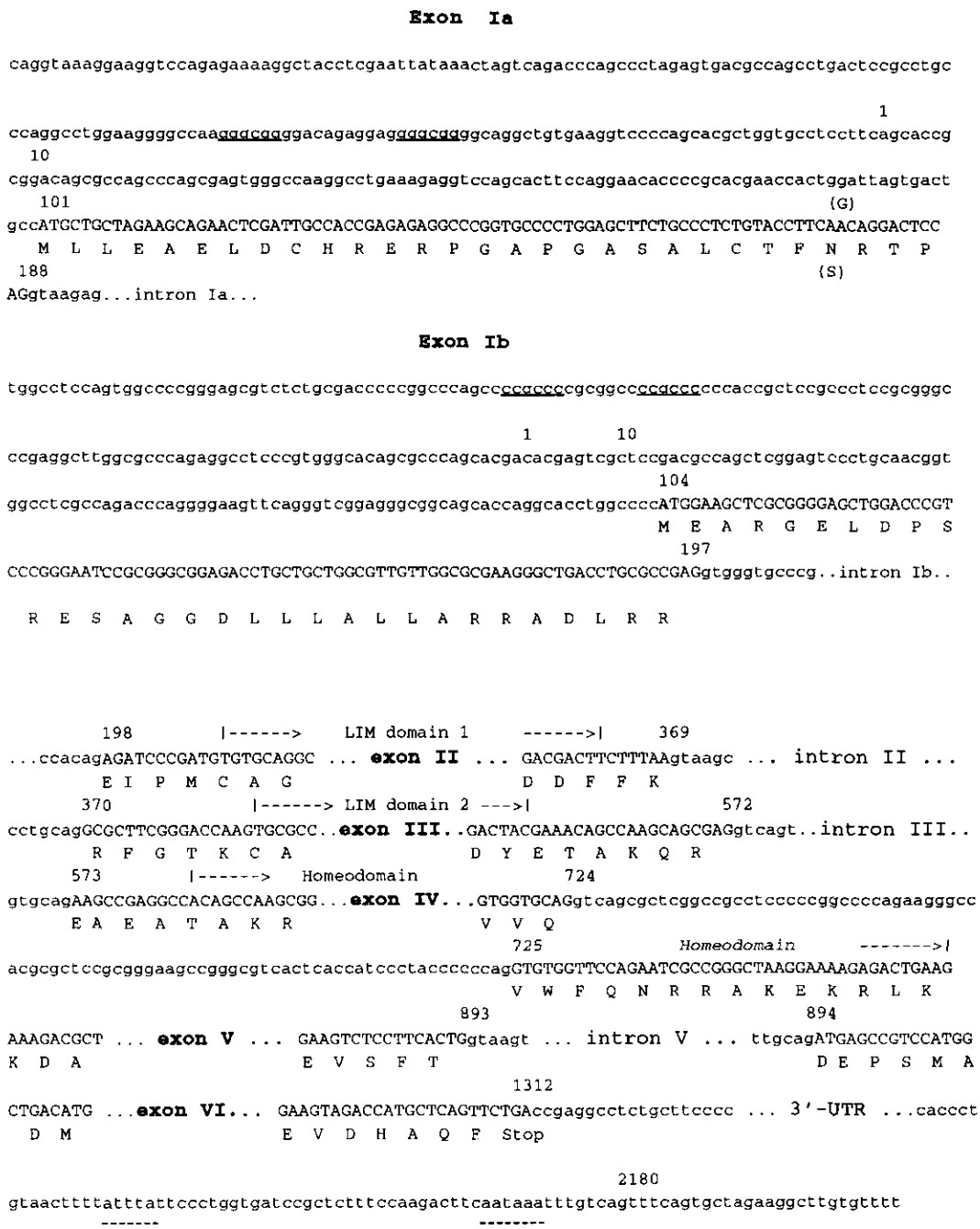

FIG. 2. Sequence of the 5′-region and the intron–exon boundaries of the *Lhx3* gene. These data can be recalled from the Genomic Sequence Data Base of the National Center of Genome Resources (Accession Nos. L38248, L38249, L40482, and L40483). The numbers are placed above the sequence and correspond to the numbered basepairs of the pRC17 clone (position 1–187) and the cDNA clone (position 1–197 and 198–2179) (Zhadanov *et al.*, 1995). Nucleotides of the coding region are in capitals, whereas those in 5′- and 3′-untranslated regions and in the introns are in lowercase. Encoded amino acids (in single-letter code) are shown below the coding region. The amino acid asparagine (N at position 26 in the amino acid sequence of exon Ia), deduced from the 5′-RACE clones of mouse strain FVB/N, was found to be replaced by serine (S) due to a substitution of adenine (position 176) by guanine in mouse strain 129^{SV}, reflecting a polymorphism between these strains. Sp1 sites in the 5′-flanking regions are underlined. A polyadenylation site aataaa and an instability motif attta in the 3′-region are marked by a dashed line.

moters when they are located close to the transcriptional start point (Pugh and Tjian, 1990) and not in cases where they are further upstream (Hariharan and Perry, 1989). The functional significance of these alternative splicing events is not known. Exon Ia of the *RBTN1* gene is predominantly expressed in the embryonic brain, while exon Ib predominates in the adult (Boehm *et al.*, 1991). Both first exons of *Lhx3* are expressed in the adult pituitary (Zhadanov *et al.*, 1995). The study of differential expression of each exon of *Lhx3* during development is ongoing. Another variation of genomic organization at the 5′-end is demonstrated by the *RBTN2* gene. *RBTN2* contains a unique exon I upstream of the first LIM domain that is encoded by

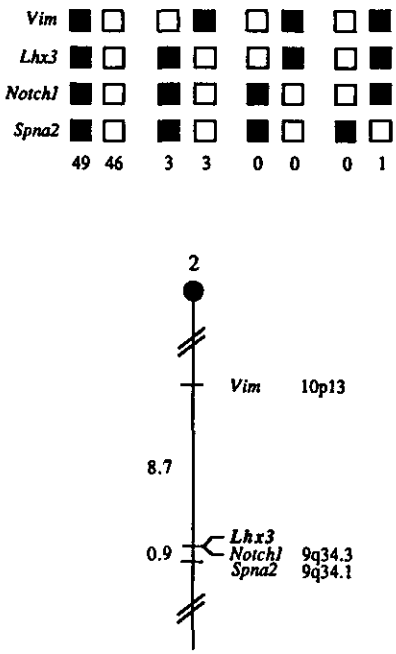

FIG. 3. *Lhx3* maps to the proximal region of mouse chromosome 2. *Lhx3* was placed on mouse chromosome 2 by interspecific backcross analysis. The segregation patterns of *Lhx3* and flanking genes in 102 backcross animals that were typed for all loci are shown at the top of the figure. For individual pairs of loci, more than 102 animals were typed (see text). Each column represents the chromosome identified in the backcross progeny that was inherited from the (C57BL/6J × *M. spretus*)F1 parent. The black boxes represent a C57BL/6J allele, and white boxes a *M. spretus* allele. The number of offspring inheriting each type of chromosome is listed at the bottom of each column. A partial chromosome 2 linkage map showing the location of *Lhx3* in relation to linked genes is shown at the bottom of the figure. Recombination distances between loci in centimorgans are shown to the left of the chromosome, and the positions of loci in human chromosomes, where known, are shown to the right. References for the human map positions of loci cited in this study can be obtained from GDB (Genome Data Base), a computerized database of human linkage information maintained by the William H. Welch Medical Library of the Johns Hopkins University (Baltimore, MD).

the second exon; the second LIM domain is encoded by a single exon (Boehm *et al.,* 1990). We suggest that the rule "one exon–one LIM domain–two finger-like structures" seems to be common for the LIM genes with the exception of the second LIM domains of the *RBTN1* and *RBTN3* genes.

The Intron–Exon Structure of Lhx3 in the Region of the Homeobox, Compared with That of Other Homeobox Genes

In most vertebrate HOX genes the homeobox is encoded by a single exon and an intron is located immediately upstream of the homeobox (Awgulewitsch *et al.,* 1990; Hewitt *et al.,* 1991). A rare intron/exon structure is found in the POU genes *Oct2* and *Pit-1.* Their homeoboxes are interrupted by an intron upstream of helices 2 and 1, respectively (Wirth *et al.,* 1990; Li *et al.,* 1990). An unusual location of an intron was also described for the homeobox of the mouse *HNF I* gene (Bach *et al.,* 1992). The intron in this case is present between the second helix and the 21 amino acid-loop or within the 18 amino acids of the alternative structure loop B. The homeobox of *Lhx3* displays an intron at the beginning of the third helix as also seen in the *Abd B* (Celniker *et al.,* 1989), *evx1* and *evx2* (Bastian and Gruss, 1990), and *lab* (Mlodzic *et al.,* 1988) genes. The position of the intron/exon boundary in the region of the homeobox of LIM-homeobox genes seems to be very conservative since the same location of a homeobox-specific intron was found in the mouse *Gsh-4* gene (Singh *et al.,* 1991) and in the zebrafish homologue of the *Lhx3* gene (Glasgow, unpublished data).

Lhx3 Is Localized in the Proximal Region of Mouse Chromosome 2

We have compared our interspecific map of chromosome 2 with a composite mouse linkage map that reports the map location of many uncloned mouse mutations (compiled by M. T. Davisson, T. H. Roderick, A. L. Hillyard, and D. P. Doolittle and provided from GBASE, a computerized database maintained at The Jackson Laboratory, Bar Harbor, ME). *Lhx3* mapped to a region of the composite map that lacks mouse mutations with a phenotype that might be expected for an alteration in *Lhx3* normally expressed in the pituitary and during prenatal development in areas corresponding to motor neurons and raphe nuclei of the hindbrain (Zhadanov *et al.,* 1995). The proximal region of mouse chromosome 2 shares regions of homology with human chromosomes 9q and 10p (summarized in Fig. 3). In particular, the human homolog of *Notch1* has been placed on human 9q34.3 (Ellisen *et al.,* 1991). The tight linkage between *Lhx3* and *Notch1* in mouse suggests that *Lhx3* will reside on 9q as well. Therefore, there are no data at present to support the role of an involvement of *Lhx3* in human or mouse diseases, but isolation of the mouse gene has enabled the construction of a vector for targeted disruption of *Lhx3* in ES cells. Generation of $Lhx3^{(-/-)}$ mice will allow analysis of physiological functions of this gene.

ACKNOWLEDGMENTS

We thank Mary Barnstead for excellent technical assistance, Dr. Eric Glasgow for unpublished data, and Dr. Masanori Taira for helpful discussions throughout the project. We are also grateful to Dr. Alan Agulnick for critical reading of the manuscript. This research was supported, in part, by the National Cancer Institute, DHHS, under Contract NO1-CO-46000 with ABL.

REFERENCES

Archer, V. E. V., Breton, J., Sanchez-Garcia, I., Osada, H., Forster, A., Thomson, A. J., and Rabbitts, T. H. (1994). Cysteine-rich LIM domains of LIM-homeodomain and LIM-only proteins contain zinc but not iron. *Proc. Natl. Acad. Sci. USA* **91:** 316–320.

Awgulewitsch, A., Bieberich, C., Bogarad, L., Shashikant, C., and Ruddle, F. H. (1990). Structural analysis of the *Hox-3.1* transcription unit and the *Hox-3.2–Hox-3.1* intergenic region. *Proc. Natl. Acad. Sci. USA* **87:** 6428–6432.

Bach, I., Pontoglio, M., and Yaniv, M. (1992). Structure of the gene

encoding hepatocyte nuclear factor 1 (HNF1). *Nucleic Acids Res.* **20:** 4199–4204.

Bastian, H., and Gruss, P. (1990). A murine *even-skipped* homologue, *Evx 1,* is expressed during early embryogenesis and neurogenesis in a biphasic manner. *EMBO J.* **9:** 1839–1852.

Boehm, T., Greenberg, J. M., Buluwela, L., Lavenir, I., Forster, A., and Rabbitts, T. H. (1990). An unusual structure of a putative T cell oncogene which allows production of similar proteins from distinct mRNAs. *EMBO J.* **9:** 857–868.

Boehm, T., Spillantini, M. G., Sofroniev, M. V., Surani, M. A., and Rabbitts, T. H. (1991). Developmentally regulated and tissue specific expression of mRNAs encoding the two alternative forms of the LIM domain oncogene rhombotin: Evidence for the thymus expression. *Oncogene* **6:** 695–703.

Celniker, S. E., Keelan, D. J., and Lewis, E. B. (1989). The molecular genetics of the bithorax complex of *Drosophila:* Characterization of the products of the *Abdominal-B* domain. *Genes Dev.* **3:** 1424–1436.

Copeland, N. G., and Jenkins, N. (1991). Development and applications of a molecular genetic linkage map of the mouse genome. *Trends Genet.* **7:** 113–118.

Del Amo, F. F., Gendron-Maguire, M., Swiatek, P. J., Jenkins, N. A., Copeland, N. G., and Gridley, T. (1993). Cloning, analysis, and chromosomal localization of *Notch-1,* a mouse homolog of Drosophila *Notch. Genomics* **15:** 259–264.

Ellisen, L. W., Bird, J., West, D. C., Soreng, A. L., Reynolds, T. C., Smith, S. D., and Sclar, J. (1991). *TAN1,* the human homolog of the Drosophila *Notch* gene, is broken by chromosomal translocations in T lymphoblastic neoplasms. *Cell* **66:** 649–661.

Feuerstein, R., Wang, X., Song, D., Cooke, N. E., and Liebhaber, S. A. (1994). The LIM/double zinc-finger motif functions as a protein dimerization domain. *Proc. Natl. Acad. Sci. USA* **91:** 10655–10659.

Foroni, L., Boehm, T., White, L., Forster, A., Sherrington, P., Liao, X. B., Brannan, C. I., Jenkins, N. A., Copeland, N. G., and Rabbitts, T. H. (1992). The rhombotin gene family encodes related LIM-domain proteins whose differing expression suggests multiple roles in mouse development. *J. Mol. Biol.* **226:** 747–761.

Freyd, G., Kim, S. K., and Horvitz, H. R. (1990). Novel cysteine-rich motif and homeodomain in the product of the *Caenorhabditis elegans* cell lineage gene *lin11. Nature* **344:** 876–879.

Green, E. L. (1981). Linkage, recombination and mapping. *In* "Genetics and Probability in Animal Breeding Experiments," pp. 77–113, Oxford Univ. Press, New York.

Hariharan, N., and Perry, R. P. (1989). A characterization of the elements comprising the promoter of the mouse ribosomal protein gene *RPS17. Nucleic Acids Res.* **17:** 5323–5337.

Hewitt, J., Clark, L. N., Ivens, A., and Williamson, R. (1991). Structure and sequence of the human homeobox gene HOX7. *Genomics* **11:** 670–678.

Jenkins, N. A., Copeland, N. G., Taylor, B. A., and Lee, B. K. (1982). Organization, distribution, and stability of endogenous ecotropic murine leukemia virus DNA sequences in chromosomes of *Mus musculus. J. Virol.* **43:** 26–36.

Karlsson, O., Thor, S., Norberg, T., Ohlsson, H., and Edlund, T. (1990). Insulin gene enhancer binding protein *Isl-1* is a member of a novel class of proteins containing both a homeo- and a Cys–His domain. *Nature* **344:** 879–882.

Klug, A., and Rhodes, D. (1987). "Zinc fingers": A novel protein motif for nucleic acid recognition. *Cold Spring Harbor Symp. Quant. Biol.* **52:** 473–482.

Letovsky, J., and Dynan, W. S. (1989). Measurement of the binding of transcription factor Sp1 to a single GC box recognition sequence. *Nucleic Acids Res.* **17:** 2639–2653.

Levin, M., and Hoey, T. (1988). Homeobox proteins as sequence specific transcription factors. *Cell* **55:** 537–540.

Li, S., Crenshaw, E. B., III, Rawson, E. J., Simmons, D. M., Swanson, L. W., and Rosenfeld, M. G. (1990). Dwarf locus mutants lacking three pituitary cell types result from mutations in the POU-domain gene *pit-1. Nature* **347:** 528–533.

McGuire, E. A., Hockett, R. D., Pollock, K. M., Bartholdi, M. F., O'Brien, S. J., and Korsmeyer, S. J. (1989). The t(11;14) (p15;q11) in a T-cell acute lymphoblastic leukemia cell line activates multiple transcripts, including *Ttg-1,* a gene encoding a potential zinc finger protein. *Mol. Cell. Biol.* **9:** 2124–2132.

Michelsen, J. W., Scheichel, K. L., Beckerle, M. C., and Winge, D. R. (1993). The LIM motif defines a specific zinc-binding protein domain. *Proc. Natl. Acad. Sci. USA* **90:** 4404–4408.

Mitchell, P. J., and Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* **245:** 371–378.

Mlodzik, M., Fjose, A., and Gehring, W. J. (1988). Molecular structure and spatial expression of a homeobox gene from the *labial* region of the Antennapedia-complex. *EMBO J.* **7:** 2569–2578.

Nagata, T., Kanno, Y., Ozato, K., and Taketo, M. (1994). The mouse *Rxrβ* gene encoding RXR beta: Genomic organization and two mRNA isoforms generated by alternative splicing of transcripts initiated from CpG island promoters. *Gene* **142:** 183–189.

Pugh, B. F., and Tjian R. (1990). Mechanism of transcriptional activation by Sp1: Evidence of coactivators. *Cell* **61:** 1187–1197.

Sachs, A. B. (1993). Messenger RNA degradation in eukaryotes. *Cell* **74:** 413–421.

Sanchez-Garcia, I., Osada, H., Forster, A., and Rabbitts, T. H. (1993). The cysteine-rich LIM domains inhibit DNA binding by the associated homeodomain in *Isl-1. EMBO J.* **12:** 4243–4250.

Sanchez-Garcia, I., and Rabbitts, T. H. (1993). LIM domain proteins in leukemia and development. *Semin. Cancer Biol.* **4:** 349–358.

Sanchez-Garcia, I., and Rabbitts, T. H. (1994). The LIM domain: A new structural motif found in zinc-finger-like proteins. *Trends Genet.* **10:** 315–320.

Schmeichel, K. M., and Beckerle, M. C. (1994). The LIM domain is a modular protein-binding interface. *Cell* **79:** 211–219.

Shapiro, M. B., and Senapathy, P. (1987). RNA splice junctions of different classes of eukaryotes: Sequence statistic and functional implications in gene expression. *Nucleic Acids Res.* **15:** 7155–7174.

Shaw, G., and Kamen, R. (1986). A conserved AU sequence from the 3′ untranslated region of GM-CSF mRNA mediates selective mRNA degradation. *Cell* **46:** 659–667.

Singh, G., Kaur, S., Stock, J. L., Jenkins, N. A., Gilbert, D. J., Copeland, N. G., and Potter, S. S. (1991). Identification of 10 murine homeobox genes. *Proc. Natl. Acad. Sci. USA* **88:** 10706–10710.

Smale, S. T., and Baltimore, D. (1989). The "initiator" as transcriptional control element. *Cell* **57:** 103–113.

Taira, M., Hayes, W. P., Otani, H., and Dawid, I. B. (1993). Expression of LIM class homeobox gene *Xlim-3* in Xenopus development is limited to neural and neuroendocrine tissues. *Dev. Biol.* **159:** 245–256.

Taira, M., Otani, H., Saint-Jeannet, J.-P., and Dawid, I. B. (1994). Role of the LIM class homeodomain protein *Xlim-1* in neural and muscle induction by the Spemann organizer in *Xenopus. Nature* **372:** 677–679.

Way, J. C., and Chalfie, M. (1988). *Mec-3,* a homeobox-containing gene that specifies differentiation of the touch receptor neurons in *C. elegans. Cell* **54:** 5–16.

Wirth, T., Priess, A., Annweiler, A., Zwilling, S., and Oeler, B. (1990). Multiple Oct2 isoforms are generated by alternative splicing. *Nucleic Acids Res.* **19:** 43–51.

Xue, D., Tu, J., and Chalfie, M. (1993). Cooperative interaction between the *Caenorhabditis elegans* homeoproteins *UNC86* and *MEC3. Science* **261:** 1324–1328.

Zhadanov, A. B., Bertuzzi, S., Masanori, T., David, I., and Westphal, H. (1995). Expression pattern of the murine LIM class homeobox gene *Lhx3* in subsets of neural and neuroendocrine tissues. *Dev. Dyn.* **202:** 354–364.