The EMBO Journal vol.7 no.7 pp.2131 – 2138, 1988

Expression of multiple homeobox genes within diverse mammalian haemopoietic lineages

Kritaya Kongsuwan, Elizabeth Webb, Philip Housiaux and Jerry M.Adams

The Walter and Eliza Hall Institute of Medical Research, Post Office Royal Melbourne Hospital, Victoria, 3050, Australia

Communicated by J.Tooze

Several mouse and human genes encoding the DNA-binding homeobox domain are implicated here in haematopoiesis, a differentiation process maintained throughout life. Four homeobox cDNA clones were isolated from bone marrow and spleen of adult mice and two from the human leukaemia cell line K562. They derive from the *Hox 1.1, Hox 2.3, Hox 6.1* genes and two previously undescribed genes, one of a type (*paired*) not found before in vertebrates. A survey of 36 cell lines of the lymphoid, myeloid and erythroid lineages revealed that certain homeobox transcripts were almost ubiquitous, while others were restricted to certain lineages or even particular cell lines. The expression pattern altered in a myeloid and an erythroid line induced to terminal differentiation, and in novel lines that had switched from a lymphoid to a myeloid phenotype. Altogether, the haemopoietic compartment may contain up to 20 homeobox transcripts. In one myeloid leukaemia, DNA rearrangement has perturbed expression. These findings suggest that homeobox genes may influence developmental decisions within the haemopoietic system.

Key words: homeobox genes/haematopoiesis/differentiation control/DNA-binding proteins/leukaemia

Introduction

It seems likely that the orchestrated changes in the expression of numerous structural genes associated with lineage commitment and maturation are directed by a modest number of regulatory genes. For vertebrates, only one such gene, involved in myogenesis, has been identified (Davis *et al.*, 1987). The more amenable genetics of *Drosophila*, however, has uncovered many of the genes governing the intricate process of insect morphogenesis. The great majority include a conserved domain of 60 amino acid residues denoted the 'homeobox' (reviewed by Gehring, 1987). Its C-terminal half exhibits some similarity to DNA-binding regulatory proteins of prokaryotes and yeast, and the box appears to recognize specific DNA sequences (Desplan *et al.*, 1985; Fainsod *et al.*, 1986). It seems likely that many homeobox proteins serve as transcriptional activators or repressors for classes of both structural and regulatory genes.

The presence in vertebrates of genes bearing the conserved 180-bp homeobox sequence has engendered the hope that many such genes participate in embryogenesis. To date ~18 mouse homeobox genes, falling within seven known loci, have been identified (see Hart *et al.*, 1987; Martin *et al.*, 1987), and many equivalent genes exist in other vertebrates (Acampura *et al.*, 1987). The circumstantial evidence that these genes influence vertebrate development is substantial: many reside in tight clusters, as in *Drosophila*; a few encode proteins closely related to specific *Drosophila* ones; and nearly all are expressed at one stage or another of embryogenesis, often within limited anatomical regions (reviewed by Snow, 1986).

Homeobox genes might also play pivotal roles in adult differentiation processes. A domain that recognizes specific DNA sequences clearly might be used outside embryogenesis, and indeed many mammalian homeobox sequences are expressed within specific adult tissues, as well as in the embryo (reviewed by Schofield, 1987; see Discussion). We are exploring whether homeobox genes could participate in haemopoiesis. The diverse blood cell elements are replenished throughout life from self-renewing haemopoietic stem cells via a hierarchy of progenitor cells (Dexter and Spooncer, 1987). As the mature cells comprise multiple lineages (B and T lymphocytes, granulocytes, macrophages, mast cells, erythrocytes, eosinophils and megakaryocytes), many decisions on lineage commitment and maturation must be made. The availability of numerous leukaemia cell lines representative of particular lineages and maturation stages allows cell-type specificity to be explored, an issue difficult to address in most embryonic systems. Moreover, certain lines can be induced to differentiate (reviewed by Hozumi, 1985).

To determine whether homeobox genes are expressed within haemopoetic cells, we have screened cDNA libraries from mouse haemopoeitic tissues and a human leukemic cell line. Evidence for expression of multiple homeobox genes emerged from the cDNA clones isolated and subsequent hybridization experiments with mRNA from different haemopoietic cell lines. The pattern of expression within lines representing diverse haemopoietic lineages suggests that the expression of certain homeobox genes may be linked to cell type, or degree of maturation. Moreover, the profile of expression changed upon terminal differentiation in a myeloid and an erythroid line, and in novel lines that have switched from a lymphoid to a myeloid phenotype. These findings raise the intriguing possibility that many homeobox genes serve as controlling elements in adult differentiation processes, independent of their probable role in early development.

Results

Multiple cDNA clones revealed with oligonucleotide probes

The most conserved portion of the homeobox spans the second helix of the helix–turn–helix motif implicated in DNA binding, from amino acid residues 44–55 (Gehring, 1987). To detect homeobox sequences, we used as hybridization probes four related oligonucleotides spanning

©IRL Press Limited, Oxford, England

this region (see Materials and methods). They were designed to detect not only homeoboxes of the predominant *Antennapedia* class and the smaller *engrailed* class, the only ones so far identified in vertebrates, but also three other classes recently identified in *Drosophila*: *paired* (Bopp *et al.*, 1986), *bicoid* (Frigerio *et al.*, 1986) and *even-skipped* (*eve*) (Macdonald *et al.*, 1986). To maximize the range of cell types sampled, we initially screened cDNA libraries from murine haemopoietic tissues rather than specific cell lines. The libraries derived from bone marrow, the predominant site of post-natal haemopoiesis (including lymphopoiesis), and from spleen, a major repository of mature B and T lymphocytes, as well as other haemopoietic lineages. These tissues also include non-haemopoietic stromal cells. A screen of $\sim 10^6$ clones yielded 40 hybridizing clones, most of which are accounted for by the four cDNA sequences described below. Hence the candidate homeobox clones are relatively rare, perhaps deriving from only $\sim 0.002-0.01\%$ of the mRNAs.

A richer source proved to be the human cell line K562, which originated from a chronic myeloid leukaemia (Lozzio and Lozzio, 1975), a neoplasm of the haemopoietic stem cell. Although K562 cells are predominantly erythroblastic, residual multi-potential character is indicated by their granulocytic, megakaryocytic and lymphocytic features (e.g. Tetteroo *et al.*, 1984; Fraser and Berridge, 1987). A screen of 2×10^5 K562 clones (constructed by Dr Iswar Hariharan) yielded 15 positives, which represent at least four different cDNA sequences, two of which we have characterized.

Clones from three distinct loci: Hox 1.1, Hox 2.3 and Hox 6.1

Nucleotide sequences from the four murine and two human cDNA clones verified that each contains a homeobox (filled or hatched box in Figure 1). All the clones except S8 fall into the *Antp* class, and two (B2 and S1) come from well-studied murine loci. The bone-marrow-derived B2 sequence (not shown) corresponds to a 300-bp segment at the 3′ end of a recently described *Hox 1.1* cDNA (Kessel *et al.*, 1987). The sequence of S1, which was isolated from the bone marrow as well as spleen libraries, corresponds to the *Hox 2.3* sequence reported during the preparation of our manuscript by Meijlink *et al.* (1987) and Lonai *et al.* (1987). *Hox 2.3* is very similar to its human homologue, HHO.c1 (Simeone *et al.*, 1987).

The murine B7 cDNA clone from bone marrow proved to be a full-length clone of the recently identified murine *Hox 6.1* gene on chromosome 14 (Sharpe *et al.*, 1988). Moreover, the closely related K562 clone K10 corresponds to the recently reported human HHO.c8 cDNA sequence (Simeone *et al.*, 1987). The entire murine *Hox 6.1* (B7) and human HHO.c8/K10 sequences are remarkably similar (Figure 2). If the most 5′ ATG in frame with the box (stippled) is the initiation codon, the predicted 153-residue mouse and human polypeptides differ at only one position (the Leu with an asterisk in Figure 2). It is striking, moreover, that the open reading frame continues for 68

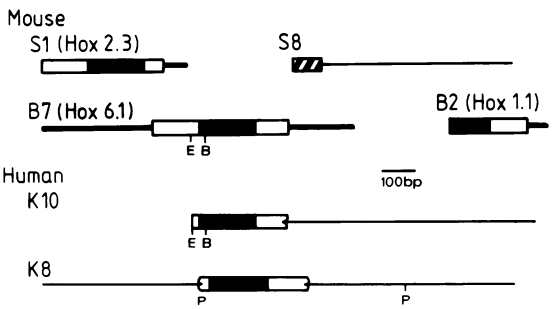

Fig. 1. cDNA clones isolated from mouse spleen (S), mouse bone marrow (B) and the human K562 cell line (K). The homeobox is indicated by a filled box for the *Antp*-type sequences and a cross-hatched one for the *prd*-type box in S8 (see text). Since K10 is the human homologue of B7 (*Hox 6.1*) (see text), K10 is aligned below the B7 sequence. Where the nucleotide sequence is known a thick line denotes non-coding sequences, while the coding region is boxed; a thin line indicates an unsequenced region. All the clones are bounded by introduced *Eco*RI sites (E), so only those thought to be natural ones are indicated; other restriction sites shown are B, *Bgl*II; and P, *Pst*I.

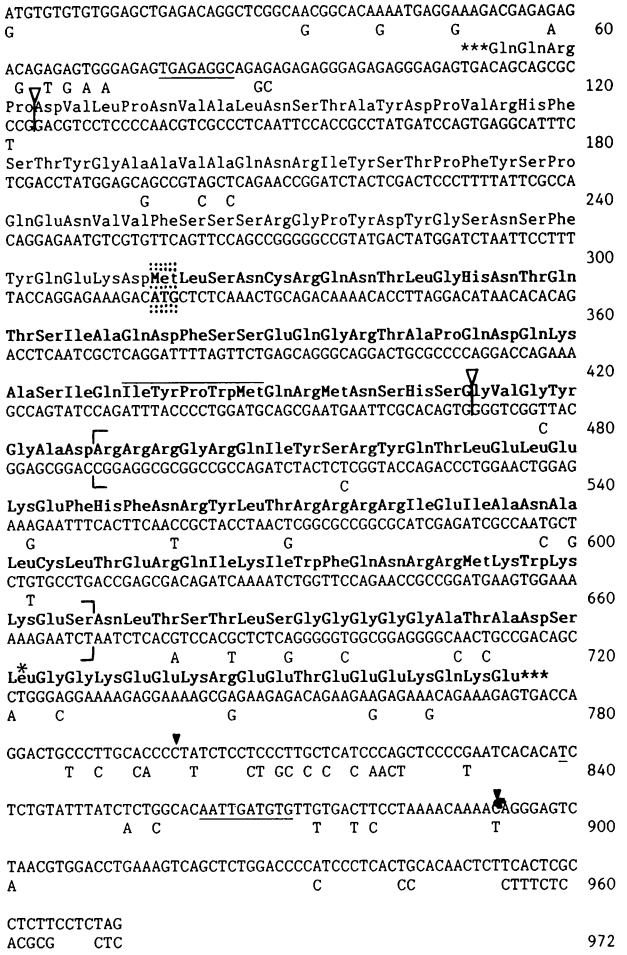

Fig. 2. Nucleotide sequence of the complete *Hox 6.1* (B7) cDNA. Nucleotides that differ in the human HHO.c8 sequence (Simeone *et al.*, 1987) are indicated; the segments underlined in the untranslated regions are not present in the human sequence, while filled arrowheads in the 3′ untranslated region mark positions where the human sequence has extra residues. Amino acid residues in light typeface indicate that the open reading frame continues for 68 codons upstream from the presumptive initiator methionine (stippled), past a splice point (first open arrowhead) identified in the human gene (Simeone *et al.*, 1987), raising the possibility that an alternative, larger polypeptide could be made on another transcript (see text). An asterisk indicates the residue that differs between the mouse and human polypeptides. The homeobox is bracketed and the pentapeptide upstream from it, which is conserved in a number of *Antp* type genes, is overlined.

codons upstream from that ATG in both species and differs at only two nucleotides out of 204. Most likely at least part of this region has a coding function. Since the open reading frame spans a splice point (first open arrowhead in Figure 2) identified by Simeone *et al.* (1987), an alternative first exon (or intron promoter) might generate a separate mRNA that encodes a homeobox polypeptide with an extended amino-terminal segment, perhaps accounting for the two prominent *Hox 6.1* transcripts (see below).

Two novel homeobox sequences

The K8 sequence from K562 does not correspond to any previously described homeobox gene. Its box sequence (Figure 3) probably falls within the *Antennapedia* class but is more divergent than *Hox 6.1*, *Hox 2.3* or *Hox 1.1*. Twenty-one of the 60 amino acid residues in the K8 box differ from *Antp*, including two changes within the second helix, albeit conservative ones (Ile/Val). The K8 box is most similar to that of the murine *Hox 1.5* gene (McGinnis *et al.*, 1984), which is related to that of *Hox 2.6* (Lonai *et al.*, 1987), but K8 differs even from *Hox 1.5* at 16 amino acid residues within the box. Thus K8 may define a new *Antp* sub-class.

The S8 box (Figure 3) is more closely related to *Drosophila prd* than to *Antp* or any *Antp*-like mouse homeobox. Although the S8 clone terminates within the box, of the 25 amino acid box residues available for comparison, 15 are identical to *prd* and several other residues involve

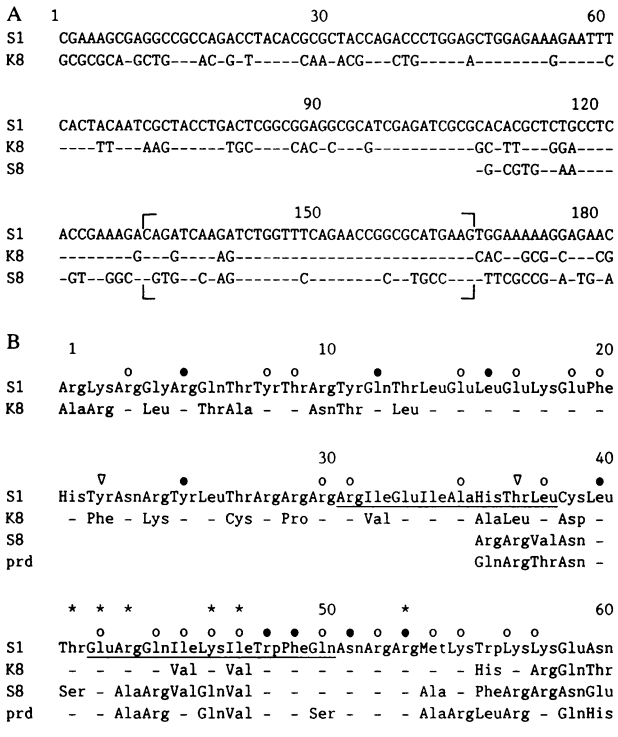

Fig. 3. Nucleotide sequence (**A**) and predicted amino acid sequences (**B**) in the homeobox region of the cDNA clones. Residues that differ from that in S1 (*Hox 2.3*) are indicated. In A, the region corresponding to the probe sequences is bracketed. The K8 clone first analysed lacks T_{59}, but that appears to be an aberration in that clone, because another (shorter) clone has that residue. In B, the *prd* sequence is included to show its similarity to S8 (see text). The residues highly conserved in other homeoboxes or implicated in DNA recognition are indicated by circles and asterisks, respectively. Open arrowheads mark the two residues in S1 that differ from *Antp*.

conservative changes (e.g. Thr/Ser, Ile/Val, Arg/Lys). It is noteworthy that S8 differs from *Antp* at five amino acid residues within the second helix region, including three of the residues thought to be associated with recognition of specific sequences. Hence, the S8 and other *prd*-type homeoboxes are likely to bind to DNA sequences that differ markedly from those recognized by *Antp*- or *engrailed*-type genes. To our knowledge, S8 provides the first evidence that homeobox classes other than *Antp* and *en* exist in vertebrates.

Varied patterns of expression in cell lines

To determine whether the homeobox genes are expressed within particular haemopoietic lineages, we analysed RNA from cloned haemopoietic cell lines, which obviate the problem of cell purity associated with populations isolated from tissues. The lines examined (see Materials and methods), which include two of human origin (K562 and HL60) and 36 of murine, represent diverse haemopoietic cell lineages and maturation stages. Within the myeloid series, early stages are represented by several pro-myelocytic lines, including HL60 (Collins *et al.*, 1978), and by pre-macrophage lines, while four mature macrophage lines were examined. The erythroid series included the complex human line K562 (see above), and two murine lines, while the lymphoid lines comprised six in the T lineage and 15 from various stages of B cell development. The two non-haemopoietic lines examined were BALB/c 3T3 fibroblasts, an embryonic mesenchymal cell line and a bone marrow adherent (stromal) cell line (BAd.2) (Li and Johnson, 1985).

Representative Northern blots of poly(A)$^+$ RNA hybridized to three of the homeobox cDNAs are shown in Figure 4A–C. Table I summarizes the lines that showed detectable hybridization with the five homeobox probes examined. The highest levels observed would be no more than 10–20 molecules per cell. As described below, the patterns of expression varied from nearly ubiquitous to highly restricted.

Preferential macrophage expression of Hox 2.3

The *Hox 2.3* cDNA probe (S1) revealed prominent transcripts of 1.2 and 1.6 kb in 3T3 fibroblasts and certain haemopoietic lines (Figure 4A). Both transcripts probably derive from *Hox 2.3*, since they both hybridized at high stringency to a genomic sequence 3′ to the *Hox 2.3* box (probe a in Figure 5), kindly provided by Dr F.Ruddle. The 1.6-kb RNA is very unlikely to be a nuclear precursor of the 1.2-kb species, because the only known *Hox 2.3* intron is 2.2 kb long (Meijlink *et al.*, 1987). The two *Hox 2.3* mRNAs presumably are generated by alternate splices, promoters, or poly(A) addition sites. Their production may be independently regulated since with various other RNA sources either two transcripts (Sebastio *et al.*, 1987) or only the 1.6-kb mRNA have been observed (Meijlink *et al.*, 1987; Simeone *et al.*, 1987). A 1.4-kb species was observed in the bone marrow adherent line (Figure 4A).

Within the haemopoietic system, *Hox 2.3* expression was remarkably restricted (Figure 4A and Table I). Aside from the complex K562 line, the 1.2- and 1.6-kb transcripts were essentially confined to the four macrophage lines examined, as illustrated by J774 and WR19 in Figure 4A. It is noteworthy that neither transcript was observed within any of eight myeloid lines (Table I), including three pre-macrophage lines. Thus, within the haemopoietic system,

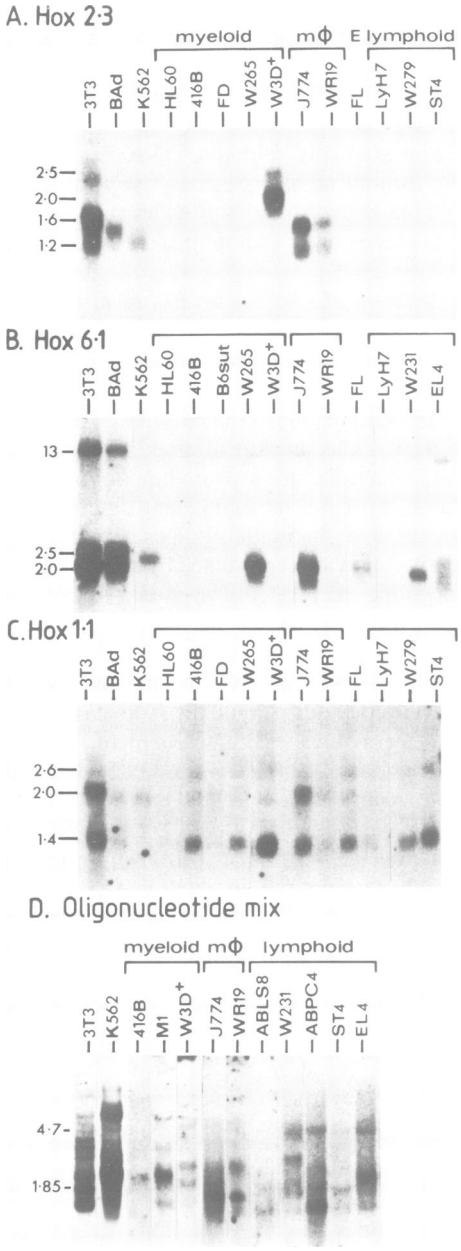

Fig. 4. Transcripts detected with three homeobox cDNA probes and a mixed oligonucleotide in different cell lines. Northern blots of poly(A)$^+$ RNA (4 μg) are shown. Panels contain 3T3 fibroblasts, a bone marrow adherent line and K562, and lines representative of early myeloid cells, of macrophages (Mϕ), an erythroleukemia (E) line and B and T lymphoid lines (see Materials and methods). Transcript sizes are in kb. The probes (see Figure 1) were (**A**) the 450-bp S1 (*Hox 2.3*) cDNA; (**B**) the 5′ 470-bp *Eco*RI fragment of B7 (*Hox 6.1*), outside the box region; (**C**) the 300-bp *Hox 1.1* (B2) probe; (**D**) the mixed oligonucleotide used for screening. The hybridization in A–C was performed at high stringency and in D at low stringency (see Materials and methods).

Hox 2.3 expression is restricted by maturation stage as well as lineage. Small amounts of the 1.6-kb species were observed in one of five pre-B lines, ABLS8; an unusual feature of ABLS8 that may be relevant is that macrophages arise in ABLS8 cultures treated with 5-azacytidine (Boyd and Schrader, 1982).

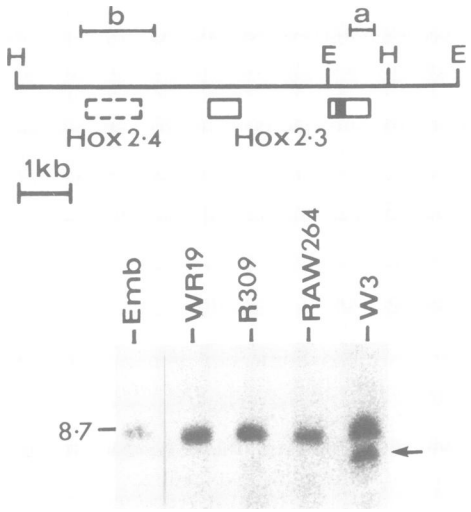

Fig. 5. Rearrangement of DNA in the vicinity of the *Hox 2.4/2.3* genes. The map of the 5′ portion of the *Hox 2* locus (top) derives from Hart *et al.* (1985) and Meijlink *et al.* (1987). The Southern blots show *Hin*dIII digests of genomic DNA from four myeloid or macrophage cell lines and BALB/c embryos (emb). The digests (10 μg DNA) were separated by electrophoresis through a 1% agarose gel and probed with the ^{32}P-labeled 450-bp *Hox 2.3* (S1) fragment. The ~8.7-kb unrearranged fragment spans the *Hox 2.4*. The arrow indicates the 8.4-kb band unique to WEHI3B. On the map, boxes denote the two exons of the *Hox 2.3* gene, the *Hox 2.3* homeobox being filled; the broken box marks the approximate location of the *Hox 2.3* homeobox. E, *Eco*RI; H, *Hin*dIII.

Table I. Homeobox transcripts in haematopoietic cell lines[a]

Cell type	*Hox 2.3*	*Hox 6.1*	*Hox 1.1*	S8	K8
Myeloid	0/8	1/8[b]	8/8	0/6	0/8
Macrophage	4/4	2/4[c]	4/4	0/2	0/4
Erythroid	(1)/3[d]	2/3[e]	3/3	0/2	(1)/3[d]
B lymphoid	1/15[f]	1/7[g]	15/15	0/4	0/8
T lymphoid	0/6	1/4[h]	5/5	0/1	0/5

[a]Cell lines are listed in Materials and methods. Poly(A)$^+$ RNA was examined on Northern blots like those in Figure 4. Lines that showed hybridization are listed in subsequent footnotes, those giving weak signals being italicized.
[b]WEHI265.
[c]J774 and *RAW309*.
[d]K562, which has a complex phenotype (see text).
[e]*K562* and *FL-F4N*.
[f]*ABLS8*.
[g]*WEHI231*.
[h]*EL-4*.

Hox 2.4 expression induced by DNA rearrangement in a myeloid leukaemia

The *Hox 2.3* (S1) probe revealed unique transcripts of 2.0 and 2.5 kb in the promyelocytic leukaemia line WEHI3B (Figure 4A), the D$^+$ subline of which can differentiate to macrophages and granulocytes (Cooper *et al.*, 1982). However, these mRNAs do not derive from *Hox 2.3* itself, because the *Hox 2.3* 3′ genomic probe did not hybridize to the WEHI3B RNAs, nor did the S1 probe at higher stringency. These transcripts arise instead from the adjacent *Hox 2.4* gene, since a *Hox 2.4* probe (probe b in Figure 5; a 1.4-kb *Sac*I fragment in Meijlink *et al.*, 1987), kindly provided by Dr F.Meijlink, hybridized at high stringency.

The anomalous *Hox 2.4* expression in WEHI3B appears to be due to a DNA rearrangement in the *Hox 2.4* vicinity. Figure 5 shows that WEHI3B DNA contains a rearrangement within the *Hin*dIII fragment spanning the *Hox 2.3* and *2.4* genes; probes from both genes revealed the altered fragment. Since the *Eco*RI fragment spanning the 3′ end of *Hox 2.3* was unrearranged (not shown), the alteration must occur within the 6-kb region 5′ to the *Hox 2.3* box. Indeed a preliminary report by Blatt *et al.* (1987) indicates that WEHI3B bears a proviral insertion just 5′ to the *Hox 2.4* gene. Since *Hox 2.4* transcripts were undetectable in any other haemopoietic cell line we examined, they must have been induced by the alteration. Because the change (and transcripts) were present in both the D^+ and D^- sublines, it probably occurred during the transformation events that generated the original WEHI3B leukaemia (see Discussion). This rearrangement is not common, since *Hox 2.4* and *2.3* were not rearranged in seven other myeloid or macrophage lines, three of which are shown in Figure 5, nor in nine lymphoid lines.

Sporadic expression of the Hox 6.1 gene

A *Hox 6.1* (B7) probe, either including the box or only the region 5′ to it, typically revealed prominent 2.0- and/or 2.5-kb species and small amounts of an ~13-kb species (Figure 4B), which may well be a nuclear precursor, as the human homologue (c8) contains an 11-kb intron (Simeone *et al.*, 1987). Since two major transcripts (1.8 and 2.2 kb) were also observed in human tissue samples with various c8 probes (Simeone *et al.*, 1987), both mRNAs probably derive from the same gene, perhaps by alternative splicing (see above). In contrast to *Hox 2.3*, the pattern of *Hox 6.1* expression was not clearly related to cell type. Transcripts were found in one of eight early myeloid lines (WEHI265), two of four macrophage lines and two of 11 lymphoid lines, as well as two of three erythroid lines examined (Table I). No clear trend is yet evident.

Widespread expression of Hox 1.1-related genes

The *Hox 1.1* probe revealed transcripts in all the lines examined (Figure 4C and Table I). In most cell lines, at least three diffferent transcripts were observed: 1.4−1.5 kb, 2.0 and 2.6 kb. A similarly complex pattern was observed in various tissues by Colberg-Poley *et al.* (1985a,b).The prominent 1.4- to 1.5-kb band (probably two species), does not derive from *Hox 1.1* itself, because that band was not revealed by an oligonucleotide just 3′ to the *Hox 1.1* box, nor by a probe 5′ to the *Hox 1.1* box, from a cDNA clone (Kessel *et al.*, 1987) kindly provided by M.Kessel and P.Gruss. Whether the 1.4-kb transcript bears a homeobox is not yet established. Nor is it clear whether both the larger transcripts are generated from the *Hox 1.1* gene. The 2.6-kb RNA is unlikely to be an unspliced precursor of the 2.0-kb transcript, because the identified *Hox 1.1* intron is 1.1 kb long (Kessel *et al.*, 1987).

Highly restricted expression of S8 and K8

S8, the *prd*-like cDNA from the mouse spleen library, probably originated from a non-haemopoietic cell. An S8 transcript of 1.4 kb was relatively abundant in 3T3 fibroblasts and the bone marrow adherent (BAd.2) cell line (not shown) but was undetectable in any of the 17 haemopoietic cell lines surveyed (Table I). It remains possible that S8 functions in cells that influence haemopoietic cell development, such as BAd. 2 (Li and Johnson, 1985).

The K8 human cDNA probe hybridized strongly to 1.6- and 2.2-kb transcripts in K562, the line from which K8 originated, and also revealed transcripts of similar size in mouse 3T3 fibroblasts. However, we have detected no more than trace hybridization in any of the other 28 lines examined, including the two mouse erythroid lines, which are similar to K562 in some respects. Perhaps the relatively abundant expression of K8 in K562 is related to some unique feature of that line (see Discussion).

Up to 20 homeobox transcripts in haemopoietic compartment

To gauge the spectrum of homeobox transcripts within the haemopoietic system, we performed low-stringency hybridization with either the oligonucleotide mixture used for screening or homeobox DNA probes. The blot with the oligonucleotides in Figure 4D shows that nearly every cell line displays unique species, although 3T3 fibroblasts yield the most complex pattern. We think that most are authentic homeobox transcripts, because our cDNA screening to date has yielded only one false positive out of the dozens of cDNA clones. An even more complex pattern was observed at low stringency with the S1 (*Hox 2.3*) probe, which is very similar to the consensus for the large *Antp* class, and multiple transcripts were also detected in every line with the *Hox 1.1* (B2) probe (data not shown). As all these patterns differed to some extent, we consider it likely that up to 20 different homeobox transcripts are expressed within the haemopoietic compartment.

Altered expression on terminal differentiation

To determine whether the pattern of expression changes as cells differentiate, we examined WEHI3B D^+ cells induced to differentiate to granulocytes and macrophages, and Friend erythroleukemia cells induced to erythrocyte differentiation; maturation of both was nearly quantitative (see Materials and methods). Since *Hox 2.3* is expressed primarily in macrophage lines (Table I), we examined differentiated WEHI3B D^+ cells for the 1.6- and 1.2-kb *Hox 2.3* transcripts, but only the 2.0- and 2.5-kb *Hox 2.4* species were found (data not shown). Perhaps the DNA alteration near *Hox 2.4* in WEHI3B (Figure 5) in some way prevents *Hox 2.3* expression. A change was, nevertheless, observed in expression of *Hox 1.1*-related transcripts: a marked drop in the level of the unidentified 1.4-kb species, and smaller decreased for the *Hox 1.1* 2.0-kb mRNA (Figure 6A). With erythroid differentiation, the *Hox 1.1*-related transcripts fell at least several-fold in abundance (Figure 6B), while *Hox 6.1* transcripts increased ~3-fold, and *Hox 2.3*, S8 and K8 remained silent (data not shown).

Altered expression in cells that have switched from the lymphoid to myeloid lineage

A recent unexpected finding is that cloned B lymphoid cell lines bearing certain oncogenes throw off variants with a complete macrophage phenotype (Klinken *et al.*, 1988). To evaluate whether homeobox expression is linked to lineage determination, we compared several such lymphoid lines and their macrophage derivatives. Figure 7A shows that two such

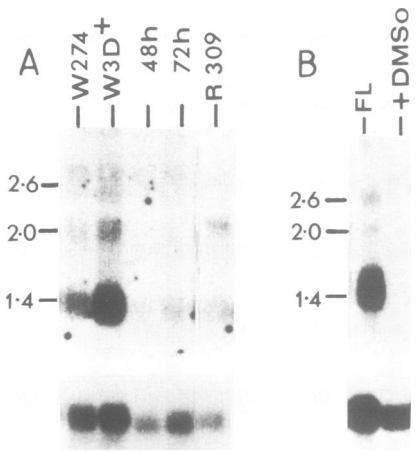

Fig. 6. Altered pattern of homeobox expression on terminal differentiation. (**A**) Blot of poly(A)$^+$ RNA from WEHI3B D$^+$ cells, either untreated or treated 48 or 72 h with G-CSF and actinomycin D (see Materials and methods), hybridized with the *Hox 1.1* probe, or a control 'housekeeping' gene, gapdh. (**B**) Blots of poly(A)$^+$ RNA from FL-F4N erythroleukaemia cells untreated or treated with dimethylsulphoxide for 72 h, hybridized with the *Hox 1.1* probe or the gapdh probe. The gapdh results suggest that less mRNA was loaded in the treated samples.

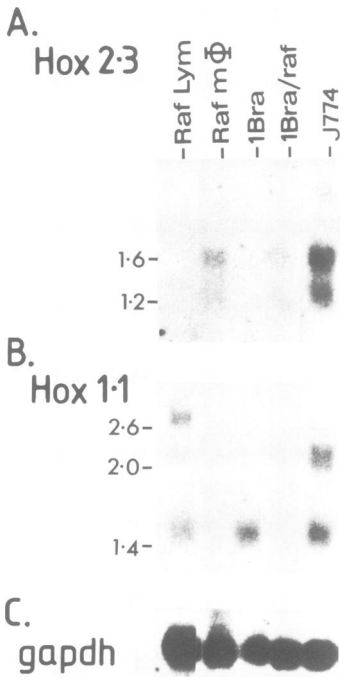

Fig. 7. Altered pattern of homeobox expression in B lymphoid lines that have converted to the myeloid lineage. Raf-Lym, and 1Bra are pre-B lymphoma lines, whereas Raf-Mϕ and 1Bra/raf are their macrophage derivatives (Klinken *et al.*, 1988). The blots were hybridized with (**A**) the *Hox 2.3* (S1) probe and (**B**) the 300-bp *Hox 1.1* probe. The gapdh hybridization is shown at the bottom.

macrophage derivatives, Raf-Mϕ and 1Bra/raf, exhibited the 1.2- and 1.6-kb *Hox 2.3* transcripts, albeit much less than in the conventional macrophage line J774, while none was detected in their parental pre-B lymphoid lines, Raf-Lym and 1Bra. Conversely, whereas the parental lines expressed *Hox 1.1*-related transcripts, their level diminished in the macrophage derivatives (Figure 7B). These results, also found with four other pairs of these unusual lines, strongly suggest that expression of certain homeobox genes in haematopoietic cells is coupled to cell phenotype rather than merely being an accident of the history of particular lines.

Discussion

Our results indicate that a number of homeobox genes are expressed within the haematopoietic system. The cDNA clones isolated so far derive from the *Hox 1.1*, *Hox 2.3* and *Hox 6.1* genes and the previously undescribed human gene K8. Thus homeobox genes from at least three chromosomes are expressed in haemopoeitic cells. In addition, the S8 cDNA clone isolated from spleen revealed that the *paired* class of genes exists in vertebrates, although S8 itself probably derives from non-haematopoietic stromal cells. From the number of poly(A)$^+$ transcripts observed in RNA from various haematopoietic cell lines (Figure 4), particularly with low stringency hybridization (e.g. Figure 4D), we estimate that the haemopoietic compartment as a whole contains up to 20 homeobox transcripts. A diverse set of homeobox gene products are thus available as candidates for the regulators of haematopoietic differentiation.

The patterns of expression revealed by different probes varied dramatically (Table I). The *Hox 1.1* probe revealed almost ubiquitous transcripts, whereas *Hox 2.3* expression was restricted by lineage and stage, K8 mRNA was confined to a single haemopoietic cell line and *Hox 6.1* was expressed sporadically among the lines. These diverse patterns indicate that most of these genes are under independent control and may reflect different types of functions for homeobox genes. Extending this analysis to the equivalent normal cells would be desirable but will be complicated by the problems of cell purity, particularly given the very low abundance of the transcripts found in the cell lines.

Within the haematopoietic system, *Hox 2.3* may have a role in cell type determination, as its expression was essentially confined to macrophage lines (Table I). Since earlier cell types within the myeloid lineage lack the *Hox 2.3* transcripts, their expression seems to be turned on late in macrophage development, perhaps because they participate in that maturation step. Consistent with that notion, the unusual lines that have switched from a lymphoid to a macrophage phenotype have acquired the *Hox 2.3* transcripts (Figure 7A). On the other hand, they were not detected in one experiment with normal macrophages (unpublished results), nor in differentiated WEHI3B cells, which predominantly comprise monocytes and macrophages. The WEHI3B result might reflect the disturbed homeobox expression in that cell line (Figures 4A and 5), or the fact that they stop dividing as they differentiate (Cooper *et al.*, 1982), whereas the macrophage cell lines proliferate rapidly.

The other homeobox genes studied may have roles other than cell type determination. *Hox 1.1*-related transcripts were detectable in diverse cell types (Figure 4C and Table I), but their level dropped in terminally differentiated WEHI3B and erythroleukaemia cells (Figure 6). Since both lines cease dividing concomitantly with maturation, these genes may be associated with proliferation, as is *Hox 1.3* in cultured fibroblasts, though not in neurons (Odenwald *et al.*, 1987). The restriction of K8 expression within the haematopoietic system to the complex K562 line may reflect one of its unique features, such as its residual stem cell character, its ability

to generate megakaryocytes and lymphocytes (Tetteroo *et al.*, 1984; Fraser and Berridge, 1987), or the presence of the translocated *abl* oncogene (Shtivelman *et al.*, 1985). Although the presence of *Hox 6.1* transcripts did not correlate with cell type (Table I), it remains conceivable that more than one combination of homeobox gene products can lead to the same cell phenotype. Alternatively, the pattern of expression may have been altered in some lines during leukaemic transformation.

Our results add to the evidence (Schofield, 1987) that nearly every mouse or human homeobox gene is expressed in some adult tissue(s) as well as the embryo. Adult expression of some (e.g. *Hox 1.4* and *1.6*) is highly restricted (Duboule *et al.*, 1986; Baron *et al.*, 1987). A more complex pattern, often differing from that in the embryo, is more typical; e.g. *Hox 1.1* and *Hox 1.2* (Colberg-Poley *et al.*, 1985b), *Hox 1.3* (Odenwald *et al.*, 1987), *Hox 2.1* (Krumlauf *et al.*, 1987) and *Hox 2.3* (Meijlink *et al.*, 1987). Although it is not yet established whether the adult expression is confined to specific cell types within a tissue, other evidence with cloned lines has appeared recently. During differentiation of a human colon cell line, the ratio of the two *Hox 2.3* transcripts altered and the level of c8 (*Hox 6.1*) mRNA increased (Sebastio *et al.*, 1987). Moreover, during erythroleukemia cell differentation, a novel transcript of the presumptive *Hox 2.6* gene appeared (Lonai *et al.*, 1987). Taken together, the range of expressed sequences and tissues involved strongly suggest that these genes are regulated in a complex fashion and may mean that many participate in adult differentiation processes.

Because most neoplasms represent a clone with a reduced tendency to undergo terminal differentiation, a somatic mutation within a regulatory gene controlling that step could block programmed cell death and thereby promote neoplasia. The altered homeobox gene in the myeloid leukaemia WEHI3B (Figure 5 and Blatt *et al.*, 1987) may indicate that homeobox genes can participate in the development of neoplasia.

Materials and methods

Oligonucleotide probes and cDNA library screening

A collation of 32 known homeobox sequences of the *Antennapedia* and *engrailed* classes revealed that 28 of the 36 nucleotides encoding residues 44–55 were almost invariant. Hence the complement of CAGATCAAGATCTGGTTCCAGAAC$^{C}_{A}$GG$^{C}_{A}$G$^{C}_{G}$ATGAAG provided a probe which matches most individual box sequences at 25 or more residues, with runs of at least 11 residues. Because *paired* (Boop *et al.*, 1986), *bicoid* (Frigerio *et al.*, 1986) and *even-skipped (Macdonald et al.*, 1986) are significantly more divergent in this region, we included equivalent 36mers from them.

Mouse cDNA libraries from spleen and bone marrow at λgt11 were purchased from Clonetech. A K562 cDNA library in λgt10 was provided by Iswar Hariharan. Phage were plated on either Y1090 (for λgt11) or NM514 (for λgt10) at 10^4–10^5 plaques/120-mm plate. Phage DNA was transferred to duplicate nitrocellulose filters (Schleicher and Schuell) and hybridized at low stringennecy with a mixture of four ^{32}P end-labeled homeobox oligonucleotides (see above). Hybridization was carried out in 50% formamide, 5 × SSPE (1 × SSPE = 10 mM sodium phosphate, 0.15 M NaCl, 1 mM EDTA, pH 7.4), 5 × Denhardt's solution 0.1% SDS and 250 μg/ml denatured sheared herring sperm DNA at 30°C. Filters were washed in 2 × SSC, 0.1% SDS at 50°C (1 × SSC = 0.15 M NaCl, 0.015 M sodium citrate).

Cell lines and cell culture

The cell lines not cited in a previous study (Kemp *et al.*, 1980) are referenced here or in Results. The murine myeloid cell lines included: B6SutA (Greenberger *et al.*, 1983); 416B, a bipotential cell line (Dexter *et al.*, 1979); FDC-P1, an IL3-dependent progenitor cell (Dexter *et al.*, 1980); WEHI3B D^+ and D^- (Cooper *et al.*, 1982); and WEHI265 and WEHI274. Macrophages lines were WR19, RAW264 and RAW309 and J774 (Ralph *et al.*, 1983). Erythroid lines included an erythropoietin-responsive erythroblast line (J2E) established by Dr P.Klinken and the Friend erythroleukemia line FL-F4V. B lymphoid lines included the pro-B line LyH7, kindly provided by Dr R.Palacios; conventional pre-B lines ABLS8 and 18-48 (Siden *et al.*, 1979) and Avrij-1 and B lymphomas (WEHI231, WEHI279); pre-B and B cell lines from Eμ-*myc* transgenic mice (1Bra, 2Mes, 3Mes, 5Thy, 13BM; Adams *et al.*, 1985); plasmacytomas ABPC4 and ABPC103 (from Dr M.Potter), MPC11 (Laskov and Scharff, 1970) and P3. T lymphomas included Yac-1 (Klein and Klein, 1964), ST4, EL4, Yac-1 and WR19L (Raschke *et al.*, 1978) and WEHI701. RNA from the lymphoid lines that have switched to macrophages (Klinken *et al.*, 1988) as kindly provided by P.Klinken and W.Alexander.

Murine cell lines were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum, and human cells in RPMI 1640 plus 10% fetal calf serum. For IL3-dependent cell lines, the factor was provided by 10% cell supernatant from WEHI3B cultures. Differentation of WEHI3B D^+ cells was induced by the procedure of Cooper *et al.* (1982), as modified by Dr Ralph Boehmer. Cells growing at 5 × 10^5 ml were treated for 6 h with 15 ng/ml actinomycin D and 1 ng/ml of recombinant human G-CSF (Amgen) and then diluted with 2 vol of medium plus serum containing 1 ng/ml G-CSF and actinomycin D at 7 ng/ml (final). Cultures (240 ml) were harvested after 48 h and 72 h. Cytocentrifuge preparations stained with May–Grünwald/Giemsa confirmed that all the cells had differentiated, with cells of the moncytic series predominating over granulocytes (Cooper *et al.*, 1982). The Friend erythroleukaemia cells (FL-F4N) were differentiated by treatment with 1% dimethylsulphoxide (Dube *et al.*, 1975). After 48 h benzadine treatment (performed by Dr P.Klinken) confirmed that >85% of the cells synthesized haemoglobin.

RNA isolation and Northern blot hybridization

Total cellular poly(A)$^+$ RNA was isolated from cells by a proteinase K method, followed by oligo(dT)–cellulose chromatography. RNA samples (4 μg/lane) were electrophoresed on a 2.2 M formaldehyde/1% agarose gel in 1 × MOPS buffer at 150 V for 4 h (1 × MOPS = 20 mM MOPS, 1 mM EDTA, 5 mM sodium acetate). After ethidium bromide staining, the gel was photographed under UV illumination to check the amounts and integrity of the mRNA, and the mobilities of RNA size markers (BRL). The RNA was then transferred to nitrocellulose (Thomas, 1980). The filters were hybridized overnight at high stringency (50% formamide, 5 × SSC at 42°C) with ^{32}P-labeled DNA probes, then washed twice in 2 × SSC, 0.1% SDS at 65°C for 2 h and in 0.2 × SSC, 0.1% SDS at 65°C for 30 min. Low-stringency hybridization and washing were similar to the phage screening (see above). The quantity and integrity of the mRNA loaded on most filters was subsequently checked by hybridization with a 'housekeeping' gene such as actin, c-*myc* or, most often, glycerol 3-phosphate dehydrogenase (gapdh).

Sequence analysis

For some clones, the presence of a homeobox was first verified by direct sequencing on double-stranded plasmid DNA using an *Antp* homeobox oligonucleotide primer. Subsequently, the cDNA inserts or fragments thereof were subcloned in both orientations into M13 mp8 or mp9, and sequenced by chain termination (Sanger *et al.*, 1977) using the M13 universal primer or appropriate specific primers.

Acknowledgements

We are grateful to Gina Barri and Linda Byrt for excellent tecnical assistance; to Drs P.Klinken, W.Alexander, I.Hariharan and J.Hamilton for materials; to Drs S.Cory, D.Metcalf, R.Saint, A.Harris and R.Boehmer for discussions; and to Drs M.Kessel, P.Gruss, D.Duboule, F.Meijlink and F.Ruddle for clones. This research was supported by the National Health and Medical Research Council (Canberra), the US National Cancer Institute (R01-CA12421) and the American Heart Association. K.K. holds a National Research Fellowship.

References

Acampura,D., Pannese,M., D'Esposito,M., Simeone,A. and Boncinelli,E. (1987) *Human Reprod.*, **2**, 407–414.

Adams,J.M., Harris,A.W., Pinkert,C.A., Corcoran,L.M., Alexander,W.S., Corry,S., Palmiter,R.D. and Brinster,R.L. (1985) *Nature*, **318**, 533–538.

Baron,A., Featherstone,M.S., Hill,R.E., Hall,A., Galliot,B. and Duboule,D. (1987) *EMBO J.*, **6**, 2977–2986.
Blatt,C., Goldberg,R., Aberdam,D. and Sachs,L. (1987) Third Annual Oncogene Meeting, Abstract 364.
Boyd,A.W. and Schrader,J.W. (1982) *Nature*, **297**, 691–694.
Bopp,D., Burri,M., Baumgartner,S., Frigerio,G. and Noll,M. (1986) *Cell*, **47**, 1033–1040.
Colberg-Poley,A.N., Voss,S.D., Chowdhury,K. and Gruss,P. (1985a) *Nature*, **314**, 713–718.
Colberg-Poley,A.M., Voss,S.D., Chowdury,K., Stewart,C.L., Wagner,E.F. and Gruss,P. (1985b) *Cell*, **43**, 39–45.
Collins,S.J., Ruscotti,F., Gallagher,R. and Gallo,R. (1978) *Proc. Natl. Acad. Sci, USA*, **75**, 2458–2463.
Cooper,P.C., Metcalf,D. and Burgess,A.W. (1982) *Leukemia Res.*, **6**, 313–327.
Davis,R.L., Weintraub,H. and Lasser,A.B. (1987) *Cell*, **51**, 987–1000.
Desplan,C., Theis,J. and O'Farrell,P.H. (1985) *Nature*, **318**, 19–26.
Dexter,T.M. and Spooncer,E. (1987) *Annu. Rev. Cell. Biol.*, **3**, 423–441.
Dexter,T.M., Allen,T.D., Scott,D. and Teich,N.M. (1979) *Nature*, **277**, 471–474.
Dexter,T.M., Garland,J., Scott,D., Scolnick,E. and Metcalf,D. (1980) *J. Exp. Med.*, **152**, 1036–1047.
Dube,S.K., Pragnell,I.B., Kluge,N., Gaedick,G., Steinheider,G. and Ostertag,W. (1975) *Proc. Natl. Acad. Sci. USA*, **72**, 1863–1867.
Duboule,D., Baron,A., Mähl,P. and Galliot,B. (1986) *EMBO J.*, **5**, 1973–1980.
Fainsod,A., Bogarad,L.D., Ruusala,T., Lubin,M., Crothers,D.M. and Ruddle,F.H. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 9532–9536.
Fraser,J. and Berridge,M. (1987) *Exp. Haematol.*, **15**, 406–413.
Frigerio,G., Burri,M., Bopp,D., Baumartner,S. and Noll,M. (1986) *Cell*, **47**, 735–746.
Gehring,W.J. (1987) *Science*, **236**, 1245–1252.
Greenberger,J.S., Sakakeeny,M.A., Humphries,R.K., Eaves,C.J. and Eckner,R.J. (1983) *Proc. Natl Res. Sci.*, **80**, 2931–2935.
Hart,C.P., Awgulewitsch,A., Fainsod,A., McGinnis,W. and Ruddle,F.H. (1985) *Cell*, **43**, 9–18.
Hart,C.P., Fainsod,A. and Ruddle,R. (1987) *Genomics*, **1**, 182–195.
Hozumi,M. (1985) *CRC Crit. Rev. Oncol. Haematol.*, **3**, 235–277.
Kemp,D.J., Harris,A.W., Cory,S. and Adams,J.M. (1980) *Proc. Natl Acad. Sci. USA*, **77**, 2876–2880.
Kessel,M., Schulze,F., Fibi,M. and Gruss,P. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 5306–5310.
Klein,E. and Klein,G. (1964) *J. Natl. Cancer Inst.*, **32**, 547–568.
Klinken,S.P., Alexander,W.S. and Adams,J.M. (1988) *Cell*, in press.
Krumlauf,R., Holland,P., McVey,J. and Hogan,B. (1987) *Development*, **99**, 603–617.
Laskov,R. and Scharff,M.D. (1970) *J. Exp. Med.*, **13**, 515–541.
Li,C.L. and Johnson,G.R. (1985) *Nature*, **316**, 633–636.
Lonai,P., Arman,E., Czosnek,H., Ruddle,F.H. and Blatt,C. (1987) *DNA*, **6**, 409–418.
Lozzio,C. and Lozzio,B. (1975) *Blood*, **45**, 321–334.
Macdonald,P.M., Ingham,P. and Struhl,G. (1986) *Cell*, **4**, 721–734.
Martin,G.R., Boncinelli,E., Duboule,D., Gruss,P., Jackson,I., Krumlauf,R., Lonai,P., McGinnis,W., Ruddle,F. and Wolgemuth,D. (1987) *Nature*, **325**, 21–22.
McGinnis,W., Garber,R.L., Wirz,J., Kuroiwa,A. and Gehring,W.J. (1984) *Cell*, **37**, 403–408.
Meijlink,F., Laat,R., Verrijzer,P., Destree,O., Kroezen,V., Hilkens,J. and Deschamps,J. (1987) *Nucleic Acids Res.*, **15**, 6773–6786.
Odenwald,W.F., Taylor,C.F., Palmer-Hill,F.J., Friedrich,F.Jr, Tani,M. and Lazzarini,R.A. (1987) *Genes Dev.*, **1**, 482–496.
Ralph,P., Ho,M.-K., Litcofsky,P.B. and Springer,T.A. (1983) *J. Immunol.*, **130**, 108–114.
Rabin,M., Ferguson-Smith,A., Hart,C.P. and Ruddle,F.H. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 9104–9108.
Raschke,W.C., Baird,S., Ralph,P. and Nakoinz,I. (1978) *Cell*, **29**, 615–618.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
Schofield,P.N. (1987) *Trends Nuerosci.*, **10**, 3–6.
Sebastio,G., D'Esposito,M., Montanucci,M., Simeone,A., Auricchio,S. and Boncinelli,E. (1987) *Biochem. Biophys. Res. Commun.*, **146**, 751–756.
Sharpe,P.T., Miller,J.R., Evans,E.P., Burtenshaw,M.D. and Gaunt,S.J. (1988) *Development*, **102**, 397–407.
Shtivelman,E., Lifshitz,B., Gale,R.P. and Canaani,E. (1985) *Nature*, **315**, 550–554.
Siden,E., Baltimore,D., Clark,D. and Rosenberg,N.E. (1979) *Cell*, **16**, 389–396.
Simeone,A., Mavilio,F., Acampora,D., Giampaolo,A., Faiella,A., Zappavigna,V., D'Esposito,M., Pannese,M., Russo,G., Boncinelli,E. and Peschle,C. (1987) *Proc. Natl. Acad. Sci. USA*, **34**, 4914–4918.
Snow,M.H.L. (1986) *Nature*, **324**, 618–619.
Tetteroo,P., Massaro,A., Mudler,R., Schreuder-van-Gelder,R. and von dem Borne,A. (1984) *Leukemia Res.*, **8**, 197–206.
Thomas,P. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 5201–5205.

Received on April 8, 1988