Am. J. Hum. Genet. 41:1–15, 1987

Human Homeo Box–containing Genes Located at Chromosome Regions 2q31→2q37 and 12q12→12q13

L. A. CANNIZZARO,*[,1] C. M. CROCE,† C. A. GRIFFIN,‡ A. SIMEONE,§ E. BONCINELLI,§ AND K. HUEBNER†

*Division of Genetics, Milton S. Hershey Medical Center, Pennsylvania State University, Hershey, PA 17033; †The Wistar Institute of Anatomy and Biology, Philadelphia 19104; ‡Department of Pediatrics, University of Pennsylvania School of Medicine, Philadelphia 19104; and §International Institute of Genetics and Biophysics, V. Marconi 10, 80125 Naples, Italy

SUMMARY

Four human homeo box–containing cDNAs isolated from mRNA of an SV40-transformed human fibroblast cell line have been regionally localized on the human gene map. One cDNA clone, c10, was found to be nearly identical to the previously mapped Hox-2.1 gene at 17q21. A second cDNA clone, c1, which is 87% homologous to Hox-2.2 at the nucleotide level but is distinct from Hox-2.1 and Hox-2.2, also maps to this region of human chromosome 17 and is probably another member of the Hox-2 cluster of homeo box–containing genes. The third cDNA clone, c8, in which the homeo box is ~84% homologous to the mouse Hox-1.1 homeo box region on mouse chromosome 6, maps to chromosome region 12q12→12q13, a region that is involved in chromosome abnormalities in human seminomas and teratomas. The fourth cDNA clone, c13, whose homeo box is ~73% homologous to the Hox-2.2 homeo box sequence, is located at chromosome region 2q31→q37. The human homeo box–containing cluster of genes at chromosome region 17q21 is the human cognate of the mouse homeo box–containing gene cluster on mouse chromosome 11. Other mouse homeo box–containing genes of the Antennapedia class (class I) map to mouse chromosomes 6 (Hox-1, proximal to the IgK locus) and 15 (Hox-3). A mouse gene, En-1, with an engrailed-like homeo box (class II) and flanking region maps to mouse chromosome 1 (near the domi-

1. Present address: The Genetics Center, 6401 East Thomas Road, Scottsdale, AZ 85251.
Address for correspondence and reprints: Dr. K. Huebner, The Wistar Institute of Anatomy and Biology, 36th Street at Spruce, Philadelphia, PA 19104.

 0002-9297/87/4101-0001$02.00

nant hemimelia gene). Neither of the class I homeo box–containing genes—c8 and c13—maps to a region of obvious homology to chromosomal positions of the presently known mouse homeo box–containing genes.

INTRODUCTION

Drosophila homeo box sequences are 183-bp elements within exons of genes involved in the determination of body-segment number and identity. Homeo box–containing genes include members of the Antennapedia (ANT-C) and bithorax (BX-C) complexes as well as others involved in *Drosophila* morphogenesis, for example, the engrailed locus (Laughon and Scott 1984; McGinnis et al. 1984*a*, 1984*c;* Scott and Weiner 1984; Fjose et al. 1985; Poole et al. 1985). Although functions of the homeo box sequences have not been determined, they are present in open reading frames within genes whose protein products are localized in the nucleus (White and Wilcox 1984; Beachy et al. 1985; Laughon et al. 1985).

In *Drosophila*, many homeotic genes are grouped in two gene clusters on the right arm of the third chromosome, the ANT-C and BX-C complexes (Lewis 1978; Kaufman et al. 1980). More than 10 *Drosophila* genes, mapping mainly in these complexes, share the conserved 183-bp homeo box sequence (McGinnis et al. 1984*a*, 1984*c;* Gehring 1985). The ANT-C has at least five genes—including Antennapedia (Antp) (McGinnis et al. 1984*a*) and fushi tarazu (ftz) (McGinnis et al. 1984*a;* Scott and Weiner 1984)—containing the homeo box, whereas the BX-C contains at least three homeo boxes, corresponding to three lethal complementation groups (McGinnis et al. 1984*a*). Outside the two main clusters in *Drosophila*, other genes containing the homeo box have also been cloned. Among them, a homeo box has been identified within the engrailed (en) gene and a closely linked gene designated engrailed related (en-r) or invected (inv) (Fjose et al. 1985; Poole et al. 1985).

More than 11 copies of the homeo box have been characterized in flies (for review, see Gehring 1985; Manley and Levine 1985). Mutations in any of these gene complexes alter the normal embryonic segmentation pattern. It has been suggested that these loci specify particular pathways of segment morphogenesis by directly regulating larger "batteries" of genes (Garcia-Bellido 1977; Manley and Levine 1985).

With *Drosophila* homeo box DNA probes, cross-hybridizing sequences have been identified in frogs (Carrasco et al. 1984; Mueller et al. 1984), mice (McGinnis et al. 1984*b;* Colberg-Poley et al. 1985; Hart et al. 1985; Joyner et al. 1985*a*, 1985*b;* Rabin et al. 1985), and man (Levine et al. 1984; Boncinelli et al. 1985). The vertebrate and *Drosophila* sequences exhibit from 70% to >90% sequence conservation, and in mouse and man the number of copies of homeo boxes detected is of the same order of magnitude as that found in *Drosophila*. As pointed out in a recent review (Manley and Levine 1985), there is increasing

evidence consistent with the view that mammalian homeo box–containing genes are analogous to *Drosophila* homeo box–containing genes: they are, at least in some cases, clustered in mouse and in man, and they are expressed during development and thus may have, as in *Drosophila*, important functions in controlling development.

Boncinelli et al. (1985) have isolated cDNA clones c1, c10, c8, and c13 containing class I homeo boxes from an SV40-transformed human fibroblast cDNA library and have sequenced these genes (Boncinelli et al. 1985) and studied their expression in human embryos (Mavilio et al. 1986; Simeone et al. 1986, 1987).

Previously described human homeo box–containing genes Hox-2.1 and Hox-2.2 (see Martin et al. [1987] for Hox gene nomenclature) have been assigned to human chromosome region 17q21, whereas mouse homeo box–containing genes have been assigned to four unlinked loci: the Hox-1 locus containing three or more homeo box–containing genes is on mouse chromosome 6 proximal to the 1qK locus (McGinnis et al. 1984*b;* Rabin et al. 1985); the Hox-2 locus containing at least four homeo boxes within ~30 kb is on mouse chromosome 11 and is the mouse equivalent of the human homeo box cluster (Hox-2) at 17q21 (Joyner et al. 1985*b;* Rabin et al. 1985); the mouse Hox-3 locus, containing thus far one homeo box, EA, has been mapped to mouse chromosome 15 (Ruddle et al. 1985); and a fourth murine homeo box–containing gene—the murine engrailed-like homeo box gene, En-1—maps to mouse chromosome 1 (near the dominant hemimelia gene) (Joyner et al. 1985*a*). All mammalian homeo box–containing genes thus far examined are expressed—albeit not exclusively—in embryos and/or teratocarcinoma cells.

To determine the relationship of the homeo box–containing genes represented by the cDNA clones c1, c10, c8, and c13 to previously described mammalian homeo box–containing genes, Boncinelli et al. (1985) have sequenced homeo box and flanking sequences and we have determined the chromosomal location of the cDNA clones by somatic-cell hybrid analysis and in situ hybridization to metaphase chromosomes. Each cDNA probe was used to detect the presence of homologous sequences in a panel of rodent-human hybrid cells retaining defined complements of human chromosomes, and presence in hybrids of each homeo box–containing cDNA was correlated with the presence of a specific human chromosome region. In situ hybridization of the same individual homeo box–containing cDNA probes to human metaphase chromosomes confirmed chromosomal localization and refined the regional localization on specific human chromosomes.

MATERIAL AND METHODS

Homeo Box–containing cDNA Clones

Antp/ftz homologue containing cDNA clones obtained from a cDNA library from SV40-transformed human fibroblast, GM637B (Genetic Mutant Cell Repository, Camden, NJ), have been described (Boncinelli et al. 1985).

In brief, c1, c10, c8, and c13 represent transcripts from four different genes.

The c1 homeo box shares 85% direct sequence homology with the Antp homeo box and 86.9% nucleotide-sequence homology with the human Hox-2.2 homeo box. The cDNA clone is an ~900-bp insert in the Okayama-Berg vector (Okayama and Berg 1983). The insert of clone c10 shares 80% nucleotide-sequence homology with the Antp homeo box region, is nearly identical to the human Hox-2.1 genomic sequence (Levine et al. 1984), and is probably a transcript from the Hox-2.1 sequence. The predicted peptide sequence of the homeo domain in clone c8 is identical to that of the homeo domain of the *Xenopus* clone AC1 (Mueller et al. 1984). The cDNA insert of c8 is ~1,100 bp in the Okayama-Berg vector. The homeo box contained in clone c13 (~1,300 bp inserted in the Okayama-Berg vector) is uniformly related to most homeo boxes since it shares 70% nucleotide-sequence homology with every class I homeo box analyzed.

All homeo box–containing cDNA clones (c1, c8, and c13) were used as entire plasmid both for Southern blot analysis of hybrid DNAs and for in situ chromosome hybridization. The homeo domains within the cDNAs represented only ~10% of the entire insert, and cross-hybridization with other homeo box–containing genes was not a problem.

Probes for Southern blot analysis or for in situ hybridization were radiolabeled by means of nick-translation with α-^{32}P or α-^{3}H deoxynucleotide triphosphates (dNTPs), respectively.

Molecular Probes Used as Chromosomal Markers

DNA probes mapping to each human chromosome and in some cases both arms of a human chromosome have been used to characterize the somatic-cell hybrid panel (Huebner et al. 1985*a*, 1985*b*, 1986*a*, 1986*b*, 1986*c;* Isobe et al. 1985, 1986; Nagarajan et al. 1986). For characterization of regions of chromosomes 2, 12, and 17 present in the panel we have used the following: a Leu-2 (T8) probe that maps in the 2p1 region near the IgK locus (Sukhatme et al. 1985) and α1 (III) type III collagen chain probe that maps to 2q24.3→2q31 (Emanuel et al. 1985); a T4 cDNA probe that maps to 12p12→12p13 (Isobe et al. 1986) and a K-ras-2 fourth-exon probe (McBride et al. 1983) that detects the K-ras-2 locus on chromosome 12 that has been mapped to the short arm of 12 (Ryan et al. 1983; Popescu et al. 1985); and an erbA cDNA probe that maps to 17q12→q21 (Dayton et al. 1984) proximal to the characteristic acute promyelocytic leukemia (APL) chromosome 17 breakpoint seen in the t(15;17)(q22;q21) translocation and a nerve growth-factor receptor (NGFR) cDNA clone that maps to 17q21 distal to the APL chromosome 17 breakpoint (Huebner et al. 1986*b*). Total plasmid DNA was radiolabeled with α-^{32}P dNTPs and hybridized to Southern blots containing DNA from the somatic-cell hybrid panel.

Cells

Isolation, propagation, and characterization of parental cells and somatic-cell hybrids used in this study have been described elsewhere (Dalla Favera et al. 1982; Dayton et al. 1984; Huebner et al. 1985*a*, 1985*b*, 1986*a*, 1986*b*, 1986*c;* Isobe et al. 1985, 1986; Sukhatme et al. 1985; Nagarajan et al. 1986). Hybrids

were characterized for expression of enzyme markers assigned to each of the human chromosomes (Dalla Favera et al. 1982). Some hybrid clones were karyotyped by means of trypsin/Giemsa and/or G-11-banding methods as described elsewhere (Dalla Favera et al. 1982). In addition, the presence of specific human chromosomes in many of the mouse-human hybrids has been confirmed by means of DNA hybridization using probes for genes assigned to specific human chromosomes.

Chromosomes or partial chromosomes present in most of the hybrid cells used in this study are diagrammatically depicted in figure 1. Partial chromosomal regions present in hybrids were defined by chromosomal banding methods and/or by determining presence or absence in hybrid DNA of specific gene sequences that previously have been localized to specific chromosome regions.

For regional localization of the c1 homeo box–containing gene on chromosome 17, two hybrids retaining translocation chromosomes were used. Hybrid 275S retains a $21q^+$ chromosome—t(17;21) (17q21→17qter::21pter→21q22)—in the absence of other human chromosomes (Dayton et al. 1984). Hybrid 570 retains a $15q^+$ chromosome (from an APL tumor)—t(15;17)(15pter→15q22::17q21→17qter)—in the absence of the reciprocal (containing 17pter→17q21) and the normal 17 (J. Finan, P. Nowell, K. Huebner, and C. M. Croce, unpublished data).

Southern Blot Analysis

DNAs from human peripheral blood lymphocytes (PBL) or human cell lines, from mouse cell lines, and from rodent-human hybrid cell lines were extracted by means of cell lysis, proteinase K digestion, phenol extraction, and ethanol precipitation. Cellular DNAs were digested with an excess of restriction enzyme *Hin*dIII, sized in 0.8% agarose gels, and transferred to nitrocellulose or nylon filters in the manner described by Southern (1975). Hybridization was carried out in 50% formamide, 4 × NaCl/Cit (1 × NaCl/Cit = 0.15 M NaCl/0.015 M sodium citrate, pH 7.0), 0.2 mg of sonicated salmon sperm DNA/ml, 1 × Denhardt's solution (0.02% bovine serum albumin/0.02% Ficoll/0.02% polyvinylpyrrolidone) at 42 C for 15 h. Some hybridizations were performed at 68 C without formamide. After hybridization, filters were washed (0.1 × SSC, 0.1% sodium dodecyl sulfate at 65 C) and exposed to Kodak XAR-5 film with intensifying screens.

Chromosomal In Situ Hybridization

Metaphase chromosome preparations were obtained by culturing PBL from a normal male subject (46, XY) for 96 h in RPMI medium supplemented with 15% fetal bovine serum. Cultures were harvested according to standard procedure.

Probes were radiolabeled with tritium [^{3}H] to a specific activity of 4×10^7 cpm/μg DNA. In situ hybridization was performed using a modification of the standard protocol (Harper and Saunders 1981; Cannizzaro and Emanuel 1984). Slides were aged 10–14 days at 4 C and treated with ribonuclease A (Sigma) for

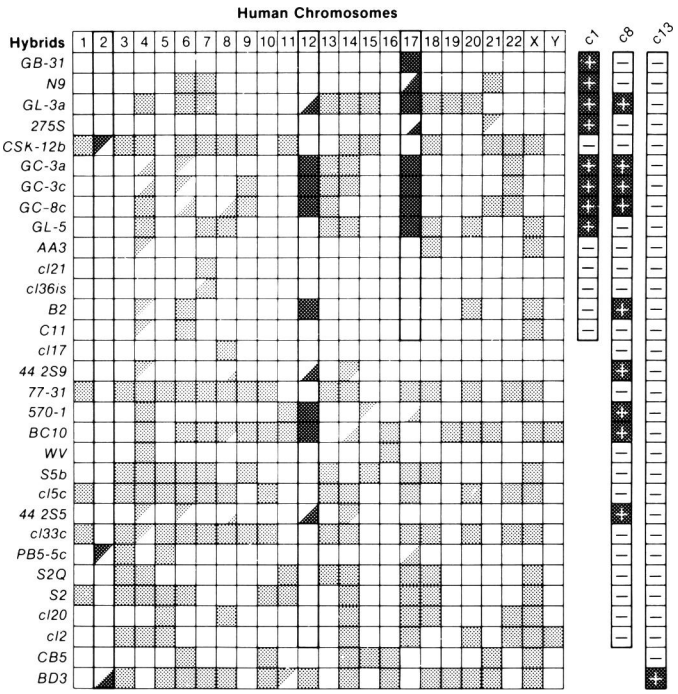

FIG. 1.—Presence of human homeo box–containing genes c1, c8, and c13 in a panel of 32 rodent-human hybrids. ▣ Indicates that the hybrid named in the left column contains the chromosome indicated in the upper row. □ Indicates presence of the long arm (or in some cases a part of the long arm, indicated by a smaller fraction of stippling) of the chromosome shown above the column. ▣ Indicates presence of the short arm of the chromosome listed above the column. □ Indicates absence of the chromosome listed above the column. The columns for chromosomes 2, 12, and 27 are boldly outlined and stippled to emphasize correlation of presence of these chromosomes (or region of these chromosomes) with presence of the c13, c8, and c1 sequences. The pattern of retention of the c1, c8, and c13 hybrids in the panel is shown in the three columns to the right of the figure, where presence of the sequences in the hybrid is indicated by a stippled box with a plus sign and absence of the gene is indicated by an open box enclosing a minus sign.

1 h at 37 C. Chromosomal DNA was denatured at 70 C for 2 min in a mixture of 70% formamide and 2 × SSC (pH 7.0). Probe DNA was denatured in a hybridization mixture of 25% formamide, 2 × SSC, and 10% dextran sulfate (pH 7.0). The final concentration of probe DNA placed on each slide was 0.035 μg DNA/ml. Hybridization was carried out at 37 C for 18 h. Slides were subsequently rinsed at 39 C, dehydrated in an alcohol series, and air dried. Slides were then dipped in nuclear track emulsion (Kodak NTB-2), fan dried, and stored in light-tight boxes at 4 C. At different time intervals, slides were developed and fixed at 15 C, air dried, then stained according to a modified Wright's-Giemsa staining protocol (Cannizzaro and Emanuel 1984). Each slide was then dipped for 10–16 s in a pH 9.2 borate buffer (50 mM $NaSO_4$, 2.5 mM $Na_2B_4O_7$) at 35 C and stained for 4–5 min in a staining dish containing 2 parts pH 9.2 buffer and 1 part Wright's-Giemsa stain (J. T. Baker and Co.). Grains situated on nonoverlapping chromosome regions were counted and scored.

RESULTS

Human Homeo Box–containing Gene c1 Maps to Chromosome Region 17q21→q22

Since human homeo box–containing genes Hox-2.1 and Hox-2.2 (and now Hu-5 [Hauser et al. 1985]) had already been mapped to human chromosome 17, we first tested a small panel of hybrid cells containing entire and broken chromosome 17's for presence of the three homeo box–containing cDNAs (c1, c8, and c13); c10, by virtue of its identity to Hox-2.1, was known to map in the region of chromosome 17 near the thymidine kinase gene. The loci detected by the c8 and c13 cDNAs did not segregate with chromosome 17 (fig. 1; also see below), but the gene detected by the c1 probe, which in *Hin*dIII-digested DNAs detects a mouse band of ~9.5 kbp (fig. 2, lane 8) and a human-specific fragment of ~4.5 kbp (fig. 2, lane 1), segregated concordantly with human chromosome region 17q21→17qter (see figs. 1, 2). Lane 3 of figure 2 retains only a $21q^+$ chromosome—t(17;21)(17q21→17qter:21pter→21q22)—and is positive for c1. In situ hybridization of the c1 probe to metaphase chromosomes (not shown) was consistent with a localization of this gene to 17q21→q22, distal to the 17q21 breakpoint seen in the t(15;17)(q22;q21) chromosome translocation in APL. This region of chromosome 17 also carries the thymidine kinase gene and the gene for the NGFR (Huebner et al. 1986*b*). Although this has not been proven conclusively, c1 most probably represents a member of the cluster of homeo box–containing genes in the Hox-2 complex that includes Hox-2.1, Hox-2.2, and Hu-5.

Human Homeo Box–containing Gene c8 Maps to Chromosome Region 12q12→q13

*Hin*dIII-digested DNAs from a panel of 29 rodent-human hybrid cells was tested for presence of the c8 gene (fig. 1). The 31-kbp *Hin*dIII-digested human

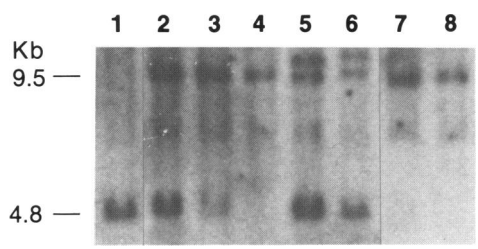

FIG. 2.—The c1 gene is at 17q21→qter. DNA (~10 μg/lane) from PBL (lane 1); hybrid N9 retaining human chromosomes 6, 7, 17q, and 21 (lane 2); hybrid 275S retaining $21q^+$ chromosome (21pter→21q22::17q21→17qter) (lane 3); hybrid CSK-12b retaining 1, 2p, 3, 4, 6–9, 11, 14, 15, 18, 21, 22, and X (lane 4); hybrid GC-3a retaining 4q, 6p, 12–14, 17, and 22 (lane 5); hybrid GC-3c retaining 4q, 6p, 9, 12–14, 17, and 22 (lane 6); hybrid AA3 retaining 4p, 18, and X (lane 7); and mouse cell line (lane 8) was cleaved with an excess of restriction enzyme *Hin*dIII, fractionated on an agarose gel, transferred to nitrocellulose filter, and hybridized to ^{32}P-labeled c1 DNA. Approximate molecular weights of mouse and human bands detected by the c1 probe are shown on the left. Additional bands for which molecular weights are not given are mouse cross-hybridizing sequences that are polymorphic in some strains of mice; e.g., the mouse parent of the hybrid in lane 4 was of the C3H strain, whereas the mouse parent of hybrids in lanes 5 and 6 was of C57BL/6 origin.

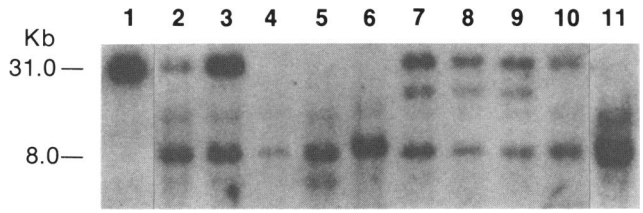

FIG. 3.—The c8 gene segregates with chromosome 12 in rodent-human somatic-cell hybrids. DNA (~10 μg/lane) from human PBL (lane 1); hybrid 44 2S5 retaining human chromosomes 4p, 6p, 8q24→8qter, 12q, and 14pter→q32 (lane 2); hybrid 44 2S9 retaining 4p, 8q24→8qter, 12q, and 14pter→q32 (lane 3); hybrid c133c retaining 1, 3, 4p, 5–10, 13, 14, 17, 18, 20, 22, and X (lane 4); hybrid 77-31 retaining 1, 3–10, 13, 14, 17, 18, 20, 22, and X (lane 5); hybrid PB5-5c retaining 2p, 3, 5, and 17q (lane 6); hybrid GC-3a retaining 4q, 6p, 12–14, 17, and 22 (lane 7); hybrid GC-3c retaining 4q, 6p, 9, 12–14, 17, and 22 (lane 8); hybrid GC-8c retaining 4, 6q, 8q, 9, 12, 13, 17, 21, and 22 (lane 9); hybrid B2 retaining 4p, 6, 12, 20, and X (lane 10); and mouse cell line (lane 11) was analyzed for presence of the c8 gene as described in the legend to fig. 2. Molecular weights of the human and the major mouse bands detected by the c8 probe are indicated on the left. Variations in size of minor mouse bands detected by the c8 probe are due to polymorphisms within the various mouse strains used as parental cells for hybrids.

band (fig. 3, lane 1) was present in nine hybrids (fig. 1) and absent in the others; see figure 3 for a representative Southern blot. Inspection of figure 1 shows that presence of c8 in the hybrid cells correlates with presence of the long arm of chromosome 12 and not with any other human chromosome region. Southern blots that were hybridized to c8 were stripped of probe and rehybridized either with a T4 probe (which maps to 12p12→pter [Isobe et al. 1986]) or with a K-ras-2 probe (McBride et al. 1983), which maps to region 12p12 (Popescu et al. 1985). c8 Segregated 100% concordantly with the K-ras-2 gene in the hybrid panel and discordantly with the T4 probe (data not shown). Thus Southern blot analysis of somatic-cell hybrids indicates that c8 (as well as K-ras-2) maps to chromosome region 12p12→12qter.

In situ hybridization of the c8 probe to metaphase chromosomes from a normal male confirmed this assignment and narrowed the localization of the c8 gene to 12q12→q13 (see fig. 4). A total of 192 metaphases were analyzed, and 753 grains situated on chromosomes were counted. Figure 4 illustrates the grain distribution over all chromosomes. The primary site of grain localization was the long arm of chromosome 12 (12q), containing 115 (15.3%) of the 753 grains counted. The majority (79%) of these 115 grains situated on 12q were within the q12-q13 region.

Human Homeo Box–containing Gene c13 Maps to Chromosome Region 2q31→2q37

In the screening of 31 rodent-human hybrid cells (see fig. 1) for the presence of the c13 gene, only one hybrid cell line (BD3) was found to retain c13 sequences (fig. 1; also see fig. 5, lane 3). This hybrid was the only hybrid found to retain a marker for the long arm of chromosome 2—the α1 (Type III) collagen chain gene (not shown), which maps to chromosome region 2q24.3→2q31

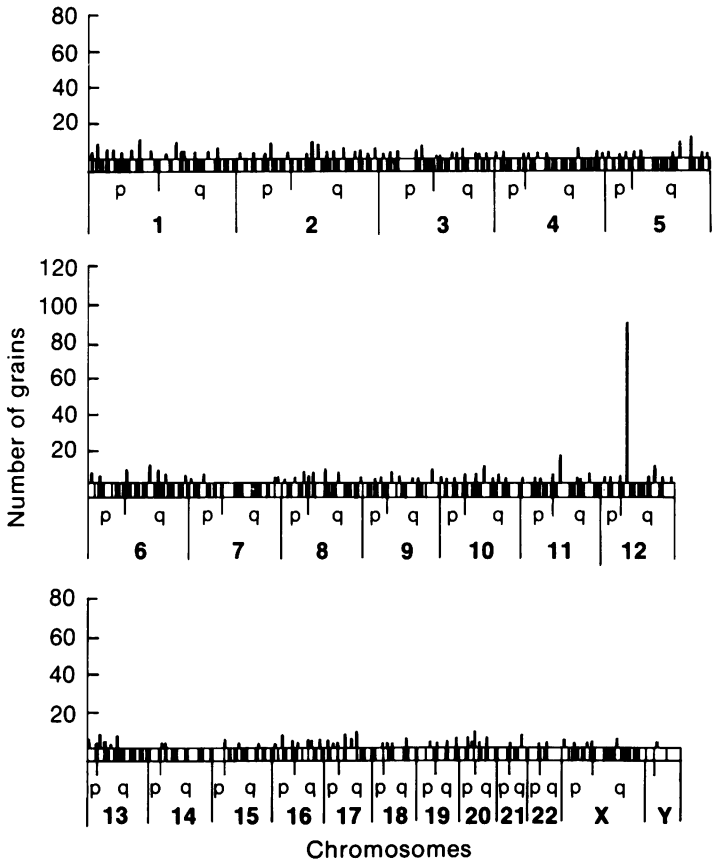

FIG. 4.—Localization of the c8 gene to human chromosome region 12q12→12q13 by in situ hybridization to metaphase chromosomes. The diagram shows the distribution of 753 autoradiographic grains on 192 metaphases. The abscissa represents the chromosome banding pattern of each human chromosome in relative size proportion; the ordinate shows the number of silver grains. Ninety-one grains were found over the 12q12→13 region.

(Emanuel et al. 1985); and the long arm of chromosome 2 was the only chromosome region that segregated concordantly with the presence of the c13 locus (fig. 1). Several hybrids that retained both the short arm of chromosome 2 and the T8 (Leu-2) gene did not retain the c13 sequence (see fig. 1, col. 2).

Chromosomal in situ hybridization confirmed the assignment of the c13 gene to chromosome 2 and sublocalized the c13 gene to chromosome region 2q31→2q37. A total of 128 metaphases were analyzed, and 353 grains were counted over chromosomes. Figure 6 depicts the grain distribution to all chromosomes, with the primary site of localization being the long arm of chromosome 2 (2q). Twenty-one percent of the total grains counted were located over 2q, and the majority (78%) of grains on 2q hybridized to the distal portion between the q31 to q37 region.

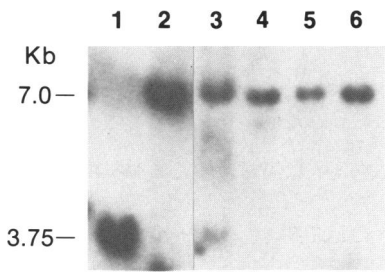

FIG. 5.—Segregation of the c13 gene in somatic-cell hybrids. DNA (~10 μg/lane) from human cell line (lane 1); hybrid CB5 retaining human chromosomes 6, 10, 14–16, 21, and X (lane 2); hybrid BD3 retaining 2q, 3, 5–8, 10, 11p, 12, 14, 16, 18–21, and X (lane 3); hybrid S5b retaining 3–7, 9, 13, 15, 17, 18, and X (lane 4); hybrid AA3 retaining 4p, 18, and X (lane 5); and mouse cell line (lane 6) was analyzed for presence of the c13 gene by Southern blot analysis as described in the legend to fig. 2. Molecular weights of the mouse (7-kb) and human (3.75-kb) gene fragments detected by the c13 probe are shown on the left. Only one hybrid (BD3, lane 3; also see fig. 1) is positive for the c13 human sequence, and this is the only hybrid that retains the α1 (III) collagen chain gene (not shown), which maps to the long arm of chromosome 2.

DISCUSSION

Mutations affecting development and morphogenesis have been regionally mapped in the mouse, and murine homeo box–containing genes have been mapped in the vicinity of some of these genes (Joyner et al. 1985*a;* Ruddle et al. 1985). It will eventually be possible to determine whether any of the murine homeo box–containing genes overlap with some of the genes known to affect development. Since similar mutations affecting development are not known in man, it would be very useful to know whether the murine homeo box–containing genes have human cognates. At least for one cluster of mouse and human homeo box–containing genes, this is the case: murine Hox 2.1 and Hox 2.2 on mouse chromosome 11 are cognates of Hox 2.1 and Hox 2.2, respectively, on human chromosome 17.

After submission of the present report, Bucan et al. (1986) and Rabin et al. (1986) reported that the human Hox-1 locus, the cognate of the murine Hox-1 locus on mouse chromosome 6, maps to the short arm of chromosome 7 at 7p14→7p21. Rabin et al. (1986) also reported localization of the human Hox-3 locus, cognate of the murine Hox-3 on mouse chromosome 15, to human chromosome region 12q11→q21. Thus, the c8 clone, which maps to 12q12→q13, belongs to the Hox-3 cluster. A human cognate of the mouse En-1 gene on mouse chromosome 1 has not yet been reported; nor has a mouse homologue of the c13 Hox gene at 2q31 been reported.

The region of chromosome 12 where c8 maps is a potentially interesting region. There is a fragile site at 12q13.1, and 12q13 is involved in abnormalities (deletions and translocations) in human seminomas, teratomas, and teratocarcinoma cell lines (Atkin and Baker 1982; Andrews et al. 1984, 1985). It will be particularly interesting to study expression of c8 in stem and differentiated human teratocarcinoma cells and to investigate the structure of the c8 gene in human teratocarcinoma cells with 12q13 abnormalities.

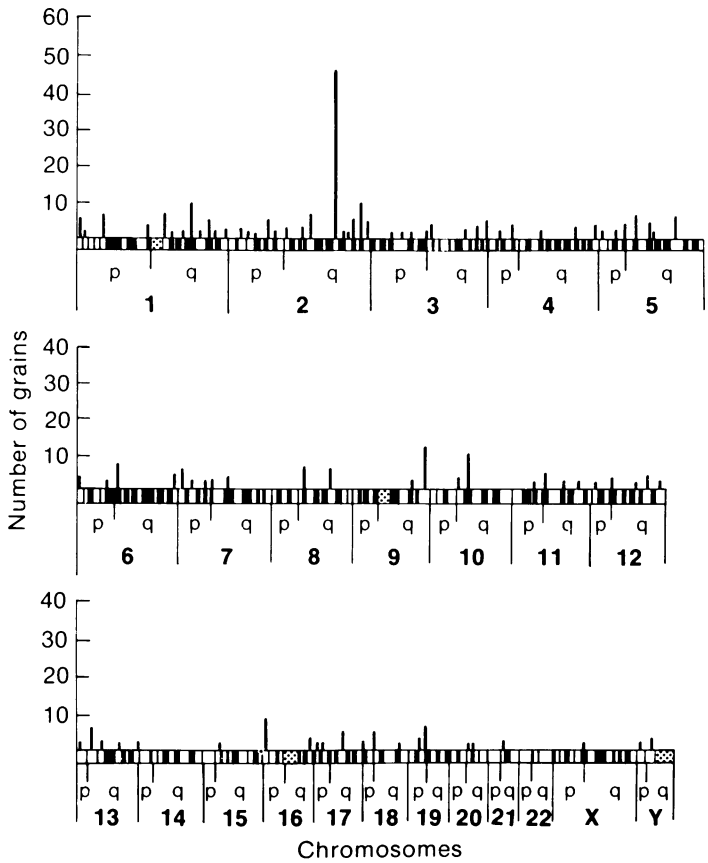

FIG. 6.—Localization of the c13 gene to human chromosome region 2q31→q37 by chromosomal in situ hybridization. The diagram shows the distribution of 353 grains over chromosomes of 128 metaphases. The abscissa represents the chromosome banding pattern of each human chromosome in relative size proportion; the ordinate shows the number of silver grains. Fifty-eight grains were found over the 2q31→q37 region.

Perhaps the most intriguing feature brought to light by the localization of human homeo box genes to human chromosome regions 17q21, 12q12→q13, and 2q31→q37 is that there is a collagen chain gene nearby in each case: COL1A1 at 17q21, COL2A1 at 12q13, and COL3A1 and COL5A2 at 2q31 (summarized in McAlpine et al. 1985). Additionally, the human Hox-1 cluster at 7p15 is near a collagen-like gene (Retief et al. 1985) and the murine α1 (I) collagen gene has been mapped to mouse chromosome 11 (Münke et al. 1985). Since the collagen protein family in vertebrates consists of a minimum of nine types of collagen molecules whose constituent chains are coded by a minimum of 17 genes (Solomon et al. 1985) and since the Hox family is of the same order of magnitude with some members not yet isolated in both mouse and man, it is possible that, in both mouse and man, the collagen and Hox loci are closely linked in each case. Presently, in man there are several collagen gene loci that

have not been shown to be near a Hox gene, and in the mouse very few collagen genes have been mapped. Both gene families are ancient ones, and it would be interesting to know whether this linkage of Hox and collagen genes is unique to mammals or vertebrates or is seen in flies and other invertebrates as well. Both Hox and collagen genes show complex patterns of expression during development, and it would be very interesting to determine whether there is coordinate expression of a Hox gene and its nearby collagen gene. In man there are numerous diseases involving inappropriate expression of collagen genes, and not all of these diseases have been shown to be due to structural mutations in involved collagen loci. It is conceivable that Hox genes could have a regulatory influence on nearby collagen genes that could be the cause of some of these developmental deficiences in specific types of collagen. It will be important to determine how close the Hox and collagen genes are to each other at the various human loci and to determine the location of the relevant murine collagen genes before the significance of the observed Hox/collagen pairs can be understood.

ACKNOWLEDGMENTS

We thank Felicia Watson and Deborah Geiman for excellent technical assistance and John Hurwitz, a student who helped in plasmid preparation and Southern blotting. We are indebted to Dr. F. Mavilio for helpful discussion and to the following colleagues for generously providing probes for use as chromosome markers: Dr. Steve Tronick (for K-ras-2), Dr. Jeanne Myers (for α1 [III] collagen), Dr. Jane Parnes (for T8), and Drs. Paul Maddon and Richard Axel (for T4). This work was supported by grants CA-10805, CA-21124, and CA-39860 from the National Institutes of Health; by Progetti Finalizzati CNR 'Ingegneria Genetica e Basi Molecolari delle Malattie Ereditarie' and 'Oncologia'; and by the Italian Association for Cancer Research (AIRC).

REFERENCES

Andrews, P. W., I. Damjanov, D. Simon, G. S. Banting, C. Carlin, N. C. Dracopoli, and J. Fogh. 1984. Pluripotent embryonal carcinoma clones derived from the human teratocarcinoma cell line Tera-2: differentiation in vivo and in vitro. Lab. Invest. **50**:147–162.

Andrews, P. W., I. Damjanov, D. Simon, and M. Dignazio. 1985. A pluripotent human stem-cell clone isolated from the Tera-2 teratocarcinoma line lacks SSEA-3 and SSEA-4 in vitro, but expresses these antigens when grown as a xenograft tumor. Differentiation **19**:127–135.

Atkin, N. B., and M. C. Baker. 1982. Specific chromosome change, i(12p), in testicular tumors? Lancet **2**:1349.

Beachy, P. A., S. L. Helfand, and D. S. Hogness. 1985. Segmental distribution of bithorax complex proteins during *Drosophila* development. Nature **313**:545–551.

Boncinelli, E., A. Simeone, A. La Volpe, A. Faiella, V. Fidanza, D. Acampora, and L. Scotto. 1985. Human cDNA clones containing homeo box sequences. Cold Spring Harbor Symp. Quant. Biol. **50**:301–306.

Bucan, M., T. Yang-Feng, A. M. Colberg-Poley, D. J. Wolgemuth, J.-L. Guenet, U. Francke, and H. Lehrach. 1986. Genetic and cytogenetic localization of the homeo box containing genes on mouse chromosome 6 and human chromosome 7. EMBO J. **5**:2899–2905.

Cannizzaro, L. A., and B. S. Emanuel. 1984. An improved method for G-banding chromosomes after in situ hybridization. Cytogenet. Cell Genet. **38**:308–309.

Carrasco, A. E., W. McGinnis, W. J. Gehring, and E. M. DeRobertis. 1984. Cloning of an *X. laevis* gene expressed during early embryogenesis coding for a peptide region homologous to *Drosophila* homeotic genes. Cell **37**:409–414.

Colberg-Poley, A. M., S. D. Vass, K. Chowdbury, and P. Gruss. 1985. Structural analysis of murine genes containing homeo box sequences and their expression in embryonal carcinoma cells. Nature **314**:713–718.

Dalla Favera, R., M. Bregni, J. Erikson, D. Patterson, R. C. Gallo, and C. M. Croce. 1982. Human c-myc onc gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. Proc. Natl. Acad. Sci. USA **79**:7824–7827.

Dayton, A. I., J. R. Selden, G. Laws, D. J. Dorney, J. Finan, P. Tripputi, B. S. Emanuel, G. Rovera, P. C. Nowell, and C. M. Croce. 1984. A human c-erbA oncogene homologue is closely proximal to the chromosome 17 breakpoint in acute promyelocytic leukemia. Proc. Natl. Acad. Sci. USA **81**:4495–4499.

Emanuel, B., L. A. Cannizzaro, J. M. Seyer, and J. C. Myers. 1985. The α1 (III) and α2 (V) procollagen genes are located on the long arm of chromosome 2. Proc. Natl. Acad. Sci. USA **82**:3385–3389.

Fjose, A., W. J. McGinnis, and W. J. Gehring. 1985. Isolation of a homeo box-containing gene from the engrailed region of *Drosophila* and the spatial distribution of its transcripts. Nature **313**:284–289.

Garcia-Bellido, A. 1977. Homeotic and atavic mutations in insects. Am. Zool. **17**:613–629.

Gehring, W. J. 1985. The homeo box: a key to the understanding of development? Cell **40**:3–5.

Harper, M. E., and G. F. Saunders. 1981. Localization of single copy DNA sequences on G-banded human chromosomes by in situ hybridization. Chromosoma **83**:431–439.

Hart, C. P., A. Awgulewitsch, A. Fainsod, W. McGinnis, and F. H. Ruddle. 1985. Homeo box gene complex on mouse chromosome 11: molecular cloning, expression in embryogenesis and homology to a human homeo box locus. Cell **43**:9–18.

Hauser, C. A., A. L. Joyner, R. D. Klein, T. K. Learned, G. R. Martin, and R. Tijian. 1985. Expression of homologous homeo box-containing genes in differentiated human teratocarcinoma cells and mouse embryos. Cell **43**:19–28.

Huebner, K., H. Ar-Rushdi, C. A. Griffen, M. Isobe, C. Kozak, B. Emanuel, L. Nagarajan, J. L. Cleveland, T. I. Bonner, M. D. Goldsborough, C. M. Croce, and U. Rapp. 1986*a*. Actively transcribed genes in the raf oncogene group, located on the X chromosome in mouse and human. Proc. Natl. Acad. Sci. USA **83**:3934–3938.

Huebner, K., M. Isobe, M. Chao, M. Bothwell, A. H. Ross, J. Finan, J. A. Hoxie, A. Sehgal, C. R. Buck, A. Lanahan, P. C. Nowell, H. Koprowski, and C. M. Croce. 1986*b*. The nerve growth factor receptor gene is at human chromosome region 17q12-17q22, distal to the chromosome 17 breakpoint in acute leukemias. Proc. Natl. Acad. Sci. USA **83**:1403–1407.

Huebner, K., M. Isobe, C. M. Croce, D. W. Golde, S. E. Kaufman, and J. C. Gasson. 1985*a*. The human gene encoding GM-CSF is at 5q21-q32, the chromosome region deleted in the $5q^-$ anomaly. Science **230**:1282–1285.

Huebner, K., M. Isobe, J. C. Gasson, D. W. Golde, and C. M. Croce. 1986*c*. Localization of the gene encoding human erythroid-potentiating activity to chromosome region Xp11.1→Xp11.4. Am. J. Hum. Genet. **38**:819–826.

Huebner, K., A. P. Palumbo, M. Isobe, C. A. Kozak, S. Monaco, G. Rovera, C. M. Croce, and P. J. Curtis. 1985*b*. The α-spectrin gene is on chromosome 1 in mouse and man. Proc. Natl. Acad. Sci. USA **82**:3790–3793.

Isobe, M., K. Huebner, J. Erikson, R. C. Peterson, F. J. Bollum, L. M. S. Chang, and C. M. Cross. 1985. Chromosome localization of the gene for human deoxynucleotidyltransferase to region 10q23→10q25. Proc. Natl. Acad. Sci. USA **82**:5836–5840.

Isobe, M., K. Huebner, P. J. Maddon, D. R. Littman, R. Axel, and C. M. Croce. 1986. The gene encoding the T-cell surface protein T4 is located on human chromosome 12. Proc. Natl. Acad. Sci. USA **83**:4399–4402.

Joyner, A. L., I. Kornberg, K. G. Coleman, D. R. Cox, and G. R. Martin. 1985*a*. Expression during embryogenesis of a mouse gene with sequence homology to the engrailed gene. Cell **43**:29–37.

Joyner, A. L., R. V. Lebo, Y. W. Kan, R. Tijian, D. R. Cox, and G. R. Martin. 1985*b*. Comparative chromosome mapping of a conserved homeo box region in mouse and human. Nature **314**:173–175.

Kaufman, T. C., R. Lewis, and B. Wakimoto. 1980. Cytogenetic analysis of chromosome 3 in *Drosophila melanogaster:* the homeotic gene complex in polytene chromosome interval 84 A-B. Genetics **94**:115–133.

Laughon, A., S. B. Carroll, F. A. Storfer, P. D. Riley, and M. P. Scott. 1985. Common properties of proteins encoded by the Antennapedia complex genes of *Drosophila melanogaster*. Cold Spring Harbor Symp. Quant. Biol. **50**:253–262.

Laughon, A., and M. P. Scott. 1984. Sequence of a *Drosophila* segmentation gene: protein structure homology with DNA-binding proteins. Nature **310**:25–31.

Levine, M., G. N. Rubin, and R. Tijian. 1984. Human DNA sequences homologous to a protein coding region conserved between homeotic genes of *Drosophila*. Cell **38**:667–673.

Lewis, E. B. 1978. A gene complex controlling segmentation in *Drosophila*. Nature **276**:565–570.

McAlpine, P. J., T. B. Shows, R. L. Miller, and A. J. Pakstis. 1985. The 1985 Catalog of Mapped Genes and Report of the Nomenclature Committee: Eighth International Workshop on Human Gene Mapping. Cytogenet. Cell Genet. **40**:9–66.

McBride, O. W., D. C. Swan, S. R. Tronick, R. Gol, D. Klimanis, D. E. Moore, and S. A. Aaronson. 1983. Regional chromosomal localization of N-ras, K-ras-1, K-ras-2 and myb oncogenes in human cells. Nucleic Acids Res. **11**:8221–8235.

McGinnis, W., R. L. Garber, J. Wirz, A. Kurolwa, and W. J. Gehring. 1984*a*. A homologous protein-coding sequence in *Drosophila* homeotic genes and its conservation in other metazoans. Cell **37**:403–408.

McGinnis, W., C. P. Hart, W. J. Gehring, and F. H. Ruddle. 1984*b*. Molecular cloning and chromosome mapping of a mouse DNA sequence homologous to homeotic genes of *Drosophila*. Cell **38**:675–680.

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

Manley, J. L., and M. S. Levine. 1985. The homeo box and mammalian development. Cell **43**:1–2.

Martin, G. R., E. Boncinelli, D. Duboule, P. Gruss, I. Jackson, R. Krumlauf, P. Lonai, W. McGinnis, F. Ruddle, and D. Wolgemuth. 1987. Nomenclature for homeo-box-containing genes. Nature **325**:21–22.

Mavilio, F., A. Simeone, A. Giampaolo, A. Faiella, V. Zappavigna, D. Acampora, G. Poiana, G. Russo, C. Peschle, and E. Boncinelli. 1986. Tissue and stage-specific expression of multiple transcripts from a new homeo box gene in human embryonic development. Nature **324**:664–668.

Mueller, M. M., A. E. Carrasco, and E. M. De Robertis. 1984. A homeo box-containing gene expressed during oogenesis in *Xenopus*. Cell **39**:157–162.

Münke, M., K. Harbers, R. Jaenisch, D. R. Cox, and U. Francke. 1985. Assignment of the alpha 1 (I) collagen gene to mouse chromosome 11 (abstract). Cytogenet. Cell Genet. **40**:706.

Nagarajan, L., E. Louie, Y. Tsujimoto, A. Ar-Rushdi, K. Huebner, and C. M. Croce. 1986. Localization of the human pim oncogene (PIM) to a region of chromosome 6 involved in translocations in acute leukemia. Proc. Natl. Acad. Sci. USA **83**:2556–2560.

Okayama, H., and P. Berg. 1983. A cDNA cloning vector that permits expression of cDNA inserts in mammalian cells. Mol. Cell. Biol. **3**:280–289.

Poole, S. J., L. M. Kauver, B. Drees, and T. Kornberg. 1985. The engrailed locus of *Drosophila:* structural analysis of an embryonic transcript. Cell **40:**37–43.

Popescu, N. C., S. C. Amsbaugh, J. A. DiPaolo, S. R. Tronick, and S. A. Aaronson. 1985. Chromosomal localization of three human Ras genes by in situ molecular hybridization. Somat. Cell. Mol. Genet. **11:**149–155.

Rabin, M., A. Ferguson-Smith, C. P. Hart, and F. H. Ruddle. 1986. Congrate homeobox loci mapped on homologous human and mouse chromosomes. Proc. Natl. Acad. Sci. USA **83:**9104–9108.

Rabin, M., C. P. Hart, A. Ferguson-Smith, W. McGinnis, M. Levine, and F. Ruddle. 1985. Two homeo box loci mapped in evolutionarily related mouse and human chromosomes. Nature **314:**175–178.

Retief, A. E., M. I. Parker, M. F. Fox, E. Dietzsch, D. L. du Toit, M. Kotze, J. Grobbelaar, and J. Brusnicky. 1985. Chromosomal mapping of collagen genes by low stringency in situ hybridization conditions (abstract). Cytogenet. Cell Genet. **40:**730.

Ruddle, F. H., C. P. Hart, A. Awgulewitsch, A. Fainsod, M. Utset, D. Dalton, N. Kerk, M. Rabin, A. Ferguson-Smith, A. Fienberg, and W. McGinnis. 1985. Mammalian homeo box genes. Cold Spring Harbor Symp. Quant. Biol. **50:**277–284.

Ryan, J., P. E. Barker, K. Shimuzu, M. Wigler, and F. H. Ruddle. 1983. Chromosomal assignment of a family of human oncogenes. Proc. Natl. Acad. Sci. USA **80:**4460–4463.

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.

Simeone, A., F. Mavilio, D. Acampora, A. Giampaolo, A. Faiella, V. Zappavigna, M. D'Esposito, G. Russo, E. Boncinelli, and C. Peschle. 1987. Human homeo box-containing genes: analysis of two unrelated cDNA clones and expression of the corresponding gene in embryonic development (submitted).

Simeone, A., F. Mavilio, L. Bottero, A. Giampaolo, G. Russo, A. Faiella, E. Boncinelli, and C. Peschle. 1986. A human homeo box gene specifically expressed in spinal cord during embryonic development. Nature **320:**763–765.

Solomon, E., L. R. Hiorns, N. Spurr, M. Kurkinen, D. Barlow, B. L. M. Hogan, and R. Dalgleish. 1985. Chromosomal assignments of the genes coding for human types II, III, and IV collagen: a dispersed gene family. Proc. Natl. Acad. Sci. USA **82:**3330–3334.

Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98:**503–517.

Sukhatme, V. P., A. C. Vollmer, J. Erikson, M. Isobe, C. M. Croce, and J. R. Parnes. 1985. Gene for the human T cell differentiation antigen Leu-2/T8 is closely linked to the κ light chain locus on chromosome 2. J. Exp. Med. **161:**429–434.

White, R. A. H., and M. Wilcox. 1984. Protein products of the bithorax complex in *Drosophila*. Cell **39:**163–171.