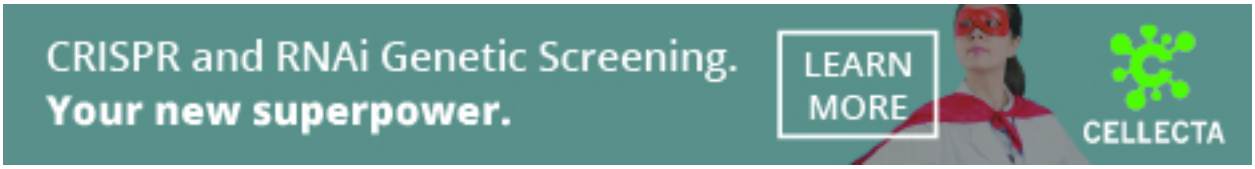
Identification of a conserved family of Meis1-related homeobox genes.

S Steelman, J J Moskow, K Muzynski, et al.

Genome Res. 1997 7: 142-156
Access the most recent version at doi:10.1101/gr.7.2.142



LETTER

Identification of a Conserved Family of *Meis1*-Related Homeobox Genes

Scott Steelman,[1] John J. Moskow,[1] Karen Muzynski,[2] Charlotte North,[1] Teresa Druck,[1] Jeffry C. Montgomery,[1] Kay Huebner,[1] Ira O. Daar,[3] and Arthur M. Buchberg[1,4]

[1]Kimmel Cancer Center, Jefferson Medical College, Philadelphia, Pennsylvania 19107; [2]Intramural Research and Support Program, Science Applications International Corporation (SAIC), and [3]Laboratory of Leukocyte Biology National Cancer Institute–Frederick Cancer Research and Development Center, Frederick, Maryland 21702

The *Meis1* locus was isolated as a common site of viral integration involved in myeloid leukemia in BXH-2 mice. *Meis1* encodes a novel homeobox protein belonging to the TALE (three amino acid loop extension) family of homeodomain-containing proteins. The homeodomain of Meis1 is the only known motif within the entire 390-amino-acid protein. Southern blot analyses using the *Meis1* homeodomain as a probe revealed the existence of a family of *Meis1*-related genes (*Mrgs*) in several diverged species. In addition, the 3′ untranslated region (UTR) of *Meis1* was remarkably conserved in evolution. To gain a further understanding of the role *Meis1* plays in leukemia and development, as well as to identify conserved regions of the protein that might reveal function, we cloned and characterized *Mrgs* from the mouse and human genomes. We report the sequence of *Mrg1* and *MRG2* as well as their chromosomal locations in murine and human genomes. Both *Mrgs* share a high degree of sequence identity with the protein coding region of *Meis1*. We have also cloned the *Xenopus laevis* ortholog of *Meis1* (*XMeis1*). Sequence comparison of the murine and *Xenopus* clones reveals that *Meis1* is highly conserved throughout its coding sequence as well as the 3′ UTR. Finally, comparison of *Meis1* and the closely related *Mrgs* to known homeoproteins suggests that *Meis1* represents a new subfamily of TALE homeobox genes.

[The sequence data described in this paper have been submitted to GenBank under accession nos. U68383 (*Mrg1a*), U68384 (*Mrg1b*), U68385 (*MRG2*), U68386 (*XMeis1–1*), U68387 (*XMeis1–2*), U68388 (*XMeis1–3*), and U68389 (*XMeis1–4*).]

Homeobox genes represent a large diverse family of transcription factors that share a common helix–turn–helix DNA-binding motif designated the homeodomain. Homeoproteins were originally identified in *Drosophila* (McGinnis et al. 1984; Scott and Weiner 1984) and have since been found in the genomes of species as evolutionarily diverse as yeast and humans (Scott et al. 1989; Kappen et al. 1993). The *HOX* genes are perhaps the most well -haracterized category of homeobox genes and have been demonstrated to show remarkable biological specificity during normal development despite the ability of individual *HOX* genes to bind similar DNA consensus sequences (Desplan et al. 1988; Gehring et al. 1994). In addition to crucial roles in normal development, several homeoproteins are involved in neoplasia. Specifically, a subset of homeoproteins, including *HoxA10* (Lawrence et al. 1995), *HoxB8* (Kongsuwan et al. 1989), *PBX1* (Kamps et al. 1990; Nourse et al. 1990), and *Meis1* (Moskow et al. 1995; Nakamura et al. 1996), have been found to play important roles in leukemia.

The *Meis1* locus was identified as a common site of viral integration involved in murine myeloid leukemia (Moskow et al. 1995). Proviral integrations at the *Meis1* locus were found to be clustered in two regions located ~90 kb apart (Moskow et al. 1995; Nakamura et al. 1996). A single gene was found to reside between these two clusters and the expression of the *Meis1* gene was shown to be altered in tumors with viral integrations (Moskow et al. 1995). *Meis1* encodes a 3.8-kb transcript with a predicted open reading frame (ORF) that encodes a homeodomain–containing protein. Proviral integrations in the 3′-untranslated region (UTR) of *Meis1* results in the production of a truncated transcript (Moskow et al.

[4]Corresponding author.
E-MAIL buchberg@lac.jci.tju.edu; FAX (215) 923-4153.

7:142–156 ©1997 by Cold Spring Harbor Laboratory Press ISSN 1054-9803/97 $5.00

1995). Although the precise mechanism by which these proviral integrations lead to leukemia is unclear, perturbation of normal *Meis1* expression may be a factor.

Comparison of the *Meis1* homeodomain to known homeobox genes identifies *Meis1* as having homology with several members of the HAC–ATYP family of homeobox genes (Burglin 1995), as well as the homeoproteins *TGIF* (Bertolino et al. 1995) and *PBX1* (Kamps et al. 1990; Burglin 1995). The HAC–ATYP family includes human, *Arabidopsis thaliana*, and *Caenorhabditis elegans* homeobox genes that share the common feature of an atypical homeodomain structure in which the homeodomain is composed of 63, rather than 60, amino acids (Burglin 1995). To date, the functions of HAC–ATYP members are largely uncharacterized (Burglin 1995). *TGIF* has been demonstrated to recognize the previously characterized retinoid response motif CRBPII–RXRE (Mangelsdorf et al. 1991; Nakshatri and Chambon 1994) and inhibits retinoid X receptor (RXR)-mediated transcriptional activation by competing for the CRBPII–RXRE response element (Bertolino et al. 1995). *PBX1* was originally identified as a fusion protein with *E2A* in a t(1;19) translocation associated with pre-B cell acute lymphoblastic leukemia (pre-B ALL) (Kamps et al. 1990; Nourse et al. 1990). The *E2A–PBX1* fusion has also been shown experimentally to induce the formation of acute myeloid leukemias (AMLs) in BALB/c mice (Kamps and Baltimore 1993).

Southern blot analyses using the *Meis1* homeodomain as a probe suggested the presence of a family of genes in the genomes of diverse species. In this paper we describe the identification and cloning of two such *Meis1*-related genes (*Mrgs*), *Mrg1* and *MRG2* in the murine and human genomes, respectively. Similarly, Southern analysis using the *Meis1* 3′ UTR as a probe revealed that the *Meis1* 3′ UTR is surprisingly well conserved in mammalian species and in *Xenopus laevis*. Therefore, we have also cloned *XMeis1*, the *X. laevis Meis1* ortholog, to identify regions that may represent conserved domains involved in *Meis1* function and/or regulation.

Based on sequence similarity, it was presumed that *Meis1* was a member of the HAC–ATYP homeoprotein family. The identification of *Mrgs* and *Meis1* orthologs and analysis of their homology to the individual HAC–ATYP family members, however, suggests that a reorganization of the HAC–ATYP family is necessary. We therefore propose that *Meis1*, *Mrgs*, and certain HAC–ATYP family members represent a novel family of homeoproteins defined by their homology to the *Meis1* homeodomain.

RESULTS

Evolutionary Conservation Analysis of the *Meis1* Gene

Previous data from our laboratory demonstrated that several proviral integrations occur in the *Meis1* 3′ UTR, suggesting that this region has an important function in *Meis1* regulation (Moskow et al. 1995). To further address the significance of the *Meis1* 3′ UTR, probe PBc1 was hybridized to genomic DNA from diverse species to determine whether the *Meis1* 3′ UTR was conserved, as evolutionary conservation is a characteristic indicative of functional domains. Southern blot analysis using probe PBc1 detected a restriction fragment in *Xenopus*, mouse, guinea pig, rat, dog, monkey, and human genomic DNAs (Fig. 1). There was no detectable hybridization of the PBc1 probe to yeast, *Drosophila*, zebra fish, or chicken genomic DNA. These results indicate that the *Meis1* 3′ UTR is well conserved throughout evolution and suggests an important role in *Meis1* expression.

Additionally, conservation of the *Meis1* homeodomain was assessed, as the homeodomain represents the only known functional motif within the predicted *Meis1* ORF (Moskow et al. 1995). Under relatively high stringency wash

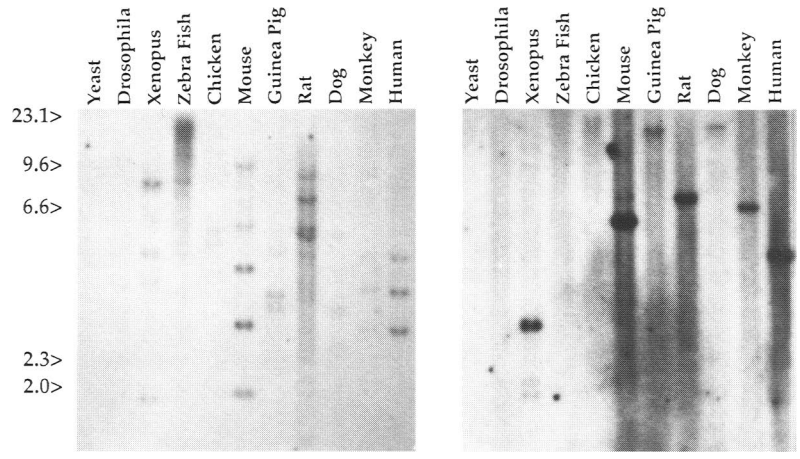


Figure 1 Analysis of conservation of the *Meis1* homeodomain (probe P32-33; *left*) and 3′ UTR (probe PBc1; *right*). Five micrograms of genomic DNA from each of the species indicated above was digested with *Eco*RI. The same Southern blot was hybridized to both probes. Molecular mass markers are given in kilobases at *left*.

conditions (i.e., 65°C; 0.2× SSCP, 0.1%SDS), the *Meis1* homeodomain probe hybridized to all genomic DNAs studied, from yeast to humans (Fig. 1). Furthermore, the *Meis1* homeodomain probe detected at least six restriction fragments in the mouse genome and multiple restriction fragments in the other species examined (Fig. 1). This finding indicates that the *Meis1* homeodomain is well conserved and suggests that *Meis1* may represent a family of *Meis1*-related genes (*Mrgs*) that share a highly similar homeodomain.

Interestingly, although the *Meis1* homeodomain-specific probe detects multiple restriction fragments in all species examined, the *Meis1* 3′ UTR specific probe hybridized to a single restriction fragment in mammalian genomes. This finding suggests that the *Meis1* 3′ UTR is unique in these species and supports the hypothesis that the additional restriction fragments detected using the *Meis1* homeodomain probe represent *Mrgs*.

Identification and Cloning of *Mrg1*, a Murine *Meis1*-Related Gene

To isolate *Mrgs* identified by Southern blot analysis, probe P32-33, which represents the *Meis1* homeodomain, was used to screen an adult mouse kidney cDNA library. Filters were counterscreened with the *Meis1* 3′ UTR probe (PBc1) to eliminate *Meis1*-specific clones, increasing the likelihood that the remaining positive clones were indeed *Mrgs*. The first clone identified, K-41, was 1368 bp in length, 412 bp of which was 3′ UTR. Clone K-41 was found to contain a predicted ORF 92% similar (82% identical) to *Meis1* at the amino acid level (data not shown). A comparison of the *Meis1* and K-41 3′ UTRs, however, revealed only 37% identity at the nucleotide level (data not shown), indicating that clone K-41 was a *Mrg*. Furthermore, no significant homologies to the K-41 3′ UTR nucleotide sequence were detected in the GenEMBL (release 94.0) database, suggesting that clone K-41 represented a novel sequence. The K-41 clone was therefore designated *Mrg1*, for *Meis1*-related gene 1

Cloning of *Mrg1* Alternative Transcripts

Previous studies demonstrated that two isoforms of the *Meis1* gene exist (Moskow et al. 1995) and are defined by the presence (*Meis1a*) or absence (*Meis1b*) of a 95-bp exon (J.J. Moskow and A.M. Buchberg, unpubl.). Inclusion of this exon in the *Meis1a* isoform leads to a shortened ORF with a distinct carboxyl terminus (Fig. 2) due to an in-frame stop codon (Moskow et al. 1995). Alternative mRNA processing resulting in the removal of this exon gives rise to the *Meis1b* isoform in which the ORF contains an additional 75 amino acids (Fig. 2) (Moskow et al. 1995). Sequence analysis indicated that the *Mrg1* (K-41) sequence isolated was 76.3% and 83.2% identical at the nucleotide (data not shown) and amino acid level, respectively, to the *Meis1b* isoform (Fig. 2), and has subsequently been determined to represent the *Mrg1b* isoform.

Northern blot analysis indicated that *Mrg1* encodes a 3.5- to 4.0-kb mRNA; thus, K-41 represented a partial *Mrg1* transcript (see below). To clone addi-

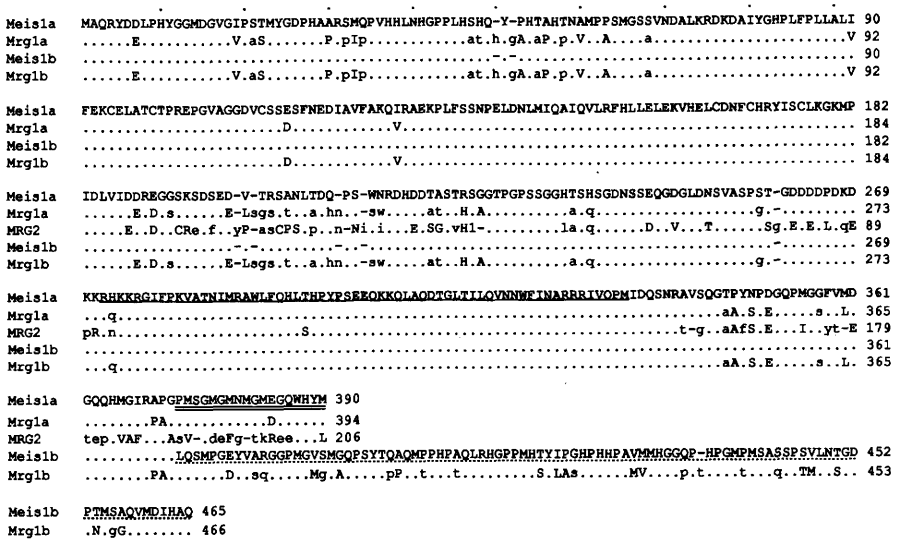


Figure 2 Sequence alignment of *Meis1*, *Mrg1*, and *MRG2*. The amino acid sequences of Meis1a, Mrg1a, MRG2, Meis1b, and Mrg1b proteins are shown with the homeodomain motif underlined. Identical amino acids are denoted by dots. Similar amino acid residues are given in uppercase. Gaps within the ORF are indicated by dashes. Dissimilar amino acids are given in lowercase. The carboxy-terminal end of the Meis1a, Mrg1a, and MRG2 proteins resulting from the inclusion of the 95 bp alternatively processed exon is double underlined. The carboxy-terminal end of Meis1b and Mrg1b proteins resulting from the splicing out of the 95-bp exon is highlighted with a dotted underline.

tional *Mrg1* sequence, an adult brain cDNA library was screened with probes M8-9 and M2-5. A single 2642-bp cDNA clone that hybridized to both probes was identified (cl8-9), sequenced, and determined to contain a predicted ORF of 394 amino acids, including a consensus ATG start site (Kozak 1987).

Sequence analysis revealed that cl8-9 contained a 95-bp insertion not found in *Mrg1b* and encoded an ORF 86.9% and 78.5% identical at the nucleotide (data not shown) and amino acid levels, respectively, to *Meis1a*, with a 3′ UTR 100% identical to *Mrg1b*. Therefore, cl8-9 was determined to represent the *Mrg1a* isoform (Fig. 2). Interestingly, although there are differences between *Meis1a* and *Mrg1a* at both the nucleotide and amino acid levels within the 95-bp alternatively processed exon (Fig. 2), the location and size of the alternatively spliced exon within the cDNAs is identical, suggesting a similar genomic structure despite localization to two different chromosomes (see below).

A comparison of the *Meis1* and *Mrg1* homeodomains at the amino acid level shows that they are 98.4% identical, differing by only 1 of 63 amino acids (Fig. 2), whereas the respective 3′ UTRs are only 37% identical (data not shown). These data are consistent with the Southern blot analyses (Fig. 1), which suggested the presence of a family of genes in mammalian genomes containing similar homeodomains, but divergent 3′ UTRs.

Chromosomal Localization of Murine *Mrg1*

The identification of the *Mrg1a* isoform suggests that *Meis1* and *Mrg1* share a similar genomic organization, perhaps because of an evolutionary duplication of the *Meis1* locus. This duplication would presumably lead to a clustering of *Mrgs* around the *Meis1* locus on mouse chromosome 11 (Moskow et al. 1995). To determine whether the *Mrg1* locus was near the *Meis1* locus, the *Mrg1* 3′ UTR probe M2-5 was hybridized to a panel of parental AEJ/Gn and *Mus spretus* genomic DNAs digested with various restriction endonucleases to identify an informative polymorphism (Table 1). Probe M2-5 was then hybridized to an interspecific backcross mapping panel, and a comparison of the segregation pattern of the *M. spretus* and AEJ/Gn *Mrg1* alleles with the patterns of known genes and microsatellite markers (L.D. Siracusa and A.M. Buchberg, unpubl.) revealed that *Mrg1* resides on central mouse chromosome 2 (Fig. 3). The order of the genes typed in the cross and the ratio of recombinants to N_2 mice examined are as follows: centromere–*D2Mit42*–2/93–*Mrg1*–1/93–*D2Mit63*–telomere. The genetic distances given in centiMorgans (±S.E.) are as follows: centromere–*D2Mit42*–2.2 ± 1.5–*Mrg1*–1.1 ± 1.0–*D2Mit63*–telomere.

Mapping the *MRG1* Locus in Humans

Segregation analysis on a panel of rodent–human somatic cell hybrids was used to localize the human homolog of *Mrg1* to determine whether human *MRG1* might be associated with any known cytogenetic aberrations. *Mrg1* was shown to map to a region of mouse chromosome 2 that is syntenic to human chromosome 7, 11, or 15 (Siracusa et al. 1996). By scanning the GenEMBL STS database (release 94.0) with the *Mrg1* cDNA sequence, the *D15S555* (WI-1222) amplimer was identified as a portion of human *MRG1*, as it was 94% identical to a 96-bp stretch of mouse *Mrg1* at the nucleotide level (100% identical across 32-amino acid residues). *D15S555* has been mapped to human chromosome 15 but has not been regionally localized. Two of three somatic cell hybrids retaining portions of human chromosome 15 were positive for the *MRG1* locus based on PCR amplification using the WI-1222 oligonucleotides, indicating that *MRG1* localizes to 15q22–q25 (Fig. 3).

Like human *MEIS1*, which maps to human chromosome 2p13–14 near three translocation breakpoints involved in human leukemia (Moskow et al. 1995), human *MRG1* also maps to a region of human chromosome 15 associated with various cytogenetic abnormalities associated with AML, chronic myeloid leukemia (CML), and human astrocytomas (Le Beau et al. 1986; Ishihara and Minamihisamatsu 1988; Ransom et al. 1992; Griffin et al. 1992; Willman et al. 1993).

Expression of *Mrg1* Is Distinct from *Meis1*

Previous Northern blot data has shown that *Meis1* expression is ubiquitous (J.J. Moskow and A.M. Buchberg, unpubl.). To determine whether *Mrg1*, like *Meis1*, was expressed ubiquitously, Northern blotting was performed. Hybridization of the *Mrg1* 3′ UTR probe M2-5 to total RNA from various adult mouse tissues reveals that *Mrg1* expression is represented as two distinct transcripts, 4.0 and 3.5 kb in size (Fig. 4) in contrast to the single 3.8-kb *Meis1* transcript (data not shown). Furthermore, as shown in Figure 4, *Mrg1* expression is more restricted than *Meis1* expression. *Mrg1* expression is most abundant in adult brain, is present to lesser degrees in most other tissues examined, including the myelomono-

Table 1. Probes and RFLPs Used for Mapping

Locus	Probe or Primers	Gene Name	Enzyme	Size (kb) AEJ/Gn	Size (kb) *Mus spretus*[a]	Reference
Mrg1	M2-5	*Meis1*-related gene 1	*Eco*RI	11.22	8.9	
Mrg2-rs1	H3-4	*Meis1*-related gene 2-related sequence-1	*Sac*I	4.47 3.31	6.31 3.16	
Tpi-rs3	pHTPI-5A	Triosephosphate isomerase, related sequence 3	*Hind*III	19.95	18.62	Brilliant et al. 1996
D2Mit42	ATTACTGGGCAGGAACATTTG GCCAAACTTCCAGACTCCTC	DNA segment, Chr 2, MIT-42	*b*	0.136	0.112	Dietrich et al. 1993
D2Mit63	GCAGTCTACCAGGAGCAACC TGGATGTAGGCATGTGCCT	DNA segment, Chr 2, MIT-63	*b*	0.212	0.204	Goldstein et al. 1994 Dietrich et al. 1993
D7Mit21	GGGTTGAACCTTACAGGGGT ATCAAACCAGCCCAAGTGAC	DNA segment, Chr 7, MIT-21	*b*	0.128	0.166	Dietrich et al. 1993
D7Mit27	TGAACTGGGGAGGAAAGTTG AACATGAAAAGACATTCCCCC	DNA segment, Chr 7, MIT-27	*b*	0.248	0.226	Dietrich et al. 1993

[a]The underlined restriction fragments identify the segregating *M. spretus* alleles followed in the N_2 progeny.
[b]These loci were typed by PCR.

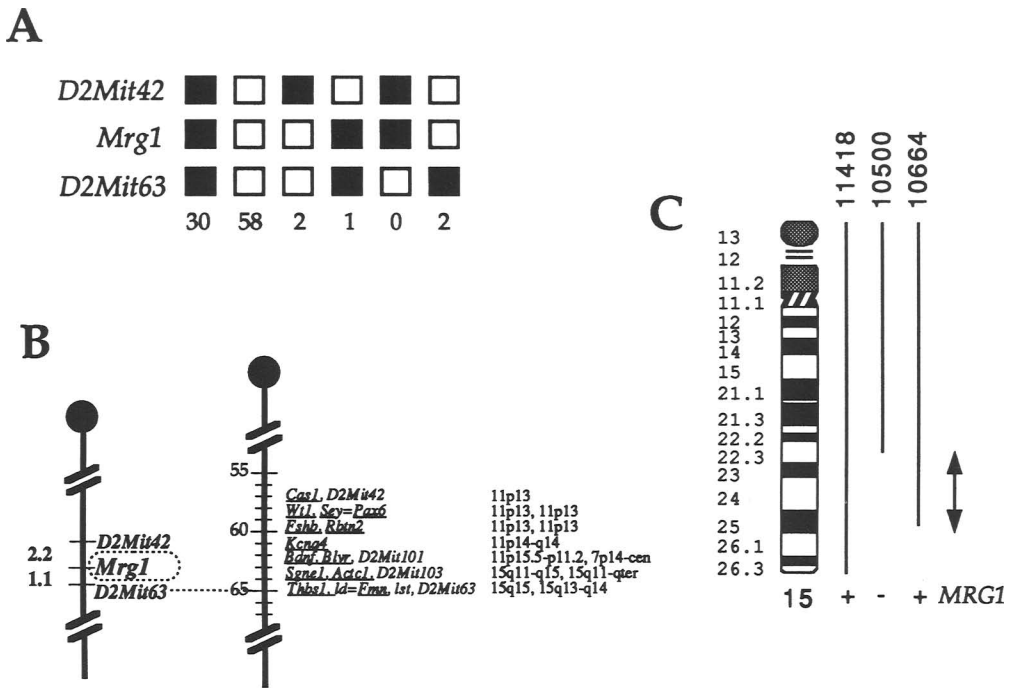


Figure 3 Chromosomal localization of *Mrg1* in the murine genome. (*A*) Haplotype analysis of 93 N_2 offspring from the (AEJ/Gn × *M. spretus*)F_1 × AEJ/Gn cross. (■) The AEJ/Gn allele; (□) the *M. spretus* allele. Loci typed are to the *left*. The number of N_2 progeny carrying each allele is given at the *bottom*. (*B*) Linkage map of the *Mrg1* locus. The *left* chromosome shows the loci typed in the backcross (Table 1), with distances between the markers used given in centiMorgans. The *right* chromosome illustrates a partial consensus linkage map of mouse chromosome 2 (Siracusa et al. 1996). The maps were aligned at *D2Mit63* (dotted line). Loci mapped in humans are underlined; syntenic regions are indicated to the *right*. (*C*) A diagram of human chromosome 15 is shown at *left*. The portion of chromosome 15 present in three of the somatic cell hybrids are shown along with a score of positive or negative for *MRG1* as determined by PCR. These results indicate that *MRG1* maps to human chromosome 15q22–25.

cytic cell line WEHI3B, but is not detectable in testes. Additionally, *Mrg1* expression is undetectable via Northern analysis in the human promyelocytic cell line HL-60 and murine 32Dcl3 cells (data not shown). The precise difference between the two *Mrg1* transcripts has not yet been determined, although alternative use of polyadenylation signals, alternative splicing, or cross-hybridization to a related 3′ UTR are all potential mechanisms to explain the 500-bp size difference.

Cloning of Human *Meis1*-Related Genes

Southern blot analysis indicated that *Meis1* is well conserved between mouse and human within the homeodomain and 3′ UTR (Fig. 1). Additional regions of evolutionary conservation within the ORF, however, could not be determined by Southern analysis alone. Determining additional specific regions of conservation that may indicate unique motifs critical for *Meis1* function and/or regulation required cloning of an evolutionarily distant *Meis1* homolog. Screening of a human fibroblast cDNA library with a pool of *Meis1* probes (P18-19 and PBc1) led to the identification of a single clone. Se-

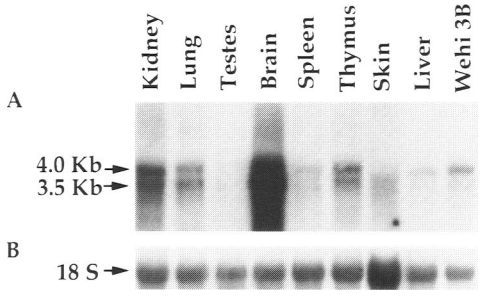


Figure 4 Expression of *Mrg1* in adult mouse tissues. (*A*) *Mrg1* expression is identified by two transcripts, 4.0 and 3.5 kb in size (arrows at *left*) and is detected in all tissues studied, except testes. (*B*) Relative levels of total RNA, as determined by hybridization to a probe for 18S RNA (American Type Culture Collection, Rockville, MD).

quence analysis identified this clone as being 82% similar to both *Meis1* and *Mrg1* within the ORF (Fig. 2), but there is only 37% identity between the respective 3′ UTRs (data not shown). Therefore, we have designated this gene *MRG2*, for *Meis1*–related gene 2, as it appears to represent a related gene that is distinct from both *Meis1* and *Mrg1*.

Chromosomal Localization of *MRG2* in Human and Murine Genomes

Segregation analysis of a panel of rodent–human somatic cell hybrids digested with *Eco*RI was performed using a probe representing the 3′ UTR of *MRG2* (HF–UTR). The *MRG2* 3′ UTR probe, however, was found to hybridize to three restriction fragments in human genomic DNA. The two most intense restriction fragments detected by Southern analysis were localized to human chromosome 17 (data not shown). The assignment of *MRG2* to human chromosome 17 was confirmed by PCR analysis (data not shown). Interestingly, the third restriction fragment was localized to human chromosome 19 (data not shown), suggesting the existence of another human *MRG* with sequence similarity to the *MRG2* 3′ UTR.

Having observed that the *Meis1* and *Mrg1* loci were dispersed in the murine genome, it was of interest to localize murine *MRG2* as well. Hybridization of the *MRG2* 3′ UTR probe (HF–UTR) to murine genomic DNA failed to detect any restriction fragments by Southern analysis; therefore, a probe representing the entire *MRG2* cDNA (H3-4) was utilized. The H3-4 probe, however, contains ORF sequence, including the homeodomain, and was found to hybridize to two cosegregating *M. spretus* restriction fragments (Table 1). Because it could not be determined definitively whether the locus mapped with the H3-4 probe was the murine homolog of *MRG2*, we have designated this locus *Mrg2-rs1*, for *Meis1*-related gene 2-related sequence 1. The chromosomal location of *Mrg2-rs1* was determined in the same manner as described above for *Mrg1* (Table 1) and found to reside on proximal mouse chromosome 7 (Fig. 5) in a region syntenic with human chromosome 19q13 (Brilliant et al. 1996). The lack of linkage synteny between proximal mouse chromomsome 7 and human chromosome 17, where *MRG2* maps, suggests that the *Mrg2-rs1* locus is not the murine *MRG2* locus but, rather, a *MRG2*-related locus. The order of the genes typed in the cross and the ratio of recombinants to N_2 mice examined are as follows: centromere–*D7Mit21*–1/80–*Mrg2-rs1*–0/80–*Tpi-rs3*–12/80–*D7Mit27*–telomere. The genetic distances given in centiMorgans (±S.E.) are as follows: centromere–*D7Mit21*–1.3 ± 1.2–(*Mrg2-rs1*, *Tpi-rs3*)–15.0 ± 11.1–*D7Mit27*–telomere.

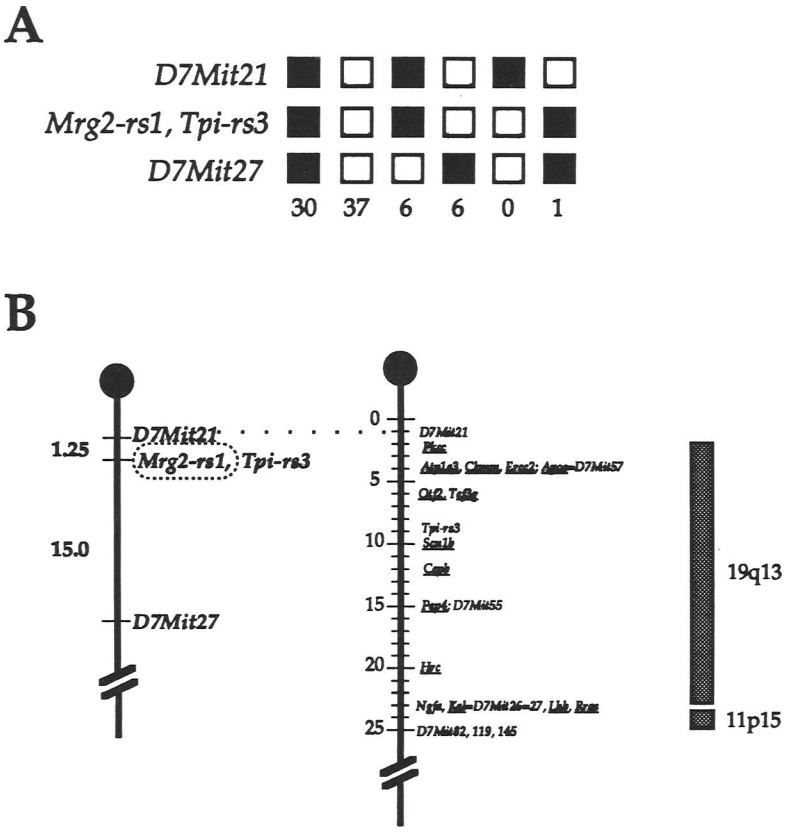


Figure 5 Chromosomal localization of murine *Mrg2-rs1*. (*A*) Haplotype analysis of 93 N_2 offspring from the (AEJ/Gn × *M. spretus*)F_1 × AEJ/Gn cross. (■) The AEJ/Gn allele; (□) the *M. spretus* allele. Loci typed are to the *left*. The number of N_2 progeny carrying each allele is given at the *bottom*. (*B*) Linkage map of the *Mrg2-rs1* locus. The *left* chromosome shows the loci typed in the backcross (Table 1), with distances between the markers used given in centiMorgans. The *right* chromosome illustrates a partial consensus linkage map of mouse chromosome 7 (Brilliant et al. 1996). The maps were aligned at *D7Mit21* (dotted line). Loci mapped in humans are underlined; syntenic regions are indicated at *right*.

Cloning of the *Xenopus laevis* Homolog of *Meis1* and Related Genes

To further identify conserved regions of *Meis1* that may define additional functional domains, it was necessary to identify *Meis1* in distant species. Of the 11 species examined by Southern blot, *Xenopus* was the most distant species in which hybridization to the *Meis1* 3′ UTR probe was detected (Fig. 1), suggesting the presence of a *Meis1* ortholog in the *Xenopus* genome. Using probe P8-11, which represents the *Meis1a* 3′ UTR, screening of embryonic *Xenopus* libraries led to the identification of four cDNA clones, two of which contain the entire coding region of the *Xenopus Meis1a* and *Meis1b* homologs (*XMeis1-1* and *XMeis1-2*, respectively). Sequence alignment at the amino acid level indicated that *XMeis1-1* and *XMeis1-2* are 95% and 90% identical to *Meis1a* and *Meis1b*, respectively (Fig. 6). The extensive evolutionary conservation throughout the entire *Meis1* ORF is surprising and suggests a fundamental role for the Meis1 protein in normal development, even in evolutionarily distant species.

In addition to the extensive evolutionary conservation observed throughout the entire *Meis1* ORF, the cloned portions of the 3′ UTRs of *XMeis1-1* and *XMeis1-2* are ~80% and 66% identical at the nucleic acid level, respectively, to the *Meis1* 3′ UTR, with distinct stretches of nucleotide identity (Fig. 7). The striking homology within the 3′ UTRs may indicate an evolutionarily conserved regulatory function for this region.

DISCUSSION

As reported here, a number of cDNA clones homologous to *Meis1* have been identified. Among these clones are the *Meis1*-related genes *Mrg1* and *MRG2*. Additionally, the *Xenopus Meis1* homologs *XMeis1-1* and *XMeis1-2* have been cloned and show surprisingly high sequence identity at the nucleotide and amino acid level to *Meis1* within both the ORF and 3′ UTR. In fact, sequence analysis reveals that *XMeis1* is more similar to *Meis1* than either *Mrg1* or *MRG2*, despite evolutionary distance. By cloning *Mrgs* and *Meis1* orthologs, conserved regions could be used to determine functional motifs that are essential to *Meis1* function and/or regulation and may identify previously uncharacterized motifs that distinguish *Mrgs* from other homeobox genes.

The first *Mrg* identified, *Mrg1*, is a novel murine homeobox gene that is highly homologous to *Meis1* within the predicted ORF. Two isoforms, *Mrg1a* and *Mrg1b*, have been cloned and determined to be 83% and 87% identical (at the amino acid level), respectively, to *Meis1a* and *Meis1b*. The identification of alternatively spliced forms of *Mrg1* furthers the similarities between *Meis1* and *Mrg1*, because despite differences at the nucleotide level, the location and size of the splice within the cDNAs of both genes is identical, suggesting a similar genomic organization even though the loci are dispersed in the murine genome (Moskow et al. 1995; Fig. 3).

The second *Mrg* identified, *MRG2*, was isolated from the human genome, and translation of the largest ORF in the *MRG2* cDNA shows 66.5% identity to *Meis1* across 209 amino acids. The identification of a less-conserved *Mrg* suggests that a spectrum of *Mrgs* may exist. Interestingly, within this spectrum, ORF variation is contrasted by preservation of the homeodomain motif (Figs. 8 and 9).

More strikingly, the *Xenopus*

```
Meis1a      MAQRYDDLPHYGGMDGVGIPSTMYGDPHAARSMQPVHHLNHGPPLHSHQYPHTAHTNAMPPSMGSSVNDALKRDKDAIYG 80
Xmeis1-1    ..................L..S...........ks...........p.....S.....................s... 80
Meis1b      ..................L..S...........s...........p.....S.....................s... 80
Xmeis1-2    ..................L..S...........s...........p.....S.....................s... 80

Meis1a      HPLFPLLALIFEKCELATCTPREPGVAGGDVCSSESFNEDIAVFAKQIRAEKPLFSSNPELDNLMIQAIQVLRFHLLELE 160
Xmeis1-1    ..........................................s..................................... 160
Meis1b      ..........................................s..................................... 160
Xmeis1-2    ..........L.................................s..................................... 160

Meis1a      KVHELCDNFCHRYISCLKGKMPIDLVIDDREGGSKSDSEDVTRSANLTDQPSWNRDHDDTASTRSGGTPGPSSGGHTSHS 240
Xmeis1-1    ................E............D.........L....p.......S.....a..i..................... 240
Meis1b      ..........................................s..................................... 240
Xmeis1-2    ..............................D.........L....p.......S.....a..i..................... 240

Meis1a      GDNSSEQGDGLDNSVASPSTGDDDDPDKDKKRHKKRGIFPKVATNIMRAWLFQHLTHPYPSEEQKKQLAQDTGLTILQVN 320
Xmeis1-1    ..............I.............E...n.g.......................................... 320
Meis1b      ..........................................s..................................... 320
Xmeis1-2    ..............I.............E.......................................... 320

Meis1a      NWFINARRRIVQPMIDQSNRAVSQGTPYNPDGQPMGGFVMDGQQHMGIRAPGPMSGMGMNMGMEGQWHYM 390
XMeis1-1    ..................................I.................................... 390
Meis1b      ..................................................IQSMPGEYVARGGPMGVSMGQPSYTQAQ 400
XMeis1-2    ..................................I.....................g...s.Isps....M..A.....Ts. 400

Meis1b      MPPHPAQLRHGPPMHTYIPGHPHHPAVMMHGGQPHPGMPMSASSPSVLNTGDPTMSAQVMDIHAQ 465
XMeis1-2    ..L.h......tsV.......h....M.....p.q....I.............S..GhX.X.... 465
```

Figure 6 Sequence alignment of the Meis1a, Meis1b, XMeis1-1, and XMeis1-2 proteins with the homeodomain motif underlined. Identical amino acids are denoted by dots. Similar amino acid residues are given in uppercase. Gaps within the ORF are indicated by dashes. Dissimilar amino acids are given in lowercase. The carboxy-terminal end of the Meis1a and XMeis1-1 proteins resulting from the inclusion of the 95-bp alternatively processed exon is double underlined. The carboxy-terminal end of Meis1b and XMeis1-2 proteins resulting from the splicing out of the 95-bp exon is highlighted with a dotted underline. (*X*) Amino acid residues that cannot be determined given the available sequence.

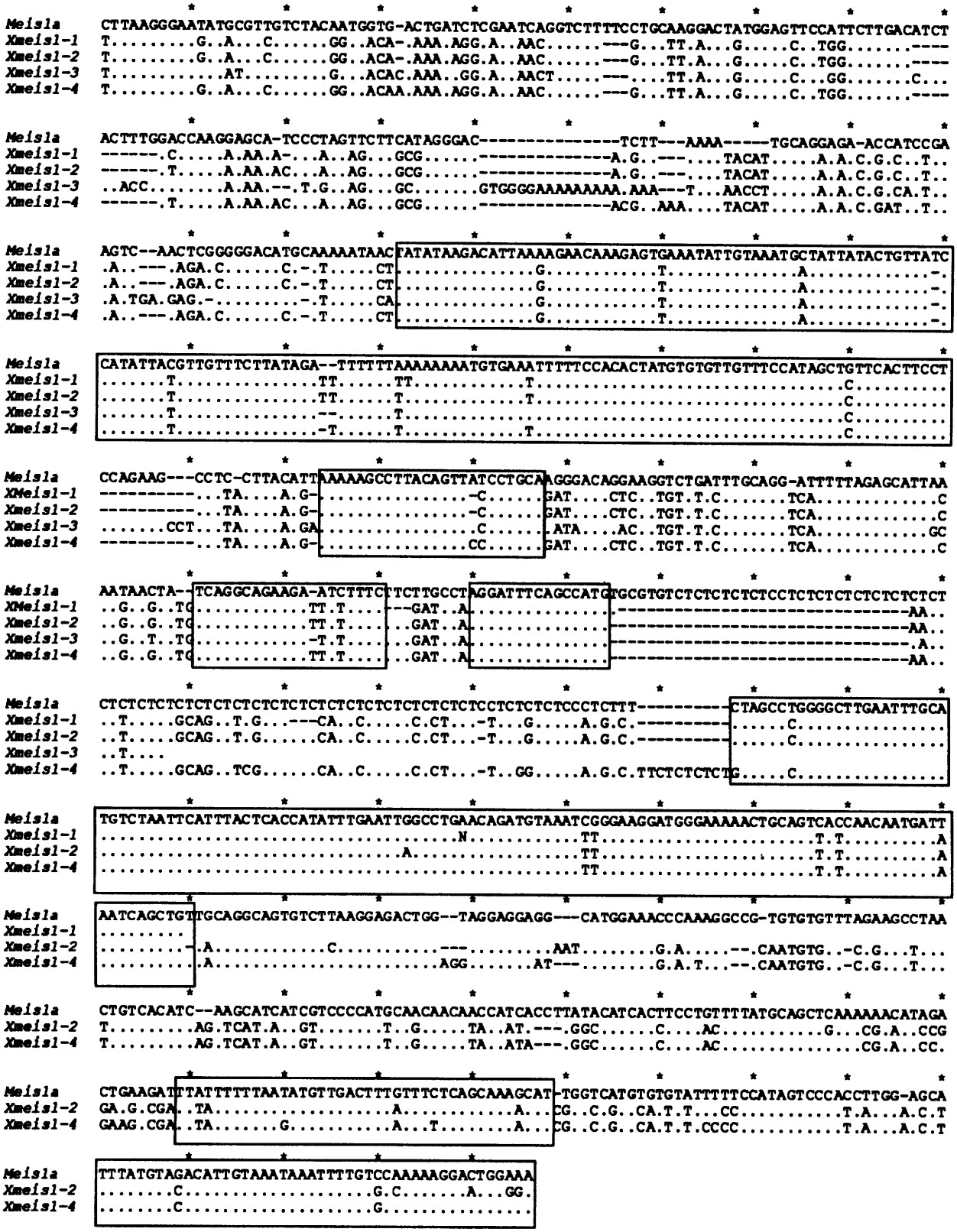
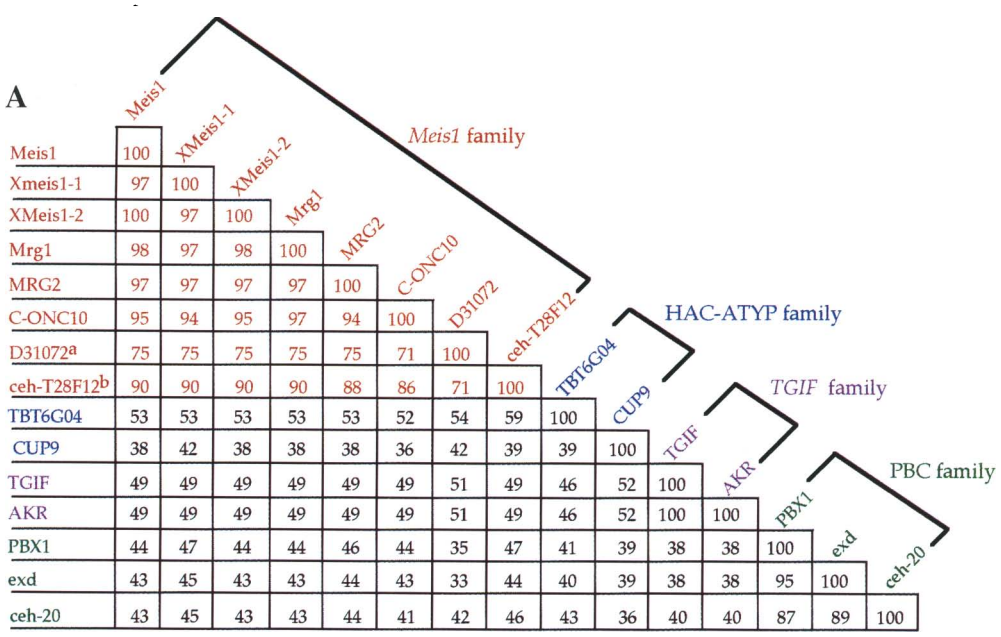


Figure 7 Sequence alignment of *Meis1* and *XMeis1* 3′ UTRs. The 3′ UTRs of *Meis1a*, *XMeis1-1*, *XMeis1-2*, *XMeis1-3*, and *XMeis1-4* are shown. Identical nucleotides are denoted by dots. Gaps within the sequence alignment are denoted by dashes. Regions of conservation are boxed.

A

	Meis1	XMeis1-1	XMeis1-2	Mrg1	MRG2	C-ONC10	D31072	ceh-T28F12	TBT6G04	CUP9	TGIF	AKR	PBX1	exd	ceh-20
Meis1	100														
Xmeis1-1	97	100													
XMeis1-2	100	97	100												
Mrg1	98	97	98	100											
MRG2	97	97	97	97	100										
C-ONC10	95	94	95	97	94	100									
D31072[a]	75	75	75	75	75	71	100								
ceh-T28F12[b]	90	90	90	90	88	86	71	100							
TBT6G04	53	53	53	53	53	52	54	59	100						
CUP9	38	42	38	38	38	36	42	39	39	100					
TGIF	49	49	49	49	49	49	51	49	46	52	100				
AKR	49	49	49	49	49	49	51	49	46	52	100	100			
PBX1	44	47	44	44	46	44	35	47	41	39	38	38	100		
exd	43	45	43	43	44	43	33	44	40	39	38	38	95	100	
ceh-20	43	45	43	43	44	41	42	46	43	36	40	40	87	89	100



B

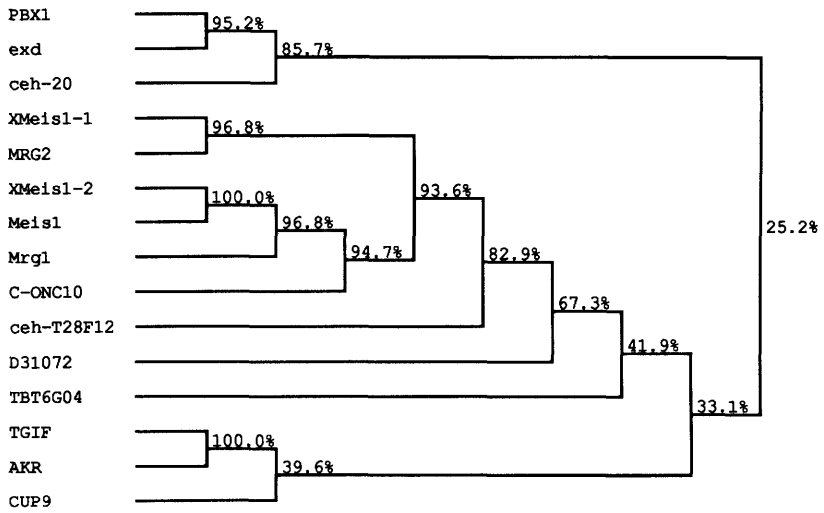


Figure 8 (*A*) Matrix showing the percent identity among the homeodomains of various homeobox genes. The *Meis1* family (red) includes *Meis1*, *XMeis1-1*, *XMeis1-2*, *Mrg1*, *MRG2*, *c-onc10*, *D31072*, and *ceh–T28F12*. The HAC–ATYP family (blue) includes *TBT6G04* and *CUP9* (Burglin 1995). The *TGIF* family (purple) includes *TGIF* and *AKR*. The PBC family (green) includes *PBX1*, *exd*, and *ceh-20* (Burglin 1995). [a]Percent identities are based on the 55-amino-acid residues of the homeodomain currently available. [b]Percent identities are based on the 59-amino-acid residues of the homeodomain currently available. (*B*) Phylogenetic tree generated using the MacDNASIS PRO V3.6 program. Percent amino acid similarity between individual or among multiple homeodomains is given at each of the branch points.

Meis1 orthologs *XMeis1-1* and *Xmeis1-2*, are 95% and 91% identical across the entire 390-amino-acid ORF to *Meis1a* and *Meis1b*, respectively, and are 96.8% identical within the homeodomain. Surprisingly, the *XMeis1-1* and *XMeis1-2* 3′ UTRs have been determined to be 66% and 80% identical at the nucleotide level, respectively, to the *Meis1a* and *Meis1b* 3′ UTRs. Sequence alignment of the respective 3′ UTRs has also identified regions of conservation (Fig. 7), suggesting that these regions may play a crucial role in *Meis1* function and/or regulation.

The extensive conservation among *Mrgs*,

```
                          Helix 1        TALE      Helix 2            Helix 3
              1        10        20 ***      30        40        50        60
Meis1         RHKKRGIFPKVATNIMRAWLFQHLTHPYPSEEQKKQLAQDTGLTILQVNNWFINARRRIVQPM
Xmeis1-1      .n.g...........................................................
Xmeis1-2      ...............................................................
Mrg1          .q.............................................................
MRG2          .n......................S......................................
C-ONC10       .q......................................................X.....X
D31072*       --------..h...V..s.....Ig....T.De...I.aq.n..L..............L...
ceh-T28F12*   ----......q...rL......n...............kE.......................

TBT6G04      AwRpq..-L.Er.vSVL.....e.Fl....kDsd.Hm..kq....rs..S.......v.LWk..
CUP9          nsgR.snL..etvQ.Lnt..LN..Nn...Tqqe.Re.LiK.n..kI.LS....aV...kIfsd

TGIF          KRRR..nL..esvQ.L.d..ye.rYnA....qe.al.s.q.h.St...C.........LLpd.
AKR           KRRR..nL..esvQ.L.d..ye.rYnA....qe.al.s.q.h.St...C.........LLpd.

PBX1          aRR..rn.n.q..e.LneyFyS..Sn......a.Ee..kKC.I.Vs..S...G.k.i.ykknI
exd           aRR..rn.S.q.Se.LneyFyS..Sn......a.Ee..rKC.I.Vs..S...G.k.i.ykknI
ceh-20        aRR..rn.S.q..eVLneyFyg..Sn......a.Ed..rqCnI.Vs..S...G.k.i.ykkn.

* Complete homeodomain sequence is not available
```

Figure 9 Sequence alignment of the various homeodomains to the *Meis1* homeodomain. Identical residues are denoted by dots. Similar amino acid residues are shown in uppercase. Gaps within the homeodomain are denoted by dashes. Dissimilar residues are shown in lowercase. Residues comprising the TALE domain are indicated by three asterisks (***). (*X*) Amino acid residues that cannot be determined given the available sequence.

XMeis1, and *Meis1* outside the homeodomain suggests a fundamental role for the Meis1 and Mrg proteins in normal growth and development and is unusual for homeoproteins in general, as homology is usually limited to small motifs, such as POU and paired-box domains (Bopp et al. 1986; Herr et al. 1988). The only other homeoproteins identified thus far with extensive ORF identity are the *PBX* genes, which are 92%–94% identical across a 266-amino-acid region, including the homeodomain (Monica et al. 1991).

The *PBX* genes are a multigene homeobox family consisting of *PBX1*, *PBX2*, and *PBX3* (Monica et al. 1991). The founding member of the family, *PBX1*, has been shown to be involved in pre-B ALL (Kamps et al. 1990; Nourse et al. 1990). The *PBX* genes have been classified as members of the PBC homeobox family, which includes *Drosophila extradenticle* (*exd*) and *C. elegans ceh-20* (Burglin 1995). The PBC family, among others, belongs to the TALE (three amino acid loop extension) superfamily of homeoproteins (Burglin 1995). All members of the TALE superfamily share the common characteristic of a 3-amino-acid insertion between helix 1 and helix 2 of the helix–turn–helix homeodomain motif. Interestingly, the *PBX* genes share a homeodomain that is 97% identical at the amino acid level (Monica et al. 1991) and is 44% identical at the amino acid level to the *Meis1* homeodomain (Fig. 8). Previous reports have identified the *Meis1* homeodomain as being most homologous with the *PBX1* homeodomain (Moskow et al. 1995; Nakamura et al. 1996), suggesting a similar protein function. Although the data presented here do not refute the hypothesis that the Meis1 and PBX1 proteins share a functional commonality, it is now clear that Meis1 is much more homologous with a distinct family of homeoproteins than with PBX1 (see below).

Although *Meis1* is a common site of viral integration (Moskow et al. 1995; Nakamura et al. 1996), integrations have not been detected at the *Mrg1* locus. Screening of BXH-2 tumors with a portion of the *Mrg1* 3′ UTR has failed to identify rearrangements indicative of a proviral insertion. Because the probe used represents a small portion of the *Mrg1* locus, this does not necessarily preclude the possibility that viral integrations may be occurring within the coding region, 5′ UTR, or extreme 3′ UTR of *Mrg1*. However, cytogenetic abnormalities have been associated with the q22–25 region of human chromosome 15 where *MRG1* is located. Given the finding that *Mrg1* expression is abundant in brain (Fig. 4), it is interesting to note that human astrocytomas resulting from an addition or deletion [add(15)(q24) or del(15)(q22–26), respectively] have been documented (Griffin et al. 1992; Ransom et al. 1992), suggesting a potential involvement of *MRG1* in tumor development.

Although *Mrg1* may not be a common site of viral integration, the data presented here suggest that *Meis1* and *Mrg1* may have similar functions. Differences in the flanking regions of the homeodomain, however, may be responsible for binding partner selection or heterodimer affinity to a DNA consensus sequence. The complexity of such interactions is furthered by the alternative splicing events identified within the *Meis1* and *Mrgs* (Fig. 2).

Sequence alignment has shown that *Meis1* and *Mrgs* are similar to specific members of the HAC–ATYP family of homeoproteins (Burglin 1995). The HAC–ATYP family belongs to the TALE superfamily and consists of human, *Arabadopsis*, *Caenorhabditis*, and *Saccharomyces* homeobox-containing sequences (Burglin 1995).

A comparison of the HAC–ATYP members with *Meis1*, *Mrgs*, and *Meis1* orthologs presented here indicate that the creation of a separate *Meis1* family

branch of the TALE superfamily is in order (Figs. 8 and 9). Specifically, sequence alignment has demonstrated that the *c-onc10* (GenEMBL accession no. F05816) and *D31072* (accession no. D31072) human ESTs are 95% and 75% identical at the amino acid level to the *Meis1* homeodomain, respectively (Fig. 8). The remaining members of the HAC–ATYP family, *Arabadopsis TBT6G04* (accession no. Z35398) and *Saccharomyces CUP9* (accession no. P41817), however, are only 53% and 38% identical to the *Meis1* homeodomain, respectively (Fig. 8). Furthermore, *TBT6G04* and *CUP9* are only 36% to 54% identical to either *c-onc10* or *D31072* (Fig. 8). These data indicate that *c-onc10* and *D31072* are more homologous to *Meis1* and *Mrgs* than to the remaining HAC–ATYP family members. Furthermore, *Meis1*, *Mrgs*, *XMeis1*, *D31072*, *c-onc10*, and *C. elegans ceh–T28F12* share a homeodomain structure that clearly distinguishes these genes from other homeoprotein families (Fig. 9). Surprisingly, in addition to the exceptional conservation of the *Meis1* homeodomain between mouse and *Xenopus*, these data show that the homeodomain of the *C. elegans* cosmid *ceh–T28F12* is 90% identical to the *Meis1* homeodomain (Fig. 9). *ceh–T28F12* was identified as a homeodomain-containing gene within cosmid T28F12 of the *C. elegans* sequence project using the TBLASTN program (Devereux et al. 1984) to search the St. Louis *C. elegans* BLAST server (http://genome.wustl.edu/htbin/Blast_Server) using the *Meis1a* cDNA as a query sequence. The assignment of the various homeoproteins to their respective families was confirmed using the MacDNASIS PRO V3.6 (Hitachi Software Engineering Co., Ltd., San Bruno, CA) program (Fig. 8B).

Therefore, we propose that *Meis1* represents a novel homeoprotein family whose members to date are *Meis1*, *XMeis1-1*, *XMeis1-2*, *Mrg1*, *MRG2*, *c-onc10*, *D31072*, and *ceh-T28F12*. A summation of the percent identity among various homeoproteins, their classification into previously defined families, and the presentation of the novel *Meis1* homeobox family is given in Figure 8. The respective amino acid sequences are given in Figure 9.

Although the precise function of *Meis1* or *Mrgs* has not been defined, the data presented here suggest a role as a specifically regulated, remarkably conserved family of transcription factors. Taken together, the extensive sequence identity within the *Meis1* family (Fig. 8), the exceptional evolutionary conservation of the Meis1 protein (Fig. 9), and the potential for *Meis1*, *Mrg1*, and *MRG2* to act as transcriptional activators or repressors, suggests an important function for the *Meis1* family of genes in normal as well as neoplastic development.

METHODS

Probes

PBc1 is a 528-bp *Eco*RI fragment derived from a cDNA clone representing the *Meis1* 3′ UTR extending from bp 2632 to 3160 (Moskow et al. 1995). Probe P337-1 is a 890-bp *Pst*I fragment derived from the *Meis1* λ clone 337 (Moskow et al. 1995). Probe H3-4 is a 1.7-kb *Eco*RI fragment derived from the *MRG2* cDNA clone HF-14. Probes M8-9 and M2-5 were generated by PCR using the *Mrg1* cDNA clone K-41 (see below) as a template and represent a 5′ segment of the *Mrg1* ORF and a portion of the *Mrg1* 3′ UTR, respectively. M8-9 was generated using oligonucleotides 5′-GGAGTTAGAAAAGGTCCACG-3′ and 5′-CTTGTCTGGATCGTCGTC-3′. M2-5 was generated using oligonucleotides 5′-CAGGTTATGGACATTCATGC-3′ and 5′-AACAACACACATAGTGTGG-3′. Probes P32-33, P18-19, and P8-11 were generated by PCR using the *Meis1* cDNA clone C-21 (Moskow et al. 1995). P32-33 is a homeodomain-specific probe generated with oligonucleotides 5′-CCAGCACAGGTGACGATG-3′ and 5′-TGGCTGACTGCTCGGTTG-3′. P18-19 represents the entire *Meis1* ORF and was generated with oligonucleotides 5′-CACTGCTGTCTTGGTGGAAC-3′ and 5′-GGTTACATGTAGTGCCACTG-3′. P8-11 is a probe containing the *Meis1a* 3′ UTR and was generated with oligonucleotides 5′-GCACATGGCTGAAATCC-3′ and 5′-GTGATGATGCATGGAGGA-3′. Probe HF–UTR represents the *MRG2* 3′ UTR and was generated using the *MRG2* cDNA clone HF-14 as a template with the oligos 5′-CCCAGAAGATGGCAGCTAGG-3′ and 5′-CCTACCTCAGAACTGGAG-3′. PCR conditions used throughout this work were as previously described (Ma et al. 1993) with a thermocycling protocol consisting of an initial denaturation at 94°C for 4 min, followed by 35 cycles of 94°C for 30 sec, 58°C for 30 sec, and 72°C for 30 sec, unless noted otherwise. All oligomers used were synthesized on an Applied Biosystems model 394 DNA synthesizer.

Cloning of cDNAs

A total of 3×10^5 PFU of an adult mouse kidney cDNA library (kindly provided by Dr. Franklin G. Berger, University of South Carolina, Columbia) was screened with P32-33 using standard techniques (Sambrook et al. 1989). The cDNA clone K-41 was isolated from the kidney library and identified as a *Meis1*-related gene (*Mrg1*) by sequence comparison to *Meis1*. Using probes M8-9 and M2-5 to screen 3×10^5 PFU from an adult mouse brain cDNA library (Stratagene, La Jolla, CA) with standard techniques (Sambrook et al. 1989), a 2.6-kb cDNA clone (cl8-9) containing the entire *Mrg1a* predicted ORF was identified.

MRG2 cDNA was obtained by screening 5×10^5 PFU of a human fibroblast cDNA library (Stratagene) with probes P18-19 and P337-1 using standard techniques (Sambrook et al. 1989). Hybridizations were performed at 58°C with washes in 1× SSCP, 0.1% SDS, through 0.2× SSCP, 0.1% SDS, at 58°C. A single 1.7-kb cDNA clone (HF-14) was isolated.

To obtain *Xenopus Meis1* cDNAs, a *Xenopus* oocyte cDNA library was created in lambda ZAP II vector. Recombinants (5×10^5 PFU) were screened with the *Meis1* 3′ UTR probe P8-

11. Hybridizations were conducted using the Quik-Hyb kit as recommended by the manufacturer (Stratagene). Washes were performed in 1× SSC, 0.1% SDS, at 50°C. A 1.3-kb (*XMeis1-3*) and a 1.4-kb (*Xmeis1-4*) clone were identified. An 800-bp *Sac*I fragment containing the coding sequence of *XMeis1-3* was used to screen 5×10^5 recombinants from another oocyte cDNA library (gift of Dr. Douglas Melton, Harvard University, Cambridge, MA) constructed in λ gt10. Two cDNA clones, 2.3 kb (*XMeis1-1*) and 3.0 kb (*XMeis1-2*), were isolated and sequenced.

Sequence analysis of all plasmid inserts was performed on an Applied Biosystems model 373A DNA sequencing system using the Genetics Computer Group (Madison, WI) software package (Devereux et al. 1984).

Southern Blot Analysis

Southern blots were prepared by digesting 5 μg of genomic DNA with *Eco*RI, followed by size fractionation on a 0.8% agarose gel and transfer to Zetabind membranes (AMF Cuno, Meriden, CT) using standard techniques (Sambrook et al. 1989). Southern blots were probed with [α-^{32}P]dCTP-labeled PBc1 or P32-33 at 65°C as described (Church and Gilbert 1984) and washed in 1× SSCP, 0.1% SDS, through 0.2× SSCP, 0.1% SDS, at 65°C.

RNA Isolation and Northern Blot Analysis

Total RNA was isolated from fresh C57BL/6J mouse tissues as described previously using guanidinium isothiocyanate (Chomczynski and Sacchi 1987). Fifteen micrograms of total RNA was size fractionated on 1.0% formaldehyde agarose gel and transferred onto Zetabind membrane (AMF Cuno) as described (Moskow et al. 1995). Northern blots were hybridized at 65°C as described (Church and Gilbert 1984) with [α-^{32}P]dCTP-labeled M2-5 to detect *Mrg1* and washed in 1× SSCP, 0.1% SDS through 0.2× SSCP, 0.1% SDS at 65°C.

Genetic Mapping by Interspecific Backcross Analysis

The (AEJ/Gn-*a* bp^H/*a* bp^H × *M. spretus*)F_1 × AEJ/Gn-*a* bp^H/*a* bp^H interspecific backcross has been described previously (Marini et al. 1993). Five micrograms of genomic DNA from AEJ/Gn and *M. spretus* was digested with a variety of restriction endonucleases to detect informative RFLPs for mapping the murine loci under study (see Table 1). Southern blots were prepared by digesting genomic DNA from N_2 animals with the appropriate restriction endonuclease. High molecular weight DNA was size fractionated on a 0.8% agarose gel and transferred onto Zetabind membranes (AMF Cuno) using standard techniques (Sambrook et al. 1989). Blots were hybridized at 65°C with [α-^{32}P]dCTP-labeled M2-5, H3-4, or pHTPI-5A to localize *Mrg1*, *Mrg2-rs1*, and *Tpi-rs3*, respectively, and were washed in 1× SSCP, 0.1% SDS, through 0.2× SSCP, 0.1% SDS, at 65°C. PCR amplification was also used to detect simple sequence length polymorphisms for the markers *D2Mit42*, *D2Mit63*, *D7Mit21*, and *D7Mit27* under PCR conditions discussed above with an annealing temperature of 55°C. References for each of the markers typed in the cross are given in Table 1.

The segregation pattern of the *M. spretus* and AEJ/Gn alleles for the *Mrg1* or *Mrg2-rs1* locus were monitored in a random subset of 195 N_2 mice from the interspecific backcross. Linkage of each locus typed in the interspecific backcross was analyzed by calculating the maximum-likelihood estimates of linkage parameters as described (Green 1981), using the computer program Spretus Madness: Part Deux (developed by Karl Smalley, Jim Averback, Linda D. Siracusa, and Arthur M. Buchberg, Kimmel Cancer Institute, Philadelphia, PA).

Rodent–Human Hybrids

Rodent-human hybrid DNAs were obtained from the National Institutes of General Medical Sciences (NIGMS) Human Genetic Mutant Cell Repository (Camden, NJ). To localize *MRG1*, PCR amplification of the hybrid template DNAs (100 ng) was performed under PCR conditions stated above. The WI-1222 primers (5′-TGATCCGGATAAGGACAAAA-3′ and 5′-TCAAGAATAAGAAAATAGGCACC-3′; accession no. G03481, Whitehead Institute/MIT Center for Genome Research; Random Genome Wide STSs) were used with 55°C annealing.

Sequence Comparisons

Percent similarities and identities were determined using the GAP program (creation penalty = 5.00; extension penalty = 0.30) of the GCG (Madison, WI) software package (Devereux et al. 1984). All alignments shown were performed manually. Searches for existing sequences similar to *Meis1* were performed using the BLAST program of the GCG (Madison, WI) software package (Devereux et al. 1984), unless noted otherwise. To generate phylogenetic trees, the Higgins–Sharp algorithm within the MacDNASIS V3.6 program (Hitachi Software Engineering Co., Ltd., San Bruno, CA) was used with the following parameters: gap penalty = 5; number of top diagonals = 5; fixed gap penalty = 10; K-tuple = 2; window size = 5; floating gap penalty = 10.

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health grants CA 21124 (A.M.B.) and CA 58586 (A.M.B), National Cancer Institute, and National Research Service Award training grant T32HI07780 (J.J.M.). We thank Dr. Linda D. Siracusa for helpful discussion and critical reviews of this manuscript. We also thank Julie B. Engiles, Kelly K. Nelson, and Kimberly A. Chianese for technical assistance.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES

Bertolino, E., B. Reimund, D. Wildt-Perinic, and R.G. Clerc. 1995. A novel homeobox protein which recognizes a TGT core and functionally interferes with a retinoid-responsive motif. *J. Biol. Chem.* **270:** 31178–31188.

Bopp, D., M. Burri, S. Baumgartner, G. Frigerio, and M. Noll. 1986. Conservation of a large protein domain in the

segmentation gene *paired* and in functionally related genes of *Drosophila*. *Cell* **47:** 1033–1040.

Brilliant, M.H., R.W. Williams, B.C. Holdener, and J.M. Angel. 1996. Mouse chromosome 7. *Mamm. Genome* **6:** S135–S150.

Burglin, T.R. 1995. The evolution of homeobox genes. In *Biodiversity and Evolution* (ed. R. Arai, M. Kato, and Y. Doi), pp. 291–336. The National Science Museum Foundation, Tokyo, Japan.

Chomczynski, P. and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidium thiocyanoate-phenol-chloroform extraction. *Anal. Biochem.* **162:** 156–159.

Church, G.M. and W. Gilbert. 1984. Genomic sequencing. *Proc. Natl. Acad. Sci.* **81:** 1991–1995.

Desplan, C., J. Theis, and P.H. O'Farrell. 1988. The sequence specificity of homeodomain-DNA interaction. *Cell* **54:** 1081–1090.

Devereux, J., P. Haeberli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12:** 387–395.

Dietrich, W., J. Miller, H. Katz, D. Joyce, R. Steen, S. Lincoln, M. Daly, M.P. Reeve, A. Weaver, P. Anagnostopoulos, N. Goodman, N. Dracopoli, and E.S. Lander. 1993. SSLP genetic map of the mouse (*Mus musculus*). In *Genetic maps: Locus maps of complex genomes, Book 4: Nonhuman vertebrates* (ed. S.J. O'Brien), pp. 4110–4142. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Gehring, W.J., Y.Q. Qian, M. Billeter, K. Furukubo-Tokunaga, A.F. Schier, D. Resendez-Perez, M. Affolter, G. Otting, and K. Wuthrich. 1994. Homeodomain-DNA recognition. *Cell* **78:** 211–223.

Goldstein, C., P. Liaw, S.A. Jimenez, A.M. Buchberg, and L.D. Siracusa. 1994. Of mice and Marfan: Genetic linkage analyses of the fibrillin genes, *Fbn1* and *Fbn2*, in the mouse genome. *Mamm. Genome* **5:** 696–700.

Green, M.C. 1981. Gene mapping. In *The mouse in biomedical research* (ed. H. L. Foster, J.D. Smith, and J.G. Fox), pp. 105–117. Academic Press, New York, NY.

Griffin, C.A., P.P. Long, B.S. Carson, and H. Brem. 1992. Chromosome abnormalities in low-grade central nervous system tumors. *Cancer Genet. Cytogenet.* **60:** 67–73.

Herr, W., R.A. Sturm, R.G. Clerc, L.M. Corcoran, D. Baltimore, P.A. Sharp, H.A. Ingraham, M.G. Rosenfeld, M. Finney, G. Ruvkun, and H.R. Horowitz. 1988. The POU domain: A large conserved region in the mammalian *pit-1*, *oct-1*, *oct-2*, and *Caenorhabditis elegans unc-86* gene products. *Genes & Dev.* **2:** 1513–1516.

Ishihara, T. and M. Minamihisamatsu. 1988. The Philadelphia chromosome. Considerations based on studies of variant Ph translocations. *Cancer Genet. Cytogenet.* **32:** 75–92.

Kamps, M.P., C. Murre, X.-H. Sun, and D. Baltimore. 1990. A new homeobox gene contributes the DNA binding domain of the t(1;19) translocation protein in pre-B ALL. *Cell* **60:** 547–555.

Kamps, M.P. and D. Baltimore. 1993. *E2A-Pbx1*, the t(1;19) translocation protein of human pre-B-cell acute lymphocytic leukemia, causes acute myeloid leukemia in mice. *Mol. Cell. Biol.* **13:** 351–357.

Kappen, C., K. Schughart, and F.H. Ruddle. 1993. Early evolutionary origin of major homeodomain sequence classes. *Genomics* **18:** 54–70.

Kongsuwan, K., J. Allen, and J.M. Adams. 1989. Expression of *Hox-2.4* homeobox gene directed by proviral insertion in a myeloid leukemia. *Nucleic Acids Res.* **17:** 1881–1892.

Kozak, M. 1987. An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* **15:** 8125–8148.

Lawrence, H.J., G. Sauvageau, N. Ahmadi, A.R. Lopez, M.M. LeBeau, M. Link, K. Humphries, and C. Largman. 1995. Stage- and lineage-specific expression of the *HOXA10* homeobox gene in normal and leukemic hematopoietic cells. *Exp. Hematol.* **23:** 1160–1166.

Le Beau, M.M., J.D. Rowley, N. Sacchi, D.K. Watson, T.S. Papas, and M.O. Diaz. 1986. Clinical and cytogenetic correlations in 63 patients with therapy-related myelodysplastic syndromes and acute nonlymphocytic leukemia: Further evidence for characteristic abnormalities of chromosomes no. 5 and 7. *J. Clin. Oncol.* **4:** 325–345.

Ma, Q., H. Alder, K.K. Nelson, D. Chatterjee, Y. Gu, T. Nakamura, E. Canaani, C.M. Croce, L.D. Siracusa, and A.M. Buchberg. 1993. Analysis of the murine *All-1* gene reveals conserved domains with human *ALL-1* and identifies a motif shared with DNA methyltransferases. *Proc. Natl. Acad. Sci.* **90:** 6350–6354.

Mangelsdorf, D.J., K. Umesono, S.A. Kliewer, U. Borgmeyer, E.S. Ong, and R.M. Evans. 1991. A direct repeat in the cellular retinol-binding protein type II gene confers differential regulation by RXR and RAR. *Cell* **66:** 555–561.

Marini, J.C., K.K. Nelson, J. Battey, and L.D. Siracusa. 1993. The pituitary hormones arginine vasopressin-neurophysin II and oxytocin-neurophysin I show close linkage with interleukin-1 on mouse chromosome 2. *Genomics* **15:** 200–202.

McGinnis, W., M.S. Levine, E. Hafen, A. Kuroiwa, and W.J. Gehring. 1984. A conserved DNA sequence in homeotic genes of the *Drosophila Antennapedia* and *bithorax* complexes. *Nature* **308:** 428–433.

Monica, K., N. Galili, J. Nourse, D. Saltman, and M.L. Cleary. 1991. *PBX2* and *PBX3*, new homeobox genes with

extensive homology to the human proto-oncogene *PBX1*. *Mol. Cell. Biol.* **11:** 6149–6157.

Moskow, J.J., F. Bullrich, K. Huebner, I.O. Daar, and A.M. Buchberg. 1995. *Meis1*, a *PBX1*-related homeobox gene involved in myeloid leukemia in BXH-2 mice. *Mol. Cell. Biol.* **15:** 5434–5443.

Nakamura, T., D.A. Largespada, J.D. Shaughnessy, Jr., N.A. Jenkins, and N.G. Copeland. 1996. Cooperative activation of *Hoxa* and *Pbx1*-related genes in murine myeloid leukaemias. *Nat. Genet.* **12:** 149–153.

Nakshatri, H., and P. Chambon. 1994. The directly repeated RG(G/T)TCA motifs of the rat and mouse cellular retinol-binding protein II genes are promiscuous binding sites for RAR, RXR, HNF-4, and ARP-1 homo- and heterodimers. *J. Biol. Chem.* **269:** 890–902.

Nourse, J., J.D. Mellentin, N. Galili, J. Wilkinson, E. Stanbridge, S.D. Smith, and M.L. Cleary. 1990. Chromosomal translocation t(1;19) results in synthesis of a homeobox fusion mRNA that codes for a potential chimeric transcription factor. *Cell* **60:** 535–545.

Ransom, D.T., S.R. Ritland, C.A. Moertel, R.J. Dahl, J.R. O'Fallon, B.W. Scheithauer, D.W. Kimmel, P.J. Kelly, O.I. Olopade, M.O. Diaz, and R.B. Jenkins. 1992. Correlation of cytogenetic analysis and loss of heterozygosity studies in human diffuse astrocytomas and mixed oligo-astrocytomas. *Genes Chrom. Cancer* **5:** 357–374.

Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Scott, M.P. and A.J. Weiner. 1984. Structural relationships among genes that control development: Sequence homology between the *Antennapedia*, *Ultrabithorax* and *fushi tarazu* loci of *Drosophila*. *Proc. Natl. Acad. Sci.* **81:** 4115–4119.

Scott, M.P., J.W. Tamkun, and G.W. Hartzell. 1989. The structure and function of the homeodomain. *Biochem. Biophys. Acta* **989:** 25–48.

Siracusa, L.D., J.L. Morgan, J.K. Fisher, C.M. Abbott, and J. Peters. 1996. Mouse chromosome 2. *Mamm. Genome* **6:** S51–S63.

Willman, C.L., C.E. Sever, M.G. Pallavicini, H. Harada, N. Tanaka, M.L. Slovak, H. Yamamoto, K. Harada, T.C. Meeker, A.F. List, and T. Taniguchi. 1993. Deletion of *IRF-1*, mapping to chromosome 5q31.1, in human leukemia and preleukemic myelodysplasia. *Science* **259:** 968–971.

Received August 30, 1996; accepted in revised form December 17, 1996.