Expression during Embryogenesis of a Mouse Gene with Sequence Homology to the Drosophila engrailed Gene

Alexandra L. Joyner,* Thomas Kornberg,†t
Kevin G. Coleman,‡ David R. Cox,‡‡
and Gail R. Martin†t
* Department of Anatomy
† Division of Genetics
‡ Department of Biochemistry and Biophysics
University of California, San Francisco
San Francisco, California 94143

Summary

Regions of the mouse and human genomes with strong homology to the Drosophila engrailed gene have been identified by Southern blot analysis. One mouse engrailed-like region, Mo-en.1, has been cloned and partially sequenced; homology with the engrailed gene is localized to a 180 bp engrailed-like homeo box and 63 nucleotides immediately 3' to it. The protein sequence this region can encode includes 61 amino acids, of which 60 (75%) are identical with those of the putative translation product of the corresponding engrailed sequence. These data suggest that Mo-en.1 represents a mouse homolog of a gene of the Drosophila engrailed gene complex. Mo-en.1 has been mapped to chromosome 1, indicating that it is not linked to other homeo box sequences that have been mapped in the mouse genome. Analysis of poly(A)⁺ RNA extracted from teratocarcinoma cells and whole mouse embryos demonstrates that the conserved homeo box region of Mo-en.1 is expressed differentially during mouse embryogenesis.

Introduction

Attempts to identify genes that regulate mammalian development have been hindered by the long gestation periods of mammalian embryos, their relatively small numbers, and their inaccessibility during development in utero. These problems are minimized in studies of insect development, and genetic experiments in the model organism Drosophila melanogaster have identified a number of loci with a basic role in the control of pattern formation. The recent discovery of a region of significant homology between some of these loci and regions of significant homology in the mouse genome (McGinnis et al., 1984a; Muller et al., 1984) has suggested that genes that control mammalian development may be identified and cloned by using sequences from Drosophila as probes.

The Drosophila genes that are known to regulate early embryonic development each fall into one or more of three general categories: maternal-effect genes, which are expressed during oogenesis and specify the structure and spatial coordinates of the egg (Nusslein-Volhard, 1979); segmentation genes, which determine the number and polarity of the body segments (Nusslein-Volhard and Wieschaus, 1980); and homeotic genes, which specify segment identity (Ouweneel, 1976). Several of the genes in the latter two categories share a conserved 180 bp DNA sequence that has been termed the "homeo box" (McGinnis et al., 1984a, 1984c; Scott and Weiner, 1984). The presence of the homeo box sequence in these genes suggests that they may derive from a common ancestor and perhaps carry out their respective functions by similar mechanisms.

The Drosophila homeo-box-containing genes that have been identified thus far can be grouped into two classes. The smaller class is represented only by the two genes (Poole et al., 1985; Fjose et al., 1985) in a region, the engrailed gene complex (EN-C), that includes the engrailed gene (Garcia-Bellido and Santamaria, 1972). The larger class includes at least seven genes (Regulski et al., 1985) in the Antennapedia gene complex (ANT-C; Kaufman et al., 1980) and the Bithorax gene complex (BX-C; Lewis, 1978). There are several features that distinguish the two classes of genes. EN-C is located on the right arm of chromosome 2 and ANT-C and BX-C are on the right arm of chromosome 3. The homeo boxes found in the EN-C genes are interrupted by an intervening sequence, whereas those in the ANT-C/BX-C genes are not. The polypeptides encoded by the homeo boxes found in these two classes share approximately 87% amino acid homology; the homeo boxes in the ANT-C/BX-C genes also share 75% or more amino acid homology; however, there is only approximately 50% amino acid homology between homeo boxes from genes of the two classes. Thus, homeo boxes themselves can be generally classified as EN-C or ANT-C/BX-C on the basis of their degree of amino acid sequence homology with members of a given class. The most striking difference between the genes in the two classes is that the two genes in the EN-C show DNA sequence and putative amino acid homology outside their homeo boxes. This unique region of homology can encode 31 amino acids immediately 3' to the homeo box. In contrast, no homology outside the homeo box has yet been found among any of the ANT-C/BX-C homeo-box-containing genes of Drosophila.

Several different DNA segments containing homeo boxes have been isolated, using ANT-C/BX-C homeo box sequence probes, from the genomes of the frog (Carrasco et al., 1984; Muller et al., 1984), mouse (McGinnis et al., 1984b; Colberg-Poley et al., 1985; Hauser et al., 1985), and man (Levine et al., 1984). The putative amino acid sequence encoded by each of these vertebrate homeo boxes shares at least 70% homology with all members of the ANT-C/BX-C class of homeo box. In the mouse genome the ANT-C/BX-C homeo box sequences appear to reside in two clusters, one on chromosome 11 that contains at least two, and probably more, homeo boxes (Joyner et al., 1985; Rabin et al., 1985; Hauser et al., 1985) and one on chromosome 6 that contains at least three homeo boxes within a 40 kb region (McGinnis et al., 1984b; D. Duboule, personal communication). A cluster of at least three homeo boxes that appears to have homology...
with the one found on mouse chromosome 11 maps to a region on human chromosome 17 (Joyner et al., 1985; Rabin et al., 1985; Hauser et al., 1985). Recent experiments have also demonstrated that these homeo box sequences reside in transcribed regions of the genome. RNAs containing several of the ANT-C/BX-C homeo box segments have been found to be differentially expressed during frog (Carrasco et al., 1984; Muller et al., 1984) and mouse embryonic development (Hauser et al., 1985), and during differentiation of mouse (Colberg-Poley et al., 1985) and human (Hauser et al., 1985) teratocarcinoma cells. Vertebrate representatives of the smaller EN-C class of homeo box have not been previously reported.

There is increasing evidence to support the concept that genes with homologous functions in vertebrates and invertebrates may share common DNA sequences despite the vast evolutionary distance that separates them. For example, the sequence homology between bovine and Drosophila rhodopsin has been exploited in the isolation of the gene from the fly genome (O'Tousa et al., 1985; Zucker et al., 1985), and the identification of the human and hamster genes for RNA polymerase was facilitated by the use of Drosophila sequences (Ingles et al., 1983). The studies presented here were designed to identify genes in mammals that control pattern formation in a manner analogous to that of the Drosophila engrailed gene by identifying mammalian genes with sequence homology to the Drosophila EN-C genes. Using a Drosophila engrailed gene cDNA clone as a probe, two regions of the mouse genome and a region of the human genome were found to share homology with sequences in the Drosophila engrailed gene; and one of the homologous mouse sequences, Mo-en.1, was cloned. As a first step in determining the function of this mouse engrailed-like sequence, we determined the sequence of the region of homology, the chromosomal location of Mo-en.1, and its temporal pattern of expression during mouse embryonic development.

Results

Isolation of Mouse Genomic Clones with Sequence Homology to the Drosophila engrailed Gene

To determine whether the mouse genome contains sequences homologous to those in the Drosophila engrailed gene, Southern blot analysis was carried out using mouse genomic DNA digested separately with four different restriction endonucleases. The Southern blot was probed under conditions of low stringency with a 1.4 kb Drosophila engrailed cDNA clone that contains a homeo box (see Experimental Procedures). Two prominent restriction fragment bands hybridized to this probe in each of the four mouse DNA digests (data not shown, but see Figure 4). To explore the possibility of a relationship between these homologous sequences and those that can be detected with an ANT-C/BX-C homeo box probe, the same Southern blot was rehybridized to a probe, H.1, that contains the human Hu1 homeo box, which is a member of the ANT-C/BX-C class (Joyner et al., 1985). Six to eight restriction fragments were detected in each DNA digest; none of these comigrated with the two detected with the engrailed probe (data not shown).

To isolate the DNA sequences with engrailed homology, a mouse genomic library was screened under conditions of low stringency using the 1.4 kb engrailed cDNA probe. Five clones that cross-hybridize with the cDNA probe were isolated. Southern blot analysis, using the engrailed cDNA probe, of restriction endonuclease digests of one of these clones, termed Mo-en.1, indicated that it represents one of the two mouse genomic regions containing engrailed homology. The engrailed-homologous region in Mo-en.1 localizes to a 700 bp Bam HI/Eco RI fragment. A partial restriction map of the Mo-en.1 clone, which was further analyzed as described below, is shown in Figure 1.

Sequence Analysis of the Mo-en.1 engrailed-Homologous Region

The nucleotide sequence of the Mo-en.1 region that cross-hybridized with the engrailed probe was determined. Figure 2 shows 313 nucleotides of Mo-en.1, which include a homeo box, 37 bases of 5' sequence, and 96 bases of 3' sequence ending at a putative translation stop codon. To facilitate comparison with the homologous portion of the Drosophila engrailed cDNA, its sequence is also shown.
Mouse engrailed-like Gene

and the two sequences are aligned to achieve the best match of their respective homeo box sequences. No homology was found over a 60 bp stretch immediately 5' of the homeo box, nor over a 160 bp stretch beginning 63 bp 3' of the homeo box; but the homeo box (delineated by a solid line) and the 63 bases immediately 3' of it (delineated by a dashed line) share approximately 73% homology. Although the Mo-en.1 homeo box sequence also shares approximately 65%–75% nucleotide homology with ANT-C/BX-C homeo boxes from Drosophila and vertebrates, the putative amino acid sequence of the Mo-en.1 homeo box is much more closely related to those of the EN-C genes than to those of the ANT-C/BX-C genes (see below).

The homologous mouse and Drosophila genomic regions differ in that Mo-en.1 encodes a continuous open reading frame whereas the homologous regions of the two EN-C genes are each interrupted by an intervening sequence. Comparison of the amino acid sequences of the putative homeo box peptide domains specified by Mo-en.1 and engrailed, as well as the homologous sequence from the other EN-C gene, invected (previously called engrailed-anterior), reveals that the Mo-en.1 homeo box is more closely related to the EN-C than to the ANT-C/BX-C homeo boxes (Figure 4). This probe, which extends from the Sst I site in the middle of the homeo box to approximately 50 bp 3' of the homeo box, hybridized to two regions in mouse genomic DNA that were indistinguishable from those detected with the 1.4 kb engrailed cDNA probe. Thus, the second mouse genomic region, Mo-en.2, also contains sequences homologous with the conserved EN-C region found in Mo-en.1.

To determine whether sequences in the EN-C homeo box region also are conserved in the human genome, restriction-enzyme-digested human genomic DNA was probed with the a-s probe. At least one restriction fragment in each digest hybridized to the Mo-en.1 homeo box region probe (Figure 4). To determine whether this human region contains additional homology to Mo-en.1, a similar Southern blot was hybridized to a single-copy probe, mp, derived from Mo-en.1, which lacks any of the sequences conserved in Drosophila (data not shown). This probe, which extends from an Eco RI site 360 bp 3' of the homeo box to a Sst I site approximately 4 kb 3' of the homeo box, detected a single band in each human DNA digest, and these bands were found to comigrate with the most prominent bands detected with the a-s probe. Thus, the human genome contains a region homologous to Mo-en.1 that includes both a homeo box region and 3'-flanking sequences.

Assignment of Mo-en.1 to Mouse Chromosome 1

The Mo-en.1 locus was assigned to a mouse chromosome...
analysis of the segregation of MO-en.1 and mouse chromosomes except for chromosomes 11 and Y (Figure 4 previously described (Cox et al., 1982). These nine hybrid cell hybrids segregate each of the other mouse chromosomes (Figure 5), thus assigning MO-en.1 to mouse chromosomes in the hybrid cell clones showed concordant segregation with each of the other mouse chromosomes (Figure 5), thus assigning MO-en.1 to mouse chromosome 1. Furthermore, since the 5 kb mouse Mo-en.1 Eco RI fragment is not detectable in DNA from hybrid clone III-16, we conclude that Mo-en.1 maps distal to band C2 on chromosome 1.

Regulated Expression of Mo-en.1 in Differentiating Teratocarcinoma Cells and Mouse Embryos

The strong conservation of coding capacity over a 243 bp region of Mo-en.1 and the genes in the Drosophila EN-C suggests a conservation for this region of the proteins. We would therefore predict that this region is transcribed and translated during mouse embryogenesis, as it is during embryonic development in Drosophila. As a first step in testing this hypothesis, RNA from mouse teratocarcinoma cells at several stages of differentiation in vitro and from normal mouse embryos between 9.5 and 17.5 days of gestation was analyzed for Mo-en.1 sequences.

The PSA-1 teratocarcinoma stem cell that was used provides a model system for the peri-implantation stages of mouse embryogenesis and thus makes possible the study of embryonic cells at stages at which it would otherwise not be feasible to obtain experimental material (Martin, 1980). Undifferentiated PSA-1 cells, which appear to be equivalent to the inner cell mass of the mouse blastocyst, can be stimulated to differentiate by allowing them to form aggregates in suspension culture. These cell aggregates subsequently develop into structures known as embryoid bodies (Martin and Evans, 1975). The formation of many of the cell types in mid-gestation embryos (9.5-11.5 days), such as neuroendocrine, cartilage, striated muscle, and nerve, can be obtained by allowing the embryoid bodies to attach to a tissue culture surface and culturing them for an additional 7 days or more (stage 3) (Martin and Evans, 1975b). After 6-8 days of suspension culture these embryoid bodies closely resemble the fetal portion of mouse embryos at 4.5-6.5 days of gestation (stage 1). After approximately 10 days or more in suspension the embryoid bodies form "balloon-like" structures that have some features in common with embryos at approximately 15 days of gestation (stage 2) (Martin et al., 1977). The formation of many of the cell types in mid-gestation embryos (9.5-11.5 days), such as keratinizing epithelium, cartilage, striated muscle, and nerve, can be obtained by allowing the embryoid bodies to attach to a tissue culture surface and culturing them for an additional 7 days or more (stage 3) (Martin and Evans, 1975b).

The PSA-1 teratocarcinoma stem cell that was used provides a model system for the peri-implantation stages of mouse embryogenesis and thus makes possible the study of embryonic cells at stages at which it would otherwise not be feasible to obtain experimental material (Martin, 1980). Undifferentiated PSA-1 cells, which appear to be equivalent to the inner cell mass of the mouse blastocyst, can be stimulated to differentiate by allowing them to form aggregates in suspension culture. These cell aggregates subsequently develop into structures known as embryoid bodies (Martin and Evans, 1975). The formation of many of the cell types in mid-gestation embryos (9.5-11.5 days), such as neuroendocrine, cartilage, striated muscle, and nerve, can be obtained by allowing the embryoid bodies to attach to a tissue culture surface and culturing them for an additional 7 days or more (stage 3) (Martin and Evans, 1975b).
Figure 4. Mouse Genomic DNA Contains Two Copies of the Conserved EN-C Region and Human DNA Contains Conserved Mo-en.1 Sequences

Ten micrograms of mouse and human genomic DNAs were digested with the enzymes indicated, separated by electrophoresis on a 1% agarose gel, transferred to GeneScreen and irradiated with UV light. The blot was hybridized to the Mo-en.1 homeo box probe, a-s, diagramed below the autoradiogram, and washed under moderate stringency conditions. The rectangle outlined by a double solid line in the diagram of the probe delineates the Mo-en.1 homeo box and the rectangle outlined by the dashed double line delineates the 3' region of homology between Mo-en.1 and the genes of EN-C. B, Bam HI; E, Eco RI; H, Hind III; S, Stu I.

Discussion

This study was initiated on the premise that if a homolog of the Drosophila engrailed gene exists in the mouse genome, it should be possible to identify and isolate it by virtue of conserved DNA sequences. We have identified a mouse gene, Mo-en.1, that can code for a protein that includes 81 amino acids sharing 75% homology with sequences in the two genes in EN-C. These 81 amino acids include a 60 amino acid homeo box domain that is significantly more homologous to homeo boxes of the EN-C genes than to any other homeo box described thus far, and 21 amino acids that lie 3' to the homeo box. The finding that Mo-en.1 contains a homeo box demonstrates that it is one of a growing number of possible vertebrate homologs of the Drosophila genes that control pattern formation. The conserved sequence 3' to the homeo box further identifies Mo-en.1 as a possible homolog of a gene in EN-C. Interestingly, our data provide evidence that EN-C gene homologs also may exist in the human and hamster genomes. In addition, a chicken genomic clone has been isolated that contains DNA sequences highly homologous to the same EN-C gene sequences that are conserved in Mo-en.1 (D. Darnell and C. Ordahl, personal communication).

Although the existence in Mo-en.1 of conserved EN-C gene sequences, particularly those found outside the homeo box, suggest that the gene we have isolated is a homolog of a Drosophila EN-C gene, this point will not be certain until it is known that the Mo-en.1 gene carries out a function in the mouse that is analogous to that performed by the EN-C genes in Drosophila. The engrailed gene is involved in the processes that subdivide the insect embryo into separate developmental units called compartments. Each segment of the fly is composed of an an-
Figure 5. Chromosomal Mapping of Mo-en.1 Using Somatic Cell Hybrids

The autoradiogram of the Southern blot shows mouse (left) and Chinese hamster (right) genomic DNA digested with Eco RI and hybridized to the mapping probe, mp, shown in Figure 1. A similar Southern blot of Eco RI digests of genomic DNA from nine mouse-hamster somatic cell hybrids was hybridized to the mp probe. The results are summarized in the table shown on the right. The presence of mouse chromosomes was determined by isozyme analysis of hybrid cell extracts and/or by karyotype analysis. (+) indicates that a chromosome was present by isozyme analysis but not visualized by karyotype analysis. [-] indicates that only part of the chromosome is present by karyotype analysis. Cell line III-16 contains only a portion of chromosome 1 with the material distal to band 1C2 deleted. Cell line III-23 contains a rearrangement of chromosomes 2 and X, and clone VI-25 contains a 2:16 translocation. M, mouse; H, hamster.

<table>
<thead>
<tr>
<th>Clone</th>
<th>Mo-en</th>
<th>Mouse Chromosome</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19</td>
</tr>
<tr>
<td>1-2A-2</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>1-2B-4</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>III-14</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>I-3A-2</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>II-31</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>III-16</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>III-12</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>III-23</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>VI-25</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

In the Drosophila embryo, the engrailed gene shows a cell- and stage-specific pattern of expression that is displayed as a zebra-like array of bands of engrailed-expressing and nonexpressing cells. Our data show that Mo-en.1 encodes a major 3.1 kb homeo-box-containing transcript that is expressed by differentiating PSA-1 teratocarcinoma cells. Given that the differentiating PSA-1 cells are an in vitro model for embryonic cells in the early postimplantation stages of development, these results imply that the 3.1 kb transcript, containing the Mo-en.1 homeo box region, is first detectable between 4.5 and 6.5 days of development. We also demonstrate that this 3.1 kb transcript continues to be expressed throughout embryogenesis, with a peak of expression between 10.5 and 12.5 days of gestation. A second, less abundant 2 kb homeo-box-containing transcript is expressed at all stages without apparent regulation. Although the available data do not provide any means of drawing specific parallels between the expression of Mo-en.1 and the genes in the CN-C, the fact that this mouse gene is expressed differentially during embryogenesis is at least consistent with the idea that it may be involved in the control of embryonic development.

In considering the ways in which Drosophila and mouse engrailed-like genes might have similar functions, one difficulty is that the fundamental similarities between development in higher invertebrates and vertebrates are not obvious. If a parallel can be drawn between the oogmons...
of mouse genomic DNA that had been partially digested with Sau 3A. A recombinant A EMBL3A (Frischauf et al., 1983) library, kindly provided by Drs. J. Vogel and R. Goodenow, was screened with the 1.4 kb engraleid cDNA fragment. The library contains 12-20 kb fragments of mouse genomic DNA. The single-stranded antisense probe, a-s, was approximately 140 bp in length and was derived from the M13 clone 2 shown in Figure 1. The single-stranded sense probe, s, was approximately 120 bp in length and was derived from the M13 clone 1 shown in Figure 1. The probes were made according to the methods in Church and Gilbert (1984) with the addition of 15% formamide, and the filters were washed twice at 30 min each in 2x SSC, 0.1% SDS, at 50°C followed by one wash in 2x SSC at 65°C.

Southern Blot Analysis

The mapping probe, mp, is a 3.6 kb Eco RI/Sal I fragment derived from Mo-en.1, which extends 3' from an Eco RI site 300 bp 3' of the homeobox. The mp fragment was subcloned from a Mo-en.1 into Eco RI and Sall digested SP65 plasmid DNA (Promega Biotech, Madison, Wisconsin). The single-stranded antisense probe, a-s, was approximately 140 bp in length and was derived from the M13 clone 2 shown in Figure 1. The single-stranded sense probe, s, was approximately 120 bp in length and was derived from the M13 clone 1 shown in Figure 1. The probes were made according to the methods in Church and Gilbert (1984) with the addition of 15% formamide, and the filters were washed twice at 30 min each in 2x SSC, 0.1% SDS, at 50°C followed by one wash in 2x SSC at 65°C.

Isolation of engraleid Homologous Mouse Genomic DNA

A recombinant λ EMBL3A (Frischauf et al., 1983) library, kindly provided by Drs. J. Vogel and R. Goodenow, was screened with the 1.4 kb engraleid cDNA fragment. The library contains 12-20 kb fragments of mouse genomic DNA. The single-stranded antisense probe, a-s, was approximately 140 bp in length and was derived from the M13 clone 2 shown in Figure 1. The single-stranded sense probe, s, was approximately 120 bp in length and was derived from the M13 clone 1 shown in Figure 1. The probes were made according to the methods in Church and Gilbert (1984) with the addition of 15% formamide, and the filters were washed twice at 30 min each in 2x SSC, 0.1% SDS, at 50°C followed by one wash in 2x SSC at 65°C.

Southern Blot Analysis

The mapping probe, mp, is a 3.6 kb Eco RI/Sal I fragment derived from Mo-en.1, which extends 3' from an Eco RI site 300 bp 3' of the homeobox. The mp fragment was subcloned from a Mo-en.1 into Eco RI and Sall digested SP65 plasmid DNA (Promega Biotech, Madison, Wisconsin). The single-stranded antisense probe, a-s, was approximately 140 bp in length and was derived from the M13 clone 2 shown in Figure 1. The single-stranded sense probe, s, was approximately 120 bp in length and was derived from the M13 clone 1 shown in Figure 1. The probes were made according to the methods in Church and Gilbert (1984) with the addition of 15% formamide, and the filters were washed twice at 30 min each in 2x SSC, 0.1% SDS, at 50°C followed by one wash in 2x SSC at 65°C.

Isolation of engraleid Homologous Mouse Genomic DNA

A recombinant λ EMBL3A (Frischauf et al., 1983) library, kindly provided by Drs. J. Vogel and R. Goodenow, was screened with the 1.4 kb engraleid cDNA fragment. The library contains 12-20 kb fragments of mouse genomic DNA. The single-stranded antisense probe, a-s, was approximately 140 bp in length and was derived from the M13 clone 2 shown in Figure 1. The single-stranded sense probe, s, was approximately 120 bp in length and was derived from the M13 clone 1 shown in Figure 1. The probes were made according to the methods in Church and Gilbert (1984) with the addition of 15% formamide, and the filters were washed twice at 30 min each in 2x SSC, 0.1% SDS, at 50°C followed by one wash in 2x SSC at 65°C.

Southern Blot Analysis

The mapping probe, mp, is a 3.6 kb Eco RI/Sal I fragment derived from Mo-en.1, which extends 3' from an Eco RI site 300 bp 3' of the homeobox. The mp fragment was subcloned from a Mo-en.1 into Eco RI and Sall digested SP65 plasmid DNA (Promega Biotech, Madison, Wisconsin). The single-stranded antisense probe, a-s, was approximately 140 bp in length and was derived from the M13 clone 2 shown in Figure 1. The single-stranded sense probe, s, was approximately 120 bp in length and was derived from the M13 clone 1 shown in Figure 1. The probes were made according to the methods in Church and Gilbert (1984) with the addition of 15% formamide, and the filters were washed twice at 30 min each in 2x SSC, 0.1% SDS, at 50°C followed by one wash in 2x SSC at 65°C.

Isolation of engraleid Homologous Mouse Genomic DNA

A recombinant λ EMBL3A (Frischauf et al., 1983) library, kindly provided by Drs. J. Vogel and R. Goodenow, was screened with the 1.4 kb engraleid cDNA fragment. The library contains 12-20 kb fragments of mouse genomic DNA. The single-stranded antisense probe, a-s, was approximately 140 bp in length and was derived from the M13 clone 2 shown in Figure 1. The single-stranded sense probe, s, was approximately 120 bp in length and was derived from the M13 clone 1 shown in Figure 1. The probes were made according to the methods in Church and Gilbert (1984) with the addition of 15% formamide, and the filters were washed twice at 30 min each in 2x SSC, 0.1% SDS, at 50°C followed by one wash in 2x SSC at 65°C.

Southern Blot Analysis

The mapping probe, mp, is a 3.6 kb Eco RI/Sal I fragment derived from Mo-en.1, which extends 3' from an Eco RI site 300 bp 3' of the homeobox. The mp fragment was subcloned from a Mo-en.1 into Eco RI and Sall digested SP65 plasmid DNA (Promega Biotech, Madison, Wisconsin). The single-stranded antisense probe, a-s, was approximately 140 bp in length and was derived from the M13 clone 2 shown in Figure 1. The single-stranded sense probe, s, was approximately 120 bp in length and was derived from the M13 clone 1 shown in Figure 1. The probes were made according to the methods in Church and Gilbert (1984) with the addition of 15% formamide, and the filters were washed twice at 30 min each in 2x SSC, 0.1% SDS, at 50°C followed by one wash in 2x SSC at 65°C.

Isolation of engraleid Homologous Mouse Genomic DNA

A recombinant λ EMBL3A (Frischauf et al., 1983) library, kindly provided by Drs. J. Vogel and R. Goodenow, was screened with the 1.4 kb engraleid cDNA fragment. The library contains 12-20 kb fragments of mouse genomic DNA. The single-stranded antisense probe, a-s, was approximately 140 bp in length and was derived from the M13 clone 2 shown in Figure 1. The single-stranded sense probe, s, was approximately 120 bp in length and was derived from the M13 clone 1 shown in Figure 1. The probes were made according to the methods in Church and Gilbert (1984) with the addition of 15% formamide, and the filters were washed twice at 30 min each in 2x SSC, 0.1% SDS, at 50°C followed by one wash in 2x SSC at 65°C.

Southern Blot Analysis

The mapping probe, mp, is a 3.6 kb Eco RI/Sal I fragment derived from Mo-en.1, which extends 3' from an Eco RI site 300 bp 3' of the homeobox. The mp fragment was subcloned from a Mo-en.1 into Eco RI and Sall digested SP65 plasmid DNA (Promega Biotech, Madison, Wisconsin). The single-stranded antisense probe, a-s, was approximately 140 bp in length and was derived from the M13 clone 2 shown in Figure 1. The single-stranded sense probe, s, was approximately 120 bp in length and was derived from the M13 clone 1 shown in Figure 1. The probes were made according to the methods in Church and Gilbert (1984) with the addition of 15% formamide, and the filters were washed twice at 30 min each in 2x SSC, 0.1% SDS, at 50°C followed by one wash in 2x SSC at 65°C.

Isolation of engraleid Homologous Mouse Genomic DNA

A recombinant λ EMBL3A (Frischauf et al., 1983) library, kindly provided by Drs. J. Vogel and R. Goodenow, was screened with the 1.4 kb engraleid cDNA fragment. The library contains 12-20 kb fragments of mouse genomic DNA. The single-stranded antisense probe, a-s, was approximately 140 bp in length and was derived from the M13 clone 2 shown in Figure 1. The single-stranded sense probe, s, was approximately 120 bp in length and was derived from the M13 clone 1 shown in Figure 1. The probes were made according to the methods in Church and Gilbert (1984) with the addition of 15% formamide, and the filters were washed twice at 30 min each in 2x SSC, 0.1% SDS, at 50°C followed by one wash in 2x SSC at 65°C.
primer (Pharmacia). Mo-en.1 fragments for sequencing were cloned into M13 vectors (Messing, 1983; Pharmacia), and the recombinant phages that were subsequently sequenced were identified by hybridization with bacterial host cells (JM109) that had been transfected with the appropriate ligation mixtures to radioactively labeled 1.4 kb en grailed cDNA fragment under the low stringency conditions used to screen the mouse genomic library. DNA for sequencing was then isolated from cross-hybridizing recombinant phages (Banger et al., 1980). All sequences shown were determined by independently sequencing both strands. For each M13 clone, only the length of DNA that was sequenced is shown in Figure 1.

Northern Blot Analysis

Total cellular RNA was isolated from PSA-1 cells or random bred mouse embryos (100 9.5 day embryos, 110 10.5 day embryos, 30 11.5 day embryos, and 10 to 20 embryos each from 12.5 to 17.5 day embryos) using essentially the guanidinium thiocyanate-CsCl method of Chirgwin et al. (1977). The PSA-1 cells were cultured as described by Martin et al. (1977). Mouse embryos were obtained by mating mice of various strains. The day on which the vaginal plug was detected was termed 0.5 days of gestation. Embryos at the designated stages were dissected from the implantation site and separated from the extraembryonic membranes. One gram of tissue or pellets of 105 cells were added to 5 ml of buffer (4 M guanidinium thiocyanate, 100 mM Tris [pH 7.5], 0.1 M (mercaptoethanol) and homogenized using a Polytron homogenizer (Brinkman Instruments) for 1 min, then 0.5 ml of 5% Sarkosyl was added, and the solution was mixed and then centrifuged at 10,000 rpm for 10 min. The mixture was sometimes stored at −70°C prior to layering 25 ml of the solution onto 15 ml of 5.7 M CsCl, 100 mM EDTA in SW27 tubes and centrifugation in an SW27 rotor at 25,000 rpm for 24 hr at 20°C. The resulting clear pellets were resuspended in 4 ml water, 100 mM Tris (pH 7.5), 5 mM EDTA, and the RNA mixture was extracted twice with phenol:chloroform (1:1), extracted once with chloroform, and precipitated with ethanol. The RNA was dissolved in H2O and loaded onto oligo(dT)-cellulose columns (Collaborative Research), and poly(A)+ was eluted. RNA was separated on 1% agarose-formaldehyde gels and blotted onto GeneScreen in 20× SSC as described by Maniatis et al. (1982) except that the running buffer was 10 mM EDTA, 50 mM Na-acetate, 0.6% formaldehyde, and the GinoSoronz membranes were irradiated with UV light according to Church and Gilbert (1984). The hybridizations using single-stranded probes were carried out at 65°C for 24 hr in 1% bovine serum albumin, 0.2 M NaHPO4 (pH 7.2), 1 mM EDTA, 7% SDS, and 45% formamide. The filters were washed twice in 2× SSC, 1% SDS at 65°C for 30 min and once in 50% formamide, 0.2× SSC at 65°C for 30 min. The hybridizations using the msg or pβ-actin probes were carried out using the conditions described for the "moderate" hybridization and washing of Southern blots. The GeneScreen filters were stripped of hybridized probes by washing in 75% formamide, 0.1% SDS, 0.2× SSC, at 65°C for 30 min.

Acknowledgments

We are grateful to Drs. Michael Lovett and Patricia Ashley for helping to provide mouse-hamster somatic cell DNAs, and Drs. Thomas Cooper, George Church, Stephen Poole, Barry Deves, and Leslie Lock for helpful advice and encouragement. We thank Marianne Gallup and Joy Graesser for excellent technical assistance. We are grateful to Drs. Michael Lovett and Patricia Ashley for helping to provide mouse-hamster somatic cell DNAs, and Drs. Thomas Cooper, George Church, Stephen Poole, Barry Deves, and Leslie Lock for helpful advice and encouragement. We thank Marianne Gallup and Joy Graesser for excellent technical assistance.

Received August 9, 1985

References


