MOLECULAR AND CELLULAR BIOLOGY, Dec. 1991, p. 6149–6157
0270-7306/91/126149-09$02.00/0

Vol. 11, No. 12

PBX2 and *PBX3*, New Homeobox Genes with Extensive Homology to the Human Proto-Oncogene *PBX1*

KATHERINE MONICA, NAOMI GALILI, JAMISON NOURSE, DAVID SALTMAN, AND MICHAEL L. CLEARY*

Laboratory of Experimental Oncology, Department of Pathology, Stanford University Medical Center, Stanford, California 94305

Received 23 July 1991/Accepted 23 September 1991

Two new homeobox genes, *PBX2* and *PBX3*, were isolated on the basis of their extensive homology to *PBX1*, a novel human homeobox gene involved in t(1;19) translocation in acute pre-B-cell leukemias. The predicted Pbx2 and Pbx3 proteins are 92 and 94% identical to Pbx1 over a large region of 266 amino acids within and flanking their homeodomains, but all three proteins diverge significantly near their amino and carboxy termini. Chromosome in situ hybridizations demonstrated that the *PBX* genes are not clustered but map to separate chromosomal loci: *PBX1*, 1q23; *PBX2*, 3q22-23; *PBX3*, 9q33-34. Expression of *PBX2* or *PBX3* was not restricted to particular states of differentiation or development, as mRNA transcripts of these genes were detected in most fetal and adult tissues and all cell lines, unlike *PBX1*, which is not expressed in lymphoid cell lines. Similar to *PBX1* RNA, *PBX3* RNA is alternatively spliced to yield two translation products with different carboxy termini, a feature not observed for *PBX2*. Their extensive sequence similarity and widespread expression suggest a generalized, overlapping role for Pbx proteins in most cell types. Differences in their amino and carboxy termini may modulate their activities, mediated in part by differential splicing and, for *PBX1*, protein fusion following t(1;19) chromosomal translocation.

A sequence of 180 bp termed the homeobox was initially discovered as a conserved segment of the Antennapedia, Ultrabithorax, and fushi tarazu transcription units (40, 41, 53). Subsequently, homeoboxes have been observed in the genomes of many species by cross-hybridization and have also been found fortuitously in genes isolated by other means. Homeoboxes code for a conserved protein motif, termed the homeodomain, and the degree of similarity between different homeodomains has been used to group them into related classes (52). While most of the *Drosophila* homeodomain proteins have been implicated in developmental processes, the functions of the majority of non-*Drosophila* homeodomain proteins remain unknown. However, because of the presence of homeodomains in known transcription factors and the structural similarity between the homeodomain and DNA-binding motifs in bacterial proteins, the current view is that most if not all of these proteins function as sequence-specific DNA-binding proteins that likely play an important role in transcriptional regulation (for reviews, see references 17 and 52).

Recently, a novel human homeobox gene, *PBX1*, was identified because of its involvement in t(1;19) chromosomal translocation in acute pre-B-cell leukemias (29, 47). This translocation results in the formation of a fusion transcript that codes for E2A-Pbx1 chimeric proteins, in which the C-terminal region of E2A, which contains its basic and helix-loop-helix (bHLH) DNA-binding and dimerization motifs, is replaced by sequences from Pbx1 that contain a highly divergent homeodomain. The role of E2A-Pbx chimeric transcription factors in the pathogenesis of acute pre-B-cell leukemias remains unknown.

Considering the high level of divergence between Pbx1 and previously reported homeodomains, we hypothesized that there may be additional, as-yet-uncharacterized homeobox genes which encode homeodomains similar to that of Pbx1. We describe here the characterization of two such genes, designated *PBX2* and *PBX3*, which were isolated as a result of their cross-hybridization with the *PBX1* homeobox. A striking feature of the Pbx1, Pbx2, and Pbx3 proteins is their extensive sequence identity both within and outside of their homeodomains. The similarities in their structures and patterns of expression suggest a generalized, overlapping function for Pbx homeodomain proteins in most cell types.

MATERIALS AND METHODS

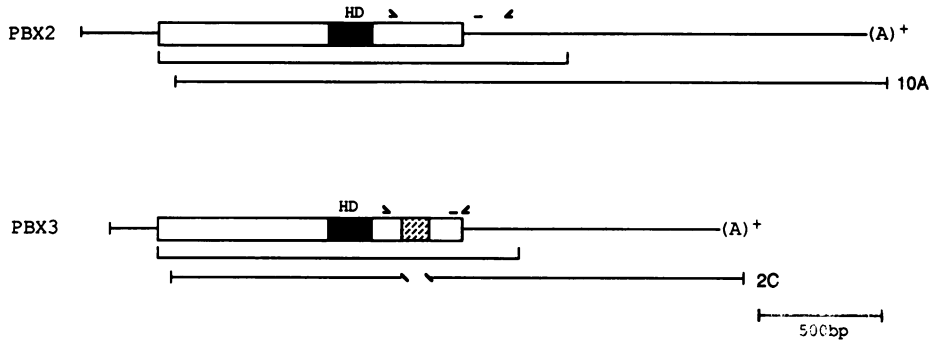
Isolation and sequencing of *PBX2* and *PBX3* cDNA clones. *PBX2* and *PBX3* cDNA clones were identified by screening a HeLa cDNA library under nonstringent conditions with a *PBX1* homeobox probe (nucleotides 1951 to 2210) (47). Library filters containing phage DNA were prehybridized overnight at 37°C in 43% formamide, 5× SSC (1× SSC is 0.15 M NaCl with 0.015 M sodium citrate), 5× Denhardt's solution, 1 mM sodium pyrophosphate, 1% sodium dodecyl sulfate (SDS), and 100 μg of denatured, sheared salmon sperm DNA per ml. Filters were hybridized overnight at 37°C in an identical solution containing 4.8% dextran sulfate and the *PBX1* homeobox probe radiolabeled with [α-^{32}P]dCTP by hexamer priming (12). Filters were washed for 10 min at room temperature and for 30 min at 52°C in 2× SSC–0.1% SDS.

Hybridizing phage were purified by using the *PBX1* homeobox probe and nonstringent hybridization and wash conditions as described above. cDNA fragments were subcloned into M13 vectors, and single-stranded DNA templates were sequenced by using commercially prepared reagents (USB, Cleveland, Ohio) and the M13 universal primer or appropriate synthetic oligonucleotide primers homologous to *PBX2* and *PBX3* sequences. Sequences were determined from at least two overlapping clones for each of the *PBX2*, *PBX3a*, and *PBX3b* cDNAs.

In situ hybridization mapping. Chromosome spreads were prepared from a lymphoblastoid cell line with a normal male

* Corresponding author.

A.



B.

```
      1                                                  50
1a    MDEQ    PRLMHSHAGVGM AGHP   GLSQHLQDGAGGTEGEGG     RKQDIGDILQQIMTITDQSLDEAQARKHALNCHRMKPALFNVLCEIKEKTVLSIRGAQEEE
2     ---RLLGP-PPGGGRG-L-LVS-E-   -GPGEPPG-GDPGG-S--VPGGRG----------------------K----------S---------G----SS----
3a    --D-    S-MLQTL---NL ---SVQG-MALPPPPHGHEGADGD-     ---------H----------K----------S---------G---------D
1b    ----    ------------ ----   ------------------     ----------------------------------------------------------
3b    --D-    S-MLQTL---NL ---SVQG-MALPPPPHGHEGADGD-     ---------H----------K----------S---------G---------D

      100                                 150                                 200
1a    PTDPQLMRLDNMLLAEGVAGPEKGGGSAAAAAAAAASGGAGS DNSVEHSDYRAKLSQIRQIYHTELEKYEQACNEFTTHVMNLLREQSRTRPISPKEIERMVSIIHRKFSS
2     -V-----------------------------------GV-P---I------S--A---H---S---------------------VA---M----------A
3a    -P----------------S-------------------- S- ---I---------T-------------------------------------G--------
1b    ---------------------------------------- ----------------------------------------------------------------
3b    -P----------------S-------------------- S- ---I---------T-------------------------------------G--------

                                  250                                 300
1a    IQMQLKQSTCEAVMILRSRFLDARRKRRNFNKQATEILNEYFYSHLSNPYPSEEAKEELAKKCGITVSQVSNWFGNKRIRYKKNIGKFQEEANIYAAKTAVTATNVSA   H
2     ------------------------------S-----V-------------------------------------------------------V----SV-QGGH
3a    ------------------------------S-----------------------------S-------------------------L---------AHAV-AAVQ
1b    ---------------------------------------------------------------------------------------------------------  -
3b    ------------------------------S-----------------------------S-------------------------L---------AHAV-AAVQ

                                  350                                 400
1a    GSQANSPSTPNSAGSSSSFNMSNSGDLFMSVQSLNGDSYQGAQVGANVQSQVDTLRHVISQTGGYSDGLAASQMYSPQGISANGGWQDATTPSSVTSPTEGPGSVHSDTSN*
2      -RTS--TP-S----GG---L-G---M-LGMPG------SA        ---ES---SMG P---G-N-GGG-I---REMR---S--E-V----------------------*
3a    NN-T---T---- ---G---LP----M--NM----------S----------------N---------GGNSL---HNLN-----------------------------*
1b    --------------GYPSPCYQPDRRIQ*
3b    NN-T---T---- ----PS---S-G-L-*
```

FIG. 1. (A) *PBX2* and *PBX3* cDNAs and probes. Physical structures of consensus cDNAs are shown. Brackets denote cDNA fragments subcloned into expression vectors. cDNA fragments used as hybridization probes for in situ and Northern analyses are marked 10A and 2C. Approximate locations of PCR primers and detection probes are indicated above cDNAs. Symbols: □, coding sequences; ▩, region deleted in splice variants of *PBX3*; ■, homeodomains (HD); ——, 5′ and 3′ untranslated sequences. (B) Alignment of Pbx amino acid sequences. The complete amino acid sequences for Pbx1b, Pbx2, Pbx3a, and Pbx3b are compared with that of Pbx1a, which is represented on the upper line in single-letter code. Amino acid numbering corresponds to that of Pbx1a (16a). Dashes denote amino acid identity with Pbx1a, and spaces allow for insertions or deletions to achieve optimal alignment. Pbx3b is compared with Pbx1a up to amino acid 332, after which it is compared with Pbx1b. The homeodomain is overlined, and asterisks denote carboxy termini.

karyotype. Cells were arrested with 0.05 μg of Colcemid (Gibco) and harvested according to standard cytogenetic procedure. In situ hybridization was performed essentially as described previously (18). DNA probes (*PBX2* fragment 10A and *PBX3* fragment 2C in Fig. 1A) were labeled with [^{3}H]dATP, [^{3}H]dCTP, and [^{3}H]dTTP to a specific activity of 0.5×10^8 to 1×10^8 cpm/μg by random hexanucleotide priming (12). Slides were hybridized at 42°C overnight. Washes were carried out initially in 50% formamide–2× SSC and then in 2× SSC alone, all at 44°C. Exposure times for autoradiography varied from 5 to 10 days. Chromosomes were R banded by using chromomycin A3 and distamycin A (11) and viewed on a Zeiss fluorescence microscope by using a 490-nm excitation and 520-nm barrier filter.

Northern (RNA) blots. Poly(A)$^+$ RNA (2 μg) was isolated from tissues and cell lines by using commercially prepared reagents (Invitrogen, San Diego, Calif.) and size fractionated in formaldehyde-agarose gels as described previously (10). RNA was transferred by diffusion to a nylon membrane overnight in 10× SSC, rinsed briefly in 2× SSC, and UV cross-linked to the membrane. Membranes were hybridized as described previously (8) with *PBX2*-specific (fragment 10A, Fig. 1A) or *PBX3*-specific (fragment 2C, Fig. 1A) probes radiolabeled by random-hexamer priming. *PBX1* expression was determined as described by Galili et al. (16a). Membranes were washed briefly at room temperature in 2× SSC–0.1% SDS and for 1 h in 0.1× SSC–0.1% SDS at 63°C. Under these conditions, no cross-hybridization was observed between the *PBX1*, *PBX2*, or *PBX3* probes.

PCR analyses of *PBX* mRNA. Differentially spliced variants of *PBX* mRNA were detected by reverse polymerase chain reaction (PCR) using total (2 μg) or poly(A)$^+$ (5 μg) RNA under conditions described previously (47). The 5′ oligonucleotide primers for *PBX2* and *PBX3* were 5′GCCA CAGCCGCACCAGCTCC3′ and 5′GCACACGCAGTAGC AGCAGC3′, respectively; the 3′ oligonucleotide primers

were 5'CCTTAGAGGCCCCATTCTTCC3' and 5'TGTG GCCAGAGATTAGTTAG3', respectively. PCR products were subjected to agarose gel electrophoresis, transferred to nylon membranes in 0.4 N NaOH, and hybridized by using internal oligonucleotide detection probes (5'GTCGCCT TGTGAGAGCCCCC3' for *PBX2* and 5'CAGAAGGCCC AGGAAGTGTGCAC3' for *PBX3*) end labeled with ^{32}P by using polynucleotide kinase (32). Detection oligonucleotides corresponded to regions of *PBX2* and *PBX3* predicted to be present in both differential splice products on the basis of analyses of *PBX1* and *PBX3* cDNAs. Membranes were washed at room temperature for 15 min and then at 60 to 63°C for 15 min in 6× SSC. Under these conditions, the various detection oligonucleotides were observed to hybridize specifically with their respective *PBX* cDNAs with no detectable cross-hybridization to other *PBX* cDNAs.

In vitro translations. cDNAs containing the open reading frames of Pbx1a, Pbx1b, Pbx2, Pbx3a, and Pbx3b were ligated into pSP64 or pSP65 expression vectors (Promega, Madison, Wis.). RNA was transcribed from linearized plasmids by using SP6 polymerase, purified, and then translated in vitro with a nuclease-treated rabbit reticulocyte lysate under conditions recommended by the supplier (Promega). Translation products (2 μl) were analyzed in 10% SDS-polyacrylamide gels which were fixed, treated with Amplify (Amersham, Arlington Heights, Ill.), and vacuum dried prior to autoradiography.

Immunoprecipitations. In vitro-translated proteins were immunoprecipitated by suspending 2 μl of the translation reaction mixture in 500 ml of RIPA buffer (0.1% Triton X-100, 50 mM HEPES [*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid], 150 mM NaCl, 5 mM EDTA, 0.5 μg of antipain per ml, 2 μg of leupeptin per ml, 1 mM phenylmethylsulfonyl fluoride) and incubating it for 3 h in the presence of anti-Pbx1a antibodies (16a) and Sepharose-protein A beads (Pharmacia, Piscataway, N.J.). Immune pellets were washed extensively in RIPA buffer and subjected to analysis in 10% SDS-polyacrylamide gel electrophoresis (PAGE) gels. Gels were fixed, treated with Amplify, dried, and subjected to autoradiography.

Immunoblotting. Cell extracts were prepared by using RIPA buffer (see above). Protein fractions containing 2×10^5 cell equivalents were subjected to 10% SDS-PAGE and electrotransferred to nitrocellulose by using transfer buffer as described elsewhere (44). Filters were then subjected to Western blot (immunoblot) analysis with anti-Pbx1a antibodies as previously described (6).

Nucleotide sequence accession numbers. Nucleotide sequences have been deposited in the EMBL, GenBank, and DDBJ nucleotide sequence data bases under the accession numbers X59841 (PBX3) and X59842 (PBX2).

RESULTS

Isolation and characterization of *PBX2* and *PBX3* cDNAs. To identify cDNAs which contained homeoboxes similar to that of *PBX1*, a HeLa cDNA library (3×10^5 plaques) was screened at reduced stringency by using a DNA probe containing the 180-bp homeobox of *PBX1*. Phages that hybridized under low- but not high-stringency conditions were purified and initially characterized by nucleotide sequence analysis of their homeodomains. All cDNA clones analyzed contained one of two divergent homeobox sequences different from *PBX1* and were designated *PBX2* (seven clones) or *PBX3* (five clones).

Complete nucleotide sequence analyses of the cDNAs showed structures for the *PBX2* and *PBX3* transcripts as shown in Fig. 1A. Two forms of the *PBX3* cDNA that differed by the inclusion or exclusion of a 113-bp segment 3' of the homeodomain were obtained. Positions for ATG initiation codons were assigned on the basis of homology to the eukaryotic translational start consensus (35) and similarity to the ATG start site in *PBX1* (Fig. 1B). *PBX2* has an in-frame stop codon upstream of the designated ATG start. No upstream in-frame stop codon was observed for *PBX3*; however, the sequence of this 5' untranslated region has not been completely determined because of the exceedingly high GC content.

Homology of amino acid sequences. The predicted proteins encoded by *PBX* cDNAs are shown in Fig. 1B. Two forms of Pbx3, designated a and b, result from differences in splicing in the 3' portion of this transcript similar to that observed for *PBX1* (47). Earlier studies (47) have shown that the homeodomain of Pbx1 is highly divergent and is most similar (36% identity) to that of yeast MATa1. Compared with the homeodomain of Pbx1, those of Pbx2 and Pbx3 show only two amino acid changes (Fig. 1B), which are not at positions believed to be involved in DNA recognition (31). Although showing 97% identity at the amino acid level within their homeodomains, the *PBX2* and *PBX3* homeoboxes are only 79 to 83% identical to *PBX1* at the nucleotide level. The stringency of screening used to detect *PBX2* and *PBX3* cDNAs may have excluded other *PBX*-like homeoboxes with slightly more divergent nucleotide sequences.

The high level of identity of these three proteins extends well beyond the homeodomains, as shown in Fig. 1B and 2. There is 92 to 94% amino acid identity between Pbx1, Pbx2, and Pbx3 from amino acid 38 to 311, a region which includes the homeodomain, and 100% identity at the extreme C termini. Overall, Pbx1 and Pbx3 are more similar to each other (84%) than Pbx1 and Pbx3 are to Pbx2 (77%), with two blocks of sequence divergence found at the amino termini and near the carboxy termini of the Pbx proteins.

A search of the Protein Identification Resource (NBRF) of regions conserved outside of the homeodomains of Pbx1, Pbx2, and Pbx3 (amino acids 45 to 100, 100 to 145, 150 to 220, and 400 to 430) showed no significant similarity to other proteins. They share no sequence similarity to the PRD repeat, paired box, M repeat, POU domain, or Y-P-W-M peptide, which is conserved upstream of many homeodomains (5, 16, 23, 39, 57). However, a repeated trinucleotide, GCX, is conserved between Pbx1, Pbx2, and Pbx3 and codes for a stretch of alanines at amino acids 126 to 135 (Fig. 1B). Similar GCX repeats of various lengths encoding alanines are present in other homeobox genes such as engrailed, even-skipped, *H2.0*, and Ultrabithorax (2, 34, 38, 49).

There are no regions which are proline or glutamine rich and no stretches of acidic amino acids that might indicate a transcriptional activation domain (43). However, like the HNF1, vHNF1, and Hlx homeodomain proteins (1, 15, 50), the C termini of Pbx1a, Pbx2, and Pbx3a from amino acids 320 to 430 are somewhat serine/threonine rich (29, 27, and 26%, respectively, with 20 of the 32 serine/threonine of Pbx1a conserved in Pbx2 and Pbx3a). The C-terminal region of HNF1 has been implicated in transcriptional activation, and the mutational deletion of this region reduces the activity to about 50% of that of wild type (46).

The predicted secondary structure (7) is conserved over most of the proteins, as was expected from their overall amino acid identity. There is a highly charged hydrophilic

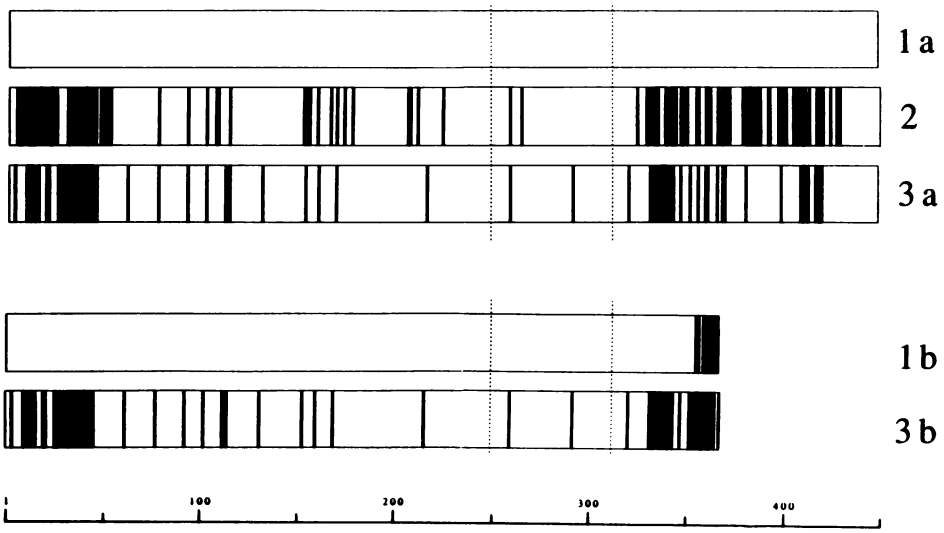


FIG. 2. Schematic comparison of Pbx protein sequences. Each sequence is displayed as a box, and amino acid positions of nonidentity with Pbx1a (upper, open box) are indicated by black vertical lines. Vertical dashed lines denote homeodomain boundaries. The scale at the bottom represents the number of amino acid residues.

region between amino acids 145 to 208, with 15% of the amino acids acidic and 23% basic; all of the charged amino acids are conserved between Pbx1, Pbx2, and Pbx3. The predicted secondary structure in this region consists of an alpha helix positioned between two beta sheets. Because of the hydrophilic nature and predicted secondary structure of this region, it is likely that it is exposed at the surface of the protein and may serve as a site for interactions with other proteins.

Chromosomal localizations. *PBX2* and *PBX3* were localized to their respective chromosomes by in situ hybridization using cDNA probes. *PBX2* mapped to chromosome 3q22-23 (Fig. 3A), with 16% of the grains localized to chromosome 3 and 57% of these to 3q22-23. *PBX3* mapped to chromosome 9q33-34 (Fig. 3B), with 34% of the grains localized to chromosome 9 and 79% of these to 9q33-34. In previous studies, *PBX1* was mapped to chromosome 1q23, the site of t(1;19) in pre-B-cell acute lymphoblastic leukemia (ALL) (47). Unlike many *Drosophila*, murine, and other human homeobox genes (4, 17, 19) but similar to human homeobox genes that are not implicated in developmental processes, *PBX1*, *PBX2*, and *PBX3* are not clustered at a single chromosomal locus (37). Our data do not exclude the possible existence of more distantly related homeobox genes flanking the *PBX* genes. Chromosomal translocations do occur at the 3q21-25 and 9q34 sites in ALL (22); however,

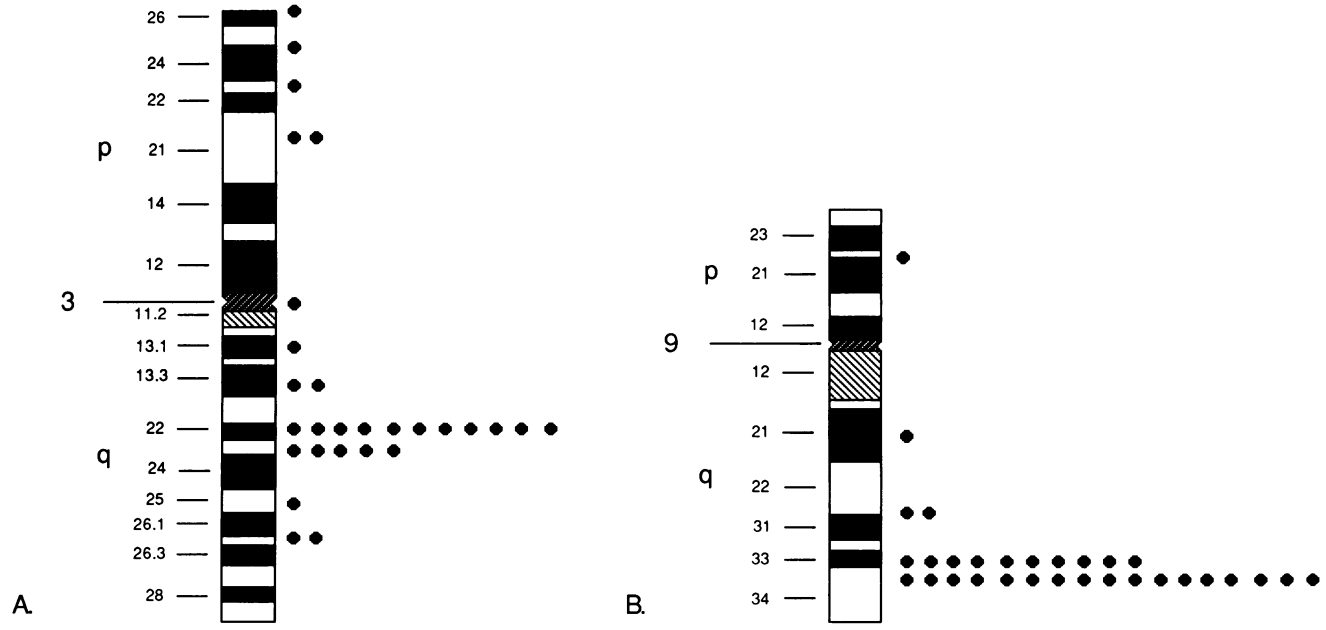


FIG. 3. Chromosomal localizations of *PBX2* and *PBX3* genes. The localizations of silver grains on banded chromosomes demonstrate the mapping of *PBX2* (A) and *PBX3* (B) genes by in situ hybridization to chromosome subbands 3q22-23 and 9q33-34, respectively.

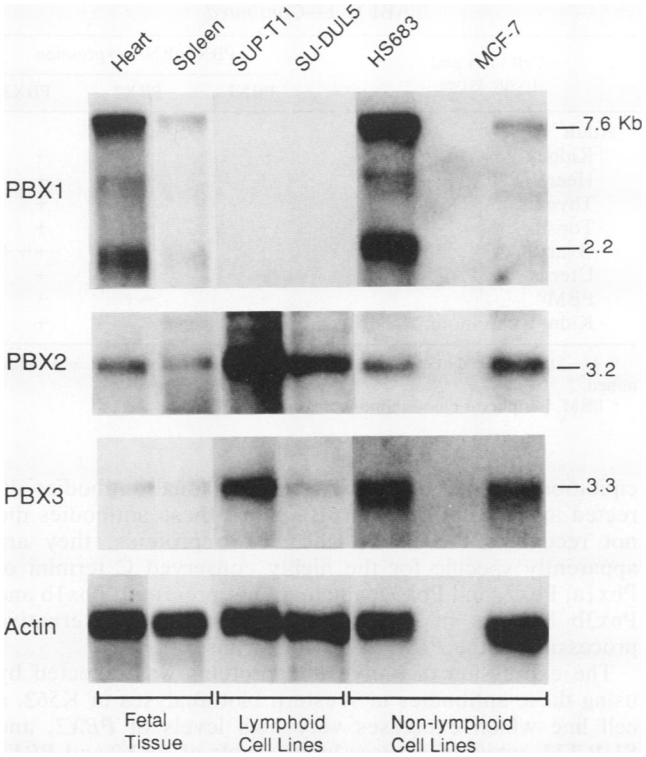


FIG. 4. Northern analysis of *PBX1*, *PBX2*, and *PBX3* expression in lymphoid and nonlymphoid cell lines and fetal tissue. Poly(A)$^+$ mRNA from various cell lines and tissues was analyzed by sequential Northern blot hybridizations for *PBX1*, *PBX2*, *PBX3*, and actin transcripts. The hybridization probe for *PBX1* transcripts was as described elsewhere (16a); the probe for *PBX2* transcripts consisted of fragment 10A (Fig. 1A), and the probe for *PBX3* transcripts consisted of fragment 2C (Fig. 1A). Blots were exposed for 1 to 3 days (background does not reflect length of exposure time).

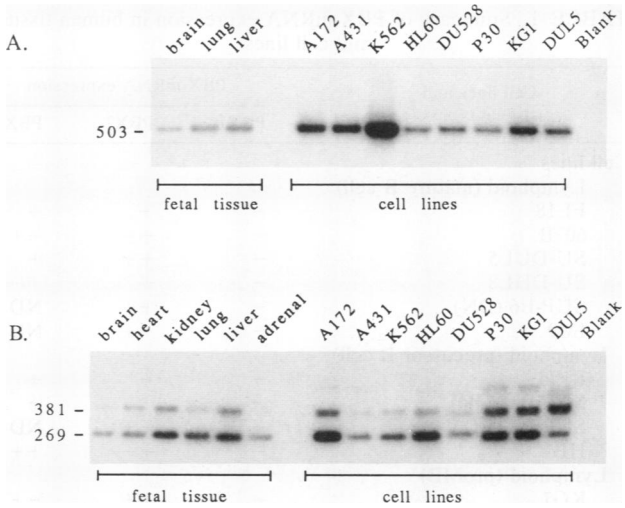


FIG. 5. PCR analysis of *PBX* mRNA. Total or poly(A)$^+$ RNA was converted to cDNA, and PCR was carried out by using *PBX2*- or *PBX3*-specific primers as shown in Fig. 1A. (A) *PBX2* PCR of poly(A)$^+$ RNA isolated from fetal tissues and cell lines. A single DNA fragment of 503 bp was observed when a *PBX2*-specific detection probe was used. (B) *PBX3* PCR of fetal tissues (total RNA) and cell lines [poly(A)$^+$]. Two DNA fragments of 381 and 269 bp were observed when a *PBX3*-specific detection probe was used (see Materials and Methods for nucleotide sequences of primers and probes). Lanes marked Blank show results of PCR reactions lacking RNA (negative control). DNA size markers consisted of a *Hin*cII digest of ϕX174.

further studies are required to determine whether *PBX2* or *PBX3* are involved.

Expression of *PBX2* and *PBX3*. Expression of the *PBX* genes was studied by Northern blot analysis using probes and high-stringency hybridization conditions that distinguished between the various transcripts. RNA transcripts for *PBX2* (3.2 kb) and *PBX3* (3.3 kb) were expressed in all lymphoid and nonlymphoid cell lines tested (Fig. 4 and 5 and Table 1). The expression of *PBX2* and *PBX3* differed significantly from that of *PBX1*, which is not expressed in cultured cells of the lymphoid lineage (Table 1) but is expressed in many nonlymphoid lineage cells as reported elsewhere (16a, 47). The predominant *PBX1* transcript was 7.6 kb, although an additional transcript of 2.2 kb was observed for *PBX1* in certain tissues and cell lines, as described elsewhere (16a, 47).

No significant tissue-specific differences were observed in the expression of *PBX* genes, as shown in Table 1. Although most tissues were observed to express each of the *PBX* genes at some level, fetal brain tissues were most notable for low levels of detectable *PBX2* and *PBX3* transcripts in contrast to abundant *PBX1* transcripts. *PBX3* expression was not detected by Northern analysis in fetal adrenal tissue, but this may reflect the overall lower level of expression of *PBX3* in most tissues.

Differential splicing. Isolation of two species of *PBX3* cDNAs suggested that differential splicing of precursor transcripts similar to that observed for *PBX1* may be a general feature of *PBX* gene expression. A reverse PCR assay was used to study the splicing patterns of *PBX* RNAs in various tissues and cell lines. Amplification and detection primers specific for potential regions of variability within the *PBX* transcripts were constructed as shown in Fig. 1A. A single amplification product was observed in all RNA preparations for *PBX2* (Fig. 5A) and corresponded to that predicted from analyses of *PBX2* cDNAs. Two predominant amplification products (Fig. 5B) which corresponded to the bands predicted from cloned *PBX3* cDNAs were observed for *PBX3*. The relative proportions of *PBX3* products varied somewhat in the different cell lines and tissues, although the smallest form predominated in most cases. A third, larger band observed in some of the *PBX3* PCR analyses represents amplification of non-PBX mRNA due to cross-hybridization of the oligonucleotide primers (data not shown) These data confirm that *PBX3* but not *PBX2* RNAs may be alternatively processed to result in translation products that differ at their carboxy termini, although the tissue-specific differences in splicing patterns of *PBX3* are not as marked as those observed for *PBX1* (16a). Consistent with these data, *PBX1* and *PBX3* but not *PBX2* contain nucleotides matching the consensus splice sites (54) at the ends of the deleted sequence, a feature similar to one noted in the differentially spliced products of vHNF1 (50).

Detection of Pbx proteins. The various Pbx proteins were synthesized in vitro and shown to have migrations in SDS gels corresponding to their respective predicted molecular weights: Pbx1a, 46,500; Pbx1b, 38,400; Pbx2, 45,900; Pbx3a, 47,100; and Pbx3b, 38,800 (Fig. 6A). The highly conserved nature of these proteins was evidenced by the immunopre-

TABLE 1. Summary of PBX mRNA expression in human tissues and cell lines

Cell lines and tissue types	PBX mRNA expression		
	PBX1	PBX2	PBX3[a]
Cell lines			
Lymphoid (mature B cell)			
FL18	−	+	+
607B	−	++	++
SU-DUL5	−	++	+
SU-DHL8	−	+	+/−
SUP-B6 (DN)	−	++	ND
IM9	−	++	ND
Lymphoid (precursor B cell)			
REH	−	+	+
SUP-B13 (JD)	−	+	+
SUP-B2 (EI)	−	++	ND
HB	−	++	+++
Lymphoid (proMB)			
KG1	−	++	+++
Lymphoid (mature T cell)			
MOLT4	−	+	+
Lymphoid (early T cell)			
P30	−	++	++
SUP-T13 (LAC)	−	+	+/−
SUP-T11 (NL)	−	+++	++
SUP-T7 (JH)	−	+	+/−
RPMI8402	−	++	++
Lymphoid (monocytic)			
HL60	−	+	+/−
U937	−	++	+/−
Indeterminate			
DU528	−	++	+/−
SU-DHL1	−	+	+
SUP-M1 (TS)	−	+/−	+/−
SUP-T8 (EA)	−	++	+
Erythroid			
K562	++	+++	+
Fibroblast			
293S	+	++	++
Epithelial			
A431	+++	+	+/−
HeLa	+	++	+
ES2-3	−	++	+
PA-1	+	++	++
MCF-7	+	+	+
HepG2	−	+	+
Neuroectodermal			
HS683	+++	+	+
A172	+	+	+
Sarcoma			
DX-5	+	++	+++
Mes-sa	−	++	+
Tissue types			
Fetal (18–22 wk)			
Kidney	++	+	+/−
Heart	++	+	+
Thymus	+	+	+
Liver	+	+/++	+/−
Lung	++	++	+/−
Gastrointestinal	+	+	+
Adrenal	++	+	−
Brain	+++	+	+/−
Spleen	+	+	+
Adult			
Kidney	−	+	+
Heart	+	−	+
Thymus	−	+	+
Tonsil	+	++	+
Ovary	+	++	+++
Uterus	+	++	+
PBM[b]	+/−	++	+
Kidney carcinoma	+	+	+

[a] +/−, transcript detected after extensive exposure time; ND, not determined.
[b] PBM, peripheral blood mononuclear cells.

cipitation of Pbx2 and Pbx3a by polyclonal antibodies directed against Pbx1a (Fig. 6B). Since these antibodies did not recognize the Pbx1b and Pbx3b proteins, they are apparently specific for the highly conserved C termini of Pbx1a, Pbx2, and Pbx3a which are not present in Pbx1b and Pbx3b because of frameshifts resulting from alternative processing of the *PBX1* and *PBX3* transcripts.

The expression of native Pbx proteins was detected by using these antibodies in Western blot analyses of K562, a cell line which expresses very high levels of *PBX2*, and SUP-T13, which expresses lower levels of *PBX2* and *PBX3* transcripts (Table 1). Since all human cell lines express *PBX2* mRNA, the murine pre-B-cell line I-8 was examined as a negative control, and I-8 transfected with a retroviral construct expressing Pbx1a served as a positive control. Immunoreactive bands detected in K562 and SUP-T13 migrated at approximately 47 kDa, which is similar to the migration of retrovirally expressed Pbx1a. The protein detected in these cell lines most likely corresponds to Pbx2, since these lines express low or undetectable levels of RNA for the comparably sized proteins Pbx1a and Pbx3a (data not shown). Although these cell lines do express transcripts for one or both of the 38-kDa Pbx1b and Pbx3b proteins, our anti-Pbx antibodies do not react with these proteins, as shown above. The immunoreactive Pbx protein detected in K562 and SUP-T13 cells is likely to be the same as that observed by Kamps et al. (28) in various lymphoid cell lines that had a reported molecular weight of 42,000.

DISCUSSION

The human *PBX2* and *PBX3* genes encode proteins with extensive similarity to Pbx1, a homeodomain protein involved in t(1;19) chromosomal translocation in acute pre-B-cell ALL. The Pbx proteins are particularly notable for their extensive sequence conservation outside of the homeodomains (92 to 94% over 266 amino acids). This extent of sequence identity has not been observed previously in homeodomain proteins, although smaller conserved regions have been reported. The POU proteins Pit-1, Oct-1, Oct-2, Unc-86, Brn-1, Brn-2, Brn-3, and Tst-1 share a conserved region of 75 to 82 amino acids (POU-specific box) immediately upstream of the homeodomain (9, 13, 21, 23, 25, 55). The paired box (5) 128 to 135 amino acids in length is located amino terminal of the homeodomains of the Drosophila paired, gooseberry proximal, and gooseberry distal proteins (3, 30). The murine *En-1* and *En-2* and the *Drosophila en* and *inv* share three conserved regions other than their homeo-

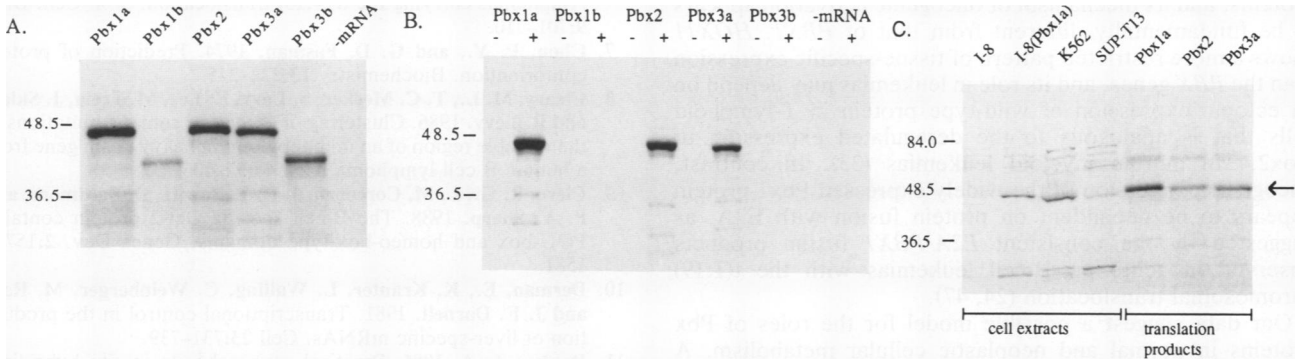


FIG. 6. In vitro translation, immunoprecipitation, and immunoblotting of Pbx proteins. (A) In vitro-translated Pbx1a, Pbx1b, Pbx2, Pbx3a, and Pbx3b were analyzed by SDS-PAGE. Lane −mRNA contains translation products obtained in the absence of input mRNA (negative control). (B) Immunoprecipitation of in vitro-translated proteins from reactions shown in panel A using rabbit anti-Pbx1 antiserum (16a). Immunoprecipitations were carried out with preimmune sera (lanes −) or anti-Pbx1 immune sera (lanes +). Lanes −mRNA are immunoprecipitations of in vitro translation products obtained in the absence of input mRNA (negative control). (C) Identification of native Pbx proteins by immunoblotting. Whole cell extracts were subjected to SDS-PAGE, and immunoblotting with anti-Pbx1 antibodies was performed. Cell lines analyzed: I-8, a mouse pre-B; I-8(Pbx1a), I-8 transfected with retrovirus expressing Pbx1a; K562, human erythroid; SUP-T13, human early T. In vitro-translated Pbx1a, Pbx2, and Pbx3a were run as positive controls for Western blotting and for comparisons of migrations of native proteins in whole-cell extracts. M_r values were determined by comparisons with prestained protein standards.

domains (26, 27), two located N terminal and one immediately C terminal of their homeodomains. Most analogous to the Pbx proteins are the Antennapedia-type proteins, which have significant homology over their entire lengths (see reference 58 for comparisons), although they are not as similar as the Pbx proteins.

Their extensive homology suggests that the Pbx proteins may share functional similarities which are mediated by protein motifs within the conserved regions. The limited divergence within and flanking the homeodomains is consistent with the possibility that the Pbx proteins bind to a common DNA sequence, particularly since the few amino acid differences observed in the homeodomains are not at positions believed to be involved in DNA recognition (31). HNF1 and vHNF1, which are highly similar homeodomain proteins, bind an identical sequence and are both capable of transactivating the albumin promoter in transfection assays (50). However, the important determinants for in vivo sequence-specific DNA binding by homeodomain proteins have not been completely elucidated. Although POU proteins with similar homeodomains have been reported to bind different DNA sequences (55), these studies may not predict the properties of Pbx proteins, since their similarities are more extensive than those of the POU proteins. The Pbx proteins may exhibit different affinities for the same target DNA sequence which are influenced by slight structural alterations due to their amino acid differences. In this case, their relative ratios and affinities would determine which Pbx protein was bound to the target site and would be consistent with a competition model for protein activity. The extensive homology of Pbx1, Pbx2, and Pbx3 could prove useful in determining the crucial regions for DNA and protein-protein interactions of homeodomain proteins.

Although remarkably similar throughout much of their lengths, the Pbx proteins are distinguished by extensive divergence at their C and N termini. C-terminal diversity results both from coding differences between the three *PBX* genes and also posttranscriptional mechanisms due to differential splicing (*PBX1* and *PBX3*). Differential splicing is a feature of many homeobox transcripts that has unknown functional consequences on the resultant proteins. However, the transcription factor NF-1 (a nonhomeodomain protein) also exhibits differential splicing at the C-terminal region similar to that observed for *PBX1* and *PBX3*, which results in long and short forms of the protein (51). As demonstrated elsewhere (42), this region of NF-1 is important for transcriptional activation, and the shorter form of the NF-1 protein has reduced activity by 40%. It is unclear from our data whether alternative forms of Pbx1 and Pbx3 differ in their biologic properties, but the smaller forms of these proteins appear more prevalent according to transcript analyses by PCR. Since splice variants of *PBX2* were not detected, a single Pbx2 translation product most comparable to the Pbx1a and Pbx3a polypeptides is predicted and supports a functional role for the long forms of the Pbx proteins. Recent data suggest that C-terminal differences of the E2A-Pbx1 fusion proteins may influence their abilities to induce morphologic transformation of NIH 3T3 cells; however, these differences did not affect tumorigenicity in nude mice (28).

Expression of *PBX* genes was observed in a wide variety of adult and fetal tissues, except that *PBX1* does not appear to be expressed in cells of the lymphoid lineage (16a). Ubiquitous expression is unusual for homeobox genes, which typically show temporally and/or spatially restricted expression during development. The observed widespread distribution of *PBX* transcripts is consistent with a more generalized role for Pbx proteins that is possibly analogous to the role of the ubiquitous homeodomain protein Oct-1. Oct-1 has been implicated in the regulation of housekeeping genes such as histone H2b (14) and U4B small nuclear RNA genes (59) and also in DNA replication (45, 48, 56). Oct-1 can also participate in lineage-specific gene expression, as Zhou and Benedict Yen (60) have demonstrated that Oct-1 and the liver-specific transcription factor HNF1 were both necessary for liver-specific transcription of the hepatitis B virus pre-S1 promoter.

Recently, a new homeobox gene, *HOX11*, was reported in T-cell ALL (20). The sequence of the predicted Hox11 protein is significantly divergent from those of the Pbx

proteins, and its mechanism of oncogenic activation appears to be fundamentally different from that of *PBX1*. *HOX11* shows a more restricted pattern of tissue-specific expression than the *PBX* genes, and its role in leukemias may depend on an ectopic expression of wild-type protein in T-lymphoid cells that is analogous to the deregulated expression of Hox2.4 in murine myeloid leukemias (33). In contrast, oncogenic conversion of the widely expressed Pbx1 protein appears to be dependent on protein fusion with E2A, as suggested by the consistent *E2A-PBX1* fusion products observed in acute pre-B-cell leukemias with the t(1;19) chromosomal translocation (24, 47).

Our data suggest a possible model for the roles of Pbx proteins in normal and neoplastic cellular metabolism. A small family of widely expressed, highly conserved Pbx homeodomain proteins suggests a generalized activity in most tissues. Extensive homology within and outside of their homeodomains may indicate competition for a core DNA-binding site, with resultant regulation of an overlapping set of target genes. Regions of sequence divergence may constitute domains that mediate associations with accessory factors that modulate their regulatory function, as demonstrated for Oct-1 (36). These interactions may be further refined and regulated by additional variability that arises through differential splicing of the primary transcripts. Within the context of this model, the fusion of Pbx1 to E2A resulting from t(1;19) can be viewed as replacement of an interaction domain with profound consequences on gene regulation. The differential expression of *PBX1* versus *PBX2* and *PBX3* in lymphoid cells may contribute to the transforming effects of E2A-Pbx1 in pre-B-cell leukemias. This model is consistent with emerging concepts of the role of homeodomain proteins in gene regulation and is ultimately testable by a gene transfer approach.

ACKNOWLEDGMENTS

We thank Pratima Jog for the I-8 cells, Pam McFall for technical assistance with tissue culture, Phil Verzola and Judith Quenvold for photography, and especially Mark Krasnow for critical reading of the manuscript.

This work was supported by National Institutes of Health grants CA42971 and CA49605. M.L.C. is a Scholar of the Leukemia Society of America.

REFERENCES

1. **Allen, J. D., T. Lints, N. A. Jenkins, N. G. Copeland, A. Strasser, R. P. Harvey, and J. M. Adams.** 1991. Novel murine homeo box gene on chromosome 1 expressed in specific hematopoietic lineages and during embryogenesis. Genes Dev. **5**:509–520.
2. **Barad, M., T. Jack, R. Chadwick, and W. McGinnis.** 1988. A novel, tissue-specific, *Drosophila* homeobox gene. EMBO J. **7**:2151–2161.
3. **Baumgartner, S., D. Bopp, M. Burri, and M. Noll.** 1987. Structure of two genes at the *gooseberry* locus related to the *paired* gene and their spatial expression during *Drosophila* embryogenesis. Genes Dev. **1**:1247–1267.
4. **Boncinelli, E., R. Sommo, D. Acampora, M. Pannese, M. D'Esposito, A. Faiedla, and A. Simeone.** 1988. Organization of human homeobox genes. Hum. Reprod. (Oxford) **3**:880–886.
5. **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.
6. **Chen-Levy, Z., J. Nourse, and M. L. Cleary.** 1989. The *bcl-2* candidate proto-oncogene product is a 24-kilodalton integral-membrane protein highly expressed in lymphoid cell lines and lymphomas carrying the t(14;18) translocation. Mol. Cell. Biol. **9**:701–710.
7. **Chou, P. Y., and G. D. Fasman.** 1974. Prediction of protein conformation. Biochemistry **13**:222–245.
8. **Cleary, M. L., T. C. Meeker, S. Levy, E. Lee, M. Trela, J. Sklar, and R. Levy.** 1986. Clustering of extensive somatic mutations in the variable region of an immunoglobulin heavy chain gene from a human B cell lymphoma. Cell **44**:97–106.
9. **Clerc, R. G., L. M. Corcoran, J. H. LeBowitz, D. Baltimore, and P. A. Sharp.** 1988. The B-cell specific Oct-2 protein contains POU-box and homeo-box-type domains. Genes Dev. **2**:1570–1581.
10. **Derman, E., K. Krauter, L. Walling, C. Weinberger, M. Ray, and J. E. Darnell.** 1981. Transcriptional control in the production of liver-specific mRNAs. Cell **23**:731–739.
11. **Donlon, T. A.** 1986. Practical approaches to in situ hybridization. Karyogram **12**:3–10.
12. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:6–13.
13. **Finney, M., G. Ruvkun, and H. R. Horvitz.** 1988. The C. elegans cell lineage and differentiation gene *unc-86* encodes a protein with a homeodomain and extended similarity to transcription factors. Cell **55**:757–769.
14. **Fletcher, C., N. Heintz, and R. G. Roeder.** 1987. Purification and characterization of OTF-1, a transcription factor regulating cell cycle expression of a human histone H2b gene. Cell **51**:773–781.
15. **Frain, M., G. Swart, P. Monaci, A. Nicosia, S. Stampfli, R. Frank, and R. Cortese.** 1989. The liver-specific transcription factor LR-B1 contains a highly diverged homeobox DNA binding domain. Cell **59**:145–157.
16. **Frigerio, G., M. Burri, D. Bopp, S. Baumgartner, and M. Noll.** 1986. Structure of the segmentation gene *paired* and the Drosophila PRD gene set as part of a gene network. Cell **47**:735–746.

16a. **Galili, N., et al.** Unpublished data.

17. **Gehring, W. J., and Y. Hiromi.** 1986. Homeotic genes and the homeobox. Annu. Rev. Genet. **20**:147–173.
18. **Harper, M. E., and G. F. Saunder.** 1981. Localization of single copy DNA sequences on G-banded human chromosomes by in situ hybridization. Chromosoma **83**:431–439.
19. **Hart, C. P., A. Fainsod, and F. H. Ruddle.** 1987. Sequence analysis of the murine *Hox*-2.2, -2.3, and -2.4 homeo boxes: evolutionary and structural comparisons. Genomics **1**:182–195.
20. **Hatano, M., C. W. M. Roberts, M. Minden, W. M. Crist, and S. J. Korsmeyer.** 1991. Deregulation of a homeobox gene, *Hox11*, by the t(10;14) in T cell leukemia. Science **253**:79–82.
21. **He, X., M. N. Treacy, D. M. Simmons, H. A. Ingraham, and H. W. Swanson, and M. G. Rosenfeld.** 1989. Expression of a large family of POU-domain regulatory genes in mammalian brain development. Nature (London) **340**:35–42.
22. **Heim, S., and F. Mitelman.** 1987. Cancer cytogenetics, p. 70. Alan R. Liss, Inc., New York.
23. **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. Horvitz.** 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.
24. **Hunger, S. P., N. Galili, A. J. Carroll, W. M. Crist, M. P. Link, and M. L. Cleary.** 1991. The t(1;19) (q23;p13) results in consistent fusion of E2A and PBX1 coding sequences in acute lymphoblastic leukemias. Blood **77**:687–693.
25. **Ingraham, H. A., R. Chen, H. J. Mangalam, H. P. Elsholtz, S. E. Flynn, C. R. Lin, D. M. Simmons, L. Swanson, and M. G. Rosenfeld.** 1988. A tissue-specific transcription factor containing a homeodomain specifies a pituitary phenotype. Cell **55**:519–529.
26. **Joyner, A. L., K. Herrup, B. A. Auerbach, C. A. Davis, and J. Rossant.** 1991. Subtle cerebellar phenotype in mice homozygous for a targeted deletion of the *En-2* homeobox. Science **251**:1239–1243.
27. **Joyner, A. L., and G. R. Martin.** 1987. *En-1* and *En-2*, two mouse genes with sequence homology to the *Drosophila en-*

grailed gene: expression during embryogenesis. Genes Dev. **1**:29–38.
28. **Kamps, M. P., A. T. Look, and D. Baltimore.** 1991. The human t(1;19) translocation in pre-B ALL produces multiple nuclear E2A-Pbx1 fusion proteins with differing transforming potentials. Genes Dev. **5**:358–368.
29. **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.
30. **Kilchherr, F., A. Baumgartner, D. Bopp, E. Frei, and M. Noll.** 1986. Isolation of the *paired* gene of *Drosophila* and its spatial expression during early embryogenesis. Nature (London) **321**: 493–499.
31. **Kissinger, C. R., B. Liu, E. Martin-Blanco, T. B. Kornberg, and C. O. Pabo.** 1990. Crystal structure of an engrailed homeodomain-DNA complex at 2.8 Å resolution: a framework for understanding homeodomain-DNA interactions. Cell **63**:579–590.
32. **Kohler, S., N. Galili, J. L. Sklar, T. A. Donlon, K. G. Blume, and M. L. Cleary.** 1990. Expression of *bcr-abl* fusion transcripts following bone marrow transplantation for Philadelphia chromosome-positive leukemia. Leukemia **4**:541–547.
33. **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.
34. **Kornfeld, R., R. B. Saint, P. A. Beachy, P. J. Harte, D. A. Peattie, and D. S. Hogness.** 1989. Structure and expression of a family of Ultrabithorax mRNAs generated by alternative splicing and polyadenylation in Drosophila. Genes Dev. **3**:243–258.
35. **Kozak, M.** 1987. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res. **15**:8125–8132.
36. **Kristie, T. M., and P. A. Sharp.** 1990. Interactions of the Oct-1 POU subdomains with specific DNA sequences and with the HSV α-trans-activator protein. Genes Dev. **4**:2383–2396.
37. **Lonai, P., and A. Orr-Urtreger.** 1990. Homeogenes in mammalian development and the evolution of the cranium and central nervous system. FASEB J. **4**:1436–1443.
38. **Macdonald, P. M., P. Ingham, and G. Struhl.** 1986. Isolation, structure, and expression of *even-skipped*: a second pair-rule gene of Drosophila containing a homeo box. Cell **47**:721–734.
39. **Mavilio, F., A. Simeone, A. Giampaolo, A. Faiella, V. Zappavigna, D. Acampora, G. Poiana, G. Russo, C. Peschle, and E. Boncinelli.** 1986. Differential and stage-related expression in embryonic tissues of a new human homeobox gene. Nature (London) **324**:644–668.
40. **McGinnis, W., R. L. Garber, J. Wirz, A. Kuroiwa, and W. J. Gehring.** 1984. A homologous protein-coding sequence in Drosophila homeotic genes and its conservation in other metazoans. Cell **37**:403–408.
41. **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 (London) **308**:428–433.
42. **Mermod, N., E. A. O'Neill, T. J. Kelly, and R. Tjian.** 1989. The proline-rich transcriptional activator of CTF/NF-1 is distinct from the replication and DNA binding domain. Cell **58**:741–753.
43. **Mitchell, P. J., and R. Tjian.** 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science **245**:371–378.
44. **Monica, K., Z. Chen-Levy, and M. L. Cleary.** 1990. Small G proteins are expressed ubiquitously in lymphoid cells and do not correspond to Bcl-2. Nature (London) **346**:189–191.
45. **Mul, Y. M., C. P. Verrijzer, and P. C. Van Der Vliet.** 1990. Transcription factors NFI and NFIII/oct-1 function independently, employing different mechanisms to enhance adenovirus DNA replication. J. Virol. **64**:5510–5518.
46. **Nicosia, A., P. Monaci, L. Tomei, R. De Francesco, M. Nuzzo, H. Stunnenberg, and R. Cortese.** 1990. A myosin-like dimerization helix and an extra-large homeodomain are essential elements of the tripartite DNA binding structure of LFB1. Cell **61**:1225–1236.
47. **Nourse, J., J. D. Mellentin, N. Galili, J. Walkinson, 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.
48. **O'Neill, E. A., C. Fletcher, C. R. Burrow, N. Heintz, R. G. Roeder, and T. J. Kelly.** 1988. Transcription factor OTF-1 is functionally identical to the DNA replication factor NF-III. Science **241**:1210–1213.
49. **Poole, S. J., L. M. Kauvar, B. Drees, and T. Kornberg.** 1985. The *engrailed* locus of Drosophila: structural analysis of an embryonic transcript. Cell **40**:37–43.
50. **Rey-Campos, J., T. Chouard, M. Yaniv, and S. Cereghini.** 1991. vHNF1 is a homeoprotein that activates transcription and forms heterodimers with HNF1. EMBO J. **10**:1445–1457.
51. **Santoro, C., N. Mermod, P. C. Andrews, and R. Tjian.** 1988. A family of human CCAAT-box-binding proteins active in transcription and DNA replication: cloning and expression of multiple cDNAs. Nature (London) **334**:218–224.
52. **Scott, M. P., J. W. Tamkun, and G. W. Hartzel III.** 1989. The structure and function of the homeodomain. Biochim. Biophys. Acta **989**:25–48.
53. **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. USA **81**:4115–4119.
54. **Shapiro, M. B., and P. Senapathy.** 1987. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. Nucleic Acids Res. **15**:7155–7174.
55. **Sturm, R. A., G. Das, and W. Herr.** 1988. The ubiquitous octamer-binding protein Oct-1 contains a POU domain with a homeo box subdomain. Genes Dev. **2**:1582–1599.
56. **Verrijzer, C. P., A. J. Kal, and P. C. Van der Vliet.** 1990. The DNA binding domain (POU domain) of transcription factor oct-1 suffices for stimulation of DNA replication. EMBO J. **9**:1883–1888.
57. **Wharton, K. A., B. Yedvobnick, V. G. Finnerty, and S. Artavanis-Tsakonas.** 1985. *opa*: a novel family of transcribed repeats shared by the *Notch* locus and other developmentally regulated loci in D. melanogaster. Cell **40**:55–62.
58. **Wright, C. V. E., K. W. Y. Cho, G. Oliver, and E. M. DeRobertis.** 1989. Vertebrate homeodomain proteins: families of region-specific transcription factors. Trends Biochem. Sci. **14**: 52–56.
59. **Zamrod, Z., and W. E. Stumph.** 1990. U4B snRNA gene enhancer activity requires functional octamer and SPH motifs. Nucleic Acids Res. **18**:7323–7330.
60. **Zhou, D., and T. S. Benedict Yen.** 1991. The ubiquitous transcription factor Oct-1 and the liver-specific factor HNF-1 are both required to activate transcription of a hepatitis B virus promoter. Mol. Cell. Biol. **11**:1353–1359.