The EMBO Journal vol.5 no.9 pp.2209-2215, 1986

# Sequential expression of murine homeo box genes during F9 EC cell differentiation

**Georg Breier, Marija Bućan[1], Uta Francke[2], Anamaris M.Colberg-Poley[3] and Peter Gruss**

Zentrum für Molekulare Biologie, Im Neuenheimer Feld 282, 6900 Heidelberg, [1]European Molecular Biology Laboratory, Meyerhofstr. 1, 6900 Heidelberg, FRG, [2]Yale University School of Medicine, Department of Human Genetics, New Haven, Connecticut 06511, USA

[3]Present address: Central Research and Development, Experimental Station, E.I. duPont de Nemours & Co., Inc., Bldg. 328, Room B30, Wilmington, DE 19898, USA

Communicated by P.Gruss

**We have isolated and characterized a previously unknown member of the murine homeo box family. The new locus, *m31*, is located on chromosome 15 and is more homologous to sequences contained in the *Drosophila* homeotic gene, *Antp*, than to any other known murine homeo box. We show that this gene encodes a 2.7 kb mRNA which is expressed during mouse embryogenesis and during differentiation of F9 teratocarcinoma cells into parietal endoderm cells. The transcript appears late in this differentiation process and is most abundant at a time when the F9 cells begin to express tissue-specific markers and the expression of another murine homeo box gene, *m6*, has decreased.**

*Key words:* Murine homeo box/mouse embryogenesis/F9 teratocarcinoma cells

## Introduction

Understanding the regulation of cellular differentiation, that is, the changes that a precursor, embryonic cell must undergo in order to produce a particular tissue-specific cell type, requires the identification of the genes involved. Although various genes which are expressed during murine embryogenesis have been isolated, many encode proteins (e.g. alpha-fetoprotein, collagen IV, laminin) which are expressed abundantly in determined cell types (Law *et al.*, 1980; Kurkinen *et al.*, 1983; Wang and Gudas, 1983). In addition, genetic analyses of the mouse have led to the identification of genes which when mutated (Green, 1981; Gluecksohn-Waelsch, 1983) result in formational defects (Woychik *et al.*, 1985) or embryonic death (Schnieke *et al.*, 1983; Frischauf, 1985). However, due to the complexity of the mouse genome and the length of its life cycle, a systematic genetic dissection of its developmental genes, as that performed in *Drosophila*, has not been possible.

Characterization of the homeotic and homeotic-like genes of *Drosophila melanogaster* has lent to biologists studying other organisms a tool for isolating potentially important developmental genes. On the basis of DNA hybridization, it was initially found that three loci, *Antennapedia* (*Antp*), *fushi tarazu* (*ftz*) and *Ultrabithorax* (*Ubx*), share a short DNA element called a homeo box (McGinnis *et al.*, 1984a). This sequence was subsequently found to be conserved in at least another seven loci in the *Drosophila* genome and in the DNA from a wide variety of animals (McGinnis *et al.*, 1984b; Gehring, 1985; Mlodzik *et al.*, 1985). The polypeptides encoded by homeo box sequences are characteristically highly basic and presumably bind DNA. This suggestion has gained support from homology with the appropriate regions of known DNA binding proteins (Laughon and Scott, 1984; Shepherd *et al.*, 1984), localization of homeo box proteins, *Ubx*, *ftz* and *engrailed* (*en*), in the nuclei of embryonic cells (White and Wilcox, 1985; Carroll and Scott, 1985; DiNardo *et al.*, 1985) and, more directly, by binding of fusion peptides containing homeo box residues to *en* DNA (Desplan *et al.*, 1985).

The murine homeo-box-containing genes which have been cloned share with the corresponding *Drosophila* DNA several properties; namely, sequence homology, structural similarity (clustering), presence on different chromosomes and differential expression during embryogenesis (McGinnis *et al.*, 1984c; Colberg-Poley *et al.*, 1985a,b; Jackson *et al.*, 1985; Hart *et al.*, 1985; Hauser *et al.*, 1985; Joyner *et al.*, 1985b). Four regions of murine DNA (*Hox-1*, *Hox-2*, *m6* and *Mo-en.1*), containing at least 9 homeo boxes, have been isolated and characterized. At least another box, *Mo-en.2*, is known to exist in the murine genome (Joyner *et al.*, 1985b). Both *Hox-1* and the *m6* cluster map on chromosome 6 (Bućan *et al.*, in preparation); however, they share less homology with each other than they do with the fruitfly probes used to isolate them (Colberg-Poley *et al.*, 1985a,b; McGinnis *et al.*, 1985c). Two genes in the *m6* cluster, *m6* and *m5*, are known to be expressed in differentiating embryonal carcinoma (EC) cells, during embryogenesis and in certain tissues of adult mice (Colberg-Poley *et al.*, 1985a,b). The *Hox-2* cluster, which contains at least four boxes, has been mapped onto chromosome 11, hybridizes extensively to human DNA located on chromosome 17, and is transcribed during embryogenesis (Joyner *et al.*, 1985a; Jackson *et al.*, 1985; Hart *et al.*, 1985). Finally, one of the two regions containing homology with the *Drosophila en* gene, *Mo-en.1*, is located on chromosome 1 and is also expressed during murine embryogenesis (Joyner *et al.*, 1985b). It has been suggested that the genes containing homeo box sequences may play a role during embryogenesis and that their conservation might have resulted from selective pressure to retain such a function (McGinnis *et al.*, 1984b; Jackson *et al.*, 1985). Whether this is actually the case remains to be determined.

We report here the isolation of a homeo box-containing region which is expressed during embryogenesis and has not been previously identified. The inserts in two recombinant phage span some 22 kb of genomic DNA from chromosome 15 where the newly identified locus lies. The expression of these sequences was established by Northern analyses of cellular RNA from differentiating F9 cells and of RNA from embryonic tissue. The induction of this gene (*m31*) is later than that of *m6* homeo box gene and occurs at a time when the amount of *m6* RNA decreases. The implications of this finding are discussed.

## Results

### *Isolation of murine homeo box genes*

A mouse genomic library (Frischauf *et al.*, 1983) was screened by hybridization at low stringency with *Antp* probe from

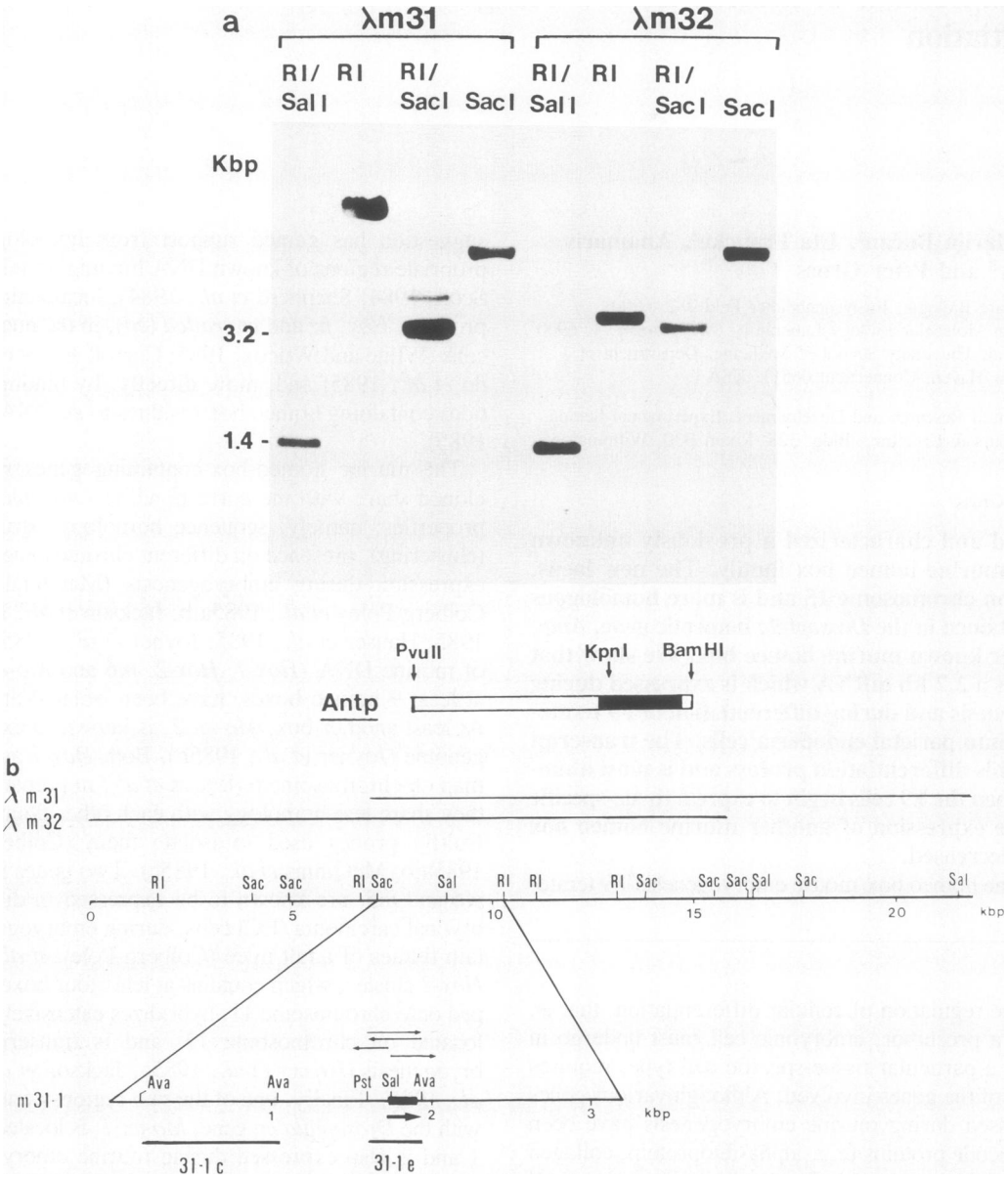


**Fig. 1.** Southern blot analysis of cloned murine DNA and restriction mapping of the *m31* locus. **(a)** Recombinant phage DNAs containing mouse genomic DNA inserts (λm31, λm32) digested with several restriction enzymes, were fractionated on an agarose gel and transferred to a Gene Screen Plus membrane filter. The blot was hybridized to the $^{32}$P-labelled *Bam*HI – *Kpn*I fragment of p903G which contains almost exclusively homeo box sequences of the *Drosophila Antp* gene. Hybridization was performed under reduced stringency conditions as described in Materials and methods. The solid box represents the homeo box sequences. **(b)** Restriction mapping of murine DNA contained in the partially overlapping inserts of λm31 and λm32. The 3.2 kb *Eco*RI – *Sac*I fragment which hybridizes strongly to the *Antp* probe was subcloned into pSP65 to give pm31-1. Finer mapping localized the region of homology to a 330 bp *Pst*I – *Ava*I subfragment, 31-1e. The thick line represents the homeo box, the arrow below shows the orientation of the box as determined by sequence analysis. The sequencing strategy is indicated by the arrows above the box.

*Drosophila* DNA (Colberg-Poley *et al.*, 1985a). This paper concerns the characterization of two recombinant phage, λm31 and 32, containing a 22 kb region of murine DNA. Southern blots of the λm DNAs cleaved with several restriction enzymes were hybridized with a *Bam*HI – *Kpn*I fragment from p903G (Figure 1a). This probe contains almost exclusively the homeo box sequences from the *Drosophila Antp* gene. Figure 1a shows the hybridization of λm31 and 32 with the *Antp* probe. The murine DNAs contained in these phage overlap partially and share a 1.4 kb *Eco*RI – *Sal*I fragment which hybridized to the *Antp* probe. This fragment is contained in a 3.2 kb *Eco*RI – *Sac*I fragment. The additional band seen in this cleavage of λm31 DNA is a partial digestion product. The restriction enzyme map of the murine DNA, constructed by partial digestion of the oligolabelled λm DNAs (Rackwitz *et al.*, 1984) is shown on Figure 1b.

The 3.2 kb *Eco*RI – *Sac*I fragment from this region was subcloned into the pSP65 vector and the resulting clone, pm31-1, was analyzed further. Hybridization located the region of homology to a 330 bp *Pst*I – *Ava*I fragment which was then sequenced.

### *Sequence analysis of the m31 homeo box*

The strategy to determine the nucleotide sequence of the murine homeo box using the chain termination technique (Sanger *et al.*,

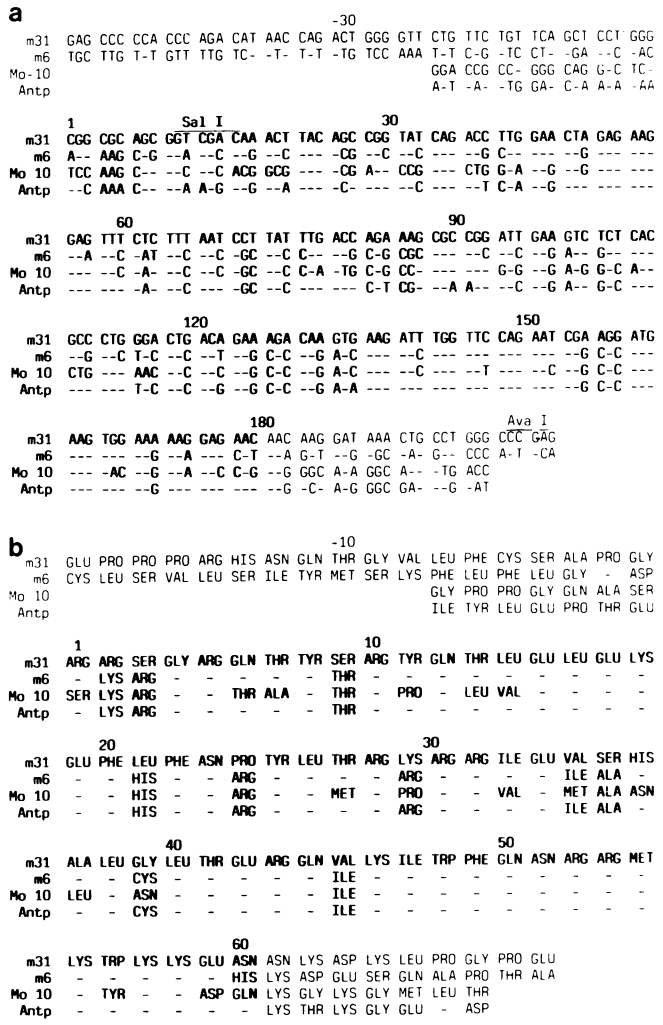


**Fig. 2.** **(a)** Nucleotide sequence of the *m31* homeo box DNA aligned 5′ to 3′ with homeo box sequences of mouse, *m6* (Colberg-Poley *et al.*, 1985a), *Mo-10* (McGinnis *et al.*, 1984c) and *Drosophila*, *Antp* (McGinnis *et al.*, 1984b). Only different nucleotides are indicated, as identical nucleotides are indicated by hyphens. The sequence of *m6* has been corrected at nucleotides 91–94. **(b)** Conceptual translation of the conserved homeo box region of *m31* and comparison to the amino acid sequence of other homeo boxes.

1977) is represented on Figure 1b. The nucleotide sequence determined for *m31* is shown on Figure 2a. Comparison of this region with homeo box sequences of the murine (*m6, Mo-10*) and of *Drosophila* (*Antp*) genomes revealed lower homologies than those observed previously for the *m6* cluster (Figure 2a). Interestingly, the highest degree of nucleotide sequence homology is to *Antp* (74%, Table I), whereas, considerably less homology to some other murine homeo box sequences is noted. Conceptual translation of *m31* homeo box nucleotide sequences is shown in Figure 2b. As previously found, in most cases, the amino acid homology to other homeo boxes is greater (40–83%, Table I) than that of the nucleotide sequences, consistent with the hypothesis that a functional protein domain is being conserved. Furthermore, the amino acid sequences putatively encoded by *m31* and other known homeo boxes show the strongest conservation in the latter portion of the homeo box (amino acid residues 42–59), the region thought to contain an α-helical domain common to DNA-binding proteins (Laughon and Scott, 1984; Shepherd *et al.*, 1984).

**Table I.** Degree of homology of the *m31* homeo box as compared to homeo boxes from mouse (*m6, m5, Mo-10, Mul, H24.1, Mo-en.1*), human (*Hul, Hu2*), *Drosophila* (*Antp, ftz*) and *Xenopus* (*MM3*)

| Homeo boxes | Nucleotide homoloy (%) | Amino acid homology (%) | References |
|---|---|---|---|
| Mouse | | | |
| *m6* | 67 | 82 | Colberg-Poley *et al.*, 1985a,b |
| *m5* | 71 | 80 | |
| *Mo-10* | 57 | 62 | McGinnis *et al.*, 1984c |
| *H24.1, Mul* | 65 | 72 | Jackson *et al.*, 1985; Hauser *et al.*, 1985 |
| *Mo-en.1* | 46 | 40 | Joyner *et al.*, 1985b |
| Human | | | |
| *Hul* | 66 | 70 | Levine *et al.*, 1984 |
| *Hu2* | 72 | 77 | |
| *Drosophila* | | | |
| *Antp* | 74 | 83 | McGinnis *et al.*, 1985b |
| *ftz* | 67 | 67 | |
| *Xenopus* | | | |
| *MM3* | 73 | 82 | Müller *et al.*, 1984 |

Only the nucleotides of positions 1–180 and the corresponding amino acid residues 1–60 of the homeo boxes were compared.

### *Chromosomal mapping*

The chromosomal location of this locus was determined using Southern blot analysis of DNA isolated from 11 mouse × Chinese hamster somatic cell hybrids probed with the 0.8 kb *Ava*I 31-1c fragment (Figure 3). These cell lines together contain the full complement of mouse chromosomes except for chromosome 11. The 31-1c probe hybridized with a 5.2 kb *Taq*I fragment in the mouse genome and to a 3 kb and a 1.7 kb fragment in Chinese hamster DNA. The pattern of hybridization of cell hybrid DNAs assigned the probe 31-1c to chromosome 15. There were multiple cases discordant with location on any chromosome other than chromosomes X and 15 (Table II). There was no discordant case for chromosome 15 and chromosome X was excluded by a single discordant hybrid. Therefore, we studied another cell hybrid, EVI-5 (gift of R.Schäfer and K.Willecke), which contained only mouse X; this hybrid DNA was negative for 31-1c DNA.

### *Expression of m31 sequences in mouse embryos and in differentiating F9 cells*

To test whether *m31* is transcriptionally active during embryogenesis, we performed Northern analyses with RNA from F9 cells and from 12-day embryos. Firstly, the 330 bp 31-1e fragment was $^{32}$P-labelled and hybridized to F9 genomic DNA cleaved with *Eco*RI, *Hin*dIII and *Kpn*I. The probe detected under stringent conditions a single 3.6 kb *Eco*RI fragment (Figure 4a). This fragment appeared in the DNA contained in λm32 (see Figure 1b). As the probe detected a single copy DNA, it was suitable for Northern analyses (Figure 4b). Poly(A)$^+$ RNA from F9 stem cells and F9 cells treated for various time periods with retinoic acid (RA) and dibutyryl cyclic AMP (cAMP) were separated on 0.9% formaldehyde agarose gels and blotted onto Gene Screen Plus. Poly(A)$^+$ RNAs from 12-day embryos and extra-embryonal tissues were also examined. A 2.7 kb RNA which is encoded by *m31* was induced in F9 cells treated for 42 h with RA and cAMP (left panel). The amount of this RNA increased until 76 h and then decreased thereafter. Although the *m31* transcript was still detectable at day 5, its detection required

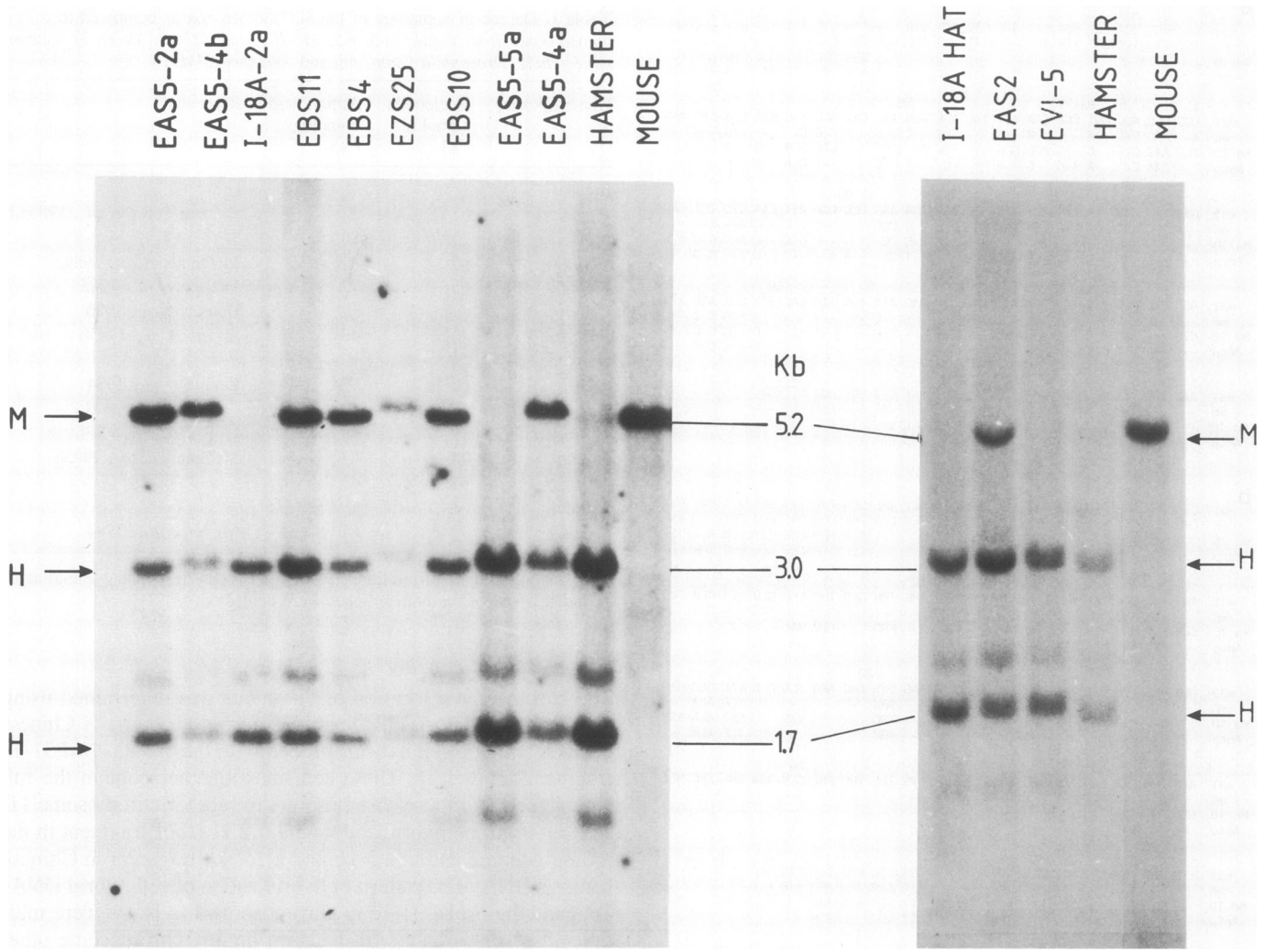


**Fig. 3.** Hybridization of the $^{32}$P-labelled fragment 31-1c (see Figure 1b) to *Taq*I-digested mouse, hamster and mouse × hamster hybrid cell DNAs.

**Table II.** Discordancy analysis of the probe 31-1c with mouse chromosomes in mouse × Chinese hamster hybrid clones

| Hybrids | Probe 31-1c | Mouse chromosomes | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | x |
| I-18A HAT | − | + | + | − | + | − | + | − | − | − | + | − | + | − | − | − | + | + | L | + | + |
| I-18A-2a aza | − | + | + | − | L | + | L | − | − | − | + | − | + | − | − | − | − | + | + | + | − |
| EAS2 | + | L | − | − | − | + | + | − | − | + | − | − | + | − | − | + | + | + | − | L | + |
| EAS5-2a | + | + | + | + | − | + | + | − | + | − | − | − | − | − | + | + | + | + | − | + | + |
| EAS5-4a | + | − | − | − | − | − | − | − | − | − | − | − | − | − | − | + | + | − | − | − | + |
| EAS5-4b | + | + | + | + | + | + | + | + | + | − | − | − | − | − | − | + | − | + | + | + | + |
| EAS5-5a | − | − | − | − | − | − | − | + | − | − | − | − | − | − | − | − | + | − | − | − | − |
| EBS4 | + | − | − | − | − | − | − | − | − | − | − | − | − | − | − | + | − | + | − | − | + |
| EBS10 | + | L | + | + | − | − | − | + | − | L | + | − | + | + | L | + | + | − | + | − | + |
| EBS11 | + | − | − | − | − | − | − | + | − | − | + | − | + | − | − | + | + | − | − | − | + |
| EZS25 | + | − | + | − | + | − | − | + | + | − | − | − | + | + | − | + | + | + | L | + | + |
| Discordant hybrids | | 6 | 6 | 5 | 7 | 6 | 6 | 5 | 5 | 6 | 8 | 8 | 6 | 6 | 6 | 0 | 4 | 5 | 6 | 6 | 1 |
| Total hybrids | | 9 | 11 | 11 | 10 | 11 | 11 | 11 | 11 | 10 | 11 | 11 | 11 | 11 | 10 | 11 | 11 | 11 | 9 | 10 | 11 |

L = Data for chromosomes present at a frequency of ≤0.1 were excluded.

the presence of three times more RNA in the sample (right panel). To standardize more accurately the amount of poly(A)$^+$ RNA in each sample, the blot was re-hybridized with β-actin DNA as a probe (lower panel). The maximal level of *m31* expression seen in differentiating F9 cells was about five times less than that observed in 12-day embryonic tissues. In 12-day extra-embryonal tissues, no expression of *m31* sequences was detectable. Analysis of the F9 poly(A)$^+$ RNA samples by hybridiza-

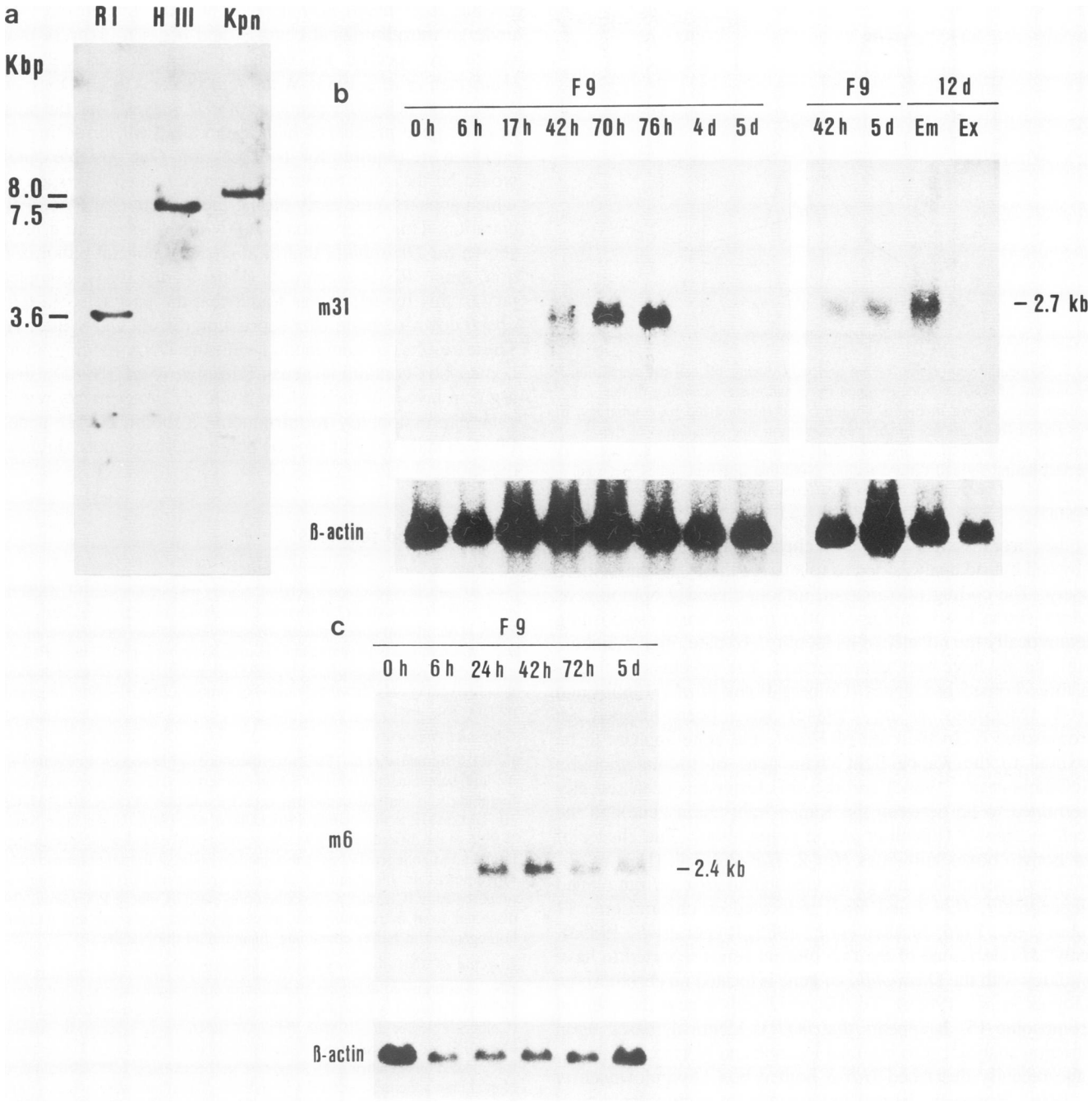


**Fig. 4.** **(a)** Southern hybridization analysis of F9 DNA samples of F9 teratocarcinoma DNA (8 μg/lane) digested with *Eco*RI, *Hin*dIII and *Kpn*I were separated on an 0.9% agarose gel, transferred to Gene Screen Plus and hybridized to the $^{32}$P-labelled *Pst*I – *Ava*I restriction fragment 31-1e containing the homeo box, as described in Materials and methods. **(b)** and **(c)** Northern blot analysis of F9 teratocarcinoma and mouse embryonal RNA. Samples of poly(A)$^+$ RNA (5 – 7.5 μg/lane) from F9 stem cells, F9 cells treated for different time periods with RA and cAMP, from 12-day mouse embryos and extra-embryonal tissues were fractionated on 0.9% formaldehyde-agarose gels and transferred to Gene Screen Plus membrane. The blots were hybridized to the purified $^{32}$P-labelled 31-1e DNA fragment **(b)**, to the 187 bp *Eco*RI – *Pvu*II m6 fragment **(c)** or to β-actin probe (lower panels) as described in Materials and methods. Exposure times were 6 days (**b**, left panel) or 10 days (**b**, right panel) for the *m31* hybridizations and 24 h for the *m6* and β-actin hybridizations.

tion with the 187 bp *Eco*RI/*Pvu*II *m6* homeo box probe showed that the 2.4 kb *m6* transcript was induced within 24 h (Figure 4c). At this time, the *m6* 2.4 kb transcript in differentiating F9 cells is about ten times more abundant than the 2.7 kb *m31* RNA. Following 72 h of differentiation, the F9 cells expressed less of the *m6* homeo box transcript. Thus, expression of *m31* reached a maximum 3 days after differentiation of the F9 cells, whereas, *m6* was maximally expressed at 1 – 2 days (Figure 5).

## Discussion

We have isolated, based upon homology to the *Antp* homeo box, a new member of the murine homeo box gene family. It shares with the other members of this family and with the homeo box genes of *Drosophila* both a conserved amino acid sequence, suggestive of DNA binding and, most importantly, expression during embryogenesis. *m31*, however, is distinct from previously described murine homeo boxes. Analysis of the restriction map

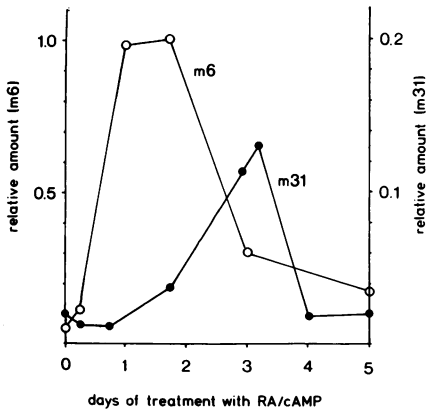


**Fig. 5.** Sequential expression of *m6* and *m31* during F9 cell differentiation. The amount of RNA was determined by densitometer scanning of the Northern blots shown in Figure 4b and c and standardized by comparison to the β-actin hybridizations. In this representation, the abundance of the transcripts is related to the maximal expression of *m6* (relative amount = 1) observed during F9 cell differentiation. Note that different scales are used for *m6* and *m31* because of the great differences in the abundance of the transcripts.

and the chromosomal location on chromosome 15, as determined by cell hybrid analysis, led us to analyze these sequences more closely. The coding information, although clearly homologous to *Antp*, shows less homology to other murine boxes establishing unequivocally the novelty of its identity. To date, nine members of this family have been characterized. The *m6* cluster containing three homeo box genes (Colberg-Poley *et al.*, 1985b) and Hox-1 (McGinnis *et al.*, 1984c) have been mapped on chromosome 6. In the case of *Hox-1*, the gene appears to be proximal to the Kappa light chain gene of immunoglobulin (McGinnis *et al.*, 1984c) and the *m6* cluster has been recently determined to lie between the Kappa light chain gene and the T cell receptor β (Bućan *et al.*, in preparation). The cluster of four homeo boxes known as *Hox-2*, *Hox-2.1-2.4*, (Hart *et al.*, 1985) and presumably the genes identical to *Hox 2.1* but cloned independently, *H24.1* and *Mu1*, is located on chromosome 11 (Rabin *et al.*, 1985; Jackson *et al.*, 1985; Joyner *et al.*, 1985a). Lastly, *Mo-en.1*, one of the two murine genes reported to have homology with the *Drosophila en* gene is located on chromosome 1 (Joyner *et al.*, 1985b). To this, *m31* adds a new member on chromosome 15. It appears that *m31* is identical, based upon restriction analysis, chromosomal location and Northern analyses, to the recently described *Hox-3* homeo box (Awgulewitsch *et al.*, 1986). However, the nucleotide sequence is different at positions 8,9 and 97 from that of *Hox-3*, which may be due to the different mouse strain used to prepare the genomic library. It is probable, based upon hybridization with the *m6* homeo box probe, that another not yet characterized homeo box lies some 8 kb upstream from *m31*. We are presently sequencing this region to determine precisely the source of homology.

Differentiation of an EC cell line, F9, by treatment with RA and cAMP, which results in the production of extra-embryonic endoderm, induced a 2.7 kb RNA from the *m31* locus. The *m31* transcript appeared within 42 h of treatment and was similar in size to a poly(A)$^+$ transcript present in 12-day mouse embryos as well. The temporal expression of *m31* in embryos concurs with the expression previously observed for two other homeo box transcripts, *m6* and *m5;* however, the results obtained in differentiating F9 cells are indeed different. Upon treatment of F9 cells with RA and cAMP, a stable *m6* 2.4 kb RNA is detectable within 6 h (Colberg-Poley *et al.*, 1985a and unpublished results). On the other hand, *m31* appears at a time when the F9 cells undergo morphological changes and begin to express proteins which are specific for certain tissues such as collagen IV (Kurkinen *et al.*, 1983; Wang and Gudas, 1983; Dony *et al.*, 1985). This delayed appearance of *m31* suggests that the presence of other gene products which are induced earlier during differentiation are required for its expression. One possible candidate would be the gene product of another homeo box such as *m6* which appears quite early during the induction of F9 cell differentiation. This observation is not without precedent: it has been shown for *Drosophila* that in absence of normal bithorax function, *Antp*$^+$ RNAs accumulate (Hafen *et al.*, 1984). Complementary to this report, it was found that the products of *abd-A* or *abd-B* down-regulate *Ubx* expression (Struhl and White, 1985). These results, in conjunction with the DNA binding ability of homeo box-containing genes (Desplan *et al.*, 1985), strongly argue in favour of their role in positively or negatively regulating gene expression. By re-introducing a modified *m31* gene into F9 stem cells we will now be able to address more directly this question.

## Materials and methods

*Screening of the mouse genomic library*

$3 \times 10^5$ members of a mouse genomic library prepared in λEMBL 3A (Frischauf *et al.*, 1983; mouse strain B10/Br) were screened as described previously (Colberg-Poley *et al.*, 1985a) under reduced stringency conditions (McGinnis *et al.*, 1984b) using inserts from both *Antp* and *ftz* homeo box clones, p903G and pFS2 (Garber *et al.*, 1983; Kuroiwa *et al.*, 1984, 1985) respectively, as probes. For the subsequent screening only the *Antp* homeo box probe, a *Bam*HI – *Kpn*I fragment from p903G, was used.

*Cells and embryos*

F9 cells (Bernstine *et al.*, 1973) were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum. The cells were induced to differentiate into parietal endoderm cells by addition of $5 \times 10^{-7}$ M RA and $10^{-3}$ M cAMP to the medium (Strickland *et al.*, 1980). Under these conditions, >90% of the differentiated F9 cells expressed the endodermal specific proteins, laminin and collagen IV, as determined by indirect immunofluorescence staining, whereas the F9 stem cells showed no reaction. Twelve-day embryos were prepared from ICR mice as described previously (Colberg-Poley *et al.*, 1985b), dissected from extra-embryonal tissues (placenta, yolk sac and amnion) and homogenized in buffer containing guanidinium thiocyanate.

*Isolation of RNA and Northern hybridization analysis*

RNA was prepared using the guanidinium thiocyanate method (Chirgwin *et al.*, 1979) and purified by centrifugation through a CsCl cushion as described previously (Colberg-Poley *et al.*, 1985a). Poly(A)$^+$ RNAs were selected by retention on oligo (dT) cellulose columns. Aliquots of poly(A)$^+$ RNA (5 – 7.5 μg/lane) were fractionated on denaturing 0.9% agarose gels containing 3.7% formaldehyde and blotted onto a Gene Screen Plus membrane. The filters were hybridized in 50% formamide, 1 M NaCl, 1% SDS, 200 μg/ml denatured calf thymus DNA and 1 ng/ml oligo-primed (Feinberg and Vogelstein, 1983) or nick-translated [$^{32}$P]DNA probe ($5 \times 10^8 - 10^9$ c.p.m./μg). Washes were twice in 2 × SSC at room temperature, twice in 2 × SSC, 1% SDS at 60°C, and finally twice in 0.1 × SSC at 48°C.

*Isolation of DNA and Southern hybridization analysis*

F9 cellular DNA was prepared by treatment of the cells with lysis buffer (50 mM Tris-HCl, 20 mM EDTA, 100 mM NaCl, pH 7.5) and digestion with proteinase K (50 μg/ml) for 24 h. DNA was further purified by phenol extraction and ethanol precipitation. DNA digested with *Eco*RI, *Hin*dIII and *Kpn*I was fractionated on a 0.9% agarose gel and transferred to Gene Screen Plus as described elsewhere (Chromczynski and Quasba, 1984). The blots were hybridized in 1 M NaCl, 1% SDS, 200 μg/ml denatured calf thymus DNA and 1 ng/ml oligo-primed [$^{32}$P]DNA (specific activity $5 \times 10^8 - 10^9$ c.p.m./μg) at 65°C. Washes were twice in 2 × SSC at room temperature, twice in 2 × SSC at 60°C and twice in 0.1 × SSC at 48°C.

Digested phage DNA was separated on 0.9% agarose gels and blotted to Gene Screen Plus. Hybridization and washes were under reduced stringency conditions as described by McGinnis *et al.* (1984b) except that SDS concentration was raised to 0.6%. The probe used for hybridization was the *Bam*HI – *Kpn*I fragment of the *Antp* cDNA insertion in p903G.

*DNA sequence analysis*

The 3.2 kb *Eco*RI – *Sac*I fragment, *m31-1*, was subcloned into pSP65 to give pm31-1. The 330 bp *Pst*I – *Ava*I subfragment 31-1e, was isolated and cloned into the M13 mp10 and mp11 vector DNAs digested with *Pst*I and *Sma*I. Furthermore, the 1.4 kb *Sal*I – *Eco*RI and 1.8 kb *Sal*I – *Sac*I lower fragments of pm31-1 were isolated and cloned into M13 mp10 DNA cut with the appropriate restriction enzymes. Single-stranded templates were prepared and sequenced using the dideoxy chain termination method of Sanger *et al.* (1977). The sequence was determined along both strands. Part of the sequence was, in addition, confirmed along one strand using different overlapping clones.

*Chromosome mapping*

Total cellular DNAs were isolated from mouse liver, Chinese hamster cell line V79/380 and various mouse × Chinese hamster hybrid cell lines (Francke *et al.*, 1977; Francke and Taggart, 1979). DNA samples (10 μg) were digested with *Taq*I, electrophoresed in an 0.7% agarose gel and transferred to Gene Screen Plus membrane in denaturation solution as previously described by Herrmann *et al.* (1986). The filter was exposed to u.v. light and hybridized according to Church and Gilbert (1984), using $2 \times 10^6$ c.p.m./ml of oligonucleotide-primed $^{32}$P-labelled 31-1c insert DNA.

## Acknowledgements

We wish to thank Michael Kessel and Stephan D.Voss for helpful discussions and critical reading of the manuscript and Uwe Drescher for help in computer analyses. We are also grateful to Erwin Wagner for obtaining the mouse embryos, to Hans-Richard Rackwitz for providing oligonucleotides and to Alexander Awgulewitsch for making available his data prior to publication. This research was supported by the Deutsche Forschungsgemeinschaft (SFB 229-86). M.B. was supported by a grant of the Deutsche Forschungsgemeinschaft (LE496/1-1) given to Hans Lehrach and U.F. by a NIH research grant (GM 26105) to U.F.

## References

Awgulewitsch,A., Utset,M.F., Hart,C.P., McGinnis,W. and Ruddle,F.H. (1986) *Nature*, **320**, 328–335.
Bernstine,E.G., Hooper,M.L., Grandchamp,S. and Ephrussi,B. (1973) *Proc. Natl. Acad. Sci. USA*, **70**, 3899–3903.
Carroll,S.B. and Scott,M.P. (1985) *Cell*, **43**, 47–57.
Chirgwin,J.M., Przybyla,A.E., MacDonald,R.J. and Rutter,W.J. (1979) *Biochemistry*, **18**, 5294–5299.
Chromczynski,P. and Quasba,P.K. (1984) *Biochem. Biophys. Res. Commun.*, **122**, 340–344.
Church,G.M. and Gilbert,W. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 1991–1995.
Colberg-Poley,A.M., Voss,S.D., Chowdhury,K. and Gruss,P. (1985a) *Nature*, **314**, 713–718.
Colberg-Poley,A.M., Voss,S.D., Chowdhury,K., Stewart,C.L., Wagner,E.F. and Gruss,P. (1985b) *Cell*, **43**, 39–45.
Desplan,C., Theis,J. and O'Farrell,P.H. (1985) *Nature*, **318**, 630–635.
DiNardo,S., Kuner,J.M., Theis,J. and O'Farrell,P.H. (1985) *Cell*, **43**, 59–69.
Dony,C., Kessel,M. and Gruss,P. (1985) *Nature*, **317**, 636–639.
Feinberg,A.P. and Vogelstein,B. (1983) *Anal. Biochem.*, **132**, 6–13.
Francke,U. and Taggart,R.T. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 5230–5233.
Francke,U., Lalley,P.A., Moss,W., Ivy,J. and Minna,J.D. (1977) *Cytogenet. Cell Genet.*, **19**, 57–84.
Frischauf,A.-M. (1985) *Trends Genet.*, **1**, 100–103.
Frischauf,A.-M., Lehrach,H., Poustka,A. and Murray,N. (1983) *J. Mol. Biol.*, **170**, 827–842.
Garber,R.L., Kuroiwa,A. and Gehring,W.J. (1983) *EMBO J.*, **2**, 2027–2036.
Gehring,W.J. (1985) *Cell*, **40**, 3–5.
Gluecksohn-Waelsch,S. (1983) In *Teratocarcinoma Stem Cells.*, Cold Spring Harbor Laboratory. pp. 3–13.
Green,M.C. (ed) (1981) *Genetic Variants and Strains of the Laboratory Mouse.* Gustav Fischer Verlag, Stuttgart, New York.
Hafen,E., Levine,M. and Gehring,W.J. (1984) *Nature*, **307**, 287–289.
Hart,C.P., Awgulewitsch,A., Fainsod,A., McGinnis,W. and Ruddle,F.H. (1985) *Cell*, **43**, 9–18.
Hauser,C.A., Joyner,A.L., Klein,R.D., Learned,T.K., Martin,G.R. and Tjian,R. (1985) *Cell*, **43**, 19–28.
Herrmann,B., Bućan,M., Mains,P.E., Frischauf,A.-M., Silver,L.M. and Lehrach,H. (1986) *Cell*, **44**, 469–476.
Jackson,I.J., Schofield,P. and Hogan,B. (1985) *Nature*, **317**, 745–748.
Joyner,A.L., Lebo,R.V., Kan,Y.W., Tjian,R., Cox,D.R. and Martin,G.R. (1985a) *Nature*, **314**, 173–175.
Joyner,A.L., Kornberg,T., Coleman,K.G., Cox,D.R. and Martin,G.R. (1985b) *Cell*, **43**, 29–37.
Kurkinen,M., Barlow,D.P., Helfman,D.M., Williams,J.G. and Hogan,B.L.M. (1983) *Nucleic Acids Res.*, **11**, 6199–6209.
Kuroiwa,A., Hafen,E. and Gehring,W.J. (1984) *Cell*, **37**, 825–831.
Kuroiwa,A., Kloter,U., Baumgartner,P. and Gehring,W.J. (1985) *EMBO J.*, **4**, 3757–3764.
Laughon,A. and Scott,M.P. (1984) *Nature*, **310**, 25–31.
Law,S., Tamoaki,T., Kreuzaler,F. and Dugaiczyk,A. (1980) *Gene*, **10**, 54–61.
Levine,M., Rubin,G.M. and Tjian,R. (1984) *Cell*, **38**, 667–673.
McGinnis,W., Levine,M., Hafen,E., Kuroiwa,A. and Gehring,W.J. (1984a) *Nature*, **308**, 428–433.
McGinnis,W., Garber,R.L., Wirz,J., Kuroiwa,A. and Gehring,W.J. (1984b) *Cell*, **37**, 403–408.
McGinnis,W., Hart,C.P., Gehring,W.J. and Ruddle,F.H. (1984c) *Cell*, **38**, 675–680.
Mlodzik,M., Fjose,A. and Gehring,W.J. (1985) *EMBO J.*, **4**, 2961–2969.
Müller,M.M., Carrasco,A.E. and DeRobertis,E.M. (1984) *Cell*, **39**, 157–162.
Rabin,M., Hart,C.P., Ferguson-Smith,A., McGinnis,W., Levine,M. and Ruddle,F.H. (1985) *Nature*, **314**, 175–178.
Rackwitz,H.-R., Zehetner,G., Frischauf,A.-M. and Lehrach,H. (1984) *Gene*, **30**, 195–200.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Schnieke,A., Harbers,K. and Jaenisch,R. (1983) *Nature*, **304**, 315–320.
Shepherd,J.C.W., McGinnis,W., Carrasco,A.E., DeRobertis,E.M. and Gehring,W.J. (1984) *Nature*, **310**, 70–71.
Strickland,S., Smith,K.K. and Marotti,K.R. (1980) *Cell*, **21**, 347–355.
Struhl,G. and White,R.A.H. (1985) *Cell*, **43**, 507–519.
Wang,S.-Y. and Gudas,L.J. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 5880–5884.
White,R.A.H. and Wilcox,M. (1985) *EMBO J.*, **4**, 2035–2043.
Woychik,R.P., Stewart,T.A., Davis,L.G., D'Eustachio,P. and Leder,P. (1985) *Nature*, **318**, 36–40.

*Received on 1 April 1986; revised on 11 June 1986*